WHEN THE (BAYESIAN) IDEAL IS NOT IDEAL

Danilo Fraga DANTAS

ABSTRACT: Bayesian epistemologists support the norms of probabilism and conditionalization using Dutch book and accuracy arguments. These arguments assume that rationality requires agents to maximize practical or epistemic value in every doxastic state, which is evaluated from a subjective point of view (e.g., the agent's expectancy of value). The accuracy arguments also presuppose that agents are opinionated. The goal of this paper is to discuss the assumptions of these arguments, including the measure of epistemic value. I have designed AI agents based on the Bayesian model and a nonmonotonic framework and tested how they achieve practical and epistemic value in conditions in which an alternative set of assumptions holds. In one of the tested conditions, the nonmonotonic agent, which is not opinionated and fulfills neither probabilism nor conditionalization, outperforms the Bayesian in the measure of epistemic value that I argue for in the paper (α-value). I discuss the consequences of these results for the epistemology of rationality.

KEYWORDS: bounded rationality, computational epistemology, Bayesian epistemology, epistemic utility theory, nonmonotonic logic

Introduction

Bayesian epistemologists propose the norms of probabilism and conditionalization.[1] The initial justifications for these norms were Dutch book arguments (DBAs). A DB is a series of bets, each of which the agent considers fair, but when taken together, they result in a guaranteed loss. A DBA aims to demonstrate that if an agent violates a specific Bayesian norm, then it is possible to create a DB against her. Ramsey (1926) and De Finetti (1937) proposed DBAs for probabilism, and Teller (1973) for conditionalization. Satisfaction of a given Bayesian norm does not guarantee invulnerability to DBs for other norms, but if a reasoner fulfills a norm (e.g., probabilism), then it can be shown that a DB for that norm cannot be constructed against her (e.g., Kemeny 1955). Since vulnerability to DBs may lead an agent to buy and sell a series of bets that amounts to a sure loss, it is associated

[1] Probabilism states that a rational agent's credence function is consistent with the axioms of probability. Conditionalization states that a rational agent who becomes certain of some new piece of evidence updates her previous unconditional credences by conditionalizing them on the evidence (see Talbott 2016).

with the minimization of 'practical value,' which, in turn, is related to the fulfillment of the agent's practical goals (e.g., having monetary gain). Some critics argue that vulnerability to DBs is not adequately related to the minimization of practical value.[2] Others argue that *practical* rationality might be associated with maximizing practical value, while *epistemic* rationality should be associated with maximizing 'epistemic value.'

Different features of a set of 'beliefs'[3] are putative sources of epistemic value: closure, coherence, amount of evidential support, etc. In the last decades, a group of epistemologists has argued for 'veritism,' the thesis that the fundamental source of epistemic value is the believing of truths rather than falsehoods.[4] These epistemologists sought to justify the Bayesian norms in purely epistemic terms (i.e., veritistic), which resulted in the epistemic utility theory (EUT). The 'accuracy arguments' of EUT purport to show that the fulfillment of the Bayesian norms leads to the minimization of (expected) epistemic disvalue (inaccuracy, see sec. 1). Joyce (1998) argues that if an agent's credence function is not probabilistic, then she is in the position to know that there is a probabilistic function that is less inaccurate than hers in all situations that she finds possible. Leitgeb and Pettigrew (2010) argue that an agent who updates her credences using a non-Bayesian strategy is in the position to know that an update using conditionalization minimizes expected inaccuracy relatively to hers, where this expectancy is calculated using her credences.

I use the expression 'maximization of value' loosely, in such a way that it includes the avoidance of minimization of value (e.g., the avoidance of a 'sure loss') and the minimization of (expected) disvalue. In this sense, the DBAs and the accuracy arguments are 'maximization arguments,' as they evaluate an agent's rationality in terms of how much they maximize practical or epistemic values. These arguments assume the following:

a1. Rationality requires maximization of value in every doxastic state;

a2. 'Maximization of value' is evaluated from a subjective point of view;

a3. Agents are opinionated.

---

[2] For example, Hájek (2009, 231) points out that violating the Bayesian norms amounts not only to the construction of DBs but also to Czech books (a series of fair bets that together amount to a sure win).

[3] I use 'belief' as a general term, which encompasses both full beliefs and credences.

[4] For example: "the fundamental source of epistemic value for a doxastic state is the extent to which it represents the world correctly: that is, its fundamental epistemic value is determined entirely by its truth or falsity" (Pettigrew 2019, 761). See Littlejohn (2015) for a critical discussion about veritism.

These arguments assume a1 because they concern the doxastic states of an agent at an arbitrary point in time (e.g., after an update). Failing to maximize value at that time is sufficient for irrationality. Nevertheless, a rational agent might risk value at the present if this may result in a gain of value in the future. A related issue is that the DBAs assume that a rational agent cannot make choices that lead to a sure loss of any practical value (monetary gain is not a priority practical value). Nevertheless, a rational agent may be forced to make such choices when facing trade-offs between goals. These arguments assume a2 because they measure value in such a way that is accessible to the agents. For example, Leitgeb and Pettigrew measure the expectation of value using the agents' credences.[5] This assumption is reasonable because rationality is usually assumed to have a subjective character (e.g., Wedgwood 2015, 221), but whether agents maximize actual value should also interest epistemology.[6] Opinionation is the holding of belief-values for every proposition in the agent's agenda (the set of propositions of interest). The accuracy arguments assume a3 because they always compare sets of beliefs with values for the same propositions (the agenda). This assumption is related to the difficulty of using inaccuracy to compare sets of beliefs with different sizes (see sec. 1). Nevertheless, it may be rational to 'withhold beliefs'[7] in some situations (e.g., in the absence of evidence).

I intend to investigate how much of the justificatory force of the maximization arguments depends on a1-a3 and whether these are necessary assumptions in a study about rationality. The investigation will proceed by comparing how a Bayesian agent and an agent who does not fulfill the Bayesian norms maximize practical and epistemic values in conditions in which a1-a3 do not hold. Several formal systems model uncertain reasoning in a way that deviates from the tenets of probability (see Genin and Huber 2020). The nonmonotonic frameworks are among those that deviate the most. There is a myriad of nonmonotonic frameworks, but not all of them are far enough from probability

---

[5] The DBAs appeal to the notion of sure loss and Joyce's argument appeals to the principle of dominance. These elements are objective in the sense of concerning guaranteed outcomes. But these arguments still conform to a2 because the agents are in the position to know that those outcomes are guaranteed.

[6] "[I]t would seem absurd to claim that it is epistemically more important to have an update rule that minimizes *expected* inaccuracy than to have one that *actually* minimizes inaccuracy" (Douven 2013, 438, his emphases).

[7] I use 'withhold beliefs' to denote the attitude of neither believing nor disbelieving a proposition. I am not using this expression to denote the attitude of holding middling credences about a proposition (see Sturgeon 2008).

(for our purposes).[8] I will focus on a three-valued version of logic programming (LP, see Doets 1994), whose third value may be used for modeling non-opinionated agents (against a3). The comparison will be performed using the tools of computational epistemology (e.g., Douven 2013), where epistemologists design computer simulations of AI agents interacting with environments that are randomly generated from fixed parameters. The measures of practical and epistemic values will be done only at the end of each trial (against a1). 'Value in a random trial' (actual value) is a contingent notion, but, if the number of trials is large enough, their mean approximates an 'objective expected value' (against a2), which results from actual values and environmental probabilities and not the agent's beliefs.

I have designed computer simulations of two AI agents facing an epistemic version of the Wumpus World, a class of environments used for investigating uncertain reasoning. The first is the 'Bayesian agent,' who holds degrees of belief (credences), is opinionated, and fulfills both probabilism and conditionalization. The second is the 'nonmonotonic agent,' based on a three-valued LP, who holds all-or-nothing beliefs (full beliefs), is not necessarily opinionated, and fulfills neither probabilism nor conditionalization. I analyze how often these agents solve the problem (practical rationality) and the epistemic value of their beliefs (epistemic rationality).[9] The idea is not to refute the conclusions of Bayesian epistemology, but to discuss the assumptions in the notion of maximization of (practical or epistemic) value, including the measure of epistemic value. In Section 1, I introduce the Wumpus World and the measures of epistemic value (α- and β-values, with linear, Brier, and log scoring rules) and of practical value (p-value) used in the investigation. In Section 2, I discuss the implementation of these models in a computer simulation and present the results from the simulation. In one condition (c3), the nonmonotonic agent outperforms the Bayesian in α-value. In Section 3, I discuss the measures of epistemic value, whether a1-a3 are necessary assumptions in a study about rationality, and how the results from EWW impact the epistemology of rationality.

[8] For example, Bourne and Parsons (1998) show how to use System P (Kraus et al. 1990) to propagate the lower bounds on conditional probabilities.

[9] The study of practical rationality most often concerns the principles of choice between actions given the beliefs and preferences of an agent (e.g., in decision theory, see Steele and Stefánsson 2020). I use 'practical rationality' in the broader sense of 'how much the agent's beliefs assist her practical goals.'

1. Setting the Stage

The Wumpus World (WW, see Russell and Norvig 2020, 210) is a class of environments used in AI for studying uncertain reasoning. WW is a cave comprising rooms connected by passageways and surrounded by walls. Somewhere in the cave, there is the Wumpus, a beast that kills anyone who enters its room. Other rooms contain bottomless pits that kill anyone who steps in. Finally, there is a heap of gold hidden somewhere in the cave. The agent's goal is to explore the cave, find the gold, and escape the cave with the gold. The agent cannot enter a room containing a pit or the Wumpus without dying, but the rooms around those contain a breeze and a stench (respectively). WW is an interesting environment for studying uncertain reasoning because it is only partially observable, in such a way the absence of direct information about pits and the Wumpus forces the agents to draw provisional conclusions and revisit them given new information. WW is more complex than the average environment used in computational epistemology, but it also is more 'realistic' than the average because it associates a practical cost (risk of death) with the gathering of evidence. This feature enables us to integrate the investigation of practical and epistemic rationality.

The only goal of WW ('grab the gold and escape the cave') is a practical goal, but I am also concerned with epistemic rationality. For this reason, I have developed an epistemic version of WW (EWW). EWW has the additional goal of forming the most comprehensive and accurate set of beliefs about the positions of the pits and the Wumpus. EWW has an additional layer of uncertainty: the agent may perceive 'random' breezes and stenches, which are independent of the pits and the Wumpus. The occurrence of random breezes and stenches may be interpreted as a feature of the agent's perception, which may return a persistent false positive (a perceptual illusion).[10] Random breezes and stenches may be used to evaluate how reasoning processes cope with a faulty perception. EWW may be represented as a matrix of dimensions $s \times s$, where the room (x, y) is in the xth row and yth col of the matrix. In the following, $p_{x,y}$ states that there is a pit in (x, y). Similarly, $w_{x,y}$ for Wumpus, $g_{x,y}$ for gold, $b_{x,y}$ for breezes, and $s_{x,y}$ for stenches. The Wumpus and the gold are randomly placed in a room other than (0, 0). Rooms

[10] The most natural model of a faulty perception is one in which each perception has some probability of returning a non-persistent false positive/negative (e.g., about breezes). This model introduces some complications for our purposes. For example, a Bayesian agent with faulty perception in this sense would need to update her credences using Jeffrey conditionalization (Jeffrey, 1983), but Leitgeb and Pettigrew (2010) have shown that Jeffrey conditionalization does not necessarily minimize expected inaccuracy. This result is exploited by Trpin and Pellert (2019), who use the natural model of faulty perception.

other than (0, 0) contain a pit with probability $pr(p)$ and a random breeze or random stench with an independent probability of $pr(rand)$.[11] The agent starts in (0, 0) ('the entrance'), facing east. She knows the description of EWW and the size of the cave, but she ignores the configuration of the cave (the position of pits, Wumpus, and gold).

The most distinguishing feature of EWW is the measure of epistemic value. The measure used in EUT is one of epistemic disvalue (inaccuracy), defined as the 'distance' between a set of beliefs and the 'ideal' set containing beliefs about the same propositions (see Carr 2015, 223),[12] where the belief that φ has the value 1 when φ is true and 0 when it is false. Inaccuracy is not an adequate measure of epistemic disvalue for the comparison between the Bayesian and the nonmonotonic agents because the latter can withhold beliefs, which is a 'cheap' way of minimizing inaccuracy. If inaccuracy was the measure of epistemic disvalue, then the nonmonotonic agent would have no reason to explore the cave and form new beliefs because, in doing so, she would risk increasing (but not decreasing) her inaccuracy (Dantas 2022). The comparison between those agents requires a measure of epistemic value that awards comprehensiveness and accuracy, which may be done by measuring the amount of truth (t) and of falsehood (f) in the agent's belief-set and then conjoining these values using an 'alethic' function $a(t, f)$.

	t_φ	f_φ
Linear score	$1 - \varepsilon_\varphi$	ε_φ
Brier score	$(1 - \varepsilon_\varphi)^2$	$(\varepsilon_\varphi)^2$
Logarithmic score	$-\ln(\varepsilon_\varphi)$	$-\ln(1 - \varepsilon_\varphi)$

Table 1: Scoring rules, where $\varepsilon_\varphi = |v(\varphi) - b(\varphi)|$, $v(\varphi)$ is φ's truth-value, and $b(\varphi)$ is the belief-value of φ for the agent.

The values of t and f may be measured using different 'scoring rules.' These values are the global amounts of truth and falsehood in the agent's set of beliefs, measured as the sum of the local amounts t_φ and f_φ, for every proposition φ in the agent's belief-set B (i.e., $t = \sum_{\varphi\in B} t_\varphi$ and $f = \sum_{\varphi\in B} f_\varphi$). Let the 'error' of the belief that φ be $\varepsilon_\varphi = |v(\varphi) - b(\varphi)|$, where $v(\varphi)$ is φ's truth-value and $b(\varphi)$ is φ's belief-value for

[11] I have also introduced some simplifications to the traditional WW. No room contains a pit and the Wumpus, or a pit and the gold, or the Wumpus and the gold. The Wumpus does not move and the agent does not have 'arrows' so that she can kill the Wumpus.

[12] "Epistemic decision theory [EUT] usually presupposes that the credence functions it compares are defined over the same algebra of propositions [agenda]. Once we abandon this presupposition, new difficulties arise" (Carr 2015, 223). I exploit some of these difficulties in the rest of this paragraph.

the agent. The scoring rules in Table 1 encode different attitudes towards epistemic risk (see Babic 2019). The Brier score awards risk-averse agents who hold middling credences because $f_\varphi = (\varepsilon_\varphi)^2 \leq \varepsilon_\varphi$, where this difference is larger when ε_φ is around 0.5. The log score invites risk-seeking agents because $t_\varphi = -\ln(\varepsilon_\varphi) \geq (1 - \varepsilon_\varphi)$ and $f_\varphi = -\ln(1 - \varepsilon_\varphi) \geq \varepsilon_\varphi$, where these differences approach $+\infty$ when ε_φ approaches 0 and 1 (respectively, but see fn. 12). The linear score is neutral in this regard because $t_\varphi = (1 - \varepsilon_\varphi)$ and $f_\varphi = \varepsilon_\varphi$. The Brier score tends to favor the Bayesian agent, who can hold credences. The log score may favor the nonmonotonic agent, who can hold all-or-nothing true beliefs and withhold those that are possibly false, or the Bayesian, who can avoid the risk by holding non-extreme credences. I will use the three scores to keep the results of the simulations independent of specific measures.[13]

The alethic function $a(t, f)$ must strictly increase with respect to (wrt) t (if $t' > t$, then $a(t', f) > a(t, f)$) and strictly decrease wrt f (if $f' > f$, then $a(t, f') < a(t, f)$). These 'basic requirements' are accepted by Douven (2013, 436): "The basic intuition underlying it is clear enough, to wit, that the higher one's degree of belief in a true proposition is, the more accurate one is, ceteris paribus, and also the lower one's degree of belief in a false proposition is, the more accurate one is, ceteris paribus". The basic requirements are put to work in a computational investigation by Trpin and Pellert (2019), who use the function $t - f$. This is the 'minimal' function that fulfills the basic requirements, but it is unbounded from above and below. This feature makes the comparison of performances in caves of different sizes difficult because the final value may be dominated by the values of t or f. This problem may be avoided by adding a denominator to this function, which may be done conscientiously in at least two ways: (i) $t + f$, as the agent's 'amount of commitment' (there is a vestigial problem with this suggestion, see below), or (ii) n, as the maximum size for the agent's agenda (the number of propositions at issue).[14]

[13] The Brier score is the most popular scoring rule in EUT, but Lewis and Fallis (2019) argue for the log score. The linear score is often dismissed for not being a proper scoring rule (see Lewis and Fallis 2019, sec. 4), but it returns the same values as the Brier score for full beliefs ($0^2 = 0$ and $1^2 = 1$). The fact that t_φ and f_φ collapse to $+\infty$ when ε_φ is 0 and 1 causes the global measure of t and f also to collapse in those cases. To avoid this, I will compute $\ln(c)$ instead of $\ln(0)$ for a small constant $c > 0$ (I will use $c = 0.01$, but there is nothing special with this value).

[14] The number of propositions in the real world is infinite, but it may be finite in toy worlds, such as EWW. It is often assumed that agendas are maximal (i.e., that agents are interested in the truth-value of 'every proposition'). This is a reasonable idealization for our purposes, but real agents may withdraw propositions from their agendas (see the 'anti-interrogative attitude' in sec. 3).

The first possibility results in the function $(t - f)/(t + f)$, which does not fulfill the basic requirements and does not measure epistemic value correctly. For example, two agents with only true beliefs about one and ten propositions (respectively) would receive the same evaluation, but the second is more well-informed than the first and should be awarded for this. The addition of a small constant $c > 0$ to the denominator avoids this problem,[15] leading to the second agent receiving more (marginal) epistemic value than the first. The resulting function is $\alpha(t, f) = (t - f)/(t + f + c)$, which is discussed in Dantas (2021). The second possibility results in the function $\beta(t, f) = (t - f)/n$, where n is the number of propositions at issue (in EWW, $n = 2 \times s^2$). This function is used in the literature about truthlikeness (e.g., Cevolani and Festa 2021, 11472). The function α awards lower amounts of commitment, which both agents can do: the Bayesian agent can hold middling credences (which lowers $t + f$ given some scores) and the nonmonotonic agent can commit herself to fewer propositions (which always lowers $t + f$). The denominator of the function β denominator is a fixed upper bound for the agents' amount of commitment and it does not award lower amounts of it. I will use both functions to keep the results of the simulations independent of specific measures. For readability, I will measure the epistemic value as 1000 x $a(t, f)$, where $a = \alpha$ (α-value) or $a = \beta$ (β-value). The practical value (p-value) will be measured as +1000 for escaping with the gold and −1000 for dying.

1.1 The Bayesian model

The Bayesian model of a rational agent has the following features:

b1. The agent's belief-values can have continuously many values between 0 and 1 (credences);

b2. The agent's credence function is consistent with the axioms of probability;

b3. The agent updates her credences using conditionalization.

The Bayesian agent holds very fine-grained beliefs (b1) and fulfills both probabilism (b2) and conditionalization (b3). It follows from b2 and b3 that the Bayesian agent is opinionated.[16] This model does not determine the initial

[15] I intend to set c = 0.01 in both uses of constants (see fn. 12), but nothing in this paper depends on this specific value. The constant c may be seen as setting the 'sensitivity' of the function α: the smaller the *c*, the greater the benefit for believing more truths and the penalty for believing more falsehoods.

[16] A (probabilistic) credence *function* (not a partial function) returns a value between 0 and 1 for every proposition in its domain (the agent's agenda) and there is no update from the absence of values about a proposition to some belief-value using conditionalization. As a result, the Bayesian agent cannot gain or lose beliefs.

credences of the Bayesian agent in the EWW (apart from their being probabilistic), but I will freely use principles such as the principal principle and the principle of indifference to determine those priors.

The following sentence schemas describe the agent's initial unconditional credences about pits and the Wumpus:

b4. $cr(p_{x,y}) = pr(p)$ if $(x, y) \neq (0, 0)$;

b5. $cr(w_{x,y}) = 1/(s^2 - 1)$ if $(x, y) \neq (0, 0)$;

b6. $cr(p_{0,0}) = cr(w_{0,0}) = 0$.

Clauses b4 state that the agent has an initial credence of $pr(p)$ that (x, y) contains a pit for every room $(x, y) \neq (0, 0)$. Clauses b5 state that, for every room $(x, y) \neq (0, 0)$, the agent has initial credence of $1/(s^2 - 1)$ that (x, y) contains the Wumpus. Clauses b6 state that the agent has an initial credence of 0 that (0, 0) contains a pit or the Wumpus.

The agent's conditional credences about pits and the Wumpus are:

b7. $cr(b_{x,y} \mid \bigvee p_{\pm x,\pm y}) = 1$ and $cr(b_{x,y} \mid \bigwedge\neg p_{\pm x,\pm y}) = pr(rand)$;

b8. $cr(s_{x,y} \mid \bigvee w_{\pm x,\pm y}) = 1$ and $cr(s_{x,y} \mid \bigwedge\neg w_{\pm x,\pm y}) = pr(rand)$;

where $\bigvee\varphi_{\pm x,\pm y}$ abbreviates $\varphi_{x+1,y} \vee \varphi_{x-1,y} \vee \varphi_{x,y+1} \vee \varphi_{x,y-1}$ and $\bigwedge\neg\varphi_{\pm x,\pm y}$ abbreviates $\neg\varphi_{x+1,y} \wedge \neg\varphi_{x-1,y} \wedge \neg\varphi_{x,y+1} \wedge \neg\varphi_{x,y-1}$.[17] Clauses b7 and b8 are a consequence of the fact that if a room is in the neighborhood of a room containing a pit or the Wumpus, then it contains a breeze or a stench with probability 1; else it contains a random breeze or a random stench with probability $pr(rand)$. The Bayesian agent updates her beliefs using conditionalization.

1.2 The nonmonotonic model

The nonmonotonic model of a rational agent has the following features:

d1. The agent's beliefs have only two potential values: 0, 1 (full beliefs);

d2. The agent can withhold beliefs (*null* values);

d3. The agent adopts and withdraws beliefs given adequate reasons and defeaters.

The nonmonotonic agent has doxastic states much less fine-grained than the Bayesian. Instead of continuously many values between 0 and 1, the model only allows for two belief-values (d1): 0 when the agent disbelieves that φ (i.e., she

[17] The Bayesian agent may not hold credences about complex propositions in EWW. In this case, $\bigwedge$ and $\bigvee$ should be seen as meta-linguistic connectives, where $\bigwedge\neg\varphi_{\pm x,\pm y}$ abbreviates the list where $\varphi_{x+1,y}$, $\varphi_{x-1,y}$, $\varphi_{x,y+1}$, and $\varphi_{x,y-1}$ appear all negated and $\bigvee\varphi_{\pm x,\pm y}$ abbreviates the lists of these same atoms where at least one of them appears in the affirmative form (in the same order).

believes that $\neg\varphi$) and 1 when she believes that φ. Another difference is that, while the Bayesian is opinionated, the nonmonotonic agent can withhold beliefs (d2). The withholding of belief is modeled by using a third value (*null*), which marks the absence of belief-values for some proposition φ (i.e., φ is not in the agent's belief-set). Nonmonotonic frameworks are used to model defeasible reasoning (d3), where agents draw and retract conclusions given new information.

The nonmonotonic agent fulfills neither probabilism nor conditionalization as she initially holds *null* values about many propositions and gains new beliefs from investigating the cave (see fn. 15). LP is a nonmonotonic framework with computational applications (Doets 1994). A 'logic program' in LP is a set of conditionals of the form $\varphi \land \neg ab \rightarrow \psi$ ('if φ and nothing is abnormal, then ψ'), where φ is a conjunction of literals, ab is a disjunction of literals ('abnormalities') indexed to a particular conditional, and ψ is a single literal. These conditionals may be seen as licenses for performing inferences (e.g., modus ponens) under certain conditions (e.g., ab is false). If something is abnormal (i.e., ab is true), then $\varphi \land \neg ab$ is false and the modus ponens is blocked. If φ is true but ψ is false, then something is abnormal (ab is true). This framework exhibits a form of negation as failure (the closed world assumption): in some cases, the absence of reasons for the truth of some proposition (e.g., ab) is a reason for its negation. LP may be used to model defeasible reasoning[18].

The nonmonotonic agent is modeled using LP. Her initial beliefs about (0, 0) are $\neg p_{0,0}$ and $\neg w_{0,0}$. For all $(x, y) \neq (0, 0)$, $p_{x,y}$ and $w_{x,y}$ are initially assigned the *null* value. The agent uses the following conditionals about pits:

d4. $b_{x,y} \land \neg ab \rightarrow p_{\pm x,\pm y}$;

d5. $\neg b_{x,y} \rightarrow \neg p_{\pm x,\pm y}$,

where $\varphi_{\pm x,\pm y}$ is either $\varphi_{x+1,y}$, $\varphi_{x-1,y}$, $\varphi_{x,y+1}$, or $\varphi_{x,y-1}$. Clauses d4 state that perceiving a breeze in a room is a defeasible reason for believing that there are pits in the adjacent rooms, whereas ab expresses the possibility that $b_{x,y}$ is a random breeze. Clauses d5 state that perceiving the absence of breezes in a room is a conclusive reason for believing that there aren't pits in the adjacent rooms. Clauses d5 express rebutting defeaters for the belief that $p_{x,y}$ based on d4.

[18] If $\varphi \land \neg ab \rightarrow \psi$ is a clause in the agent's logic program, then believing that φ is a reason for believing that ψ. If ab has zero disjuncts, then this is a conclusive reason. Otherwise, it is a defeasible reason. We could say that if φ' is a reason for $\neg\psi$, then the belief that φ' rebuts the belief that ψ on the previous basis; and that if φ' is a reason for ab, then the belief that φ' undercuts the belief that φ on that same basis. LP often blurs the distinction between rebutting and undercutting defeaters because a reason for $\neg\psi$ is often a reason for ab (by modus tollens) and a reason for ab may be a reason for $\neg\psi$ (closed world).

The agent uses the following conditionals about the Wumpus:

d6. $s_{x,y} \wedge \neg ab \rightarrow w_{\pm x,\pm y}$;

d7. $s_{x,y} \wedge \neg ab \rightarrow \neg w_{z,w}$ for $(z, w) \neq (\pm x, \pm y)$;

d8. $w_{x,y} \rightarrow \neg w_{z,w}$ for $(z, w) \neq (x, y)$;

d9. $\neg s_{x,y} \rightarrow \neg w_{\pm x,\pm y}$,

where $(z, w) \neq (\pm x, \pm y)$ abbreviates $(z, w) \neq (x+1, y) \wedge (z, w) \neq (x-1, y) \wedge (z, w) \neq (x, y+1) \wedge (z, w) \neq (x, y-1)$. Clauses d6 state that perceiving a stench in a room is a defeasible reason for believing that there are Wumpuses in each of the adjacent rooms, whereas ab expresses the possibility of $s_{x,y}$ being a random stench. Clauses d7 state that perceiving a stench in a room is a defeasible reason for believing that the Wumpus is nowhere else but in the adjacent rooms, where ab has the same interpretation. Clauses d6 and d7 may rebut each other. Clauses d8 are a consequence of the fact that there is only one Wumpus: believing that the Wumpus is in a specific room is a conclusive reason for believing that it is not elsewhere. Clauses d9 state that perceiving the absence of stench in a room is a conclusive reason for believing that the Wumpus is not in the adjacent rooms.

Integrating these conditionals in a sound pattern of inference is not trivial. Some researchers have proposed procedural semantics for LP, which are based on the fixpoints of an update operator.[19] For example, given a belief-set M and a logic program P, Stenning and van Lambalgen (2008, 194) define an application of the update operator T to M (i.e., $T(M)$) as:

(a) $T(M)(\psi) = 1$ iff there is a clause $\varphi \rightarrow \psi$ in P such that $M \vDash \varphi$;

(b) $T(M)(\psi) = 0$ iff there is a clause $\varphi \rightarrow \psi$ in P and for all such clauses, $M \vDash \neg\varphi$;

(c) $\mathrm{T}(M)(\psi) = u$ otherwise (Stenning and van Lambalgen 2008, 194),

where their third value, u (currently indeterminate), is related to my use of the *null* value. The updated belief-set $T(M)$ must be among the fixpoints of T (i.e., $T(T(M)) = T(M)$).[20] I will focus on the 'most well-supported' fixpoint, where a belief-set is more well-supported than another iff it is constructed by triggering more conditionals d6.[21]

---

[19] A fixpoint c of a function $f(x)$ is such that c belongs to the domain and the codomain of $f(x)$, and $\mathrm{f}(c) = c$.

[20] There may be different applications of T to a model M that are fixpoints of T. Suppose that $M \vDash s_{0,1}, s_{2,1}$. There are at least three applications $T(M)$ that are fixpoints of T: (1) $T(M)(w_{0,2}) = u$, $T(M)(w_{2,2}) = u$; (2) $T(M)(w_{0,2}) = 1$, $T(M)(w_{2,2}) = 0$; and (3) $T(M)(w_{0,2}) = 0$, $T(M)(w_{2,2}) = 1$.

[21] Suppose that $M \vDash s_{0,1}, s_{1,2}, s_{3,1}, \neg s_{0,0}, \neg s_{1,0}, \neg s_{1,1}, \neg s_{2,0}, \neg s_{2,1}, \neg s_{2,2}, \neg s_{3,1}$. In this case, the most well-supported model is one such that $T(M)(w_{0,2}) = 1$ and $T(M)(w_{3,2}) = 0$ because it triggers two conditionals d6 ($s_{0,1} \wedge \neg ab \rightarrow w_{0,2}$ and $s_{1,2} \wedge \neg ab \rightarrow w_{0,2}$) while the other models trigger at most one.

2. The simulation

I have implemented AI agents based on the Bayesian and nonmonotonic models in a computer simulation of EWW. The agents have four modules: perception, memory, practical and epistemic cognition. They share perception, memory, and part of practical cognition, but they have different epistemic cognitions (based on the Bayesian and nonmonotonic models respectively). The modules of perception and memory are straightforward. Perception simply receives percepts from the environment (e.g., the presence of a breeze), encodes them in the format of beliefs, and sends these beliefs to memory. Beliefs are formatted as one matrix of numbers for each of breeze, stench, pit, Wumpus, and gold, where a position in a matrix represents the corresponding room in the cave.[22] Memory simply stores these matrices and makes them available for practical and epistemic cognition.

The role of practical cognition is to construct and execute plans (sequences of actions). Practical cognition checks whether there is a plan being executed. If there is, it executes the next action in the plan. Otherwise, it constructs a plan and executes its first action. In constructing a plan, practical cognition chooses the sub-goal that she attributes higher value among 'grab the gold,' 'get out of the cave,' and 'move to a fringe room (x, y),' where fringe rooms are non-visited rooms in the neighborhood of a visited room. The value of moving to a room (x, y) is calculated differently by each agent.[23] The value of grabbing the gold (if the room contains it) is 1000 and the value of getting out of the cave is 0. The plan for moving to a room is constructed using Dijkstra's pathfinding algorithm, with cost computed as the number of actions (Russell and Norvig 2020, 84); the plan for grabbing the gold is composed solely of the action 'grab the gold;' the plan for getting out of the cave is a plan for moving to (0, 0) concatenated with the action 'get out of the cave.'

Epistemic cognition updates the agent's beliefs, given the new information. The agents differ in how they encode and update beliefs about pits and Wumpus.[24] The Bayesian agent encodes those beliefs using floating-point numbers between 0 and 1 and updates them using causal Bayesian networks and joint probability

---

[22] For example, the position (x, y) in the pit matrix represents $p_{x,y}$, where the value in that position represents the agent's belief-value for $p_{x,y}$.

[23] The Bayesian agent calculates the expected utility of moving to a room (x, y) as $-1000 \times (cr(p_{x,y}) + cr(w_{x,y})) + 1000 \times cr(g_{x,y})$. The nonmonotonic agent assigns -1000 if $b(p_{x,y}) = 1$ or $b(w_{x,y}) = 1$; she assigns $1000 \times b(g_{x,y})$ if $b(g_{x,y}) \neq null$; and an intermediate value (e.g., 500) otherwise. This is an application of the minimax principle.

[24] For simplicity, both agents encode and update beliefs about breezes, stenches, and gold in the same all-or-nothing way: if the agent perceives a breeze [stench, gold] in a room, then perception adds a 1 to the corresponding position in the breeze [stench and gold] matrix; otherwise, it adds a 0.

tables, in a standard algorithm (e.g., Russell and Norvig 2020, 134). The nonmonotonic agent encodes beliefs about pits and Wumpus using Boolean values. The algorithm for pits is very simple. If you perceive a breeze in a position (x, y), then write a 1 in all positions ($\pm x$, $\pm y$) of the pit matrix that do not already contain a 0 (d4). If you perceive the absence of breezes in a position (x, y), then write a 0 in all positions ($\pm x$, $\pm y$) of the pit matrix (d5). The algorithm for Wumpus keeps track of its viable locations and the number of cues about the Wumpus. Do the following when you perceive a new stench. If the stench matrix contains a 1 in a position (x, y), then add +1 to the value in the positions ($\pm x$, $\pm y$) of the support matrix (d6) and add -1 from the value in the other (d7 and d8). If the stench matrix contains a 0 in a position (x, y), then write a large negative number in the positions ($\pm x$, $\pm y$) of the support matrix (d9). If there is a position in the support matrix with a value that is strictly higher than the other, assign 1 to that position in the Wumpus matrix and 0 to the other. Otherwise, assign *null* to the positions with the highest value and 0 to the other.

The simulation of EWW generates a random cave for each trial and works as a loop. In each iteration, the cave outputs precepts for the agent, who updates her beliefs and returns an action to the cave, which updates its state given that action. The loop stops when the agent dies or gets out of the cave. I have simulated three conditions (c1-c3). In c1, the probability of pits is fixed at 0.1 ($pr(p) = 0.1$) and the probability of random breezes [stenches] is fixed at 0.01 ($pr(rand) = 0.01$). I have run the simulation using caves with dimensions $s \times s$ (size s), from $s = 2$ to $s = 10$ (incrementing by 1). In c2, I have run simulations with $pr(rand)$ varying from 0 to 0.5 (incrementing by 0.05), where $pr(p)$ is fixed at 0.1. In c3, I have run simulations with $pr(p)$ varying from 0 to 1 (incrementing by 0.1), where $pr(rand)$ is fixed at 0.01. In c2 and c3, I have used 'medium' caves ($s = 6$). I have run up to 50,000 trials for each configuration of the cave in each condition and averaged the results. The code was written in Java and the graphs were plotted in Grace.

2.1 Results

The results for the Bayesian and nonmonotonic agents are depicted in Figures 1 and 2 (respectively). The linear and the Brier scores return the same β-values for both agents because $t - f$ returns the same values for these scores and n is constant within each trial.[25] The linear and the Brier scores do not return the same α-values for the Bayesian agent because $t + f$ does not return the same values for these

[25] Let $y = \varepsilon_\varphi$ and $x = 1 - \varepsilon_\varphi$. The function $t - f$ returns the same values for both agents in terms of linear and Brier scores because $(x + y) \times (x - y) = x^2 - y^2$ and $x + y = 1$, which entails that $x - y = x^2 - y^2$.

scores, but this is the case for the nonmonotonic because $0 = 0^2$ and $1 = 1^2$. The log values were normalized by $-\ln(c)$, as to depict all scores in the same scale.[26] In this case, the three scores return the same α- and β-values for the nonmonotonic agent because $1 = 1^2 = (-\ln(c))/(-\ln(c))$ and $0 = 0^2 = (-\ln(1))/(-\ln(c))$. I will discuss f and the time requirements in Section 3.

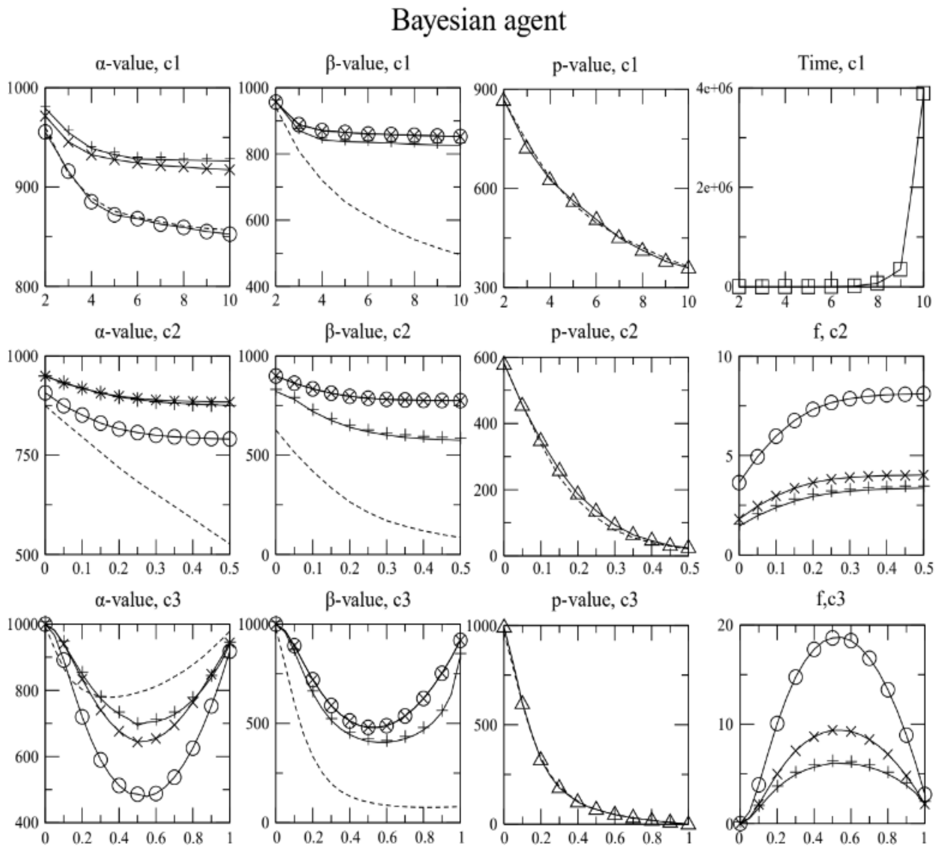

Fig. 1: Top to bottom: results of the Bayesian agent for c1, c2, and c3. Left to right: the α- and β-values, using linear (○), Brier (×), and log scores (+); and the p-value (△). The results for the nonmonotonic agent are depicted in dashed lines. The 4th column depicts the time requirements (number of variable-changes) in c1 (□) and f in c2 and c3 using the same scoring rules.

Condition c1 represents the 'normal' condition, in the sense of the evidence being relatively informative and relatively accessible for the agents. In c1, as the cave

[26] This was unnecessary for the α-values because the divisor $t + f + c$ already works as a normalizer.

grows from $s = 2$ to $s = 10$, the α-values of the Bayesian agent start stabilizing around 853 (linear), 917 (Brier), or 926 (log). The α-values of the nonmonotonic agent start stabilizing around 857. The Bayesian's α-values are higher than the nonmonotonic's when the measure uses a Brier or log score, but they are not when it uses a linear score (they are about the same, with similar curves). The β-values of the Bayesian agent start stabilizing around 852 (linear and Brier), or 827 (normalized log). The β-values of the nonmonotonic agent go down to 495 and start to stabilize at a much lower level (250).[27] The Bayesian's β-values are higher than the nonmonotonic's independently of the scores. The results of α- and β-values diverge, as the nonmonotonic agent is much closer to the Bayesian in terms of α-values (especially, in linear score) than she is in terms of β-values. However, the disagreement is not very strong as all graphs have the same general shape: they fall fast until stabilizing at different levels. Furthermore, the (higher) epistemic value of the Bayesian agent does not reflect in a higher practical value: both agents have p-values that reach around 360 when the cave reaches $s = 10$ (with very similar curves).

In c2, I have tested how the agents react when *pr*(*rand*) varies from 0 to 0.5 (the tendencies do not change from 0.5 to 1). The higher the *pr*(*rand*), the less information observed breezes [stenches] carry about the position of pits [Wumpus].[28] The α-values of the Bayesian agent stabilize around 775 (linear), 874 (Brier), or 860 (log) when *pr*(*rand*). The nonmonotonic agent performs much worse: her α-values reach 527 when *pr*(*rand*) = 0.5. The same goes for β-values. The Bayesian's β-values stabilize around 775 (linear and Brier), or 575 (normalized log), while the nonmonotonic's fall to 86 when *pr*(*rand*) = 0.5. These results highlight the strength of the Bayesian model: the ability to deal with uncertain evidence. The Bayesian agent considers the value of *pr*(*rand*) in updating beliefs, whereas the nonmonotonic can only conclude blindly that there are pits (not so much for Wumpus) in the surroundings from the presence of breezes (random or not). The disagreement between the α- and β-values is also mild in c2, where the nonmonotonic's β-values fall faster than her α-values. The higher epistemic value of the Bayesian agent affects her practical value: although both p-values converge to around 22 when *pr*(*rand*) =

---

[27] The β-values of the nonmonotonic agent are around 250 from $s = 50$ to $s = 150$ (her α-values are around 845 in those cases). These results are less reliable because they were averaged over fewer trials.

[28] For example, suppose that the agent has just arrived in the cave. If *pr*(*p*) = 0.1 and *pr*(*rand*) = 0.01, then conclusive evidence about $b_{0,1}$ transmits 0.27 bits of information about $p_{0,2}$. If *pr*(*p*) = 0.1 and *pr*(*rand*) = 0.05, then conclusive evidence about $b_{0,1}$ transmits only 0.24 bits of information about $p_{0,2}$.

0.5, the Bayesian's p-value (94) is 20% higher than the nonmonotonic's (78) when *pr*(*rand*) = 0.3.

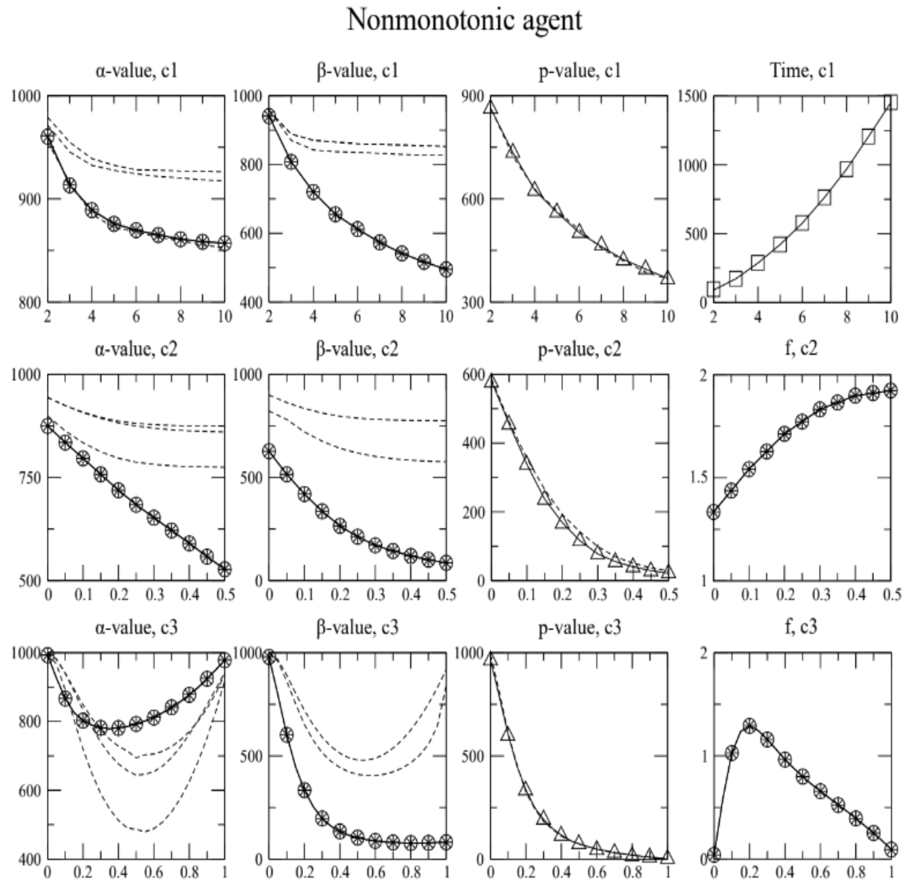

Fig. 2: Top to bottom: results of the nonmonotonic agent for c1, c2, and c3. Left to right: the α- and β-values, using linear (○), Brier (×), and log scores (+); and the p-value (△). The results for the Bayesian agent are depicted in dashed lines. The 4th column depicts the time requirements (number of variable-changes) in c1 (□) and f in c2 and c3 using the same scoring rules.

In c3, I have tested how the agents react when *pr*(*p*) varies from 0 to 1. The increase in *pr*(*p*) increases the difficulty of gathering the evidence because the higher the number of pits, the riskier it is to explore the cave. The α-values of the Bayesian reach around 918 (linear), 944 (Brier), or 939 (log) when *pr*(*p*) = 1, in a u-shape with the lowest values around 479 (linear), 644 (Brier), or 695 (log) when *pr*(*p*) = 0.5. The nonmonotonic's α-values reach around 978 when *pr*(*p*) = 1, in a u-shape with the lowest values around *pr*(*p*) = 0.35 (778). The nonmonotonic agent performs much better than the Bayesian in terms of α-values when *pr*(*p*) > 0.35, which is surprising because the Bayesian (but not the nonmonotonic) agent fulfills probabilism and conditionalization. The explanation for this result is that it is too risky to gather information when *pr*(*p*) > 0.35 when the nonmonotonic's *null* values about the unobserved rooms are worth more α-value than the Bayesian's indifferent beliefs.[29] The situation is not the same regarding β-values. The β-values of the Bayesian reach around 918 (linear and Brier) or 839 (normalized log) when *pr*(*p*) = 1, in a u-shape with the lowest values around 479 (linear and Brier) and 404 (normalized log) when *pr*(*p*) = 0.5. The nonmonotonic's β-values do not have a u-shape, as they fall to 80 when *pr*(*p*) = 1. Neither the nonmonotonic's higher α-values nor the Bayesian's higher β-values affect their p-values (both fall to 0 with similar curves).

In c1 and c2, the functions α and β only disagree on the speed at which the curves decrease and at which level they converge. In c3, they disagree on the very shape of the graphs: the α- but not the β-values of the nonmonotonic agent have a u-shape. They also disagree on which agent achieves higher epistemic value: the α-values of the nonmonotonic agent are higher than the Bayesian's when *pr*(*p*) > 0.35, but the β-values of the Bayesian are always higher than the nonmonotonic's. These results are related to the relative accessibility of evidence in c1-c3. The functions α and β generally agree in c1 and c2, where the evidence is accessible, independently of whether it is informative (c1) or misleading (c2). They disagree in c3, where the evidence is increasingly inaccessible when *pr*(*p*) > 0.35. Consequently, the functions α and β generally agree on their evaluations of how agents reason from the available evidence (c1 and c2), but they disagree about what they should do in the absence of evidence (c3). In the absence of evidence, the agents tend to maintain their initial belief-values. The initial belief-values of the Bayesian agent (e.g., about the Wumpus) are often applications of the principle of indifference; the nonmonotonic's are mostly *null*. The use of *null* values may be interpreted as the withholding of beliefs. Then the disagreement between the

[29] Suppose that the Wumpus may be in three different rooms, but that the evidence about its position is inaccessible. The Bayesian agent would hold indifferent credences of 0.33 about these positions, while nonmonotonic would withhold beliefs (*null* values). The nonmonotonic agent would have a higher α-value regarding those positions because $\alpha(t, f) > \alpha(t + 1.67, f + 1.33)$ when $t > f$, which is usually the case in EWW. The Bayesian would have a higher β-value because $\beta(t, f) < \beta(t + 1.67, f + 1.33)$. These inequalities use linear scores, but the same holds for the Brier and log scores.

functions α and β in c3 may be interpreted as they prescribing different attitudes in the absence of evidence: the function α prescribes the withholding of beliefs; the β prescribes indifference (see fn. 28).

3. Discussion

The results from EWW may affect the justificatory force of the maximization arguments because the nonmonotonic agent (who fulfills neither probabilism nor conditionalization) achieves higher α-values than the Bayesian in c3 (when $pr(p) > 0.35$) and similar p-values in the three conditions. Whether this is the case depends on the adoption of the function α as the measure of epistemic value and of a set of assumptions different from a1-a3. I will defend these choices, starting by proposing three arguments for the adoption of the function α. The first argument is pragmatic. In c3, the α-values of the nonmonotonic agent are higher than the Bayesian's when $pr(p) > 0.35$, but the β-values of the Bayesian are always higher than the nonmonotonic's. When $pr(p) > 0.35$, the Bayesian agent holds many more beliefs than the nonmonotonic (e.g., about the Wumpus). These 'extra beliefs' do not have practical value because the p-values of the Bayesian and nonmonotonic agents are similar in this condition. While the function β awards the Bayesian agent for holding beliefs without practical value, the function α awards the nonmonotonic agent for not holding those beliefs. From a pragmatic point of view, the function α is correct.[30] The pragmatic argument is controversial because it presupposes a form of pragmatic encroachment in the notion of epistemic rationality, whereas the epistemological orthodoxy most often separates epistemic rationality from practical rationality. Nevertheless, there are investigations that consider the possibility of pragmatic encroachment in the notion of epistemic rationality (e.g., Gao 2021).

The second argument is cognitive. A fundamental aspect of our cognitive situation is that human beings are in the finitary predicament of having fixed limits on their cognitive capacities and the time available to them (Cherniak 1986, 8). Epistemic rationality seems to require from 'finite reasoners' (those in the finitary predicament) a form of cognitive parsimony: to convert scarce cognitive resources (memory and time) into epistemic value efficiently. There are no interesting limits on the amount of information that we can hold in long-term memory (Dudai 1997), but the learning of new information can adversely affect our capacity to retrieve old information in a process of interference (Baddeley et al. 2020, 291).

[30] In c1, both agents are able to explore the cave and end up holding a similar number of beliefs. The Bayesian agents may hold extra beliefs in c2, but these beliefs do affect her p-value.

This cognitive limitation is modeled by adopting not an upper bound for the size of B but a diminishing reward for the amount of (truthful) commitment, which is done by the function α because $\alpha(t + 2x, f) - \alpha(t + x, f) < \alpha(t + x, f) - \alpha(t, f)$, but not by the β because $\beta(t + 2x, f) - \beta(t + x, f) = \beta(t + x, f) - \beta(t, f)$, where $x > 0$. This difference between the functions α and β explains their incompatible prescriptions in c3, where the nonmonotonic agent's withholding of beliefs in the absence of evidence is more cognitively parsimonious than the Bayesian's indifference (see Dantas 2022, for a discussion). If the notion of epistemic rationality should regard finite reasoners, then the function α should be preferred over the function β. The Bayesian model is also more cognitively demanding in time (see g. 1 and 2, 1st line and 4th column).[31]

The third argument is epistemic, as it draws on veritistic notions. The function α should be preferred over the function β because of how they relate to the most used measure of epistemic disvalue (inaccuracy). The goal of believing truths and avoiding falsehoods is two-fold, but I believe that its second part should take precedence over the first because of 'the problem of contradictory pairs.' The function β evaluates equally an agent who believes (as to the same degree) both propositions in a contradictory pair and one who believes neither because $\beta(t, f) = \beta(t + x, f + x)$ when $x > 0$, but the second agent should be evaluated higher than the first regarding these propositions. This problem could be avoided by attributing weights R and W to t and f (respectively) where $R < W$ (Fitelson and Easwaran 2015, 83), but this is to attribute priority to the goal of avoiding falsehoods. The function α deals more naturally with this problem because $\alpha(t, f) > \alpha(t + x, f + x)$ when $x > 0$ and $t > f$, but this is to attribute priority to the goal of avoiding falsehoods when $t > f$.[32] The nonmonotonic agent has a much lower inaccuracy than the Bayesian in c3, as depicted in figures 1 and 2, 2nd and 3rd lines, 4th

[31] Time requirements were measured as the number of changes in the value of variables used in update procedures. The time requirements of the nonmonotonic agent grow polynomially on the size of the cave; the Bayesian's grow exponentially (see fig. 1 and 2, 1st line and 4th column). There are algorithms for Bayesian inference that are polynomial in time (e.g., belief propagation, see Pearl 1986), but these only work for singly connected networks, whereas EWW requires multiple-connection. Inference from multiply connected Bayesian networks is NP-hard (Cooper 1990). Dantas (2017) argues that rationality demands finite reasoners to implement polynomial patterns of inference when they are available.

[32] The situation is the opposite when $t < f$, where the agent is an anti-expert about her agenda and believes a contradiction may serve as a flag for revising her beliefs (Dantas, 2022, discusses this point).

column.[33] If avoiding falsehoods takes precedence over believing truths, then the evaluation of the function α in c3 should be preferred over that of the function β.

The function β awards investigation more straightforwardly than the function α because the β-value, for example, of a new true (maximal) belief, is fixed ($1/n$), but its α-value varies inversely wrt the overall truthfulness of the agent's belief-set. If the agent holds mostly true beliefs (and many of them), the α-value of investigating will be only marginal. This feature of the function β seems welcome, but I believe that the way that the function α awards investigation is more appropriate for c3. Figures 1 and 2, 2nd and 3rd lines, 1st, 2nd, and 4th columns show that the graphs for α-values have approximately the same shape as those of inaccuracy, although vertically reflected because α is a value and inaccuracy a disvalue (the same holds in c1). This does not happen with β-values, especially in c3, where the graphs for α-values and f have a u-shape for both the Bayesian and the nonmonotonic agents, but the graphs for β-values only have a u-shape for the Bayesian. This discrepancy is not welcome in c3, where the agents are not able to investigate, and avoiding error (f) is even more important than seeking the truth (t). The conservativeness of the function α regarding the measure of inaccuracy seems to return the correct evaluation in c3. The functions α and β agree in c1 and c2. If my arguments are good, the first should be preferred when they disagree (c3). The function α should be preferred as a measure of epistemic value in general, especially regarding finite reasoners.

I have already commented on a1-a3 and EEW's alternative assumptions (a1′ and a2′), but I will return to this point before addressing the DBAs and accuracy arguments. Assumption a1 is reasonable, but it is reasonable to risk value at the present if this may result in a gain of value in the future. This remark points to the reasonableness of a1′, where the values are measured only at the end of each trial. 'Long run' evaluations of this sort may raise the concern that finite reasoners cannot always wait to fulfill their goals. They may also hide short run vices of the agents' reasoning processes. These concerns may be mitigated by paying attention to whether the agent's reasoning processes enable them to fulfill their goals. The practical performance of the Bayesian agent could be seen as the gold standard because she minimizes expected inaccuracy in every belief update (given a3) and

[33] This result does not contradict the conclusions of the accuracy arguments because the Bayesian and nonmonotonic agents are not opinionated over the same agenda. In c1 and c2, the nonmonotonic agent also has a lower inaccuracy, but both functions agree that the Bayesian achieves a higher epistemic value. In c1, as the cave grows from s = 2 to s = 10, the nonmonotonic's inaccuracy grows from 0.1 to 3, whereas the Bayesian's grows from 0.2, 0.1, or 0.3 to 15, 7, or 27 (linear, Brier, and log scores respectively).

acts to maximize expected practical value (see fn. 22). The agents achieve similar p-values in the three conditions of EWW, which suggests that their reasoning processes are equally supporting the fulfillment of their goals (except in c2, where all measures show an unmistakable advantage for the Bayesian). These results suggest that both agents are able to fulfill their golds and the measuring of values only at the end of each trial in EWW is not hiding short run vices of the nonmonotonic agent. The related assumption that rationality requires agents not to make choices that lead to a sure loss of any practical value is also reasonable, but it depends on idealizing environments that do pose trade-offs between goals (I will return to this point in discussing the DBAs).

Assumption a2 is also reasonable because rationality is usually assumed to have a subjective character (e.g., Wedgwood 2015, 221), which is related to what the agent can know from the available evidence. Philosophers often distinguish between subjective and objective norms, where subjective but not objective norms are sensitive to which evidence is available to the agents. The veritistic norm is objective because the truth-values of their beliefs are often not transparent for the agents. This is why the Bayesian epistemologists supplement their measures of value with the notion of a sure loss or principles of decision theory, to consider rationality's subjective character. The resulting Bayesian norms are subjective. The expected values of EWW (the mean of actual values) are objective and I propose that their maximization is by itself relevant to the agent's rationality (this is the assumption a2'). This focus highlights interesting aspects of rationality. Together with a1′, the objective expected values may be seen as evaluating the stable beliefs of agents, i.e., those that result from all the evidence available to her (including that available in the environment). The very notion of evidence being 'available' to a finite reasoner depends on the amount cognitive resources available to investigate. The evidence accessible through investigating the environment may also be 'available' in the relevant sense given a1′ and a2′. This discussion suggests an ecological notion of rationality (Todd and Gigerenzer 2007), where the rationality of an agent depends on how she copes with her surroundings.

If a1 and a2 are reasonable assumptions, a3 is not. Opinionation is assumed as a simplifying idealization in EUT because of the limitations of the measure of inaccuracy (see sec. 1), but the results of EWW suggest that non-opinionated epistemic practices may be worth more epistemic value (α-value) than opinionated ones (e.g., when evidence is not accessible). In this context, assuming a3 introduces a bias towards Bayesianism because it artificially eliminates these practices 'from the competition.' This assumption also defeats the purpose of veritism, of supporting norms of epistemic rationality from veritistic considerations. The function β might be used for supporting opinionation from veritistic considerations. For example, indifference is worth more β-value than the withholding of belief when the agents lack evidence about large exhaustive sets of exclusive propositions (e.g., fn. 28).[34] Since the function β abstracts from the cognitive limitations of agents, its use for supporting opinionation would vindicate the Bayesian model as describing an ideal reasoner (i.e., a reasoner without cognitive limitations). This seems correct because the Bayesian ideal reasoner exhibits a form of logical omniscience[35] and cannot forget.[36]

Choices that lead to a sure loss should be avoided, but rational agents may be forced to make such choices when the environment poses trade-offs between goals. For example, Douven (2013) simulates a Bayesian and an explanationist agent, who updates her credences using inference to the best explanation and is vulnerable to DBs. These agents watch finite sequences of coin tosses and must estimate the coin's bias. Douven sets up a game, where the first correct estimation yields a point to the estimator, and an incorrect one yields a point to the opponent (the trade-off is between speed and accuracy). The Bayesian agent always loses because the explanationist converges to an answer faster. Part of the DBAs' appeal results from non-Bayesian agents being in the position to know that their choices lead to a sure loss in DB-environments. But the same is true about Douven's game because the Bayesian ideal reasoner is in the position to know anything that we can only learn from computer simulations (it is not surprising that slow convergence reasoning processes are prone to lose in environments that award speed of convergence). Then why would DB-environments be so central to rationality? I believe they are not, especially because Dutch bookies only occur as fictional characters in philosophers' tales (Douven 2013, 431). Although (ecological) rationality is relative

[34] This is not the case for exhaustive pairs of exclusive propositions (e.g., contradictory pairs) because $\beta(t + x, f + x) = \beta(t, f)$. How large these sets must be hangs on the relative sizes of the weights $R < W$, which could be adopted to deal with the problem of contradictory pairs (see the epistemic argument).

[35] The Bayesian ideal reasoner fulfills probabilism and conditionalization. The probability's axiom of normality entails that she must be certain of (i.e., hold maximum credence about) every logical truth. If she comes to learn some evidence with certainty, then the axioms of normality and finite additivity require her to be certain of every logical consequence of that evidence (see Garber 1984, 104).

[36] If the Bayesian ideal reasoner ever reaches certainty on a proposition, then the axioms of probability assure that this certainty will be maintained after any subsequent update by conditionalization. Consequently, the Bayesian ideal reasoner cannot be certain that she is currently having spaghetti for dinner, but forget this irrelevant fact a year later (i.e., loose certainty about it, see Talbott 1991, 139).

to sets of conditions (e.g., the Bayesian model copes better with misleading evidence), the relevance of these conditions depends on how likely they are to occur to reasoners like us (i.e., how 'normal' they are).

The results of EWW impact the accuracy arguments more than the DBAs because the nonmonotonic agent (who fulfills neither probabilism nor conditionalization) achieves higher α-values than the Bayesian in c3 (when $pr(p) > 0.35$). This result suggests that the Bayesian ideal reasoner is only guaranteed to maximize epistemic value in situations where a1-a3 hold. In those situations, she minimizes (expected) inaccuracy. But this result does not hold when assumption a3 is relaxed because the nonmonotonic agent achieves lower inaccuracy in the three tested conditions (this happens even in c2, in which the Bayesian model achieves higher α- and β-values). The results of EWW impact more directly the argument for probabilism than that for conditionalization because, when $pr(p)$ is high, the agents are not able to gather evidence and update beliefs (they tend to maintain their initial beliefs). However, the argument for conditionalization is also affected. The nonmonotonic's higher α-values were explained by her withholding of beliefs in the absence of evidence (see fn. 28) and there is no update from the absence of beliefs to a belief-value by conditionalization. An agent who maximizes α-value in the absence of evidence cannot update beliefs by conditionalization. Conversely, the Bayesian ideal reasoner's failure in maximizing α-value was explained by her use of the principle of indifference, which is how she maintains opinionation in the absence of evidence. The inefficiency of the Bayesian model was located in its demand for opinionation (see fn. 15).

An epistemic interpretation of the nonmonotonic agent's *null* values is that she is suspending judgment (i.e., adopting a neutral stance toward the truth-value of a proposition). A difficulty with this interpretation is that the mere lack of belief-values about a proposition is not sufficient for suspension.[37] There are different notions of suspension, two of which may be used to interpret the nonmonotonic's use of *null* values: the interrogative view (Friedman 2015), in which suspending about a proposition involves actively inquiring about its truth, and the anti-interrogative view (Lord 2020), in which suspending about a proposition involves forgoing evidence about its truth. Intuitively, the nonmonotonic agent does the first when she attributes *null* values at the beginning of each task and when there is evidence of Wumpus in different rooms. She does the second at the end of each task when it is too risky to gather evidence about the

[37] For example, when an agent never considered a proposition, we say that she does not hold a doxastic attitude towards it and not that she holds an attitude of committed neutrality (suspension) towards it (Friedman 2013, 167).

remaining *null* propositions. There is a notion of suspension in which the Bayesian agent may be said to be suspending in c3: the credal view (Sturgeon 2008), where suspending is related to the holding of middling credences. The Bayesian agent holds middling credences about pits when $pr(p) = 0.5$, for example, because her initial (indifferent) credences are of 0.5 and it is difficult to gather evidence when $pr(p) > 0.35$. This form of suspension does not result in higher α-values (but results in higher β-values). This is a matter for another paper.

4. Conclusions

The conclusions of the maximization arguments follow from assumptions a1-a3, but the results of EWW show that these conclusions do not hold when an alternative set of reasonable assumptions is assumed. For example, the results from EWW show that there is a reasonable notion of maximization of epistemic value (α-value, without the assumptions a1-a3) in which a nonmonotonic agent, who does not fulfill the Bayesian norms, maximizes epistemic value. The nonmonotonic agent also achieves the same amount of practical value in the 'normal' condition of c1, even though she is vulnerable to DBs. These results suggest that which assumptions and measures of epistemic value should be used in an investigation of rationality is open to discussion. I have argued that a1 and a2 are reasonable assumptions, but the alternative assumptions of EWW are also reasonable. The situation is different with a3 (opinionation), which, I have argued, is not a reasonable assumption. Researchers such as Gigerenzer and Gaissmaier (2011) have already investigated conditions under which Bayesian reasoning is epistemically outperformed by simple reasoning heuristics.[38] Their claims are corroborated by our results, especially because the nonmonotonic's algorithm for Wumpus implements a tallying heuristic (Gigerenzer and Gaissmaier 2011, 469). The results of EWW are interesting because they locate the focus of epistemic inefficiency of the Bayesian model in its demand for opinionation. Forgoing this assumption is difficult for EUT because inaccuracy cannot be used to compare agents with different numbers of beliefs, which may be done with the functions α and β.

The preceding discussion motivates the distinction between two different projects about rationality. The first is the Bayesian project, which relies on the assumptions a1-a3 and uses functions such as the function β for measuring

[38] "In a number of large worlds, simple heuristics were more accurate than standard statistical methods that have the same or more information. These results became known as 'less-is-more effects': there is an inverse-U-shaped relation between the level of accuracy and amount of information, computation, or time. There is a point where more is not better, but harmful" (Gigerenzer and Gaissmaier 2011, 453).

epistemic value. These assumptions enable the focus on analytical methods within this project (see below) because they introduce several simplifications. For example, the practical side of a1 requires idealizing that environments do not pose trade-offs between goals. These environments are 'cost-free' in the sense that no practical cost is associated with the gathering of evidence. This allows the Bayesian project to separate the epistemic from the practical. This project also abstracts from the cognitive limitations of finite reasoners because the function β does not consider these limitations (the value of new true beliefs does not depend on the agent's belief-set). The assumption of opinionation consolidates this state of affairs because it forces the agents' belief-sets to be fixed in size. The goal of the Bayesian project is to describe an ideal reasoner, but it might still concern finite reasoners (indirectly). The strategy would be to propose a model as an ideal reasoner whom we should strive to approximate (Leitgeb 2014, fn. 3). Approximating the Bayesian ideal reasoner would be beneficial for finite agents, even when they cannot be 'fully' rational. For example, De Bona and Stael (2017, 2018) show that, in doing so, finite reasoners are worth more epistemic value and become less vulnerable to DBs. The Bayesian project is an axiomatic approach to rationality, where optimization is related to approximating the model.

The second project is the one that I am proposing in this paper, which measures epistemic value using the function α and drops the assumptions a1-a3. The goal of this project is to investigate the rationality of finite reasoners, where the function α considers their cognitive limitations. The dropping of a1 and a2 has consequences for the methods of the investigation because these are the assumptions that enable the use of analytical methods in the Bayesian project. For example, a2 enables the consideration of all possible truth-values of the agent's beliefs instead of their actual value. The study of actual values may hardly be carried out analytically. For example, the choice of a situation to be the actual (as to calculate actual values) would be unmotivated. This difficulty is avoided in computational epistemology, where the actual values are related to randomly generated environments. I believe that it might be a consequence of the no-free-lunch theorems (Wolpert and Macready 1997) that no general model of a rational agent will come out as maximizing practical or epistemic value in every environment. The rationality of a finite reasoner would depend not on how she approximates a model but on how she is able to exploit the features of her environment given her cognitive limitations. These are two forms of investigating two different notions of rationality. The first is an optimality-oriented axiomatic approach that abstracts the cognitive limitations of finite reasoners. The second is an ecological approach that considers those limitations. These projects have different methods and goals and should both be carried out, although the assumption of opinionation still calls for a justification from Bayesian epistemologists.

References

Babic, B. 2019. "A theory of epistemic risk". *Philosophy of Science* 86(3): 522–550.

Baddeley, A., M. Eysenck, and M. Anderson. 2020. *Memory* (3rd ed.). Routledge.

Bourne, R., and S. Parsons.1998. "Propagating probabilities in System P". *Proceedings of the Eleventh International Florida Artificial Intelligence Research Society Conference*: 440-445.

Carr, J. 2015. "Epistemic expansions". *Res Philosophica* 92(2): 217–236.

Cevolani, G., and R. Festa. 2021. "Approaching deterministic and probabilistic truth: A unified account". *Synthese* 199(3-4): 11465–11489.

Cherniak, C. 1986. *Minimal Rationality*. MIT Press.

Cooper, G. 1990. "The computational complexity of probabilistic inference using Bayesian belief networks". *Artificial Intelligence* 42: 393–405.

Dantas, D. F. 2017. "No rationality through brute-force". *Philosophy South* 18(3): 195–200.

———. 2021. "How to (blind)spot the truth: An investigation on actual epistemic value". *Erkenntnis* 88: 693–720

———. 2022. "Epistemic sanity or why you shouldn't be opinionated or skeptical". *Episteme* 20(3): 647–666

De Bona, G., and J. Staffel. 2017. "Graded incoherence for accuracy-firsters". *Philosophy of Science* 84(2): 189–213.

———. 2018. "Why be (approximately) coherent?". *Analysis* 78(3): 405–415.

De Finetti, B. 1937. "La prévision: Ses lois logiques, ses sources subjectives". *Annales de l'institut Henri Poincaré* 7 (1), 1–68.

Doets, K. 1994. *From Logic to Logic Programming*. MIT Press.

Douven, I. 2013. "Inference to the best explanation, Dutch books, and inaccuracy minimization". *The Philosophical Quarterly* 63(252): 428–444.

Dudai, Y. 1997. "How big is human memory, or on being just useful enough". *Learning & memory* 3(5): 341–365.

Fitelson, B., and K. Easwaran. 2015. "Accuracy, coherence and evidence". In *Oxford studies in Epistemology, volume 5*, edited by Tamar Gendler and John Hathorne (Eds.), 61–96. Oxford University Press.

Friedman, J. 2013. "Suspended judgment". *Philosophical Studies* 162(2): 165-181.

———. 2015. "Why suspend judging?". *Noûs* 51(2): 302–326.

Gao, J. 2021. “Self-deception and pragmatic encroachment: A dilemma for epistemic rationality”. *Ratio* 34(1); 5–84.

Garber, D. 1984. “Old evidence and logical omniscience in Bayesian confirmation theory”. In *Testing Scientific Theories*, edited by John Earman, 99–131. University of Minnesota Press.

Genin, K., and F. Huber. 2020. “Formal representations of belief”. In *The Stanford encyclopedia of Philosophy (Winter 2020)*, edited by Edward Zalta. Metaphysics Research Lab, Stanford University.

Gigerenzer, G. and W. Gaissmaier. 2011. “Heuristic decision making”. *Annual Review of Psychology* 62(1): 451–482.

Hájek, A. 2008. “Arguments for or against probabilism?”. *The British Journal for the Philosophy of Science* 59(4): 793–819.

Jeffrey, R. 1983. *The logic of decision*. University of Chicago Press.

Joyce, J. 1998. “A nonpragmatic vindication of probabilism”. *Philosophy of Science* 65(4): 575–603.

Kemeny, J. 1955. “Fair bets and inductive probabilities”. *Journal of Symbolic Logic* 20(3), 263–273.

Kraus, S., D. Lehman, and M. Magidor. 1990. “Nonmonotonic reasoning, preferential models and cumulative logics”. *Artificial Intelligence* 44(1-2): 167–207.

Leitgeb, H. 2014. “The stability theory of belief”. *The Philosophical Review*, 123(2): 131–171.

Leitgeb, H., and R. Pettigrew. 2010. “An objective justification of Bayesianism II: The consequences of minimizing inaccuracy”. *Philosophy of Science*, 77(2): 236–272.

Lewis, P. J., and D. Fallis. 2019. “Accuracy, conditionalization, and probabilism”. *Synthese* 198(5): 4017–4033.

Littlejohn, C. 2015. “Who cares what you accurately believe?”. *Philosophical Perspectives* 29(1): 217–248.

Lord, E. 2020. Suspension of judgment, rationality’s competition, and the reach of the epistemic. In *The Ethics of belief and beyond: Understanding mental normativity*, edited by S. Schmidt and G. Ernst, 126–145. Routledge.

Pearl, J. 1986. “Fusion, propagation, and structuring in belief networks”. *Artificial Intelligence*, 29: 241–288.

Pettigrew, R. 2019. “Veritism, epistemic risk, and the swamping problem”. *Australasian Journal of Philosophy* 97(4): 761–774.

Ramsey, F. 1926. “Truth and probability”. In *The foundations of mathematics and other logical essays*, edited by R. Braithwaite, 156–198. McMaster University Archive.

Russell, S., and P. Norvig. 2020. *Artificial intelligence: A modern approach (4th ed.).* Pearson.

Steele, K., and H. O. Stefánsson. 2020. “Decision theory”. In *The Stanford encyclopedia of Philosophy (Winter 2020)*, edited by Edward Zalta. Metaphysics Research Lab, Stanford University.

Stenning, K. and M. van Lambalgen. 2008. *Human reasoning and cognitive science*. MIT Press.

Sturgeon, S. 2008. “Reason and the grain of belief”. *Noûs*, 42(1): 139–165.

Talbott, W. 1991. “Two principles of Bayesian Epistemology”. *Philosophical Studies*, 62(2): 135–150.

Talbott, W. 2016. “Bayesian Epistemology”. In *The Stanford encyclopedia of Philosophy (Winter 2016)*, edited by Edward Zalta. Metaphysics Research Lab, Stanford University.

Teller, P. 1973. “Conditionalization and observation”. *Synthese*, 26(2): 218–258.

Todd, P. M. and G. Gigerenzer. 2007. “Environments that make us smart: Ecological rationality”. *Current Directions in Psychological Science*, 16(3): 167–171.

Trpin, B., and M. Pellert. 2019. “Inference to the best explanation in uncertain evidential situations”. *British Journal for the Philosophy of Science*, 70(4): 977–1001.

Wedgwood, R. 2015. “Doxastic correctness”. *Aristotelian Society Supplementary Volume* 87(1): 217–234.

Wolpert, D., and W. Macready. 1997. “No free lunch theorems for optimization”. *IEEE Transactions on Evolutionary Computation*, 1(1): 67–82.