International Journal of Approximate Reasoning 78 (2016) 1–14

Contents lists available at ScienceDirect

International Journal of Approximate Reasoning

www.elsevier.com/locate/ijar

Robust Bayesian analysis in partially ordered plausibility calculi

Stefan Arnborg

Royal Institute of Technology KTH, Sweden

ARTICLE INFO

Article history:
Received 23 January 2016
Received in revised form 7 May 2016
Accepted 20 June 2016
Available online 22 June 2016

Keywords:
Robust Bayesian analysis
Imprecise probabilities
Algebraic plausibility calculi

ABSTRACT

In Robust Bayesian analysis one attempts to avoid the 'Dogma of Precision' in Bayesian analysis by entertaining a set of probability distributions instead of exactly one. The algebraic approach to plausibility calculi is inspired by Cox's and Jaynes' analyses of plausibility assessment as a logic of uncertainty. In the algebraic approach one is not so much interested in different ways to prove that precise Bayesian probability is inevitable but rather in how different sets of assumptions are reflected in the resulting plausibility calculus. It has repeatedly been pointed out that a partially ordered plausibility domain is more appropriate than a totally ordered one, but it has not yet been completely resolved exactly what such domains can look like. One such domain is the natural robust Bayesian representation, an indexed family of probabilities.

We show that every plausibility calculus embeddable in a partially ordered ring is equivalent to a subring of a product of ordered fields, i.e., the robust Bayesian representation is universal under our assumptions, if extended rather than standard probability is used.

We also show that this representation has at least the same expressiveness as coherent sets of desirable gambles with real valued payoffs, for a finite universe.

1. Introduction

Uncertainty management schemes and their justifications is a controversial area in science and engineering, and a number of incompatible such schemes with various justifications exist. One line of investigation and justification is the attempt started by Bernstein [52], Cox [21] and Jaynes [34]. Originally formulated rather informally, it has been polished in a series of papers, for a survey see [33]. In the report [7] we showed that under reasonable assumptions a plausibility calculus based on totally ordered plausibility values must use the addition and multiplication operations of an ordered field. There are more general ordered fields than the common fields of rationals **Q** and reals **R**, including fields of hyperreal numbers that in the context of probability can be regarded as containing infinitesimals. All ordered fields are subfields of the most general ordered field of surreals, **No**. A beautiful account of this was presented by Conway [19]. This leads to extended probability defined by, among others, Wilson [60], where probabilities can contain infinitesimals, as the most general plausibility calculus with ordered plausibility values. We conjectured (in the form of a robustness assumption) that a plausibility calculus based on partially ordered plausibility values should be equivalent to a calculus based on an indexed set of extended probability distributions. Sets of probability distributions are used in (global) robust Bayesian analysis, see Berger [12,13]. In the de Finetti–Williams–Walley line of investigation [24,57,58], which is based on a betting or gambling paradigm, such

E-mail address: stefan@nada.kth.se.

http://dx.doi.org/10.1016/j.ijar.2016.06.006
0888-613X/

sets are called *imprecise probabilities*, whereas single probability distributions are called *precise*. The most general of Walley's proposals, sets of desirable gambles, has received significant attention recently. Bayesian analysis has been criticized for imposing the 'Dogma of Precision'. Robust Bayesian analysis and imprecise probability theory are answers to this criticism.

The main result of this note is a verification of the robustness conjecture, using tools developed for investigations in fundamental algebra by Birkhoff and McCoy. A preliminary sketch of this verification (with many gaps) was included in the retrospective section of [7]. An additional objective is to show the strengths of the algebraic approach which does not presuppose anything about the structure of the plausibility domain, the set of values used to describe the plausibility of an event.

In section 2 we review the background and current status of the algebraic thread of Cox/Jaynes' line of investigation and formulate the problem under investigation. In section 3 we review algebraic concepts required, summarize known results relating to the topic of this note, and pick the key observations. In section 4 we give some examples illustrating both theoretical and practical points in the relation of our result to some established uncertainty handling methods: Extended precise probability and non-monotonic reasoning, Robust Bayesian analysis illustrated by spatial/temporal tracing of Indo-European languages, upper and lower probabilities, Dempster–Shafer theory illustrated by the Boxer/Wrestler/CoinFlip problem, and sets of desirable gambles. We sum up in section 5. Appendix A contains our definitions of some of the algebraic terms used, and also two examples of embedding constructions. Appendix B contains a proof outline of the main result.

2. Background

The justifications of probability as universal plausibility calculus inspired by Cox [21] is an alternative to the approach based on coherence, attributed to de Finetti and described in his book [24]. The latter is based on a betting paradigm saying that an assignment of plausibilities, translatable to betting odds, to conditional events is *incoherent* if a Dutch book can be made against a bookmaker posting a set of, according to the assignment, fair bets. That coherent assignments are equivalent to probability assignments for a set of conditional events was shown by de Finetti and more strictly by Freedman and Purves [28]. The former approach is based on assigning conditional plausibility values to a set of symbols representing events or statements. The quantity $(A|C)$ is thus the plausibility that A is true if we know that C is true. Jaynes made a point of using statements as the units with plausibilities, but there is no reason not to use the more common event concept (a subset of the universe Ω) of probability theory which is equivalent in this context, as explained by Terenin and Draper [54]. We will not make a distinction between event and proposition and we will use both terms.

We use the notation common in probability theory: If A and B are sets, events or statements we use $\overline{A}$ for $\neg A$, $A \cup B$ for $A + B$ or $A \vee B$, and AB for $A \cap B$ or $A \wedge B$. Under additional assumptions it can be argued that functions F, S and G (the function G is not always used, but it is used by Aczél [1] and fits well with our algebraic method) must exist such that $(AB|C) = F((A|BC), (B|C))$, $(\overline{A}|C) = S(A|C)$ and, for disjoint events A and B, $((A \cup B)|C) = G(A|C, B|C)$. It is then observed that the semiring axioms of set algebra (associativity and commutativity of union and intersection, distributivity of intersection over union, existence of 0 and 1, etc.) will be, under reasonable or compelling (depending on the mood of the author) assumptions, inherited by the auxiliary functions F, S and G so that the calculus of plausibilities will satisfy the axioms of a commutative semiring, with G and F in the roles of $+$ and $\cdot$.

It is then standard to postulate that the plausibility values must be real numbers and prove that the plausibilities can be monotonically rescaled so that F translates to real number multiplication and moreover the function S translates to $S(x) = 1 - x$. This means that the common real number operations addition and multiplication will correspond to G and F, and thus the rescaled plausibility values are used exactly like standard probabilities. The existence of a strictly monotone real valued rescaling to standard probability depends however on assumptions that are not always stated: in [5] an example is given of a plausibility assignment where all algebraic laws are satisfied for a continuous F, but rescaling is not possible. So this example (described in section 4.1) must use extended probability. Such probabilities have been proposed occasionally and actually have some sway in discussions of uncertainty calculi [2,11,60]. The current research focus in investigations following the prevision-based paradigms contain significant results on the use of extended (often called non-Archimedean [27]) probability [47,61,20].

In order to handle domains not necessarily real-valued and to take care of some implicit assumptions of Cox and Jaynes in a convincing way we introduced a refinability assumption: In response to [31], we developed a very weak assumption binding the different parts of a system of events together in [5–7]. It makes no assumption on the domain of plausibility values:

Refinability assumption: In a plausibility model with a conditional event of plausibility p, it must be possible to introduce a new subcase B of a non-false event A with plausibility value p given to $B|A$. This should give a new model which is equivalent to the original one as long as no inference is made for the new event. If two new subcases B and B' of an event A are defined in this way, they can be specified to be information independent, i.e., $B|B'A = B|A$, $B'|BA = B'|A$. For two plausibility values x, y such that $x \leq S(y)$, it should be possible to define two new disjoint subcases C, C' of any non-false event A such that $x = C|A$, $y = C'|A$.

These augmentations of the model are called *refinements*. A refinement just introduces an already existing plausibility value in 'another part' of the model. This means that there is no assumption of a 'dense domain' or that a new plausibility value can be introduced. The motivation for refinability is that the same cognitive or other process that resulted in a particular plausibility value for one conditional event can always be mirrored in another part of the model to produce the