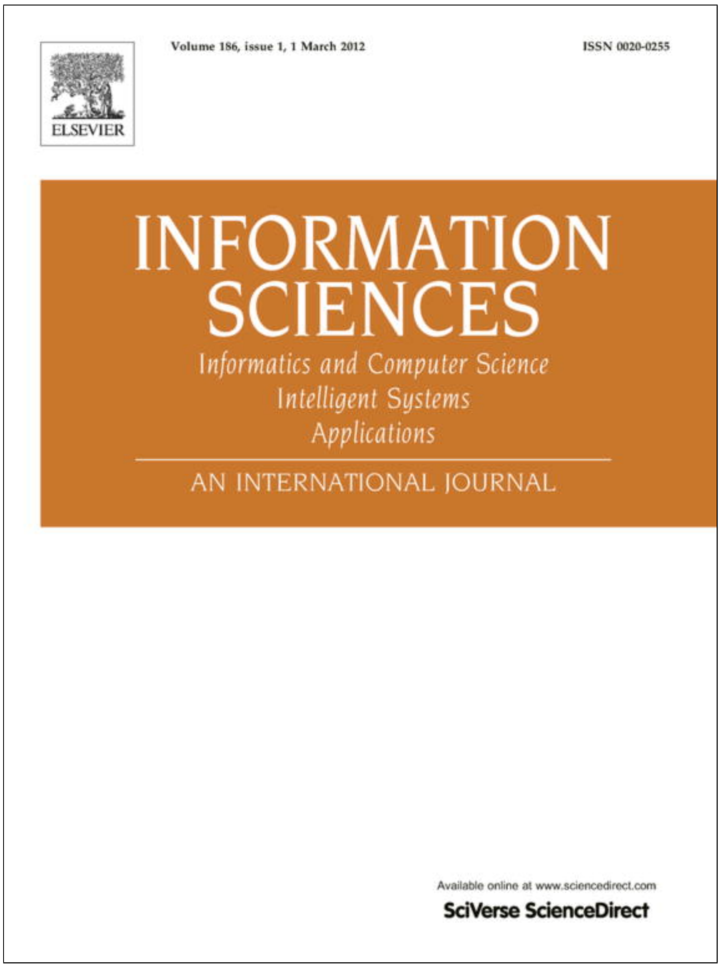
Volume 186, issue 1, 1 March 2012
ISSN 0020-0255
ELSEVIER
INFORMATION SCIENCES
Informatics and Computer Science
Intelligent Systems
Applications
AN INTERNATIONAL JOURNAL
Available online at www.sciencedirect.com
SciVerse ScienceDirect

Information Sciences 186 (2012) 105–113

Contents lists available at SciVerse ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Axiomatisation of fully probabilistic design

Miroslav Kárný, Tomáš Kroupa *

Institute of Information Theory and Automation of the ASCR, Pod Vodárenskou věží 4, 182 08 Prague 8, Czech Republic

ARTICLE INFO

Article history:
Received 3 November 2010
Received in revised form 20 March 2011
Accepted 19 September 2011
Available online 29 September 2011

Keywords:
Bayesian decision making
Fully probabilistic design
Kullback–Leibler divergence
Unified decision making

ABSTRACT

This text provides background of fully probabilistic design (FPD) of decision-making strategies and shows that it is a proper extension of the standard Bayesian decision making. FPD essentially minimises Kullback–Leibler divergence of closed-loop model on its ideal counterpart. The inspection of the background is important as the current motivation for FPD is mostly heuristic one, while the technical development of FPD confirms its far reaching possibilities. FPD unifies and simplifies subtasks and elements of decision making under uncertainty. For instance, (i) both system model and decision preferences are expressed in common probabilistic language; (ii) optimisation is simplified due to existence of explicit minimiser in stochastic dynamic programming; (iii) DM methodology for single and multiple aims is unified; (iv) a way is open to completion and sharing non-probabilistic and probabilistic knowledge and preferences met in knowledge and preference elicitation as well as unsupervised cooperation of decision makers.

1. Introduction

There is a whole range of axiomatic formulations of decision making (DM) under uncertainty and incomplete knowledge, e.g. [23]. It seems, however, that none of them fits satisfactorily to closed decision loops in which the selected actions influence distributions describing them—cf. [8]. This paper attempts to fill the gap. The text serves primarily as a formalised justification of the fully probabilistic design (FPD) of decision-making strategies—see [10,11,14]. The relationship of the FPD to the standard Bayesian DM is established too.

We consider a DM unit, called participant, that selects a T-tuple of *actions* $a^T = (a_1, \ldots, a_T)$, where T is a positive integer and each a_t belongs to a nonempty set $a_t^\star$. The actions are chosen with the aim to influence participant's environment, a thought of a part of the real world. In connection with the faced DM task, the participant considers *observations* of the environment $\Delta^T = (\Delta_1, \ldots, \Delta_T)$, where $\Delta_t \in \Delta_t^\star \neq \emptyset$, together with others *unobserved variables* $x^T = (x_1, \ldots, x_T)$, $x_t \in x_t^\star$. The triple of these vectors

$$\mathcal{Q} = (\Delta^T, a^T, x^T) \tag{1}$$

forms one possible *behaviour* of the closed loop consisting of the participant and its environment. By $\mathcal{Q}^\star$ we denote the set of all possible behaviours. For majority of applications, it suffices to identify $\mathcal{Q}^\star$ with a subset of finite-dimensional Euclidean space. For the purposes of this paper, it is enough to assume that $\mathcal{Q}^\star$ is a topological space. In the sequel we denote

$$d^T = (\Delta^T, a^T).$$

The inspected theory should help in selecting the optimal strategy among available DM *strategies* S^T, where each S^T is a T-tuple of DM rules

* Corresponding author.
E-mail addresses: school@utia.cas.cz (M. Kárný), kroupa@utia.cas.cz (T. Kroupa).

0020-0255/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2011.09.018

$$\mathsf{S}_t : d^{t-1} \mapsto a_t, \quad t = 1, \dots, T. \tag{2}$$

The choice of d^0 is governed by available prior knowledge. Throughout the remainder we write S instead of S^T if there is no ambiguity. The optimal strategy should "push" the closed-loop behaviour as much as possible towards a desired closed-loop behaviour. Of course, the assumption is that the participant has at least partial preferences among possible behaviours. The paper is organised as follows (Sections 2–6).

1. The strict partial preferential ordering $\triangleleft_{\mathcal{Q}^\star}$ on the set of all behaviours $\mathcal{Q}^\star$ is characterised and quantified by a loss function $Z : \mathcal{Q}^\star \to \mathbb{R}$. The inevitable assumptions are only made. Consequently, the loss function Z that orders possible behaviours a posteriori is non-unique.
2. Let $\mathsf{S}^\star$ be the set of all strategies. For a given strategy $\mathsf{S} \in \mathsf{S}^\star$, the possible behaviours $\mathcal{Q} \in \mathcal{Q}^\star$ are expressed as images of external unobserved influences N called *uncertainties.* We assume that the set N^* of all uncertainties is a locally compact Hausdorff space. Finite-dimensional Euclidean space or its nonempty subset is a typical example of $N^\star$. Uncertainties include anything what *a priori* prevents unambiguous determination of the closed-loop behaviour $\mathcal{Q}$ for a given S. For each $\mathsf{S} \in \mathsf{S}^\star$, the considered mapping
$$\mathsf{W}_\mathsf{S} : N^\star \to \mathcal{Q}^\star \tag{3}$$
is assumed to be bijective and Borel measurable.
3. A complete *a priori* ordering $\triangleleft_{\mathsf{S}^\star}$ is defined on $\mathsf{S}^\star$ in order to select the optimal DM strategy as the most preferred one in terms of this ordering. The ordering $\triangleleft_{\mathsf{S}^\star}$ is quantified via a local functional T acting on functions $Z_\mathsf{S} = Z(W_\mathsf{S}(N))$. In order to respect participant's preferences, the resulting ordering of losses (taken as functions of the uncertainty) must not prefer strictly dominated losses, i.e., the losses being strictly larger than another loss induced by an admissible strategy.
4. The integral representation of the local functional T developed in [22] is given by a kernel Φ and a finite nonnegative regular Borel measure μ on $N^\star$. The measure is recognised as a universal model of uncertainties $N \in N^\star$ common to all DM tasks sharing them. The non-dominance requirement is respected by using local functionals strictly isotonic with the dominance ordering of losses. The kernel Φ strictly increasing in the loss values guarantees the isotonicity.
5. The mappings Z and Φ, resulting from the quantitative characterisation of the *a priori* ordering of strategies $\triangleleft_{\mathsf{S}^\star}$, are restricted by widely acceptable conditions that make FPD the only acceptable alternative in selecting the best strategy. The essence of FPD consists in the minimisation of the Kullback–Leibler divergence (KLD) of closed loop model on its ideal counter-part. While FPD was proposed in [11] and extended into a general form in [14], the current paper tries to make the original heuristic approach more exact and to relate FPD to standard Bayesian DM.
6. A certain "closure" of the standard Bayesian DM tasks is shown to be a proper subset of FPD tasks and conclusions summarise general properties of the proposed FPD.

In this paper we build a solid mathematical background for FPD. Our approach is axiomatic, starting with the order-theoretic description of the problem, which is usual in decision theory and Bayesian DM [8,23,9]. We refrain from including examples of applications since this would substantially increase the length of the paper. Nonetheless, the original FPD has been motivated heuristically and already led to solutions of important practical problems, such as advising to human beings handling complex situations [13], non-standard knowledge sharing [12], and preference elicitation [15]. A special version of FPD (the so-called KL control) was discovered independently in [28] and FPD was re-discovered also in connection with a brain-oriented research [27]. These developments induce the demand for a rigorous framework, which will determine the boundaries of applicability and the real potential of FPD.

2. Ordering of behaviours

In this section we recall the basic notions related to strict preference orderings—see [8] for details. The participant is supposed to have a *strict preferential ordering* $\triangleleft_{\mathcal{Q}^\star}$ among behaviours $\mathcal{Q} \in \mathcal{Q}^\star$, which is an *irreflexive* and a *transitive* binary relation $\triangleleft_{\mathcal{Q}^\star}$ on $\mathcal{Q}^\star$. We write:

$$^a\mathcal{Q} \triangleleft_{\mathcal{Q}^\star} {}^b\mathcal{Q} \tag{4}$$

with the interpretation "$^a\mathcal{Q}$ is preferred against $^b\mathcal{Q}$". The irreflexivity of $\triangleleft_{\mathcal{Q}^\star}$ means that $\mathcal{Q} \triangleleft_{\mathcal{Q}^\star} \mathcal{Q}$ holds true not for every $\mathcal{Q} \in \mathcal{Q}^\star$. Transitivity of $\triangleleft_{\mathcal{Q}^\star}$ says that for every $^a\mathcal{Q},\ {}^b\mathcal{Q},\ {}^c\mathcal{Q} \in \mathcal{Q}^\star$ the implication

$$({}^a\mathcal{Q} \triangleleft_{\mathcal{Q}^\star} {}^b\mathcal{Q} \wedge {}^b\mathcal{Q} \triangleleft_{\mathcal{Q}^\star} {}^c\mathcal{Q}) \Rightarrow {}^a\mathcal{Q} \triangleleft_{\mathcal{Q}^\star} {}^c\mathcal{Q}$$

is satisfied.

The preferential ordering $\triangleleft_{\mathcal{Q}^\star}$ is in general only *partial* as the participants are often unable or unwilling to compare all pairs of possible behaviours. This is a key yet a realistic obstacle of the preference modelling. The incomparable pairs can be perceived as *indistinguishable:*

$$^a\mathcal{Q} \sim {}^b\mathcal{Q} \Longleftrightarrow_{def} \left(({}^a\mathcal{Q} \triangleleft_{\mathcal{Q}^\star} {}^b\mathcal{Q}) \vee ({}^b\mathcal{Q} \triangleleft_{\mathcal{Q}^\star} {}^a\mathcal{Q})\right).$$

In general, the relation $\sim$ is not necessarily transitive and therefore may not be an equivalence. The absence of transitivity is easily demonstrated when considering the ordering of two-dimensional integer-valued vectors $\mathcal{Q} = [\mathcal{Q}_1, \mathcal{Q}_2]$ with ${}^a\mathcal{Q} \lhd_{\mathcal{Q}^\star} {}^b\mathcal{Q}$ defined by ${}^a\mathcal{Q}_i < {}^b\mathcal{Q}_i,\ i = 1, 2$. It suffices to take ${}^a\mathcal{Q} = [0, 0],\ {}^b\mathcal{Q} = [1, 1]$ and ${}^c\mathcal{Q} = [2, -1]$, since:

$$ {}^a\mathcal{Q} \sim {}^c\mathcal{Q} \wedge {}^c\mathcal{Q} \sim {}^b\mathcal{Q} \wedge {}^a\mathcal{Q} \lhd_{\mathcal{Q}^\star} {}^b\mathcal{Q}. $$

However, a binary relation $\approx$ on $\mathcal{Q}^\star$ defined as:

$$ {}^a\mathcal{Q} \approx {}^b\mathcal{Q} \Longleftrightarrow_{\text{def}} ({}^a\mathcal{Q} \sim {}^c\mathcal{Q} \Longleftrightarrow {}^b\mathcal{Q} \sim {}^c\mathcal{Q}), \quad \text{for every } {}^c\mathcal{Q} \in \mathcal{Q}^\star, $$

turns out to be transitive. Moreover, it is an equivalence relation on $\mathcal{Q}^\star$ [8, Theorem 2.3]. By $\mathcal{Q}^\star_\approx$ we denote the set of all equivalence classes of $\mathcal{Q}^\star$ under $\approx$.

For a systematic design of DM strategies, the inspected ordering $\lhd_{\mathcal{Q}^\star}$ is represented numerically. There is a variety of conditions under which such representation exists. The next assertion is Theorem 2.5 from [8]—see [7] for an alternative formulation.

Theorem 2.1 (*Numerical representation of orderings*). *If $\lhd_{\mathcal{Q}^\star}$ is a strict preferential ordering on $\mathcal{Q}^\star$ and $\mathcal{Q}^\star_\approx$ is countable, then there is a real-valued function Z on $\mathcal{Q}^\star$ such that, for every ${}^a\mathcal{Q},\ {}^b\mathcal{Q} \in \mathcal{Q}^\star$:*

$$ {}^a\mathcal{Q} \lhd_{\mathcal{Q}^\star} {}^b\mathcal{Q} \Rightarrow Z({}^a\mathcal{Q}) < Z({}^b\mathcal{Q}), $$
$$ {}^a\mathcal{Q} \approx {}^b\mathcal{Q} \Rightarrow Z({}^a\mathcal{Q}) = Z({}^b\mathcal{Q}). $$

The function Z described in the above proposition is usually called *loss function*. In general, it is not determined uniquely.

3. Ordering of decision strategies

For design purposes, we have to specify a *complete* strict preferential ordering of strategies. The design then reduces to the selection of the optimal strategy, which is the "most preferred" one with respect to this ordering.

Let us consider the set of behaviours $\mathcal{Q}^\star$ with a strict preferential ordering $\lhd_{\mathcal{Q}^\star}$ meeting the assumptions of Proposition 2.1, and let Z be a loss function representing it according to this proposition. Every behaviour $\mathcal{Q} \in \mathcal{Q}^\star$ is the image of a considered strategy $\mathsf{S} \in \mathsf{S}^\star$ and the uncertainty $N \in N^\star$—see (3). Thus, the loss function Z can be expressed as a real function of uncertainties $\mathsf{Z}_S : N^\star \to \mathbb{R}$ defined by:

$$ \mathsf{Z}_S(N) = Z(\mathsf{W}_S(N)), \quad \text{for every } N \in N^\star \text{ and every strategy } \mathsf{S} \in \mathsf{S}^\star. $$

Considering all possible strategies $\mathsf{S} \in \mathsf{S}^\star$ together with the fixed loss function Z, we denote:

$$ \mathsf{Z}_{\mathsf{S}^\star} = \{\mathsf{Z}_S | \mathsf{S} \in \mathsf{S}^\star\}. $$

In the sequel we assume that a participant has adopted a strict complete preferential ordering $\lhd_{\mathsf{S}^\star}$ on the set $\mathsf{S}^\star$. This ordering $\lhd_{\mathsf{S}^\star}$ induces a strict complete preferential ordering $\lhd_{\mathsf{Z}_{\mathsf{S}^\star}}$ on $\mathsf{Z}_{\mathsf{S}^\star}$ given by:

$$ \mathsf{Z}_{{}^a\mathsf{S}} \lhd_{\mathsf{Z}_{\mathsf{S}^\star}} \mathsf{Z}_{{}^b\mathsf{S}} \Longleftrightarrow_{\text{def}} \quad {}^a\mathsf{S} \lhd_{\mathsf{S}^\star} {}^b\mathsf{S} \tag{5} $$

for every $\mathsf{Z}_{{}^a\mathsf{S}},\ \mathsf{Z}_{{}^b\mathsf{S}} \in \mathsf{Z}_{\mathsf{S}^\star}$.

The specification of the complete ordering $\lhd_{\mathsf{S}^\star}$ can be too complex to be used directly in selection of the optimal strategy. If this is the case, then [8, Theorem 2.5] can be applied to the ordering $\lhd_{\mathsf{Z}_{\mathsf{S}^\star}}$, which, in turn, represents the ordering $\lhd_{\mathsf{S}^\star}$. Provided $(\mathsf{Z}_{\mathsf{S}^\star})_\approx$ is countable (note that this necessarily implies the countability of $\mathsf{Z}_{\mathsf{S}^\star}$ since $(\mathsf{Z}_{\mathsf{S}^\star})_\approx$ is the set of singletons), the proposition guarantees existence of a functional

$$ \mathsf{T}_Z : \mathsf{Z}_{\mathsf{S}^\star} \to \mathbb{R}, \tag{6} $$

such that

$$ \mathsf{Z}_{{}^a\mathsf{S}} \lhd_{\mathsf{Z}_{\mathsf{S}^\star}} \mathsf{Z}_{{}^b\mathsf{S}} \Rightarrow \mathsf{T}_Z(\mathsf{Z}_{{}^a\mathsf{S}}) < \mathsf{T}_Z(\mathsf{Z}_{{}^b}\mathsf{S}). $$

4. Basis of the FPD

In order to get operational tool for the choice of the best strategy, we represent the functional T_Z from (6), whose purpose is to completely order strategies, by exploiting integral representation of local functionals [22].

In accordance with our aims, we want to make the result weakly dependent on a loss function Z chosen. Let us consider functions of the same uncertainty $N \in N^\star$, which arise by both varying possible strategies $\mathsf{S} \in \mathsf{S}^\star$ and loss functions:

$$ Z \in Z^\star = \text{losses representing quantifiable orderings } \lhd_{\mathcal{Q}^\star} (4). \tag{7} $$

This means that we consider the set of functions:

$$Z^\star_{S^\star} = \bigcup_{Z \in Z^\star} Z_{S^\star} = \bigcup_{Z \in Z^\star} \{Z \circ W_S | S \in S^\star\}.$$

For each $Z \in Z^\star$ there exists a numerical representation T_Z of $Z_{S^\star}$ given by (6). We assume that there exists a real functional T on $Z^\star_{S^\star}$ such that its restriction to every $Z_{S^\star}$ coincides with T_Z. It is reasonable to require the functional T to be continuous on the space of "nice" loss functions.

The "nice" loss functions in $Z^\star_{S^\star}$ are assumed to form the linear space $\mathcal{C}_c$ of real-valued continuous functions with a compact support in $N^\star$. From now on we assume that $Z^\star_{S^\star} = \mathcal{C}_c$. The supremum norm $\|\cdot\|$ makes $\mathcal{C}_c$ into a normed linear space. The following representation of a rich collection of functionals on $Z^\star_{S^\star}$ is described in [22, p. 479, Theorem 5], where the highly technical proof can be found.

Theorem 4.1 (*Representation of local functionals*). *Let* $T : Z^\star_{S^\star} \to \mathbb{R}$ *be a mapping such that the following three conditions are fulfilled:*

1. *(Sequential continuity) If* $(Z_n)_{n=1}^\infty$ *is a bounded point-wise convergent sequence in* $Z^\star_{S^\star}$*, then the sequence* $(T(Z_n))_{n=1}^\infty$ *is Cauchy.*
2. *(Local additivity)*

$$T({}^aZ + {}^bZ) = T({}^aZ) + T({}^bZ) \quad \text{whenever } {}^aZ\,{}^bZ = 0 \text{ for } {}^aZ,\ {}^bZ \in Z^\star_{S^\star}. \tag{8}$$

3. *(Bounded uniform continuity) For each* $\varepsilon > 0$, $\gamma > 0$, *there is a* $\delta > 0$ *such that if* $\|{}^aZ\| < \gamma$, $\|{}^bZ\| < \gamma$, ${}^aZ,\ {}^bZ \in Z^\star_{S^\star}$ *and* $\|{}^aZ - {}^bZ\| < \delta$, *then* $|T({}^aZ) - T({}^bZ)| < \varepsilon$. *Then*

$$T(Z) = \int_{N^\star} \Phi(Z(N), N)\, d\mu, \quad \text{for every } Z \in Z^\star_{S^\star}, \tag{9}$$

where μ *is a finite non-negative regular Borel measure on* $N^\star$ *and the kernel* $\Phi : \mathbb{R} \times N^\star \to \mathbb{R}$ *satisfies the following conditions:*

4. $\Phi(0,\cdot) = 0$ *and* $\Phi(\cdot,N)$ *is continuous for* μ*-almost all* $N \in N^\star$.
5. $\Phi(x,\cdot)$ *is Borel measurable for every* $x \in \mathbb{R}$.
6. *For every* $Z \in Z^\star_{S^\star}$, *the function* $\Phi(Z(N),N)$ *is bounded for* μ*-almost all* $N \in N^\star$ *and for any bounded point-wise convergent sequence* $(Z_n)_{n=1}^\infty$ *in* $Z^\star_{S^\star}$, *the sequence* $(\Phi(Z_n,\cdot))_{n=1}^\infty$ *is Cauchy in the space* $L^1(N^\star, \mu)$ *of functions* $N^\star \to \mathbb{R}$ *that are absolutely integrable w.r.t.* μ. *Conversely, if the pair* (Φ, μ) *satisfies the last three conditions and the functional* T *is defined by* (9), *then it meets the initial three conditions.*

The kernel Φ is determined by μ uniquely only outside a μ-null set. The only interpretation-sensitive assumption of the above theorem is local additivity (8) of T on the pair of loss functions with disjoint supports. It is, however, much weaker than usually required additivity.

The restriction of T on any $Z_{S^\star}$ (from now on also denoted by T) represents the ordering $\lhd_{Z_{S^\star}}$ and thus via (5) it represents the complete preferential ordering $\lhd_{S^\star}$:

$$^aS \lhd_{S^\star} {}^bS \iff T(Z_{^aS}) < T(Z_{^bS}) \quad \text{and} \quad ^aS = {}^bS \iff T(Z_{^aS}) = T(Z_{^bS}). \tag{10}$$

The definition (5) induces the complete ordering $\lhd_{Z_{S^\star}}$ of loss functions from $Z_{S^\star}$. At the same time, the ordering of behaviours induces strict partial "dominance" ordering $\lhd^d_{Z_{S^\star}}$ on the same set $Z_{S^\star}$:

$$Z_{^aS} \lhd^d_{Z_{S^\star}} Z_{^bS} \iff_{\text{def}} \begin{cases} Z_{^aS} \leqslant Z_{^bS} \text{ and} \\ \text{there is a Borel set } B \subseteq N^\star \text{ with } \quad \mu(B) > 0, \\ \text{such that } Z_{^aS}(N) < Z_{^bS}(N) \text{ for every } \quad N \in B. \end{cases} \tag{11}$$

The strategy ${}^bS$ in (11) is dominated by the strategy ${}^aS$ and any reasonable ordering introduced on $S^\star$ must not take it as the optimal one: its consequences are worse than those of ${}^aS$ irrespectively of the inaccessible uncertainties. This motivates the key requirement on the constructed ordering of strategies.

Requirement 4.1 (*Inadmissibility of dominated strategies*). If ${}^aS \lhd_{S^\star} {}^bS$, then $(Z_{^bS} \lhd^d_{Z_{S^\star}} Z_{^aS})$ for every ${}^aS,\ {}^bS \in S^\star$. Every such ordering $\lhd_{S^\star}$ is called *admissible.*

The following proposition is a straightforward consequence of the representation of T by the Lebesgue integral (9).

Theorem 4.2 (Representation of admissible strategy ordering). *For each loss function* $Z \in Z^\star$, *the functional* T *leads, due to* (10), *to an admissible ordering on* $S^\star$ *whenever the kernel* $\Phi(\cdot, N)$ *is an increasing function of the first argument for* μ*-almost all* $N \in N^\star$.

The complete ordering $\lhd_{S^\star}$ is invariant with respect to the multiplication of T by any positive real number. Thus, without a loss of generality, we can assume that μ is a regular Borel probability measure on $N^\star$.

Let $\mathsf{Z_S} \in \mathsf{Z}^{\star}_{\mathsf{S}^\star}$. Using the change of variable formula, the integral in (9) gets transformed into:

$$\mathsf{T}(\mathsf{Z_S}) = \mathsf{T}(Z \circ \mathsf{W_S}) = \int_{\mathcal{Q}^\star} \Phi(Z(\mathcal{Q}), \mathsf{W}_\mathsf{S}^{-1}(\mathcal{Q}))\,\mathrm{d}\mu_\mathsf{S}, \tag{12}$$

where $\mu_\mathsf{S}(A) = \mu(\mathsf{W}_S^{-1}(A))$, for every Borel subset A of $\mathcal{Q}^\star$. Let us suppose that there exists a probability measure ν on $\mathcal{Q}^\star$ such that, for every $S \in \mathsf{S}^\star$, the probability measure μ_S is absolutely continuous w.r.t. ν. This means that there is a probability density function (pdf) $f_\mathsf{S} = \frac{\mathrm{d}\mu_\mathsf{S}}{\mathrm{d}\nu}$ so that the last integral in (12) becomes:

$$\int_{\mathcal{Q}^\star} \Phi\big(Z(\mathcal{Q}), \mathsf{W}_\mathsf{S}^{-1}(\mathcal{Q})\big) f_\mathsf{S}(\mathcal{Q})\,\mathrm{d}\nu. \tag{13}$$

We assume that the kernel Φ does not discern between "equiprobable" elements of $N^\star$ causing the same loss.

Requirement 4.2 (*Risk attitude*). *If $f_\mathsf{S}(\mathcal{Q}_1) = f_\mathsf{S}(\mathcal{Q}_2)$, for a fixed strategy S and a pair $\mathcal{Q}_1,\ \mathcal{Q}_2 \in \mathcal{Q}^\star$, then $\Phi\big(z, \mathsf{W}_S^{-1}(\mathcal{Q}_1)\big) = \Phi\big(z, \mathsf{W}_S^{-1}(\mathcal{Q}_2)\big)$ for every $z \in \mathbb{R}$.*

It is thus correct to define a function $\Omega : \mathbb{R} \times \mathrm{Range}(f_\mathsf{S})$ to $\mathbb{R}$ by:

$$\Omega(z, f_\mathsf{S}(\mathcal{Q})) = \Phi\big(z, \mathsf{W}_S^{-1}(\mathcal{Q})\big).$$

The formula (13) then reads as:

$$\mathsf{T}(\mathsf{Z_S}) = \int_{\mathcal{Q}^\star} \Omega(Z(\mathcal{Q}), f_\mathsf{S}(\mathcal{Q})) f_\mathsf{S}(\mathcal{Q})\,\mathrm{d}\nu =_{def} \mathsf{E}_{\mu_\mathsf{S}}[\Omega(Z, f_\mathsf{S})]. \tag{14}$$

The representation (14) of the strict complete preferential ordering $\lhd_{\mathsf{S}^\star}$ of the DM strategies $\mathsf{S} \in \mathsf{S}^\star$ has the following important methodological consequences:

1. The representation by the functional T separates description of the uncertainty μ in (9) and its influence on a posterior ordering of behaviours $\mathcal{Q} \in \mathcal{Q}^\star$ expressed by values of the specific loss function $Z(\mathcal{Q})$. Thus, the probability density function f_S can be interpreted as the objective description of the uncertainty entering the closed loop formed by the considered environment and a DM strategy S. The function f_S describes distribution of behaviour realisations for a given strategy: it is thus an objective model of the closed loop.
2. The kernel Ω reflects interaction between the uncertainty, projected into $\mathcal{Q}^\star$ and a posteriori observable loss $Z(\mathcal{Q})$. It models attitude of the participant to risk (neutral, risk prone, risk aware) or more generally, a non-trivial interactions between a posteriori consequences and their distribution. There are strong indications, that such a possibility is badly needed at least in risk-facing DM (cf. [26]).

The preceding construction thus amounts to finding the optimal strategy minimising the expected value in (14). The optimised strategy influences just the pdf f_S, which enters both the function Ω and—linearly—the expectation operator $\mathsf{E}_{\mu_\mathsf{S}}$. Let us stress that the occurrence of $f_\mathsf{S}(\mathcal{Q})$ in Ω is non-standard and represents the key generalisation brought by the proposed problem formulation.

The presented results indicate that neither the kernel Ω nor the loss function Z are unique. For every pair $P = (\Omega, Z)$, let f_{S^P} be the closed-loop model with P-optimal strategy S^P such that

$$\mathsf{S}^P \in \arg\min_{\mathsf{S}\in\mathsf{S}^\star} \int_{\mathcal{Q}^\star} \Omega(Z(\mathcal{Q}), f_\mathsf{S}(\mathcal{Q})) f_\mathsf{S}(\mathcal{Q})\,\mathrm{d}\nu. \tag{15}$$

Let $P^\star$ be the set of all possible pairs $P = (\Omega, Z)$, where Ω is a kernel $\mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}$ and $Z \in Z^\star$ (7). All designs of the optimal strategies that start from different pairs $P_1, P_2 \in P^\star$ and lead to the same closed-loop description are equivalent: precisely, $P_1, P_2 \in P^\star$ are equivalent when $f_{\mathsf{S}^{P_1}}(\mathcal{Q}) = f_{\mathsf{S}^{P_2}}(\mathcal{Q})$ for ν-almost all $\mathcal{Q} \in \mathcal{Q}^\star$.

Traditionally, the design starts with the choice of the pair P. It determines the optimal strategy through (15) and consequently the optimally tuned closed loop. The fully probabilistic design changes the specification order and formulates the problem as follows: the participant specifies first some strictly positive *ideal* pdf ${}^I\!f$ on $\mathcal{Q}^\star$ describing the ideally tuned closed loop. Then the participant selects a proximity measure $D(f\|g)$ on pdfs f, g over $\mathcal{Q}^\star$ and takes the strategy minimising the proximity $D(f_\mathsf{S}\|{}^I\!f)$ on $\mathsf{S}^\star$ as the optimal one.

In order to select a reasonable proximity measure for a given ideal ${}^I\!f$, we want to recover a pair ${}^IP = ({}^I\Omega, {}^IZ)$ such that

$$f_{\mathsf{S}^{{}^IP}}(\mathcal{Q}) = {}^I\!f(\mathcal{Q}), \quad \text{for } \nu\text{-almost all } \mathcal{Q} \in \mathcal{Q}^\star. \tag{16}$$

Among all equivalent pairs ${}^IP = ({}^I\Omega, {}^IZ)$ satisfying (16), we search for a representative pair ${}^IP$ meeting the following requirement, which was inspired by a related problem inspected in [2].

Requirement 4.3 (*Representative Pair ${}^IP = ({}^I\Omega, {}^IZ)$*). *For a given strictly positive ideal pdf ${}^I\!f$, there exists a pair ${}^IP$ satisfying (16) such that:*

1. ${}^I\Omega\Big({}^IZ(\mathcal{Q}), f_{\mathsf{S}^{{}^IP}}(\mathcal{Q})\Big) = \text{constant}$ for ν-almost all $\mathcal{Q} \in \mathcal{Q}^\star$.
2. ${}^I\Omega(z, \cdot)$ is continuously differentiable for every $z \in \mathbb{R}$.

The following proposition shows that under this requirement there is a little freedom in choosing the representative pair ${}^{I}P = ({}^{I}\Omega, {}^{I}Z)$.

Theorem 4.3 (*Form of the Pair* ${}^{I}P = ({}^{I}\Omega, {}^{I}Z)$). *Let* ${}^{I}f$ *be an arbitrary, strictly positive, ideal pdf and suppose that* Requirement 4.3 *holds true. Then for every* $\mathsf{S} \in \mathsf{S}^{\star}$ *and* ν*-almost all* $\mathcal{Q} \in \mathcal{Q}^{\star}$,

$$
\begin{aligned}
&{}^{I}\Omega({}^{I}Z(\mathcal{Q}), f_{\mathsf{S}}(\mathcal{Q})) = ({}^{I}A - {}^{I}B) \ln\left(\frac{f_{\mathsf{S}}(\mathcal{Q})}{{}^{I}f(\mathcal{Q})}\right) + {}^{I}B, \quad {}^{I}A \geqslant {}^{I}B, \text{ and thus,} \\
&\mathsf{E}_{\mu_{\mathsf{S}}}[{}^{I}\Omega({}^{I}Z, f_{\mathsf{S}})] = ({}^{I}A - {}^{I}B) \int_{\mathcal{Q}^{\star}} \ln\left(\frac{f_{\mathsf{S}}(\mathcal{Q})}{{}^{I}f(\mathcal{Q})}\right) f_{\mathsf{S}}(\mathcal{Q})\, \mathrm{d}\nu + {}^{I}B.
\end{aligned}
\tag{17}
$$

Proof. By taking Gâteaux derivative [1] of the minimised functional $\mathsf{E}_{\mu_{\mathsf{S}}}[{}^{I}\Omega({}^{I}Z, f_{\mathsf{S}})]$ at $f_{\mathsf{S}} = {}^{I}f$, we get the necessary condition for minimum for almost all $\mathcal{Q} \in \mathcal{Q}^{\star}$:

$$x \frac{\partial}{\partial x} {}^{I}\Omega({}^{I}Z(\mathcal{Q}), x) + {}^{I}\Omega({}^{I}Z(\mathcal{Q}), x) = {}^{I}A, \quad \text{for } x = {}^{I}f(\mathcal{Q}) = f_{\mathsf{S}^{IP}}(\mathcal{Q}), \tag{18}$$

and some constant ${}^{I}A$. Due to the Requirement 4.3, we get

$${}^{I}\Omega({}^{I}Z(\mathcal{Q}), {}^{I}f(\mathcal{Q})) = \text{constant} = {}^{I}B. \tag{19}$$

This implies that:

$$\frac{\partial}{\partial x} {}^{I}\Omega({}^{I}Z(\mathcal{Q}), x) = \frac{{}^{I}A - {}^{I}B}{x},$$

which has the solution

$${}^{I}\Omega({}^{I}Z(\mathcal{Q}), x) = ({}^{I}A - {}^{I}B) \ln(x) + {}^{I}C(\mathcal{Q}).$$

The condition (19) determines ${}^{I}C(\mathcal{Q})$ uniquely so that we obtain

$${}^{I}\Omega({}^{I}Z(\mathcal{Q}), x) = ({}^{I}A - {}^{I}B) \ln \frac{x}{{}^{I}f(\mathcal{Q})} + {}^{I}B.$$

The expression ${}^{I}A - {}^{I}B$ must be nonnegative since the constructed functional is to be minimised. The case ${}^{I}A - {}^{I}B = 0$ is not considered in the sequel since it renders all strategies equivalent. □

Remarks 4.1. [On Proposition 4.3 and its conditions]

1. The expectation (17) is an increasing affine transformation of Kullback–Leibler divergence

$$\mathsf{D}(f \| {}^{I}f) = \int_{\mathcal{Q}^{\star}} f(\mathcal{Q}) \ln\left(\frac{f(\mathcal{Q})}{{}^{I}f(\mathcal{Q})}\right) \mathrm{d}\mathcal{Q},$$

   which is studied in [18] as the widely used proximity measure of pdfs with a range of applications in DM, statistics and information theory. It has an exceptional position within a class of so called f-divergences [29].
2. The closed-loop description f_{S} enters into the optimised functional in a non-linear way. This is a source of strength as well as weakness of the FPD discussed in subsequent sections. Related considerations of conditional expectation as a possibly non-linear mapping can be found in [21].
3. It is worth stressing that the multi-modal ideal pdf allows a straightforward quantification of multiple-aims, which otherwise is taken as a hard extension of the standard single-aim Bayesian paradigm.

5. Relation of standard Bayesian DM and FPD

The standard Bayesian DM assumes that every pdf f_{S} enters the minimised functional only linearly. This results in considering the optimal strategy:

$$\mathsf{S}^{P} \in \arg\min_{\mathsf{S} \in \mathsf{S}^{\star}} \int_{\mathcal{Q}^{\star}} \Omega(Z(\mathcal{Q}), \mathcal{Q}) f_{\mathsf{S}}(\mathcal{Q})\, \mathrm{d}\nu, \tag{20}$$

where $\Omega(Z(\mathcal{Q}), \mathcal{Q}) = \Phi\left(Z(\mathcal{Q}), W_{\mathsf{S}}^{-1}(\mathcal{Q})\right)$ is assumed to be independent of the chosen strategy $\mathsf{S} \in \mathsf{S}^{\star}$—cf. (15), (12). In order to relate the task (20) to FPD, we recall well-known properties of the standard Bayesian DM.

Below, we exploit structure of the behaviour introduced in (1). From now on, we assume that the dominating probability measure ν has the product form on the three individual constituents (1) of behaviour. Integrals of the type $\int_{\alpha^{\star}} \cdot \mathrm{d}\nu$ mean that factor of v corresponding to the considered α is used only. From now on, we also assume that the set of all behaviours $\mathcal{Q}^\star$ is $\mathbb{R}^m$ for some positive integer m.

With the considered v, the pdf $f_{\mathsf{S}}(\mathcal{Q}) = f_{\mathsf{S}}(d^T, x^T)$ describing closed decision loop factorises as follows:

$$f_{\mathsf{S}}(d^T, x^T) = \overbrace{\prod_{t=1}^{T} \underbrace{f(\Delta_t, x_t | a^t, \Delta^{t-1}, x^{t-1})}_{\text{local environment model } \mathsf{M}_t}}^{\substack{\text{environment model} \\ \mathsf{M} = \mathsf{M}(d^T, x^T)}} \times \overbrace{\prod_{t=1}^{T} \underbrace{f(a_t | d^{t-1}, x^{t-1})}_{\text{rule } \mathsf{S}_t \text{ of the strategy } \mathsf{S}}}^{\substack{\text{strategy} \\ \mathsf{S} = \mathsf{S}(d^T, x^T)}}. \tag{21}$$

This form is correct when identifying $f(a_1 d^0, x^0|)$ with $f(a_1)$. The first factor in (21) is fixed as the participant is assumed to select the optimal strategy with a fixed environment model M. The second factor describes the optimised strategy S. By definition, the participant never observes x^T directly observed. Thus, natural conditions of DM [20] have to be met:

$$f(a_t | d^{t-1}, x^{t-1}) = f(a_t | d^{t-1}). \tag{22}$$

The condition above complies with domains of decision rules S_t in (2). The optimal strategy S^P within the standard Bayesian DM is found via well-known stochastic dynamic programming [3]. The following proposition was essentially proven in [13].

Theorem 5.1 (*Stochastic dynamic programming*). *Decision rules of the optimal strategy (21) are deterministic strategies described by*

$$f^P(a_t | d^{t-1}) = \delta(a_t - a^P(d^{t-1})) = \text{formal pdf concentrated on } a^P(d^{t-1}).$$

The values $a^P(d^{t-1})$ are minimising arguments in

$$\mathcal{V}(d^{t-1}) = \min_{a_t \in a_t^\star} \int_{\delta_t^\star} \mathcal{V}(d^t) f(\Delta_t | a_t, d^{t-1}) \, dv.$$

The construction is made recursively for $t = T, T-1, \ldots, 1$ starting from

$$\mathcal{V}(d^T) = \int_{x^{T\star}} \Omega(Z(d^T, x^T), d^T, x^T) f(x^T | d^T) \, dv.$$

Under natural conditions of control (22), the predictive pdf $f(\Delta_t|a_t, d^{t-1})$ and the "filtering" pdf $f(x^T|d^T)$ of unobserved x^T result from the Bayesian filtering described by the recursive formulas:

$$f(\Delta_t | a_t, d^{t-1}) = \int_{x^{t-1\star}} f(\Delta_t | a_t, d^{t-1}, x^{t-1}) f(x^{t-1} | d^{t-1}) \, dx^{t-1}, \quad \text{(prediction)}$$

$$f(x^t | d^t) \propto f(\Delta_t, x_t | a^t, \Delta^{t-1}, x^{t-1}) f(x^{t-1} | d^{t-1}), \quad \text{(filtering)}.$$

This proposition implies that f_{S^P} resulting from the standard Bayesian design (20) cannot serve directly as the ideal pdf ${}^I f$ in FPD as it violates the positivity of ${}^I f$ enforced by Requirement 4.3. The following simple proposition provides a technical tool for coping with this problem.

Theorem 5.2 (Lower bound on entropy of deterministic rules). *Any deterministic rule $f(a) = \delta(a - a^P)$ reaches the lower bound $\underline{H}$ of the entropy $H(f) = -\int_{a^\star} f(a) \ln(f(a)) \, dv$, where*

$$\underline{H} = \begin{cases} 0, & \text{for discrete-valued action } a, \\ -\infty, & \text{continuous-valued action } a. \end{cases}$$

Proof. Direct inspection solves discrete-valued case. In continuous-valued case, the formal pdf is Dirac delta function. This generalised function can be obtained as limit of positive pdfs [30], say normal ones with the expectation a^P and diagonal covariance matrix cI, where I is unit matrix and $c > 0$ approaches zero. For them, entropy equals $\frac{\ln(|2\pi cI|)}{2} \to -\infty$ for $c \to 0$. □

Consequently, if we select $\overline{H} < -\underline{H}$ and optimise (20) over the strategies meeting the constraint

$$\int_{\mathcal{Q}^\star} \mathsf{MS} \ln(\mathsf{S}) \, dv \leqslant \overline{H} < -\underline{H}, \tag{23}$$

the constraint (23) is always active. Moreover, when the constraint becomes less severe, i.e., $\overline{H} \to -\underline{H}$, then the optimal Bayesian strategy is recovered.

Kuhn–Tucker theorem [17] implies that the minimisation (20) under the constraint (23) reduces to minimisation of

$$\begin{aligned} \mathsf{S}^P_{\overline{H}} &\in \arg\min_{\mathsf{S} \in \mathsf{S}^\star} \int_{\mathcal{Q}^\star} [\Omega(Z(\mathcal{Q}), \mathcal{Q}) + \lambda(\overline{H}) \ln(\mathsf{S})] f_{\mathsf{S}}(\mathcal{Q}) \, dv \\ &= \arg\min_{\mathsf{S} \in \mathsf{S}^\star} \mathsf{D}(f_{\mathsf{S}} \| {}^{I\overline{H}} f) \quad \text{with} \\ {}^{I\overline{H}} f(\mathcal{Q}) &= \frac{\mathsf{M}(\mathcal{Q}) \exp\left[-\frac{1}{\lambda(\overline{H})} \Omega(Z(\mathcal{Q}), \mathcal{Q})\right]}{\int_{\mathcal{Q}^\star} \mathsf{M}(\mathcal{Q}) \exp\left[-\frac{1}{\lambda(\overline{H})} \Omega(Z(\mathcal{Q}), \mathcal{Q})\right] dv}. \end{aligned} \tag{24}$$

The positive Kuhn–Tucker multiplier $\lambda(\overline{H})$ converges to zero for $\overline{H} \to -\underline{H}$ and the optimal strategy $S^{P}_{\overline{H}}$ converges to the optimal strategy of the standard Bayesian DM. The ideal pdf ${}^{I\overline{H}}f(\mathcal{Q})$ (24) is positive.

When noticing that generically the optimal strategies obtained by FPD are randomised, we can summarise the achieved results in the following expressive way.

Theorem 5.3 (FPD vs. standard Bayesian DM)

1. *Any standard Bayesian DM problem (20) can be approximated to an arbitrary precision by the FPD problem given by the ideal pdf (24) by selecting sufficiently small positive $\lambda(\overline{H})$.*
2. *There are FPD's having no standard Bayesian DM counterpart.*

6. Conclusions

The paper provides an axiomatisation of DM under uncertainty that suits to closed decision loop. It advocates that fully probabilistic design of decision strategies is the proper way to address problems of this type. The following comments add observations related to FPD.

6.1. Advantages of the FPD

- Dynamic programming shows that stochastic optimisation can be based on iterative operations such as conditional expectation and minimisation (see Proposition 5.1 or [3]). The FPD has an explicit minimiser [14], so that the (almost) inevitable approximation task [24] is substantially simplified. The evaluation complexity of the optimal strategy can be simply seen on so called data-driven FPD when no internal quantities are present and $\mathcal{Q} = d^T$. The following proposition is proved in [13].

Theorem 6.1 (Solution of the data-driven FPD). *The optimal strategy minimising the KLD of*

$$f(\mathcal{Q}) = f(d^T) = \prod_{t=1}^{T} f(\Delta_t|a_t, d^{t-1}) f(a_t|d^{t-1})$$

on the ideal pdf ${}^{I}f(\mathcal{Q}) = {}^{I}f(d^T) = \prod_{t=1}^{T} {}^{I}f(\Delta_t|a_t, d^{t-1})\, {}^{I}f(a_t|d^{t-1})$ *has the form*

$$\begin{aligned}
f(a_t|d^{t-1}) &= {}^{I}f(a_t|d^{t-1}) \frac{\exp\left[-\omega(a_t, d^{t-1})\right]}{\gamma(d^{t-1})},\\
\gamma(d^{t-1}) &= \int_{a_t^{\star}} {}^{I}f(a_t|d^{t-1}) \exp\left[-\omega(a_t, d^{t-1})\right] \mathrm{d}\nu, \quad \text{for } t < T,\\
\omega(a_t, d^{t-1}) &= \int_{\Delta_t^{\star}} f(\Delta_t a_t, d^{t-1}|) \ln\left(\frac{f(\Delta_t|a_t, d^{t-1})}{\gamma(d^t)\, {}^{I}f(\Delta_t|a_t, d^{t-1})}\right) \mathrm{d}\nu.
\end{aligned}$$

The solution runs for $t = T, T-1, \ldots, 1$ starting with $\gamma(d^T) = 1$.

Notice that the restricted support of the ideal pdf on actions implies restricted support of the chosen strategy. Thus, the ideal pdf quantifies both decision aims and constraints.

- Multi-modal ideal pdf expresses "naturally" multiple decision aims [4]. There is no conceptual jump between single and multiple aim optimisation.
- In the multiple-participant context, the well-developed art of combining pdfs [19,6,5,12], can be extended to combination of preferences expressed by ideal pdfs [16]. The similar problem is much harder in the classical setting.
- Unlike in the standard Bayesian DM, the optimal strategies are randomised. It is much more realistic as any channel implementing the designed DM strategy has a finite capacity (cf. [25]), i.e., it is unable to implement non-randomised strategy.

6.2. Drawbacks of the FPD

General limitations of the FPD follow predominantly from the fact that preferences are quantified in a non-standard way:

- Expression of the real aims by ${}^{I}f$ is non-trivial and easily it may happen that the option made does not reflect them properly.
- The usual complete separation of the a posteriori loss and description of the uncertainty is broken. This argument is, however, valid only when the neutral risk attitude is (implicitly) assumed.

Acknowledgement

This research has been partially supported by grants GA ČR 102/08/0567, MŠMT 2C06001. Discussions with T.V. Guy, M. Janžura and J. Kracík improved the original ideas and text substantially.

References

[1] J.-P. Aubin, Optima and equilibria, volume 140 of Graduate Texts in Mathematics, Springer-Verlag, Berlin, second edition, 1998, An introduction to nonlinear analysis.
[2] J.M. Bernardo, Expected information as expected utility, The Annals of Statistics 7 (3) (1979) 686–690.
[3] D.P. Bertsekas, Dynamic Programming and Optimal Control, second ed., Athena Scientific, Nashua, US, 2001.
[4] J. Böhm, T.V. Guy, M. Kárný, Multiobjective probabilistic mixture control, in: P. Horáček, M. Šimandl, P. Zítek (Eds.), Preprints of the 16th IFAC World Congress, IFAC, Prague, 2005.
[5] C.J. Butz, S. Hua, J. Chen, H. Yao, A simple graphical approach for understanding probabilistic inference in Bayesian networks, Information Sciences 179 (6) (2009) 699–716.
[6] R.G. Cowell, A.P. Dawid, S.L. Lauritzen, D.J. Spiegelhalter, Probabilistic Networks and Expert Systems, Springer-Verlag, 1999.
[7] G. Debreu, Representation of a preference ordering by a numerical function, New York, 1954.
[8] P.C. Fishburn, Utility Theory for Decision Making, J. Wiley, New York, London, Sydney, Toronto, 1970.
[9] J.W. Guan, Z. Guan, D.A. Bell, Bayesian probability on boolean algebras and applications to decision theory, Information Sciences 97 (3-4) (1997) 267–292.
[10] T.V. Guy, M. Kárný, Stationary fully probabilistic control design, in: J. Filipe, J.A. Cetto, J.L. Ferrier, (Eds.), Proceedings of the Second International Conference on Informatics in Control, Automation and Robotics, Barcelona, INSTICC, 2005, pp. 109–112.
[11] M. Kárný, Towards fully probabilistic control design, Automatica 32 (12) (1996) 1719–1722.
[12] M. Kárný, J. Andrýsek, A. Bodini, T.V. Guy, J. Kracík, F. Ruggeri, How to exploit external model of data for parameter estimation? International Journal of Adaptive Control and Signal Processing 20 (1) (2006) 41–50.
[13] M. Kárný, J. Böhm, T.V. Guy, L. Jirsa, I. Nagy, P. Nedoma, L. Tesař, Optimized Bayesian Dynamic Advising: Theory and Algorithms, Springer, London, 2006.
[14] M. Kárný, T.V. Guy, Fully probabilistic control design, Systems & Control Letters 55 (4) (2006) 259–265.
[15] M. Kárný, T.V. Guy, Preference elicitation in fully probabilistic design of decision strategies, in: Proceedings of the 49th IEEE Conference on Decision and Control, IEEE, Atlanta, 2010.
[16] J. Kracík, On composition of probability density functions, in: J. Andrýsek, M. Kárný, J. Kracík (Eds.), Multiple Participant Decision Making, Advance Knowledge International, Adelaide, 2004, pp. 113–121.
[17] H.W. Kuhn, A.W. Tucker, Nonlinear programming, in: Proceedings of 2nd Berkeley Symposium, University of California Press, Berkeley, 1951, pp. 481–492.
[18] S. Kullback, R. Leibler, On information and sufficiency, Annals of Mathematical Statistics 22 (1951) 79–87.
[19] R.B. Nelsen, An Introduction to Copulas, Springer, New York, 1999.
[20] V. Peterka, Bayesian system identification, in: P. Eykhoff (Ed.), Trends and Progress in System Identification, Pergamon Press, Oxford, 1981, pp. 239–304.
[21] J. Pfanzangl, Characterizations of conditional expectations, The Annals of Mathematical Statistics 38 (2) (1967) 415–421.
[22] M.M. Rao, Measure Theory and Integration, John Wiley, New York, 1987.
[23] L.J. Savage, Foundations of Statistics, Wiley, New York, 1954.
[24] J. Si, A.G. Barto, W.B. Powell, D. Wunsch (Eds.), Handbook of Learning and Approximate Dynamic Programming, Wiley-IEEE Press, Danvers, 2004.
[25] C.A. Sims, Implications of rational inattention, Technical report, Department of Econometrics, Princeton University, 2002, 28 pages.
[26] C. Starmer, Developments in non-expected utility theory: The hunt for a descriptive theory of choice under risk, Journal of Economic Literature XXXVIII (2000) 332–382.
[27] N. Tishby, D. Ploani, Information theory of decisions and actions, in: Vassilis, Hussain, Taylor (Eds.), Perception-Reason-Action Cycle: Models, Algorithms and Systems, Springer, 2010.
[28] E. Todorov, Linearly-solvable Markov decision problems, in: B. Scholkopf, J. Platt, T. Hoffman (Eds.), Advances in Neural Information Processing, MIT Press, New York, 2006, pp. 1369–1376.
[29] I. Vajda, Theory of Statistical Inference and Information, Kluwer Academic Publishers, Dordrecht, 1989.
[30] V.S. Vladimirov, Generalized Functions in Mathematical Physics, Mir Publishers, Moscow, 1979.