



Bounded rationality: from fast and frugal heuristics to logic programming and back

Francisco Vargas¹ · Laura Martignon¹ · Keith Stenning²

Received: 13 October 2020 / Accepted: 17 February 2023 / Published online: 15 March 2023
© The Author(s) 2023

Abstract

The notion of “bounded rationality” was introduced by Simon as an appropriate framework for explaining how agents reason and make decisions in accordance with their computational limitations and the characteristics of the environments in which they exist (seen metaphorically as two complementary scissor blades). We elaborate on how bounded rationality is usually conceived in psychology and on its relationship with logic. We focus on the relationship between heuristics and some non-monotonic logical systems. These two categories of cognitive tools share fundamental features. As a step further, we show that in some cases heuristics themselves can be formalized from this logic perspective. We have therefore two main aims: on the one hand, to demonstrate the relationship between the bounded rationality programme and logic, understood in a broad sense; on the other hand, to provide logical tools of analysis of already known heuristics. This may lead to results such as the characterization of fast and frugal binary trees in terms of their associated logic program here provided.

Keywords Non-monotonic reasoning · Bounded rationality · Ecological rationality · Logic programming · Heuristics

1 Introduction

“Multiple logics may become necessary when approaching heuristic decision making”

(Simon 1969).

✉ Francisco Vargas
fvargasmancera@gmail.com

¹ Ludwigsburg University of Education, Ludwigsburg, Germany

² Edinburgh University, Edinburgh, UK

The Enlightenment's faith in reason and in the solvability of scientific problems, at least in principle, is epitomized in the figure of Laplace, an apostle of mechanistic determinism in physics and of the confident identification of probability theory with intuitive *bon sens*. This derives from postulating the possibility of ideal reasoners and complete (or completable) knowledge, something radically challenged by the subsequent developments of science at the turn of the nineteenth century and confined to fields such as pure mathematics. In logic and mathematics there were Frege's search for ultimate foundations in logic and Hilbert's program. Even there though, the discovery of incompleteness and undecidability phenomena put inherent limitations on what was expected in terms of absolute foundations and certainty. We see furthermore, from Turing and Church onwards, a precise mathematical treatment of computational constraints and a treatment of combinatorial feasibility/infeasibility of problems.

In social science research, the optimality of idealized rationality continues to be assumed, even if human beings cannot reach that standard. When approaching human rationality, in fact these optimal ideal standards continue to be used as the main or unique norms in regard of which performance should be evaluated. This occurs in logic and probabilistic reasoning (divergences from classical logic or from probability theory are often regarded as “fallacies” or “biases” proving our irrationality) as well as in decision theory (divergences from optimality theory are regarded, at best, as imperfect approximations). From such perspectives human behaviour continues to be judged in absolute terms regarding unique normative theories: classical logic, probability theory and rational choice theory.

The bounded rationality paradigm, as proposed by Herbert Simon, constitutes an alternative conception where rationality (or irrationality) is not approached from a preconceived optimal, general, or ideal reasoner. It proposes a more realistic account, as the result of integrating both human faculties and the nature of the problems faced: “The capacity of the human mind for formulating and solving complex problems is very small compared with the size of the problems whose solution is required for objectively rational behavior in the real world or even for a reasonable approximation to such objective rationality” (Simon 1957, p. 198).¹

“Boundedness” concerns limits in availability of information, unlike “global rationality” (the normative theories considered above) that requires knowledge of all possible behavior alternatives and anticipation about all future consequences that will follow each of them. Idealized, unbounded or global rationality abstracts from concerns such as perceptual inputs, memory and processing capacity, or plain unavailability of information. These concerns are inherent to “flesh-and-blood humans” (and, as considered in AI, are relevant to any reasoner).

In the present paper we elaborate on how bounded rationality is usually conceived in psychology and what its relations with logic (especially with some non-classical logics) can be.

¹ Simon's primary concerns are in the field of administration and economics: “Individual choice takes place in an environment of ‘givens’ — premises that are accepted by the subject as bases for his choice; and behavior is adaptive only within the limits set by these ‘givens’” (Simon 1947, p. 92).

The program of bounded rationality has been developed in cognitive psychology along more than one line. The school around Kahneman and Tversky (Tversky and Kahneman 1974) viewed bounded rationality as limitations of the mind. They analysed cognitive tasks for which the mind uses heuristics, but these appear to be maladaptive, in the sense that they lead to mistakes, at least from the perspective of normative frameworks. The school around Gigerenzer and the ABC Group has described a category of heuristics, which result from the integration of mind and environment. These are the so-called fast and frugal heuristics (FFH for short, see Gigerenzer, Todd and the ABC Group 1999). These heuristics are thus conceived as ecologically rational, in the sense of their adaptation to the environment surrounding the agent's mind. The bounded rationality paradigm, in this version, postulates the existence of different heuristics applied in accordance with different environmental prompts and constraints. These heuristics are usually conceived in contraposition to more traditional accounts, logical and probabilistic: "We do not compare human judgment with the laws of logic or probability, but rather examine how it fares in real-world environments" (Gigerenzer et al. 1999, p. 22). Unlike logic, in this view, "The function of heuristics is not to be coherent. Rather, their function is to make reasonable, adaptive inferences about the real social and physical world given limited time and knowledge." In our view, this contraposition of heuristics to "logic" is highly influenced by long established traditions in psychology in relation to the latter: above all, the identification of logic with classical logic. From this position, it is fair to share Wundt's judgement² about the descriptive and normative uses of logic: "We can in fact say of such attempts, that measured by the results they have been absolutely fruitless. They have disregarded the psychological processes themselves" (cited by Gigerenzer 2008).

Modern developments of logic, however, have led to a complete diversification and reconfiguration of its aims. This includes strong interactions and applications to AI, linguistics, and psychology.³ Logic, in this wider sense, has developed a variety of concepts and tools for the analysis of information processing in contexts of incomplete information and constraints. We focus on some of these tools, particularly in logic programming (LP). In our view, LP shares the main properties usually ascribed to heuristics, and therefore should be seen beside them within the bounded rationality program. It is possible also to argue in favour of the stronger idea that LP would be a convenient framework to develop important heuristics postulated in the context of Bounded Rationality, as we will examine for some cases. This goes in the direction of (Stenning and van Lambalgen 2008) in showing that rich empirical results flow from LP formalisation of a central discourse function of 'narrative' (in a very broad sense). Thus, 'narrative logic' provides a crucial high-level account of what is involved when people construct the necessary new interpretations that found communications. And (Pijnacker et al. 2011) shows that some of the electrophysiology of the knowledgebase cast in LP is a realistic processing interpretation.

² Valid from the perspective of what was logic in 1912, but not anymore today.

³ We can see that, in addition to early work in the 70's and 80's (mainly in philosophy and AI), "modern logic is undergoing a cognitive turn" (van Benthem 2008).

This then contrasts with classical logic. Participants are rather good at classical logical countermodelling to settle disputes about interpretations (Vargas and Stenning 2020). FFHs generally fit best in the former: they require interpretations that then allow their methods to be applied.

We organize our exposition according to the following structure: In Sect. 2 we develop some concepts from LP, the main technical tool that we shall use; we propose an analysis of the concept of Bounded Rationality, which leads us to examine the relationships between heuristics and LP; we show that LP shares the main “virtues” of heuristics for the analysis of human behaviour (Sect. 3) and that some heuristics may even be approached using LP (Sect. 3.2). We make explicit a characterization of binary categorization trees in terms of his associated logic program (Result 1 below).

We conclude from this that heuristics and “logic” (in the sense of multiple logics) need not and should not be seen as excluding approaches in the bounded rationality program.

2 Logic programming and human reasoning

In this section we overview some of the properties, which will be relevant for the aims of this paper, of LP, a family of non-monotonic logics particularly adapted for those aims. A technical treatment of LP as a non-monotonic logic may be found in Doets (1994) whereas applications for the psychology of reasoning are developed in (Stenning and van Lambalgen 2008), (Kowalski 2011) and (Hölldobler 2015).

We see LP as a fundamental processing apparatus applicable across domains. Before implementing this, it is worth saying that we are not proposing to *substitute* in general the use of classical logic by the use of LP or any other logical system, not even for the specific aims of cognitive psychology. We are committed to the general tenets of logical pluralism, intended as “the view that there is more than one genuine deductive consequence relation, and that this plurality arises not merely because there are different languages, but rather arises even *within* the kinds of claims expressed in the one language” (Beall and Restall 2006, p.3). This leads, of course, to the demarcation problem, which we will not approach here, on what counts as a deductive relation, and therefore as a logic.⁴

Classical logic has its specific aims and properties and is suitable for some crucial aspects of scientific and mathematical reasoning. But even in this last case it is disputable whether, say, first order classical logic, is the only, all-encompassing logical approach, as may be seen by the use of intuitionistic logic (which captures the notion of constructible mathematics), second order logic (Shapiro 1991), or the variety of model-theoretic logics (Barwise and Feferman 2017), just to mention some

⁴ We will only mention here, as a radically different alternative to the classical case, that deduction for a wide spectrum of non-monotonic logics may be defined in terms of the “preferred models” approach, as introduced by Shoham (1987).

well-known examples. We claim that use of different logics should also be crucial in modelling human reasoning as a complex, many-faceted phenomenon.

We see LP in particular as a family of logics at the core of planning and language processing (temporal/causal relations in narrative discourse as in (van Lambalgen and Hamm 2005)). In this sense it is fundamental to many cognitive functions underlying processes of interpretation in reasoning, decision-making and problem-solving. Born from AI and computer science (including actual programming languages such as Prolog and Constraint Logic Programming), from its very origin it was concerned with computability (Kowalski 2014). We will examine next this and other fundamental properties of LP.

2.1 Syntax and computability

The search for logical systems in AI (specifically in the field of knowledge representation and reasoning) is dependent on the trade-off between expressiveness and tractability (Brachman and Levesque 2004). This motivates the syntax used in LP both in the propositional and the predicate cases. Formulas in LP are restricted to clauses (conditional statements with the antecedent being a conjunction of literals and the consequent a single positive literal). They are therefore of the form $(\neg)p_1 \wedge \dots \wedge (\neg)p_n \rightarrow q$ where p_1, \dots, p_n, q are atomic (positive) formulas, and $(\neg)p$ means that p may appear either negated or not. A set of clauses of this kind is called a definite logic program. In the predicate case, this formula is intended to be universally quantified over all the occurring variables.

This kind of sentence (clause) encapsulates the notion of “rules” of the form “if conditions then conclusion”. This syntactic restriction implies that we do not have (as in classical logic) full recursivity (we do not admit, for instance, formulas with nested implications). This decreases substantially computational complexity in comparison with classical logic. Even so, LP is still expressive enough to cover significant fragments of human reasoning and language processing, especially if considering the predicate and not only the propositional versions of LP, see (Kowalski 2011), (Stenning and van Lambalgen 2008) and (van Lambalgen and Hamm 2005).

The computational efficiency derives from the application of the resolution method to program clauses for deciding the problem of satisfiability of formulas. The syntactic restriction allows very efficient proof-search: derivations in logic programming are simultaneously constructive computations, in the sense that if one concludes that a formula $\varphi(x)$ is satisfiable given a logic program, the derivation actually produces a computable witness for x . This is done by applying the resolution method, which we will not describe here, to queries. A *query* or goal is a conjunction of atomic formulas indicated as $?p_1 \wedge \dots \wedge p_n$, which is posed to a program in order to know if it is derivable from it.

Some authors consider the syntactic restriction to formulas in clausal form not only to be convenient for dealing with tractable fragments of predicate logic, but also because of its presumed closeness to how information is represented in human minds (Kowalski 2011).

2.2 Closed world reasoning

Databases or “programs” codify the available information. From these, it is possible to extract additional information in a way which substantially differs from classical logic. Particularly important is what is known as “closed world reasoning” (CWR), which manages the problem of overcoming the information incompleteness of a program. Technically, CWR operates through the *completion* of a program. If the program is taken as if complete, the set of all the program clauses with the same consequent is a disjunctive definition of this consequent. In this sense, implications are “biconditionals in disguise” (Kowalski 2011). In order to define the completion of a program P , for every q occurring in P , all clauses $\varphi_i \rightarrow q$ in P (with q as a consequent or “head”) are taken in order to form the expression $\bigvee_i \varphi_i \rightarrow q$. In case that there is no such φ_i , then this reduces to the expression $\perp \rightarrow q$. Finally, all the occurrences of the \rightarrow ’s are replaced by \leftrightarrow ’s (here, \leftrightarrow has a classical interpretation given by: $\psi \leftrightarrow \varphi$ is true if ψ, φ have the same truth-value, and false otherwise).⁵ From this completion, inferences can be extracted which were not classically derivable from P . In particular, negation as failure (NAF) consists of assuming that if a statement cannot be proved (i.e., if there is a *failure* of the derivational procedure), then its *negation* must be the case.

We can see CWR as *assuming* that the ontological level (“the world”) is captured by the epistemic level. This entails the risk of a possibly mistaken evaluation but the advantage of at least coming to evaluations and decisions that are often required under information and time constraints. But the epistemic level may change and CWR works by assuming that during each information update, things that follow from information that has been added by the last addition are “swept”.

2.3 Dealing with abnormalities

The fact that most actual conditionals or rules in real life presuppose an unmanageable series of conditions or exceptions in order to be valid lies at the basis of different kinds of default reasoning, starting with circumscription (McCarthy 1986).

Unlike material implication, regularities in the world are exception-tolerant. This may be captured by an abnormality statement which is assumed to be false (by NAF), unless there is positive information supporting that there is something abnormal preventing the inference to the conclusion. The conditional that allows exceptions is of the form $p \wedge \neg ab \rightarrow q$, to be read as ‘if p and nothing abnormal is the case, then q '.⁶

⁵ Intuitively, this says that the whole information about a statement q is that provided by the clauses in the program P , so q is defined entirely by this information (the disjunction mentioned is in fact called the *definition of q*). The world is “closed” in the sense that nothing outside P is potentially considerable as part of the definition of q . P does not only include truths about q , but represents the *whole truth* about it.

⁶ The “abnormality clause” ab , is a list of propositions (or atomic formulas) treated as an inclusive disjunction of abnormalities, indexed to a particular conditional: If any one of them is true, then ab is true, and so the antecedent of the conditional ($p \wedge ab$) is false.

2.4 Non-monotonicity

CWR is applied in practice as a non-monotonic kind of reasoning: classical logic is said to have a *monotonic* consequence relation, meaning that if $\varphi_1, \dots, \varphi_n / \psi$ is a valid argument, then also the argument $\varphi_1, \dots, \varphi_n, \theta / \psi$ holds. That is, enlarging the premise set can never lead to the withdrawal of an inferred conclusion. This follows from the definition of validity in classical logic: an argument is valid if the conclusion is true in all situations in which the premises are true. A situation in which $\varphi_1, \dots, \varphi_n, \theta$ are true is also a situation in which $\varphi_1, \dots, \varphi_n$ are true, so that ψ still follows. By contrast, a logic is *non-monotonic* if it is possible that $\varphi_1, \dots, \varphi_n / \psi$ is valid, but $\varphi_1, \dots, \varphi_n, \theta / \psi$ is not. For instance, it may be assumed by default that nothing abnormal is the case and that, given p , from a conditional $p \wedge \neg ab \rightarrow q$ we may infer q . But a new fact (an abnormality) may be added into the program which makes ab become true. In this case q may no longer be inferred.

In brief, classical monotonic reasoning adopts a “hindsight perspective” which works when all the information is “already in”, but which makes it not suitable for situations of incomplete information and under constraints.

2.5 Non truth-functionality

An important difference between LP semantics and classical logic is that it is not truth functional: truth values assigned to atomic formulas do not determine in general the assignments of a truth value to a complex formula formed with them. This is the case for conditional statements which are not seen as formulas with truth values in themselves, but more as “licenses for inference” in logic programs — a sort of contentful inference rules. Take for instance the case when the antecedent of a conditional is known to be false. According to classical logic this makes the conditional true, but in LP it simply deactivates inferences from the conditional.

Besides this, for the purposes of modeling cognitive phenomena in situations of incomplete information, it is convenient to go beyond classical bivalence in truth values. This can be done in a parsimonious way using the Kleene semantics (Kleene 1952, p.332) as in (Stenning and van Lambalgen 2008). In this framework we can have, besides the usual truth values 0 and 1, a truth value u (for “undefined”). Acquisition of new information may modify the value of a statement from u to 0 or 1.

2.6 LP seen from the main features of bounded rationality

We focus now on exploring how the bounded rationality program and LP relate (see also Stenning et al. 2017 for other aspects of the interaction between LP and heuristics). According to the basic tenets of Bounded Rationality, we see that LP can assume in this framework (at least) three different roles. In order of strength:

- (1) LP can be seen as fulfilling the main properties of bounded rationality (we will examine them next) and in this sense can be seen as complementary to known heuristics in the description of mental processes from this perspective.
- (2) LP works as a fabric for interpretation (reasoning *to* an interpretation) which provides a context for the information available in a situation. Only *from* this interpretation can reasoning or a decision occur. In this sense, LP would accomplish the interpretational process previous to the application of reasoning or decision methods. In this process, general knowledge is a source of recruited information. This is integrated in the current-preferred model of the information so far constructed about the current developing interpretation. This interpretative process would happen, in particular, for the application of heuristics. Heuristics require a “general background” part that tells you which to use, and a lot of general knowledge about how to compute them, as well as the episodic part about the current application. LP can be seen here as a general supporting structure: it supports interpretation and it supports deduction and inference. It provides the fundamental qualitative coordinates from which we make inferences, probabilistic reasoning or come to decisions.
- (3) Some fast and frugal heuristics in themselves (or some aspects of them) can be seen as the content of LP programs.

In this section we will focus on the first of these integrative approaches, analysing how the main properties of FFH are also applicable to LP. In Sect. 3.2 we will deal with (3). As for (2), we distinguish, as argued in (Stenning and van Lambalgen 2008), between the processes of “reasoning *to* an interpretation” and “reasoning *from* an interpretation”. There is no “general purpose” logic to describe reasoning in all situations, ignoring the interpretational process demanded at a first stage. Stenning and van Lambalgen (2008, p.20) see interpretation as a “parameter setting” in the sense of fixing a logic. The variation of logics may be analysed along three different dimensions (parameters to be fixed), namely: choice of a formal language (syntax), choice of a semantics and choice of a definition of valid arguments. Most of this process of interpretation and “making sense” in our daily practice is of course unconscious. We do it all the time without needing to “know” in a conscious or systematic way any logic at all.

Regarding point (1), what are the properties of FFH which make them appropriate bounded rationality models of human reasoning and decision making? According to Gigerenzer and collaborators: “The program is to design and test computational models of heuristics that are (a) ecologically rational (i.e., they exploit structures of information in the environment), (b) founded in evolved psychological capacities such as memory and the perceptual system, (c) fast, frugal, and simple enough to operate effectively when time, knowledge, and computational might are limited, (d) precise enough to be modelled computationally, and (e) powerful enough to model both good and poor reasoning.” (Goldstein and Gigerenzer 2002, p.75). We proceed to discuss briefly these different aspects without considering the last point (we already mentioned the relevant literature regarding it at the beginning of Sect. 2).

3 Ecological rationality

Rationality, for Simon, does not happen in a vacuum, independently of the context from which emerges. With a well-known image he summarized: “Human rational behaviour is shaped by a scissors whose blades are the structure of task environments and the computational capabilities of the actor” (Simon 1990). The author specifies further how an “environment” is to be intended: “We are not interested in describing some physically objective world in its totality but only those aspects of the totality that have relevance as the “life space” of the organism considered. Hence, what we call the “environment” will depend upon the ‘needs’, ‘drives’, or ‘goals’ of the organism, and upon its perceptual apparatus” (Simon 1956).

This way, rationality is dependent on the “goals” pursued. This is part of the contextual character of reasoning, something forgotten by abstract norms but crucial, in particular, for reasoning in communicative situations. Most reasoning tasks studied in psychology posit classical logic as the normative benchmark, without being fitted into the communicational context from which these norms emerged in the first place.

Given that our cognition is determined by these ecological aspects, the search for effective and transparent communication should acknowledge them. The representation formats of information may be tuned to our psychological functioning and facilitate ecologically adapted decisions. A well-known case is the use of natural frequencies instead of probabilistic formats in order to enhance extensional interpretations of tasks (Gigerenzer and Hoffrage 1995).

We can see a similar case in logic, where the communicational “ecology” determines different kinds of goals which may vary from cooperation to adversariality. This leads to different kinds of interpretation and reasoning which can be, correspondingly, either intensional or extensional. In this way, environments may condition a form of reasoning closer to classical logic or to LP (see Vargas and Stenning 2020 for the case of syllogistic reasoning). Depending on the kind of communication at hand, one kind of reasoning may be more suitable than another. This highlights the fact that the distinction here is not between correct reasoning and a poor man’s version of it, the same way that heuristics do not necessarily produce just second-best results (in comparison with optimization). The dependency on contextual goals may lead to different norms to be considered. In this sense psychology may turn out to be not only descriptive, but also normative. This may also imply the deliberate use of forms of expression closer to the relevant norms (in context). As an example, Kowalski’s (2011) analysis of an underground emergency notice and the British Nationality Act show how forms of expression closer to LP guidelines may be more suitable for the efficiency and clarity of communication.

3.1 Reliance on memory and perception

Just like heuristics, LP as a cognitive account is defined through “evolved psychological capacities”. As a logic for planning, LP works using information organized

in knowledge bases. The knowledge base may be seen as a model of semantic memory. The accessible information at a given moment may be seen as the current model of the situation held in working memory. Memory and perception go hand in hand because when new information arrives this current model is updated. This makes reasoning non-monotonic, in the sense explained in 2.4.

This is also distinct from the usual standards for experiments on reasoning. Participants in them are required to reason according to what “logically follows”, which means, abstracted from prior knowledge beyond the prompts provided experimentally. Also, classical logic reasoning places monotonicity as one of its core components, a feature that makes it more “static” or “atemporal”.

3.2 Fast and frugal processing

As we previously noticed talking about the syntax restrictions (sentences in clausal form) and the inferential machinery (the resolution method in its variants), tractability is a major concern in LP. What is clear is a huge difference with, say, first order classical logic. Seeking information for all possible outcomes, or all possible models for a given problem is, in most of the cases overly costly and unfeasibly time consuming. As emerges in the context of heuristics, robustness may be seen in contraposition with overfitting. Fitting is useful for modelling given data whereas generalization tends to be better for predicting new data. “Robustness goes hand in hand with speed, accuracy, and especially information frugality. Fast and frugal heuristics can reduce overfitting by ignoring the noise inherent in many cues and looking instead for the ‘swamping forces’ reflected in the most important cues” (Gigerenzer et al. 1999, p.20). This is precisely what is done by default assumptions. Considering by default that “nothing abnormal is the case” unless there is good reason to believe the contrary (using NAF) may increase efficiency. It alleviates the search for positive information on a potentially unmanageable (or even infinite) variety of cases. This is in accordance with the existing evidence in heuristics that basing reasoning (in a wide sense) and decision-making on more information and computation does not always lead to more accurate inferences or decisions.

3.3 Process models

It is often observed that heuristics are not only “labels” or vague descriptions but are defined through actual algorithms. As for LP (with the corresponding programming language Prolog) it is at the same time a medium for defining algorithms, and is at least partly constituted by algorithms.

As is often remarked about heuristics, in LP we do not have only “as-if models” as in the case, for example, of Bayesian models. Using Marr’s 3-levels classification, LP offers accounts both at the computational and algorithmic levels, with additional neural network implementations at the third level (Stenning and van Lambalgen 2008; Pijnacker et al. 2011). In this sense, LP not only predicts the outcomes in given situations (the “computational” level), but specifies feasible models of how these computations may be done (the “algorithmic” level).

4 Heuristics viewed from LP

In the previous section we examined how LP shares the main properties assigned to heuristics. We can also see LP as a framework for reasoning and decision-making in which heuristics can be analysed from a logical viewpoint. In this section we analyse some examples from (Gigerenzer et al. 1999). That they can be treated using LP is no surprise, given the algorithmic, computational character of heuristics in this tradition (mentioned in 3.4). Nevertheless, providing an explicit formalization may start a structural analysis of properties shared (or not) within “the adaptive toolbox” (Gigerenzer et al. 1999) and provide tools to measure, from the logical point of view, how simple “simple heuristics” are.

4.1 Recognition

Having to choose between two possibilities or items c_1 and c_2 on which has larger value according to a given criterion this judgement is based on recognition. Indicate with $R(x)$ the predicate “ x is recognized” and with $L(x, y)$ the predicate “ x is larger than y according to the given criterion”. A standard example is “ x is a larger city than y ” (Gigerenzer et al. 1999).

The program in this case is limited to the following rules:

$$R(c_1) \wedge \neg R(c_2) \rightarrow L(c_1, c_2)$$

$$R(c_2) \wedge \neg R(c_1) \rightarrow L(c_2, c_1)$$

This is of course applicable only in cases where c_1 and c_2 are not both recognized or both unrecognized.

Psychologically, we may have exceptions to the application of this rule, and LP versions acknowledge them explicitly through abnormality clauses:

$$R(c_1) \wedge \neg R(c_2) \wedge \neg ab \rightarrow L(c_1, c_2)$$

$$R(c_2) \wedge \neg R(c_1) \wedge \neg ab \rightarrow L(c_2, c_1)$$

4.2 One reason decision-making

The formalization of the recognition heuristic may be adapted to the case of several cues relevant for a decision but only one of them is considered (Gigerenzer et al. 1999, p.75). An example is constituted by the “Take the best” heuristic.

Let $S_1(x), \dots, S_n(x)$, indicate the predicates associated to the n cues considered.

Let's consider two cues where $S_1(x)$ is privileged over $S_2(x)$. $L(x, y)$, again, stands for the predicate “ x is larger than y according to the given criterion”

$$S_1(c_1) \wedge \neg S_1(c_2) \rightarrow L(c_1, c_2)$$

$$S_1(c_2) \wedge \neg S_1(c_1) \rightarrow L(c_2, c_1)$$

$$S_1(c_1) \wedge S_1(c_2) \wedge S_2(c_1) \wedge \neg S_2(c_2) \rightarrow L(c_1, c_2)$$

$$\neg S_1(c_1) \wedge \neg S_1(c_2) \wedge S_2(c_1) \wedge \neg S_2(c_2) \rightarrow L(c_1, c_2)$$

$$S_1(c_1) \wedge S_1(c_2) \wedge \neg S_2(c_1) \wedge S_2(c_2) \rightarrow L(c_2, c_1)$$

$$\neg S_1(c_1) \wedge \neg S_1(c_2) \wedge \neg S_2(c_1) \wedge S_2(c_2) \rightarrow L(c_2, c_1)$$

Therefore, the cue represented by S_2 can be used only in case that S_1 does not decide.

The generalization for more than two predicates is straightforward.

4.3 Fast-and-frugal trees

We turn now to the problem of taking a decision according to a classification (a ‘categorization’) of an object or a situation. This may be represented in the form of decision or classification trees. A typical example from (Green and Mehr 1997) is the decision whether to assign a patient suspected of heart disease to a regular nursing bed or to emergency care (Fig. 1).

We can consider the logical structure of the decisions involved in the tree. In this example, let $S_1(x), S_2(x), S_3(x)$, indicate the predicates associated to the 3 cues or questions considered in the tree, namely: “ x has an elevated ST segment”, “ x has chest pain” and “ x has other symptoms”.

Consider the program:

$$S_1(p) \rightarrow High_Risk(p)$$

$$\neg S_1(p) \wedge \neg S_2(p) \rightarrow Low_Risk(p)$$

$$\neg S_1(p) \wedge S_2(p) \wedge S_3(p) \rightarrow High_Risk(p)$$

$$\neg S_1(p) \wedge S_2(p) \wedge \neg S_3(p) \rightarrow Low_Risk(p)$$

We may notice that in such a program we find implicit the hierarchical structure of the cues $S_1(x), S_2(x), S_3(x)$. We can recover in fact such structure examining whether the predicates (either positive or negated) on the bodies of the implications appear in conjunction with the other ones. For example, we see that in the first clause, the predicate $S_1(x)$ appears alone, while all the appearances of $S_2(x)$ occur in conjunction with (a negation of) $S_1(x)$. This entails that $S_1(x)$ has priority over $S_2(x)$.

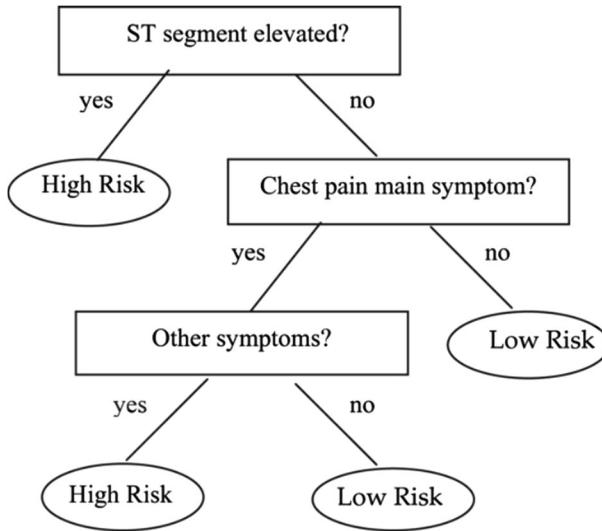


Fig. 1 A fast-and-frugal tree for the doctors decision of whether sending a patient to a regular nursing bed or the coronary care unit (Green and Mehr 1997)

Similarly, $S_2(x)$ has priority over predicate $S_3(x)$ as a decision criterion because the last one appears always in company with the former one, while $S_2(x)$ occurs without $S_3(x)$ in the second clause.

Now, generalizing this example, we have that in a classification tree each node represents a question regarding certain features of the objects to be classified. Leaves represent different answers to the question. Answers to the question are here exclusive (non-overlapping) and exhaustive (cover all objects). We consider binary trees, which are constructed from binary ramifications. A categorization tree, in this sense is *fast and frugal* if and only if it has at least one exit at each level (Martignon, Vitouch, Takezawa, and Forster 2003).

We see that we can in general assign LP programs to decision (binary) trees: we take each of the exit nodes as conclusion and codify the branch leading to it as a conjunction of literals which are either positive or negative depending on whether the corresponding question is answered positively or negatively. Thus, the body of the clause codifies the vector of 0's and 1's corresponding to each branch as in the paper just cited.

Nevertheless, we may not be able in general to reconstruct uniquely the tree from the associated program. In fact, some complete binary trees with n levels may be presented for instance in either their causal or their diagnostic versions, leading to the same associated LP programs. In fact, partitioning is commutative because we obtain the same final classification cells independently of the partitioning order as in the paper just cited.

We introduce the following two concepts which will help characterize programs associated to fast and frugal trees:

Definition 1 A set of clauses is *predicate-hierarchical* if for every pair of cues or predicates $S_i(x), S_j(x)$ appearing in them, there is one, say $S_i(x)$ which appears in a clause without the appearance of $S_j(x)$, while the appearance of $S_j(x)$ in a body is always accompanied by the appearance of $S_i(x)$.

Definition 2 A set of clauses is *body-stepwise-increasing* if for every clause in it with a body with $n > 1$ predicates appearing on it, there is a clause on the program with a body with $n-1$ among these predicates.

As an illustration, consider the set of clauses of our example, given e.g. the clause $\neg S_1(x) \wedge S_2(x) \wedge \neg S_3(x) \rightarrow Low_Risk(x)$. We can find a clause with two of the predicates $S_1(x), S_2(x), S_3(x)$ in its body, namely $\neg S_1(p) \wedge \neg S_2(p) \rightarrow Low_Risk(p)$.

If a set of clauses (a program) is body-stepwise-increasing, then it is predicate-hierarchical. We have the following result:

Result 1 A binary categorization tree is fast and frugal if and only if its associated logic program is body-stepwise-increasing.

Proof From left to right, consider a tree and its associated program. Since a fast and frugal tree has one exit at each level n , we have a corresponding clause with a body with exactly n predicates, associated with the nodes up to that level. So the set of clauses is body-stepwise-increasing by induction. From right to left, since the program is predicate-hierarchical we can linearly organize the different predicates, which correspond to the cues at the nodes at each level of the tree. Since the program is body-stepwise-increasing we have an exit at each level.

4.4 Negation as failure (NAF) in bounded search

In contexts of uncertainty, the fulfillment (or unfulfillment) of certain cues is not completely known. In cases of undetermined values, the selection follows a “positive bias” (Gigerenzer et al. 1999, p. 91), which is, precisely, the NAF principle (Sect. 2.2 above).

As an example, the case previously analyzed, $(S_1(x))$ is privileged over $S_2(x)$) Suppose we know that $\neg S_1(c_1)$ as a fact and ignore whether $S_1(c_2)$ either holds or not. Then, by CWR, $\perp \leftrightarrow S_1(c_2)$ (which means $\neg S_1(c_2)$) and we must consider the next cue ($S_2(x)$) in order to choose between c_1 and c_2 .

Similarly, if we know $S_1(c_1)$ as a fact, and again ignore if $S_1(c_2)$, our search stops because, by CWR, $\perp \leftrightarrow S_1(c_2)$ and we can apply $S_1(c_1) \wedge \neg S_1(c_2) \rightarrow L(c_1, c_2)$.

4.5 Satisficing

The structure of satisficing (Simon 1956) is determined by a threshold which, if reached, serves as a criterion for selecting an option and stopping a search. This is in contrast to strategies which search until an optimal point is reached.

Given a series of items, let t be the threshold or aspiration level for a given cue, and let l_i be the level corresponding to the i th item. The program in this case is composed by the facts:

$$\text{level}(l_1)$$

$$\text{level}(l_2)$$

$$\text{level}(l_n)$$

Satisficing can be implemented positing the following query to the program:

$$?\text{level}(x) \wedge x \geq t$$

where x is a variable and the statement is intended as existentially quantified.

The answer to the query, if successful, provides a witness for this conjunction and will be the first level value which reaches the threshold.

More complex versions of satisficing may or may not be expressible in LP. Simon himself considered the possibility of different cues and variation with time: “An alternative satisfies if it meets aspirations along all dimensions. If no such alternative is found, search is undertaken for new alternatives. Meanwhile, aspirations along one or more dimensions drift down gradually until a satisfactory new alternative is found or some existing alternative satisfies. A theory of choice employing these mechanisms acknowledges the limits on human computation and fits our empirical observations of human decision making far better than the utility maximization theory” (Simon 1969).

4.6 The gaze heuristic

A heuristic for catching a baseball, it is usually stated in the following terms: “Fixate your gaze on the ball, start running, and adjust the speed so that the angle of gaze remains constant” (Gigerenzer et al. 1999).

In order to simplify the problem, assume we have the information of the angle of gaze (to be precise we do not directly have access to this angle, but the information on the position of the ball on the retina and the head and ocular inclination). Assume also that the player must run in the opposite direction to the movement of the ball (the case in which the player and the ball go in the same direction can be treated analogously). Indicate it with α , which must be between 0° and 90° .

$$x > \alpha \rightarrow \text{run_slower}$$

$$x < \alpha \rightarrow \text{run_faster}$$

This is a schematic way which can be further specified: suppose we have a finite number of velocities, say (in increasing order) $v_0 = 0, v_1, v_2, v_3$.

Rules for running slower:

$$x > \alpha \wedge v = v_3 \wedge \neg ab \rightarrow v = v_2$$

$$x > \alpha \wedge v = v_2 \wedge \neg ab \rightarrow v = v_1$$

$$x > \alpha \wedge v = v_1 \wedge \neg ab \rightarrow v = v_0$$

The opposite rules for running faster:

$$x < \alpha \wedge v = v_0 \wedge \neg ab \rightarrow v = v_1$$

$$x < \alpha \wedge v = v_1 \wedge \neg ab \rightarrow v = v_2$$

$$x < \alpha \wedge v = v_2 \wedge \neg ab \rightarrow v = v_3$$

Rules for stopping:

$$\text{ball_landed} \rightarrow v = v_0$$

$$\text{ball_caught} \rightarrow v = v_0$$

$$\text{unreachable} \rightarrow v = v_0$$

$$x > 85^\circ \rightarrow v = v_0$$

abnormalities:

$$\text{ball_landed} \rightarrow ab$$

$$\text{ball_caught} \rightarrow ab$$

$$\text{unreachable} \rightarrow ab$$

$$x > 85^\circ \rightarrow ab$$

Of course, more abnormalities may happen in the real world.

5 Conclusions: normative and ecological perspectives

Canonical norms (such as classical logic, probability theory and rational choice theory) have been traditionally treated as fixed benchmarks to measure the validity of human reasoning and decision-making. The bounded rationality tradition opened perspectives to understand deviations from these norms. These divergences are not necessarily to be seen as errors, biases or fallacies, but rather as responding to environmental prompts and the structure of the problems faced.

We have examined the particular case of LP and have argued how it can stand beside heuristics as a framework (among others) for bounded rationality. In this sense logic, or actually multiple logics, should not be seen inherently as orthogonal to the FFH program. As we illustrated in Sect. 3, there is a coincidence in the most fundamental aims pursued by the FFH and by the LP approaches. They coincide also in their difference from other approaches such as the one based on Bayesian models.

We conceive logic as a discipline which provides tools (multiple logics) for modelling “real-world” phenomena (“logics-as-models”) particularly relevant for the cognitive study of reasoning. Specific aspects of a logical system may be more or less appropriate to describe or to predict the target phenomenon. This reminds us of what happens with mathematical models, e.g., in physics.⁷ The main features of LP have shown its cognitive relevance in modelling a variety of experimental results. Going in the same direction, Sect. 3.2 explored how some of the heuristics may be analysed from this point of view. This exploration may lead to a more explicit understanding of the logical structure operating along with FFH, without pretending to adopt a reductionistic viewpoint: we are not claiming that LP can capture all the psychological faculties operating in FFH. It is a tool designed to capture discourse processing and reasoning, and which can therefore be used in modelling these essential components of FFH. We see the simplicity of the syntax and the computational advantages of LP (see 2.1) in correspondence with the frugality and the promptness of FFH.

Authors' contributions The research in the article is part of the doctoral thesis of FV under the supervision of LM and KS. FV wrote the first version of the article. LM and KS contributed manuscript revision. All authors read and approved the submitted version.

Funding Open Access funding enabled and organized by Projekt DEAL. No funding to declare.

Availability of data and material Not applicable.

Code availability Not applicable

Declarations

Conflicts of interest The authors have no relevant financial or non-financial interests to disclose.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission

⁷ Let us remind also Barwise statement that “First-order logic is just an artificial language constructed to help investigate logic, much as the telescope is a tool constructed to help study heavenly bodies. From the perspective of the mathematician in the street, the first-order thesis is like the claim that astronomy is the study of the telescope” (Barwise & Feferman eds., 1985).

directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

6. References

- Barwise J, Feferman S (eds) (2017) Model-theoretic logics. Cambridge University Press, Cambridge, UK
- Beall JC, Restall G (2006) Logical pluralism. Oxford University Press on Demand
- Brachman RJ, Levesque HJ (2004) Knowledge representation and reasoning. Elsevier, San Francisco
- Doets K (1994) From Logic to Logic Programming. MIT Press, Cambridge, MA
- Gigerenzer G (2008) Why heuristics work. *Perspect Psychol Sci* 3(1):20–29
- Gigerenzer G, Hoffrage U (1995) How to improve Bayesian reasoning without instruction: Frequency formats. *Psychol Rev* 102(4):684
- Gigerenzer G, Todd PM (1999) The ABC research group simple heuristics that make us smart. Oxford University Press, New York
- Goldstein DG, Gigerenzer G (2002) Models of ecological rationality: the recognition heuristic. *Psychol Rev* 109(1):75
- Green L, Mehr DR (1997) What alters physicians' decisions to admit to the coronary care unit? *J Fam Pract* 45(3):219
- Hölldobler S (2015) Weak completion semantics and its applications in Human reasoning, In: Proceedings of the first workshop on bridging the gap between human and automated reasoning, CADE-25, Berlin, 2–16
- Kleene SC (1952) Introduction to metamathematics. North Holland, Amsterdam
- Kowalski R (2011) Computational Logic and Human Thinking: How to be Artificially Intelligent. Cambridge University Press
- Kowalski R (2014) "History of Logic Programming". In: Computational Logic, Vol. 9. Ed by Joerg Siekmann. Elsevier, pp 523–569
- Van Lambalgen M, Hamm F (2005) The proper treatment of events. Blackwell Publishing
- Martignon L, Vitouch O, Takezawa M, Forster MR (2003). Naive and yet enlightened: From natural frequencies to fast and frugal decision trees, in Thinking: Psychological perspective on reasoning, judgment, and decision making, 189–211
- McCarthy J (1986) Applications of circumscription to formalizing common-sense knowledge. *Artif Intell* 28(1):89–116
- Pijnacker J, Geurts B, van Lambalgen M, Buitelaar J, Hagoort P (2011) Reasoning with exceptions: an event-related brain potentials study. *J Cogn Neurosci* 23(2):471–480
- Shapiro S (1991) Foundations Without Foundationalism: A Case for Second-Order Logic. Oxford University Press
- Shoham Y (1987) A semantical approach to non-monotonic logics. In Proceedings of the Tenth International Joint Conference on Artificial Intelligence (IJCAI), 1987. A semantical approach to non-monotonic logics. Reprinted in ML Ginsberg (ed) Readings in non-monotonic reasoning (pp 227–250). Morgan Kaufmann Publishers Inc.
- Simon HA (1947) Administrative Behavior: A Study of Decision-Making Processes in Administrative Organizations. Macmillan, New York
- Simon HA (1956) Rational choice and the structure of the environment. *Psychol Rev* 63(2):129
- Simon HA (1957) Models of man; social and rational. Wiley, New York
- Simon HA (1969) The sciences of the artificial. MIT Press, Cambridge, MA
- Simon HA (1990) Invariants of human behavior. *Annu Rev Psychol* 41(1):1–20
- Stenning K, van Lambalgen M (2008) Human Reasoning and Cognitive Science. MIT Press, Cambridge, MA
- Stenning K, Martignon L, Varga A (2017) Probability-free judgment: Integrating fast and frugal heuristics with a logic of interpretation. *Decision* 4(3):171
- Tversky A, Kahneman D (1974) Judgment under uncertainty: Heuristics and biases. *Science* 185(4157):1124–1131
- Vargas F, Stenning K (2020) Communication, goals and counterexamples in syllogistic reasoning. In: Frontiers in Education, 5.

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