Bayesian epistemology through the Choice norm

First Draft of 15^{th} January 2012

Hykel Hosni

Abstract. This note offers a preview of an ongoing research project whose goal is to marry much needed formal advances in uncertain reasoning with the solid foundations provided by bayesian epistemology. To this end I will sketch an interpretation of bayesian epistemology which highlights how

(1) the bayesian norms of rational Belief (probability) and Decision norm (maximisation of expected utility) depend on specific restrictions imposed on the choice problem with which an underlying *Choice norm* is instantiated
(2) the identification of those specific restrictions allows us clarify the conditions under which *classical bayesianism* is justified – what I propose to call *first order uncertainty*
(3) and finally to set up strategies for extending bayesian epistemology to modelling *second order uncertainty*.

Contents

Key words and phrases. Uncertain reasoning, second-order uncertainty, rational choice, bayesian epistemology.

Minor changes January 20, 2012

[T]he kind of measurement of belief with which probability is concerned [. . .] is a measurement of belief qua *basis of action.* Ramsey (1931)

It is one of my fundamental tenets that any satisfactory account of probability must deal with the problem of action in the face of uncertainty. Savage (1954)

[E]very particular problem is nothing but a partial instantiation of the grand problem: to decide [. . .] in such a way as to maximise our universal objective function. de Finetti (1969).

1. The problem (is the Choice problem)

The central idea put forward in this note is that bayesian epistemology is best appreciated from the point of view of its *Choice norm*, namely the principle which normatively prevents rational agents from making dominated choices. Under some *very specific* circumstances the Choice norm justifies (subjective) probability as the norm of *rational belief* and similarly, the maximisation of expected utility as the norm of *rational decision* under uncertainty. The main question motivating the present investigation can be put as follows: How robust is this sort of justification? The centrality of the Choice norm will make the answer depend on the modelling features of the underlying choice problem.

The note is organised as follows. After recalling how the Choice norm is ubiquitous, albeit implicitly, in the development of bayesian epistemology (Section 2), I will discuss how the Choice norm naturally suggests a two-dimensional approach to modelling rationality, which is briefly sketched in Section 3. This gives us a system of coordinates to distinguish first-order and second-order uncertainty problems and to show how the classical bayesian norms apply only to the former. As many critics of classical bayesianism vehemently point out, the bayesian norms of rational behaviour find little application, if any at all, to interesting real-world decision problems. The purpose of Section 4 is to illustrate how this is best seen as consequence of the restriction of classical bayesianism to first-order uncertainty, rather than an argument to reject the bayesian foundations altogether. To the contrary, by referring to some recent formal work, I illustrate how a suitable instantiation of the Choice norm with a second-order uncertainty problem leads to the characterisation of (fuzzy) imprecise probabilities. The fact that the Choice norm allows us to extend classical bayesianism to second-order uncertainty has important methodological consequences for decision theory, some of which are discussed in Section 5. Section 6 concludes with some general observations on the potential virtues of a bayesian theory of second-order uncertainty.

Before starting, a two-fold disclaimer is in order. First of all, there are a number of important, even vital aspects of bayesianism which are not covered by this note. The most obvious omission concerns the *dynamics* of belief and decision, that is to say the spectrum of problems which spring from the concept of conditional belief measure and impact directly inductive reasoning and statistical inference. Macroscopic as it may be, this omission is dictated by the *divide and conquer* approach which makes itself necessary when tackling a problem of this dimension. I nonetheless believe that what *is* covered in this note is important enough to be relevant for bayesian epistemology at large.

Secondly, the limited scope of the present perspective implies that I will not attempt to show that the characterisation of bayesianism which underlies this project is *the right one*. The goal is rather that of exploring the consequences, and hence the potential worth, of some intuitively appealing assumptions, namely that bayesian epistemology revolves about its Choice norm and its closely related distinction between first- and second-order uncertainty may guide us in extending the foundations of the former to the formal modelling of the latter.

2. Bayesian epistemology and its choice roots

As testified by the opening quotes, bayesian epistemologists always insisted on the fact that the need for defining and measuring uncertainty arises in a context in which an individual must face the consequences of their choice behaviour[1]. It is *this* choice behaviour for which *bayesian epistemology* offers norms of rationality.

In spite of this strong underlying unity, the scope of bayesianism has proved very hard to identify with some precision[2], to the point that in his well-known provocation, I.J. Good counted $2^4 \cdot 3^6 \cdot 4 = 46656$ varieties of Bayesians.

I will take the less combinatorial approach put forward by Williamson (2010) as my starting point. In his analysis of Objective bayesianism, Williamson focusses directly on the three *norms* which are justified by the general framework of bayesian epistemology. The *Probability Norm* demands that rational agents' degrees of belief should be represented by (subjectively interpreted) probabilities. The *Calibration Norm* requires that the subjective probabilities licensed by the Probability Norm should be constrained by known physical probabilities[3]. Finally, the *Equivocation Norm* further refines the choice of subjective probabilities by excluding extreme probability values unless these are prescribed by the previous Norms, and that subject this requirement, probabilities should otherwise be maximally equivocal[4] . Objective bayesianism is the epistemological position which simultaneously endorses Probability, Calibration and Equivocation as Belief norms.

Williamson points out that the three bayesian Norms are individually justified [5] by appealing to (formal) variations of essentially the same argument which involves the minimisation of a suitably defined expected loss[6].

[1]Of course this has not prevented uncertain reasoning models from being developed more or less independently of their connections with choice. Such is the case of the class of models spanning from the theory of Belief functions to Idempotent measures (see, e.g. Dubois and Prade, 2001; Dubois et al., 2007). Joyce (2009) offers a general philosophical analysis of the "non pragmatic" justifications for taking rational belief as probability. The *Transferable Belief Model* sits somewhere in between (see, e.g. Smets, 1993; **?**).

[2]This might not so surprising after all. Bayesianism is a multifaceted enterprise encompassing theory and methods, foundations and applications of reasoning under uncertainty. There is virtually no field of scientific enquiry (broadly construed) which doesn't intersect with some aspects of bayesianism. The heterogeneity of the stakeholders, might clearly explain why there has been so far little convergence on a uniform bayesian epistemological position

[3]The best-known instantiation of this being Lewis's *Principal principle*.

[4]The best-known instantiation of the Equivocation norm is the Maximum entropy principle, which has been extensively discussed over the past three decades, often from heterogeneous points of view. For two comprehensive presentations, see Paris (1994); Jaynes (2003). See Williamson (2010) for an appraisal of the criticisms of the principle.

[5]See Chapter 3 of (Williamson, 2010) for a detailed analysis of such justifications.

[6]Note that a specific instantiation of this line of argument led to the Paris-Venkovská characterisation of Maximum Entropy reasoning Paris and Vencovská (1990); Paris (1994); Paris and Vencovská (1997); **?**.

Taking Williamson's argument one step further, it appears natural to find the root of the three, co-justified, Objective bayesian norms in the following, more general principle:

Choice norm: A rational agent must never choose dominated alternatives.

The Choice[7] norm captures a very basic aspect of bayesian epistemology and unifies the subjective component of individual uncertainty with the objective features of the underlying choice problem, thereby enabling the bayesian virtuous circle which connects choice, belief and decision[8].

The conception of rationality which underlies the Choice norm is minimal in two senses. Firstly, it is best seen as an attempt to define rationality in terms of *avoiding blatantly irrational* behaviour. This conception is ubiquitous in bayesian epistemology and it is tightly connected to the idea of rationality as maximisation. Wald's seminal "idea of associating a loss with an incorrect decision" (de Finetti, 1975, p. 253), for instance, paved the way to the analysis of subjective expected utility which, through the work of Ramsey, von Neumman and Morgenstern, culminated with Savage's Representation theorem. Similarly, as noted above, the idea of minimisation of some suitably defined expected loss, acts as a unifying principle among distinct Objective bayesian Norms of rational behaviour[9].

Secondly, the Choice norm relies on what is arguably the weakest, yet non-empty, characterisation of "purposeful behaviour", in the terminology of Bossert and Suzumura (2010)[10]. It is non-empty (and non-circular) because the occurrences of "irrational" and "dominated", appearing in the above phrasing of the Choice norm, belong to distinct epistemological spheres, so to speak. "Irrational" is meant to single out the kind of (hypothetically) observable[11] behaviour which is immediately, intuitively and pre-formally recognised as being self-defeating. As de Finetti and Ramsey clearly pointed out, if "rationality" means anything at all, it cannot be rational to behave in a way which blatantly contradicts the purpose of our own behaviour. It is the ordinary parlance meaning of "irrational" that is being used here, i.e. "stupid", "against commonsense", "illogical", etc.

[7]Alternative denominations might have included "Dominance", "Pareto" or even "Admissibility". As they all have rather specific connotations in distinct areas of the literature, it appears that "Choice" sits more comfortably at the desired level of generality of the norm.

[8] Here is one rendering of the virtuous circle

> A decision must [...] be based on probabilities: i.e. the posterior probabilities as evaluated on the basis of all information so far available. This is the main point to note. In order to make decisions, we first require a statistical theory which provides conclusions in the form of posterior probabilities. The Bayesian approach does this: other approaches explicitly refuse to do this. (de Finetti, 1974, p.252)

It can be observed in passing that this virtuous circle connecting choice, belief and decision accounts for the ubiquitous synergies of the broad concepts of "rationality", "uncertainty" and "economics". See (Rukhin, 1995; Levi, 1986; Bossert and Suzumura, 2011) for some statistical, epistemological and economic appraisals of this idea, respectively.

[9]Note that this is add odds some popular applications of "bayesian methods" which focus on *optimisation* rather than maximisation. For the important implications of this distinction, see e.g. (de Finetti, 1973) and (Sen, 1997).

[10]The potentially very fruitful connections between the present framework with the kind of minimality captured by Suzumura consistency are postponed to further investigation.

[11]See Rubinstein (2006) for a terse analysis of the role of hypothetical observations in the theory of revealed preference.

The term "dominated" which occurs in the Choice norm, on the other hand, is not meant at all in its ordinary sense. It is rather a technical concept which can be given a rigorous and *very general* mathematical definition in terms of a choice function (interpreted as a model the agent's behaviour), and a binary relation over the set of feasible alternatives (interpreted as a representation of the agent's preferences).[12] Dominance, and its cognate notion of "admissibility" (Wald, 1947), permeate the entire development of uncertain reasoning and much of economic theory to the point that it is customary to ascribe them to the *folklore*[13].

To make a rather self-explaining analogy, "irrationality" is to "dominance" what "incoherence" is to logical "inconsistency". This analogy suggests that a stupid choice is clearly *always* dominated. Hence, to justify the Choice norm we need only focus on the other direction. To see that a dominated choice is stupid, assume that i's preferences are adequately modelled by the binary relation R_i. Under *this* assumption, a dominated choice captures the idea of self-inconsistency, i.e. blatant irrationality. Put it the other way round, the Choice norm specifies the conditions under which an informal problem can be modelled as a *choice problem*, i.e. a formalised situation in which dominance cannot be rationally violated[14]. If an agent is normatively justified in making dominated choices, the situation at hand falls short of being a choice problem.

The above logical analogy suggests a further observation. Just as the formal notion of consistency depends on the specific semantics we are interested in capturing, the notion of dominance can be given a number of distinct formalisations. This simple observation captures the robustness of the approach to modelling uncertain reasoning based on the Choice norm and motivates the attempt, outlined in the remainder of this note, to apply it to model problems for which classical bayesianism is *not* justified.

3. First-order uncertainty

The field of uncertain reasoning is best understood in terms of a fact and an assumption. The fact, roughly speaking, is that whenever we face a choice problem we end up being in an epistemic state of uncertainty. The assumption is that we can make sense of such an epistemic state. Bayesian epistemology, in the rendering which I am emphasising here, is taking rationality as a function of two arguments, a choice problem and an individual facing it. Hence, as it will be apparent later on, classical bayesianism can be fruitfully seen as the solution to the following problem:

---

[12]One of the many equally good formalisation is as follows. Let X be a set of feasible alternatives, $R_i \subseteq X^2$ a binary relation such that xR_iy is interpreted as "i doesn't prefer y to x", and $\emptyset \neq C_i(X) \subseteq X$ be i's *choice set* from X (i.e. the non-empty subset of feasible alternatives selected by i). We say that i's choices are *dominated* if there exist $y \in X$ such that yR_iz, for all $z \in X$ but $y \in C_i(X)$.

[13] Rukhin (1995), for example, puts it as follows:

> How should one choose a decision rule whose performance depends upon the unknown-state of Nature? Since there is no unique recognized optimality principle that would provide a complete orderingof all statistical decision rules, this question is probably unanswerable when posed with such generality. However, it seems clear which procedures should not be used – the inadmissible ones which can be improved upon no matter what the unknown state of Nature.

[14]This is one way of interpreting the axiom of the Independence of irrelevant alternatives, according to which a dominated alternative should never be chosen from any superset of the original set of feasible alternatives.

How should a maximally idealised agent behave when facing a maximally abstract choice problem involving uncertainty?

3.1. Idealised agents. Since bayesian epistemology is uncompromisingly normative, it admits no possible relaxation of the modelling assumptions it makes about the cognitive capabilities of its agents. In particular, bayesian agents

a1 never make mistakes (neither in computing nor in recalling from memory, etc.)
a2 are fully aware of the information they possess
a3 can reason arbitrarily fast and free of any costs
a4 do exactly as they are told

An agent satisfying (a1-a4) certainly deserves to be called *maximally abstract*. Real people, e.g. experimental subjects, are obviously enough *minimally* abstract. To avoid cumbersome terminology in what follows, I will simplify "non-maximally abstract" to "non abstract".

Note that (a1-a4) above rule out the possibility that a bayesian agent might choose an alternative which they fail to realise is dominated, or similarly, that the agent sees maximal alternatives which nonetheless fails, perhaps akratically, to select. In this sense, the Choice norm finds no application in (normative) models of bounded rationality, let alone descriptive ones. There is, however, an indirect role for bayesian epistemology in descriptive models of rationality, as they provide the normative standard against which the systematic deviations or "biases" of experimental subjects are evaluated[15].

3.2. Abstract problems. The second, fundamental, modelling parameter captured by the Choice norm is the (degree of) abstraction of the choice problem which gives rise to the formal definition of dominance. As suggested by the logical analogy which closed Section 2, this is a key feature of the present proposal.

All modelling requires abstraction, and that's why in Box's infamous turn of phrase, all models are wrong, making the question of understanding the properties of useful, good, justified, etc. models, one of most important question in Science. Daunting as it may be, it is one very specific aspect of this question that we now need tackling, in the restricted domain of *choice* problems.

A maximally abstract choice problem is a mathematical description of a "real-world" problem which contains all and only the features which are relevant to its solution. When an agent is facing a maximally abstract choice problem it is *that* problem that the agent is facing, and not some other problem[16]. Thus, we can naturally think of an agent facing a maximally abstract choice problem is a *decision maker with no modelling privileges*.

The distinction between decision makers and decision modellers can be tricky and this goes some way towards explaining why so much foundational confusion arises on this point, some of which is briefly discussed in Section 5 below. It is tricky because in real life we very often play *both* roles. For definiteness, take an agent who is about to do their shopping. This

[15] A line of research which has become canonical after (Kahneman and Tversky, 1979). For a recent overview see (Wakker, 2010).

[16] This is an obvious adaptation of the so-called *Watt's assumption* of Paris (1994), which in turn builds on the related concepts of Carnap's "Principle of Total Evidence" and Keynes' "Bernoulli's maxim". The fact that experimental subject systematically violate this principle motivates the introduction of the "editing phase" in Prospect theory.

real-world problem can be given a maximally abstract representation as a consumer problem in Rational choice theory[17]. When doing our shopping, however, we make some choices as modellers (e.g. whether to include non-fair trade items as feasible alternatives) and some as decision makers (e.g preferring organic). In real-life problems we sometimes think we "don't have alternatives" or, at the other side of the spectrum, we can create alternatives where others don't see them.

Classical bayesianism developed out of a strict distinction between decision making and decision modelling, equivalently, by focussing on maximally abstract choice problems. Two obvious cases in point are de Finetti's betting problem and Savage's matrix, interpreted in the "Grand world" in which all consequences are maximally specified. I will limit myself to illustrating how the former leads to the appropriate instantiation of the Choice norm which, in essence, makes the Dutch book theorem a simple exercise in linear algebra[18]

3.3. De Finetti's betting problem. The instantiation of the Choice norm with a maximally abstract betting problem allowed de Finetti (1931) to prove the first representation theorem for rational belief. In a nusthell[19] de Finetti's betting problem is a choice problem in which a bookmaker must pick the betting odds for a set of events of interest. The notion of dominance arises as a consequence of the obligation that the bookmaker has to accept any bets on (or against) the events on the published book. It is in fact clear that under such contractual obligations, a choice of betting odds which puts the bookmaker in a position to lose money whichever events turn out to occur, undoubtedly qualifies as dominated. De Finetti's theorem shows that the bookmakers' choices can never be dominated unless they fail to conform to the laws of probability.

The contractual obligations make the problem maximally abstract. As he would later remark, the betting problem is a "device to force the individual to make conscious choices, releasing him from inertia, preserving him from whim" (de Finetti, 1974, p.76). The contract includes the following clauses:

- **Completeness:** The bookmaker's choice is forced for *series of bets* and the bookmaker, once the book has been published, is forced to accept indefinitely many bets
- **Swapping:** As the gambler has the possibility of choosing the direction of the bet *after* reading the odds, she can unilaterally impose a payoff-matrix swap to the bookmaker (i.e. the gambler can decide whether to buy or to sell the bet)
- **Rigidity:** stakes involved the betting problem correspond to actual money (in some currency). So, in order to avoid the potential complications arising from the diminishing marginal utility of money, it is assumed that stakes should be *small*[20]

[17]I use this expression to include the body of mathematical work of interest to individual and social decision making which is exemplified by (Sen, 1970) and (Kreps, 1988).

[18]I postpone the analysis of alternative choice problems, including Proper scoring rules and Savage's matrices, to future work.

[19] Since this is one of the most intensely studied aspect of bayesian epistemology, I will take many details for granted and focus on the specific aspects which are directly relevant to the present discussion.

[20]The terminology derives from de Finetti who refers to it as the *rigidity hypothesis* (de Finetti, 1974, p.77-78). On de Finetti's initial reluctance to appeal to the mathematical theory of utility see (de Finetti, 1969, especially Chapter 4).

Completeness is justified by de Finetti (1931) by noting that if the bookmaker were allowed to refuse selling a bet, or if the number of transactions were limited, the bookmaker's "whim" could not be told part from their "sincere belief"[21]. To the potential counterargument to the effect that completeness is an unreasonable normative assumption –on which more in Section 5 below– de Finetti offers the following defense:

> Among the answers that do no make sense, and cannot be admitted are the following: "I do not know", "I am ignorant of what the probability is", "in my opinion the probability does not exist". Probability (or prevision) is not something which in itself can be known or not known: *it exists in that it serves to express, in a precise fashion, for each individual, his choice in his given state of ignorance*. To imagine a greater degree of ignorance which would justify the refusal to answer would be rather like thinking that in a statistical survey it makes sense to indicate, in addition to those whose sex is unknown, those for whom one does not even know "whether the sex is unknown or not". (de Finetti, 1974, p.82, my emphasis)

Swapping and Rigidity provide the key abstraction leading to the first-order characterization of dominance. In the presence of Completeness, Swapping guarantees that a book with non-zero expectation constitutes a dominated choice for the bookmaker. In the presence of Rigidity this justifies assuming that (non-dominated) betting odds reveal (rational) degrees of belief, thus completing the Dutch book argument.

The Probability norm therefore depends essentially on the first-order modelling assumption on the choice problem. Similar arguments appear to apply to Proper scoring rules and consistent preferences among the acts of a Savage matrix. If this were indeed the case[22] the notion of first order uncertainty would be a very good candidate to delimit the class of problems for which classical bayesianism is unquestionable.

This raises immediately an obvious question. If the Probability norm is fully justified for first-order uncertainty only, what happens outside its rather narrow borders? A pictorial representation of the bi-dimensional modelling space delimited by abstraction and idealisation is given in Figure 1, whilst an example of how the Choice norm is robust to some non trivial relaxation of the underlying abstract choice problem, is discussed in the next Section.

Before moving to that, note that the distinction between decision maker and decision modeller can be used to illustrate intuitively an argument for rejecting the identification of second order uncertainty models with "higher order probability" in the sense of Good (1962). To see why, in general, second order uncertainty is not captured by rational degrees of belief about rational degrees of belief, suffice it to note that the problem faced by a decision maker is qualitatively distinct from the kind of problem faced by the decision modeller. Whilst the rationality constraints for this latter include essentially (part of) the process by means

[21] The effectiveness of the abstract choice problem with respect to the accurate elicitation of revealed belief is one of the main reason for de Finetti's later insistence on *proper scoring rules*, especially Brier's (see, e.g. de Finetti, 2008).

[22] Only a detailed formal analysis can elevate this conjecture to the status of a rigorous claim. It is also worth noting in passing, that de Finetti's two-fold derivation of the Probability norm and Savage's derivation of the MSEU norm are often regarded as being essentially similar. Yet, neither the notion of a Dutch book, nor the idea of a Scoring rule play any role in Savage's representation theorem, whereas de Finetti never seem to have never invoked directly, say the Sure thing principle (see de Finetti, 2008, pp.113-4). Both characterisations have in common the Choice norm and its restricted application to first order uncertainty.

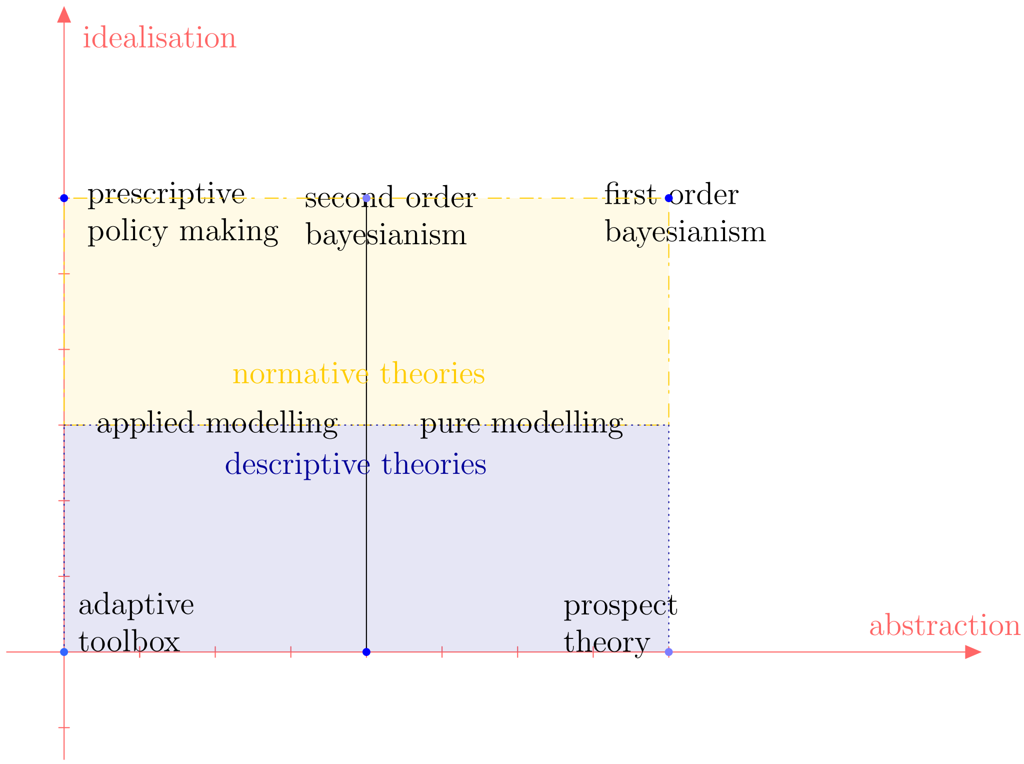

Figure 1. A stylised bi-dimensional space of models determined by the parameters intervening in the Choice norm, namely the idealisation of the agent and the abstraction of the choice problem.

of which beliefs are being constructed, the rationality constraints for the former explicitly disregard those aspects.

4. Towards second-order uncertainty

As Figure 1 illustrates, second-order uncertainty arises by reducing the abstraction of the underlying choice problem. There are a number of (largely independent) ways in which this might happen. For the purpose of illustrating the applicability of the framework I will limit myself to considering one specific case, which has been formally investigated in (Fedel et al., 2011).

4.1. Betting for profit. As recalled in the previous Section, de Finetti's betting problem is so formulated as to force the (idealised) bookmaker to choose fair betting quotients. Zero expectation of gain is the only protection that the bookmaker has against a smart gambler who, having spotted the possibility, forces their competitor into sure loss, possibly by switching payoff matrices. Moreover the values of the random variables of interest which constitute the object of the bets, are assumed to range over the binary set[23].

Fedel et al. (2011) investigate the joint relaxation of both these first-order modelling features. Firstly, they relax the (maximal) abstraction of de Finetti's betting problem by

[23]In de Finetti's intuitive characterisation, this is a crucial defining feature of events, i.e. random variables whose values range over $\{0, 1\}$.

assuming that bookmakers do their job for profit. This implies that they are no longer forced to accept payoff swaps with gamblers and, as a consequence, are allowed to differentiate between buying and selling prices for bets – a very familiar idea from the theory of Imprecise probabilities[24]. Secondly, we allow the truth value of events to range over the entire real-unit interval. This is a comparatively less investigated extension of the classical bayesian framework which has been pursued, over the past decade, mostly within the Many-valued logic community[25].

How can the Choice norm be instantiated with such a (second-order uncertainty) *for-profit betting problem*? The key idea[26] is that the bookmaker assigns two real numbers α, β to each event interest E, which are intuitively interpreted as their selling and buying prices, respectively. A book (or a system of bets) will then be a set of pairs $(E_1, [\alpha_1, \beta_1]), \ldots, (E_n, [\alpha_n, \beta_n])$ such that $0 \leq \alpha_i \leq \beta_i \leq 1$ for $i = 1, \ldots n$.

In order to define an appropriate notion of dominance note that the bookmaker knows that the gambler has a choice between:

- paying $\beta\lambda$ for the right to receive $\lambda v(E)$, with $\lambda > 0$ (betting on E)
- receiving the payment of $\alpha\lambda$ to pay back $\lambda v(E)$ (betting on the complement of E)

where $\lambda > 0$, and $v(E)$ is the truth value of E (recall that this can be any real number in the unit interval). It is easy to see that the prescription of the Choice norm cannot be to avoid incurring sure loss. Consider, to make to make a very simple example, the book $B = \{(E, [0, 1]), (E^c, [.5, 1])\}$, where E^c denotes the complement of E[27].

It is immediate to see that no sure loss is possible if the bookmaker chooses B. Yet there is a clear sense in which book B violates the Choice norm, namely the fact that the bookmaker could have made the book more attractive to gamblers, thereby potentially increasing their income, *without* incurring sure loss. Hence, dominance should require that no such improvements may be possible, or in the admittedly rather unfortunate terminology of (Fedel et al., 2011), that no *bad bet* should be made. Note that no such possibility arises in de Finetti's betting problem.

One central result of (Fedel et al., 2011) can be stated in the terminology of the present note as follows. Instantiated with for-profit betting problems (with fuzzy events) the Choice norm justifies fuzzy upper probabilities[28]. Thus, two central aspects of second-order uncertainty, namely *ignorance* and *vagueness*[29] can be formally captured in a framework which

[24] The theory received its first comprehensive systematisation in (Walley, 1991). See, e.g. (Miranda, 2008) for an outline of its most recent developments.

[25] This research strand originates with Paris (2001). See Fedel et al. (2011) for a state-of-the art and a comprehensive list of references.

[26] The simultaneous extension of classical bayesianism to imprecise *and* fuzzy probabilities is carried out by building the analytic framework of imprecise probabilities on top of a many-valued algebraic semantics. I will limit myself here to an informal discussion of the main result, and refer the interested reader to (Fedel et al., 2011) for the precise mathematical details.

[27] The set-theoretic description of events is used here to avoid unnecessary notational complications.

[28] In slightly more precise detail, a bad-bet is avoided if and only if the bookmaker's betting odds can be extended to an upper prevision over a suitable many-valued algebra of events. Dual results follow from lower previsions and probabilities.

[29] De Finetti always exhibited a casual conventionalism about vagueness. In his lectures on the foundations of probability, for instance, he insisted that

> [...] it must be borne in mind that whenever we say that an event is something that surely turns out to be either true or false, we are making a limiting assertion as it is

extends the one of classical bayesianism through an appropriate instantiation of the Choice norm.

5. The current foundational debate in perspective

In spite of the maximal idealisation required to give normative force to the bayesian model, the Choice norm captures the minimal condition of rationality to the effect that rational behaviour must not be self-defeating. As I briefly recalled in Section 2 and informally exemplified in the previous Section, bayesian epistemology aims at capturing the *necessary* conditions for rational behaviour under uncertainty which happen to be monotonically persistent via the Choice norm: Choosing a book which leads to sure loss is dominated *a fortiori* in for-profit betting problems.

Following the lead of Ellsberg (1961), whom in turn found himself on the footsteps of Knight (1921) and Keynes (1921), there is an increasing consensus among decision theorists on the idea that bayesianism fails to capture necessary conditions of rationality. If this idea is correct, bayesian epistemology would turn out to be largely irrelevant to the normative characterisation of rational reasoning under uncertainty. A good sample of the arguments put forward in support to this rather bold statement are condensed in Gilboa et al. (2011), a paper which has abundantly circulated for almost a decade before its publication.

One line of criticism, which goes back to both Keynes and Knight, albeit in rather distinct forms, is effectively summed up in the opening lines of Schmeidler (1989):

> The probability attached to an uncertain event does not reflect the heuristic amount of information that led to the assignment of that probability. For example, when the information on the occurrence of two events is symmetryc they are assigned equal probabilities. If the events are complementary the probabilities will be 1/2 independent of whether the symmetric information is meager or abundant.

Gilboa (2009) interprets Schmeidler's observation as expressing a form of "cognitive unease", namely a feeling that the theory of subjective probability which springs naturally from bayesian epistemology, is silent of one fundamental aspect of rationality (in its informal meaning). But why is it so? Suppose that some matter is to be decided by the toss of a coin. According to Schmeidler's line of argument, I should prefer tossing my own, rather than some one else's coin, on the basis, say of the fact that I have never observed signs of "unfairness" in my coin, whilst I just don't know anything about the stranger's coin[30]

The above recalled *Watt's assumption* doesn't allow this kind of reasoning to be replicated in the classical (first-order uncertainty) bayesian framework. On a single toss, there is no difference between not knowing whether the coin will land "heads", and not knowing whether

> not always possible to tell the two cases apart so sharply. Suppose, for instance, that a certain baby was born exactly at midnight on 31 December 1978 and we had to tell which was the year of his birth. How can it be decided in which year the baby was born, if his birth started in 1978 and ended in 1979? It will have to be decided by convention. [. . .] There is always a margin of approximation, which we can either take into account or not if we say '1978' or '1979'. de Finetti (2008)

[30]Observations of this sort are typically used to account for the normative rationality of ambiguity-averse preferences, as they are revealed in the so-called Ellebserg's paradoxes. See Gilboa (2009) for a comprehensive analysis and references.

the coin is biased. This is a typical limitation in expressive power which derives from the fact that the Probability norm is restricted to first-order uncertainty only. Yet, this limitation in expressive power cannot be taken as evidence that bayesian epistemology is fatally flawed. To the contrary, those limitations can be used to construct increasingly less abstract choice problems with which to instantiate the bayesian Choice norm.

Let us go back to ambiguity-averse preferences. From the (first-order) point of view of the (idealised) decision maker, there is no reason to see Ellsberg's preferences as being normatively justified and so the thesis that Ellsberg "paradox" really points at a formal incompleteness of the classical bayesian model has no normative force. Whilst Calibration is one of the Objective bayesian norms, it certainly doesn't follow that calibration is a necessary condition for the existence of a coherent probability assignment[31].

From the point of view of the decision modeller, however, we might feel that the bets for which, say no calibration is possible, should really be avoided. In a for-profit betting game, for instance, a bookmaker who felt very "unsure" about some specific event may coherently represent this "uncertainty" by assigning it the whole unit interval (i.e. buying at 0 and selling at 1). By exercising a privilege de Finetti's bookmakers *don't* have, in a second-order uncertainty setting, the agent can express their unwillingness to take bets, without violating the fundamental normative requirement of completeness[32].

Compare this with, say *Choquet expected utility*, a Decision norm which is motivated by the rejection of the completeness of revealed belief but doesn't admit of an easily interpretable Belief norm[33]. Similar difficulties arise in the context of Belief functions, Idempotent measures, Decomposable measures, and many mathematical models which developed in response to some *limitations* of the expressive power granted by classical bayesian norms. On the other hand, imprecise probabilities and their generalisation to fuzzy events, can be justified by a suitable second-order uncertainty instantiation of the Choice norm, e.g. one taking "avoiding bad bets" as the relevant formalisation of "dominance". Assigning a probability interval in an Ellsberg problem is a clear example of an epistemologically conservative way of extending the expressive power of bayesian theory to second-order uncertainty.

6. Preliminary conclusions

This note has been limited to formulating a working hypothesis and exploring some of its potentially welcome consequences for the foundations of uncertain reasoning. The hypothesis is that the Choice norm, i.e. the requirement that rational agents should not make dominated choices, may be one of the deepest roots of bayesian epistemology.

The first welcome consequence is that we can replace the philosophically challenging problem of providing a satisfactory taxonomy of allegedly distinct kinds of uncertainties, with the considerably more manageable distinction among choice problems outlined in Section 3.

[31] Unless, that is, we assume that probability is (only) objective, an assumption which can hardly be defended.

[32] A similar sort of conclusion seem to be accepted by Levi:

> it is sometimes rational to make no determinate probability judgment and, indeed, to make maximally indeterminate judgments.... [Doing so] may derive from a very clear and cool judgment that on the basis of the available evidence, making a numerically determinate judgment would be unwarranted and arbitrary. Levi (1985)

[33] See, e.g. Gilboa (2009); Wakker (2010); Ghirardato (2010)

The vast majority of the current formal developments in uncertain reasoning are motivated by the unsuitability of classical bayesianism to model rational reasoning under ambiguity, ignorance, vagueness and a number of variations on the traditional distinction between risk and uncertainty. Yet, in the absence of suitably specified choice problems, those "alternative" notions of uncertainty appear to be based on rather arbitrary distinctions, as noted some five decades ago by de Finetti (1963). Building on this intuition, I have suggested that by foccussing on the Choice norm we can overcome those difficulties. From bean jars to stock options, it is the fact that a choice must be made that forces us to measure uncertainty. How to do this, depends crucially on the modelling features of the choice problem at hand and on the assumptions on the idealisation of the agent whose uncertainty we want to measure. It is worth noticing that this two-dimensional interpretation of rationality, offers an interesting perspective on the yet unresolved tension between the subjective character of bayesian epistemology, and the objective understanding of rationality, recently revived, among others, by Gelman (2011).

I suggested that the distinction between *first-order* and *second-order* uncertainty which arises naturally in the light of the Choice norm is rich enough to account for a number of currently challenging aspects of uncertainty modelling. This gives rise to further welcome consequences. Firstly, the restriction to first-order uncertainty consolidates, as I have pointed out in Section 3, the classical bayesian standpoint. To some, this restriction is an intrinsic limitation of mathematical modelling. Binmore (2009), for instance, sees a "thin" theory as the inevitable price that we must pay if we want to prove theorems about rationality. Others, as recalled above, see the bayesian norms of first-order uncertainty as being neither necessary nor sufficient for –hence totally irrelevant to– rationality (Gilboa et al., 2011). The analysis I sketched above tells another story: The restriction of classical bayesian norms to first order uncertainty shows who uncontroversial foundations can go hand in hand with extended expressive power, as I exemplified in Section 4.1. By relaxing some saspects of the abstraction of the first-order choice problems, natural extensions of the Choice norm can be put forward to capture second-order uncertainty Belief norms, such as the (fuzzy) imprecise probability model of Fedel et al. (2011). This can be instantiated in a number of ways. As I pointed out in Section 5, the foundational solidity of classical bayesianism can be exploited to instantiate the Choice norm with increasingly less abstract choice problems, giving rise to increasingly more realistic Belief and Decision norms. Incidentally, this might also be helpful in seeing why the mutual feedback of bayesian epistemology and microeconomics should not be taken too far. The failure of first-order representations of choice problems involving second-order uncertainty should be taken simply as saying that if we are interested in normative models of rationality, we should be focussing on increasingly less abstract choice problems. What could be questioned though, is the relevance of normative models of rationality for economics.

In conclusion, the hypothesis put forward in this note appears to be well worth making and as this research project unfolds, I expect the foundational unity provided by the Choice norm to facilitate the much needed formal advance towards capturing important aspects of second order uncertainty.

References

K. Binmore. *Rational decisions*. Princeton University Press, 2009.

W Bossert and K. Suzumura. *Consistency, choice, and rationality*. Harvard University Press, 2010.

W. Bossert and K. Suzumura. Revealed preference and choice under uncertainty. *SERIEs*, pages 1–6, February 2011.

B. de Finetti. Sul significato soggettivo della probabilità. *Fundamenta Mathematicae*, 17: 289–329, 1931.

B. de Finetti. La decisione nell'incertezza. *Scientia*, 98:61–68, 1963.

B. de Finetti. *Un matematico e leconomia*. Franco Angeli, Milano, 1969.

B. de Finetti. Bayesianism: Its Unifying Role for Both the Foundations and the Applications of Statistics. *Bulletin of the International Statistical Institute*, 39:349–368, 1973.

B. de Finetti. *Theory of Probability, Vol 1*. John Wiley and Sons, 1974.

B. de Finetti. *Theory of Probability, Vol 2*. John Wiley and Sons, 1975.

B. de Finetti. Philosophical lectures on probability. Springer Verlag, 2008.

D. Dubois and H. Prade. Possibility theory, probability theory and multiple-valued logics: A clarification. *Annals of Mathematics and Artificial Intelligence*, 32(1):35–66, 2001.

D. Dubois, F. Esteva, L. Godo, and H. Prade. Fuzzy-set based logics – an history oriented presentation of their main developments. In D.M. Gabbay and J. Woods, editors, *Handbook of the History of Logic, Vol 8*, pages 325–449. Elsevier, 2007.

D. Ellsberg. Risk, Ambiguity and the Savage Axioms. *The Quarterly Journal of Economics*, 75(4):643–669, 1961.

M. Fedel, H. Hosni, and F. Montagna. A logical characterization of coherence for imprecise probabilities. *International Journal of Approximate Reasoning*, 2011. doi: 10.1016/j.ijar. 2011.06.004.

A. Gelman. Induction and Deduction in Bayesian. *Rationality, Markets and Morals*, 2:67–78, 2011.

P. Ghirardato. Ambiguity. In R. Cont, editor, *Encyclopaedia of Quantitative Finance*. Wiley and Sons, 2010.

I. Gilboa. *Theory of Decision under Uncertainty*. Cambridge University Press, 2009.

I. Gilboa, A. Postlewaite, and D. Schmeidler. Rationality of Belief. *Synthese*, To appear, 2011.

I.J. Good. Subjective probability as a measure of a non-measurable set. In T. Nagel, D. Suppes, and A. Tarsky, editors, *Logic, Methodology and Philosophy of Science*, pages 319–329. University of Calofornia Press, 1962.

E.T. Jaynes. *Probability theory: the logic of science*. Cambridge University Press, 2003.

J.M. Joyce. A Nonpragmatic Vindication of Probabilism. *Philosophy*, 65(4):575–603, 2009.

D. Kahneman and A. Tversky. Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):263–291, 1979.

J.M Keynes. *Treatease on Probability*. Harper & Row, 1921.

F.H. Knight. *Risk, uncertainty and profit*. Beard Books Inc, 1921.

David M Kreps. *Notes on the Theory of Choice*. Underground Classics in Economics, 1988.

I. Levi. Imprecision and Indeterminacy in Probability Judgment. *Philosophy of Science*, 52 (3):390–409, 1985.

I. Levi. The paradoxes of Allais and Ellsberg. *Economics and Philosophy*, 2:23–53, 1986.

E Miranda. A survey of the theory of coherent lower previsions. *International Journal of Approximate Reasoning*, 48(2):628–658, June 2008.

J.B. Paris. *The uncertain reasoner's companion: A mathematical perspective*. Cambridge University Press, 1994.

J.B. Paris. A Note on the Dutch Book Method. In Teddy Seidenfeld (Eds.): Gert De Cooman, Terrence Fine, editor, *ISIPTA '01, Proceedings of the Second International Symposium on Imprecise Probabilities and Their Applications, Ithaca, NY, USA.* Shaler, 2001.

J.B. Paris and A. Vencovská. A note on the inevitability of maximum entropy. *International Journal of Approximate Reasoning*, 4(3):183–223, 1990.

J.B. Paris and A. Vencovská. In Defense of the Maximum Entropy Inference Process. *International Journal of Approximate Reasoning*, 17(97):77–103, 1997.

F.P. Ramsey. Truth and probability (1926). In R.B. Braithwaite, editor, *The Foundations of Mathematics and other Logical Essays*, pages 156–198. Kegan, Paul & Co, London, 1931.

A Rubinstein. *Lecture Notes in Microeconomic Theory*. Princeton University Press, 2006.

A.L. Rukhin. Admissibility: Survey of a Concept in Progress. *International Statistical Review*, 63(1):95–115, 1995.

L.J. Savage. *The Foundations of Statistics*. John Wiley & Sons, 1954.

D. Schmeidler. Subjective Probability and Expected Utility without Additivity. *Econometrica*, 57(3):571–587, 1989.

Amartya Sen. Collective Choice and Social Welfare. 1970.

Amartya Sen. Maximization and the Act of Choice. *Econometrica*, 65(4):745, July 1997. ISSN 00129682.

Philippe Smets. No Dutch Book can be built against the TBM even though update is not obtained by Bayes rule of conditioning. In *Workshop on probabilistic expert systems, Societa Italiana di Statistica, Roma*, volume 6156, pages 181–204, 1993.

P.P. Wakker. *Prospect theory: for risk and ambiguity*. Cambridge University Press, 2010.

A. Wald. An essentially complete class of admissible decision functions. *The Annals of Mathematical Statistics*, 18(4):549–555, 1947.

P. Walley. *Statistical Reasoning with Imprecise Probabilities*. Wiley, 1991.

J. Williamson. *In defence of objective Bayesianism*. Oxford University Press, 2010.

Scuola Normale Superiore, Pisa
E-mail address: hykel.hosni@sns.it