# Cooperation via sharing of probabilistic elements

Miroslav Kárný•, Tatiana V. Guy•
Antonella Bodini*, Fabrizio Ruggeri*
•ÚTIA AV ČR, P.O. Box 18, 182 08 Prague 8
Czech Republic *IMATI CNR, Via E. Bassini 15, I-20133 Milan, Italy

**Summary**. Interactions of large societal groups exhibit predominantly a flat structure. It means that each member of the group has its aims, restricted perceiving, modelling, acting and evaluating abilities and interacts with a relatively small number of "neighbors". This is the fully scalable cooperation model worth of imitating. The paper introduces a formal model of this type with individual members being Bayesian decision makers who use so called fully probabilistic design of the optimal decision strategy. They are willing to cooperate with neighbors by providing them probabilistic distributions they use for their decision making (DM). At present research stage, interaction and communication structure are assumed to be given. Thus, the group DM is determined by specifying how the offered non-standard (probabilistic) fragmental information pieces should be exploited. The paper proposes a systematic procedure by formulating and solving the exploitation problem in a Bayesian way. Essentially, it: (i) takes the offered distributions as measured data; (ii) estimates an unknown global distribution describing the cooperating neighbors; (iii) projects this estimate on the domains of interest to the respective neighbors; (iv) approximate each projected distribution by the distribution, which has the form understandable to the decision maker.
The proposed procedure, whose properties are illustrated by simple examples, is of independent interest since it can be used for model and aim elicitation.

## 1. Introduction

Decision making theories help in an extreme width of activities covering estimation, prediction, pattern recognition, fault detection, control etc. Desirable theory of decision making should cope with: (i) uncertainty; (ii) incomplete information; (iii) multiple aims, and (iv) always limited abilities to sense, act and evaluate.

The requirements (i) and (ii) are well respected by the standard Bayesian decision making, e.g., DeGroot (1970); Berger (1985). The theory of fully probabilistic decision making (FPD), Kárný and Guy (2006); Kárný (2008), seems to meet the requirement (iii). The requirement (iv) can generally be met only via a distributed decision making. In spite of numerous attempts, e.g., Wolper and Tumer (1999, 2001); Vlassis (2003); Yiming et al. (2003); Stirling (2004), etc., it still lacks a systematic, widely accepted, solution. This paper proposes a promising cooperation methodology applicable under the following restrictions: (a) each decision maker, called participant within the cooperative context, is Bayesian and uses the FPD; (b) each participant is willing to share his decision elements with a small amount of neighbors, i.e., the Bayesian decision makers who have non-empty intersections of the operation domains with him; (c) each participant can extend his evaluation abilities only mildly; for a discussion see Section 2.

If (a), (b) and (c) are met, the cooperating participants have to be equipped with a computational interface that uses the offered fragmental probabilistic information pieces for improving decision elements processed by participants. This interface has to combine (merge) partial and imprecise information pieces about uncertain and random environment. This common problem has been addressed repeatedly, for instance, in connection with probabilistic expert systems Cowell et al. (2003), knowledge elicitation, O'Hagan et al. (2006), cooperation of participants, Andrýsek et al. (2007); Kárný et al. (2007), etc. None of the existing methodologies seems to be complete and automatic enough to be built in the thought interface. A proposition of such a methodology forms the core of this paper, Sections 3, 4. Its basic idea "*take all offered information pieces as outputs of noisy information channels and try to estimate parameters of the underlying source*" is not new, e.g., J.Oakley and O'Hagan (2005) and references there. The attempts made are mostly case specific and do not stress the simple methodological line: (i) take the gained information pieces as data; (ii) interpret the description of the uncorrupted information source as an unknown parameter; (iii) design suitable likelihood function and prior distribution; (iv) perform the standard Bayesian point estimation. The paper follows this line and gets plausible general results. This statement is illustrated by examples in Section 5. The design of the methodology dominates over technical and computational aspects of the problem. Consequently, the number of open problems is large, see Section 6. Complementary Section 7 provides Matlab scripts of algorithms generating the examples.

## 2. The cooperation scenario

Here, the considered decision-making scenario is introduced together with the basic notation used throughout the paper. The decision maker and participant as his instance is referred as "he" to keep language simple. This convention excludes neither females nor groups (multivariate decision maker) and mechanical decision makers as well as their combinations.

### *2.1. Bayesian decision maker*

A decision maker deals with a behavior $B$, a finite collection of random variables formed by: (i) actions $A$ that can be freely selected by him; (ii) observations that inform him about the state of the closed decision loop formed by the decision maker and his environment; (iii) internal states that he relates to the environment but he does not measure directly.

For presentation simplicity, the distribution of the behavior's realizations $b$ is described by a (regular) probability density function (pdf) $g(b)$ defined with respect to a dominating measure $\mathrm{d}b$ (typically, Lebesgue or counting one).

The decision maker designs and applies a causal decision rule $\mathsf{R}$ that maps known realizations $p \in p^*$ (a set of $p$ s) of the random "past" $P$ on admissible values $a \in a^*$ of actions $A$. By definition, the realizations of the past are known when the action $a$ is chosen. He acts under uncertainty. It means that in the formal decomposition $B = (F, A, P)$ the "future" $F$ unavailable to the causal decision rule is always non-void.

For any realization of behavior $b = (f, a, p) \equiv$(future, action,past), the pdf $g(b)$ can be factorized

$$g(b) \equiv g(f, a, p) = \underbrace{g(f|a,p)}_{\text{environment model}} \times \underbrace{g(a|p)}_{\text{randomized decision rule}} \times \underbrace{g(p)}_{\text{pdf of the past}}. \quad (1)$$

The considered decision maker tries to reach his decision aims in the best possible way under given constraints. The considered Bayesian decision maker relies on a chosen environment model $g(f|a, p)$ and selects his randomized decision rule $g(a|p)$ as a minimizer of an expected loss function expressing his aims. The Bayesian decision maker using the fully probabilistic design (FPD) expresses both his aims and constraints via an "ideal" pdf ${}^{I}g(b)$. The ideal pdf describes the desirable behavior's realizations: the higher is the value ${}^{I}g(b)$ the more desirable is $b$. He takes $\ln\left(\frac{g(b)}{{}^{I}g(b)}\right)$ as the loss function whose expectation is minimized. This defines the optimal decision rule

$$ {}^{O}g(a|p) \in \mathrm{Arg} \min_{\{g(a|p)\}} \underbrace{\int g(b) \ln\left(\frac{g(b)}{{}^{I}g(b)}\right) \mathrm{d}b}_{D(g||\,{}^{I}g) \equiv \text{ Kullback-Leibler divergence of the pdf } g(b) \text{ on the pdf } {}^{I}g(b)}, \qquad (2) $$

where $\int \cdot \, \mathrm{d}b$ means the definite multivariate integration over the domain of the integrand.

PROPOSITION 1 (SOLUTION OF THE FPD). *The optimal decision rule (2) is given by the formula*

$$ \begin{aligned} {}^{O}g(a|p) &\propto {}^{I}g(a|p) \exp[-\omega(a,p)], \ \propto \textit{ is equality without specifying unique normalizing factor}, \\ \omega(a,p) &\equiv \int g(f|a,p) \ln\left(\frac{g(f|a,p)}{{}^{I}g(f|a,p)}\right) \mathrm{d}f. \end{aligned} \qquad (3) $$

PROOF. The assertion follows from the following string of equalities and from the fact that the Kullback-Leibler divergence $D(g||h)$ reaches its smallest zero value for $g = h$, Vajda (1989). In the string, the respective $g$s and their ideal counterparts ${}^{I}g$ are regular pdfs with respect to measures $\mathrm{d}b$, $\mathrm{d}p, \mathrm{d}a$, $\mathrm{d}f$.

$$ D(g||\,{}^{I}g) = \int g(b) \ln\left(\frac{g(b)}{{}^{I}g(b)}\right) \mathrm{d}b = \int\int\int g(f,a,p) \ln\left(\frac{g(f,a,p)}{{}^{I}g(f,a,p)}\right) \mathrm{d}f\, \mathrm{d}a\, \mathrm{d}p \underbrace{=}_{\text{Fubini}} $$

$$ \int g(p) \left\{ \ln\left(\frac{g(p)}{{}^{I}g(p)}\right) + \int g(a|p) \left[ \ln\left(\frac{g(a|p)}{{}^{I}g(a|p)}\right) + \underbrace{\int g(f|a,p) \ln\left(\frac{g(f|a,p)}{{}^{I}g(f|a,p)}\right) \mathrm{d}f}_{\equiv \omega(a,p)} \right] \mathrm{d}a \right\} \mathrm{d}p $$

$$ = \underbrace{\int g(p) \ln\left(\frac{g(p)}{{}^{I}g(p)}\right) \mathrm{d}p - \ln\{\int {}^{I}g(a|p) \exp[-\overbrace{\omega(a,p)}^{\text{defined in (3)}}]\, \mathrm{d}a\}}_{\text{the term independent of the optimized } g(a|p)} + \int g(p) D(g(a|p)||\underbrace{{}^{O}g(a|p)}_{\text{defined in (3)}})\, \mathrm{d}p. $$

REMARK 1.

- *The ideal pdf ${}^{I}g(b)$ indeed describes constraints on actions as the support of the optimal decision rule (3) is contained in the support of its ideal counterpart.*
- *The explicit formula for the optimal decision rules is obtained even when designing decision strategy, i.e., a sequence of decision rules, Kárný and Guy (2006).*

### 2.2. Cooperating participants

In quest for a (relatively) realistic, normative and scalable cooperation scheme, we adopt:

Assumption 1 (Selfish Bayesian participants of given abilities).

*(a) Participants are Bayesian decision makers that use the fully probabilistic design.*
*(b) Participants, labelled by $\iota \in \{1, 2, 3, \ldots\}$, have fixed DM elements consisting of*

$b_\iota^* \equiv$ *the set of behaviors considered by the ιth participant,*

$p_\iota^* \equiv$ *the set of pasts considered by the ιth participant,*

$a_\iota^* \equiv$ *the set of actions considered by the ιth participant,*

$f_\iota^* \equiv$ *the set of futures considered by the ιth participant,*

$g_\iota(f_\iota|a_\iota, p_\iota) \equiv$ *the pdf used by the ιth participant for modelling his environment, (1),*

$g_\iota(a_\iota|p_\iota) \equiv$ *the pdf describing the decision rule considered by the ιth participant, (1),*

$g_\iota(p_\iota) \equiv$ *the pdf describing the past related to the ιth participant, (1),*

${}^I g_\iota(b_\iota) \equiv$ *the ideal pdf expressing the aims and constraints of the ιth participant.*

*(c) Any participant ι has a fixed finite collection of neighbors, i.e., participants who have non-empty intersections of behaviors' sets with $b_\iota^*$.*

Selfishness of the participants is reflected by the assumption that their DM elements are uninfluenced by the presence of their neighbors in their environment. Without additional measures, their success in DM depends heavily on their ability to model the environment even under the presence of active neighbors. It is strongly limited by their limited abilities to perceive, model, evaluate and act. The methodology described here takes these limitations as given and tries to design participants' cooperation without excessive demands on them:

Assumption 2 (Cooperation via sharing DM elements).

*(a) Each participant ι is willing to offer some projections (marginal or conditional pdfs) of pdfs $g_\iota(b_\iota)$, ${}^I g_\iota(b_\iota)$ to his neighbors and they are willing to do the same for him.*
*(b) Each participant is willing to allocate a "reasonable" part of his evaluation abilities to an interface that modifies his pdfs $g_\iota(b_\iota)$, ${}^I g_\iota(b_\iota)$ by offered probabilistic information.*
*(c) There is no external mechanism for distinguishing between reliability and importance of the pdfs offered by neighbors.*

Remark 2. *Assumption 2 (a) restricts the scheme to still selfish but* cooperating *participants. The technical core of this paper consists of creating the DM-elements-modifying interface. The term "reasonable" in Assumption 2 (b) anticipates that the participant will be able to perform the proposed evaluations. This excludes generic use of the theory of incomplete (Bayesian) games, Harsanyi (2004), which force the participant to model his environment including his neighbors.*

*Assumption 2 (c) is the most restrictive one and models insufficient reasons for making distinctions. This assumption can be relaxed in re-fined cooperation schemes in which the participant extends his actions by an active weighting of the offered information pieces.*

## 3. Interface for exploiting the offered projections

Selfish cooperating participants have a chance to improve their performance when equipped by a universal interface that processes the additional information coming from their neighbors. The constructed interface processes the information pieces offered to a specific participant $\iota$. Under Assumption 1 (c), we can consider just a fixed finite set of neighbors $\iota \in \iota^* \equiv \{1, 2, \ldots, \mathring{\iota}\}$. They act on the union of behaviors' sets

$$b^* \equiv \cup_{\iota \in \iota^*} b^*_\iota. \tag{4}$$

The information pieces concerning the closed decision loop models $g_\iota(b_\iota),\ \iota \in \iota^*$, and their ideal counterparts ${}^I g_\iota(b_\iota),\ \iota \in \iota^*$, are processed in the same way. Thus, we can focus on the former ones only.

### *3.1. Combination of pdfs as Bayesian estimation*

When designing the interface, we can think of the collection of neighbors as a multivariate participant dealing with a global closed decision loop model $g(b)$. His action realization is $a = (a_1, \ldots, a_{\mathring{\iota}})$. His past $p$ and future $f$ realizations are composed of $\{p_\iota\}_{\iota \in \iota^*}$, $\{f_\iota\}_{\iota \in \iota^*}$ in a complex way whose details unimportant within the considered context. The global pdf $g(b)$ is unknown to respective neighbors. The interface has to process projections of imprecise (noisy) versions $g_\iota(b)$ of the global pdf $g(b)$, $b \in b^*$. For a fixed $\iota \in \iota^*$, let us decompose $b$ in slightly different way than before. This decomposition reflects extension of $b_\iota$ to $b$ by introducing the part $u_\iota$ unconsidered by the participant $\iota$ and gives the common name $c_\iota$ to the action $a_\iota$ and the past $p_\iota$

$$\begin{aligned} b &\equiv (f_\iota, c_\iota, u_\iota) \\ &\equiv \text{(future, conditioning formed by action and past, unconsidered)} \\ &\quad\ \text{variables for } \iota\text{th participant.} \end{aligned} \tag{5}$$

Then, the general form of the projections dealt with is

$$g_\iota(f_\iota|c_\iota) = \frac{\int g_\iota(f_\iota, c_\iota, u_\iota)\, \mathrm{d}u_\iota}{\int g_\iota(f_\iota, c_\iota, u_\iota)\, \mathrm{d}f_\iota\, \mathrm{d}u_\iota} \equiv \mathsf{O}_\iota\left[g_\iota(b)\right]. \tag{6}$$

The introduced operator $\mathsf{O}_\iota$, acting on the pdf $g_\iota(b)$ and generating environment model used by $\iota$th participant, is non-linear but deterministic. This simple observation helps us in keeping complexity of the constructed interface low. We apply this operator to the final combination only.

The combination is constructed initially assuming that the joint pdfs $g_\iota(b),\ \iota \in \iota^*$, are available for all $b^*$. Then, we address the general case when only projections $g_\iota(f_\iota|c_\iota)$ derived from the joint pdfs $g_\iota(b)$ are available.

### *3.2. Point estimate of the global pdf*

The considered construction of the pdfs' combination (merging) casts the problem into the standard Bayesian framework. The pdfs $g_\iota(b),\ b \in b^*,\ \iota \in \iota^*$, are the "measured" data that are "noisy" versions of the unknown problem parameter, which is the global pdf $g(b),\ b \in b^*$. The point estimate $\hat{g}(b)$ of the unknown parameter $g(b)$ based on the data $g_\iota(b),\ b \in b^*,\ \iota \in \iota^*$, is the "natural" Bayesian merger searched for.

We search for a point estimate $g(\cdot|\hat{\mathcal{V}})$ of $g(\cdot)$ within a class of estimators $g(\cdot|\mathcal{V})$ given by a statistic $\mathcal{V} \in \mathcal{V}^*$ under:

ASSUMPTION 3 ("LOCALNESS PRINCIPLE"). *The relation of values $g(b)$, $g_\iota$, $\iota \in \iota^*$, matters in estimation of $g(b)$ and construction of $g(b|\hat{\mathcal{V}})$.*

If we: (i) stay within Bayesian framework; (ii) consider proper smooth loss functions; (iii) adopt Assumption 3, then the best point estimate $g(b|\hat{\mathcal{V}})$ of $g(b)$ is to be the minimizer of the expected value of the Kullback-Leibler divergence, Bernardo (1979). Its minimization is equivalent to the minimization of the expected value of the Kerridge inaccuracy, Kerridge (1961). It gives the rule for selecting the best $\hat{\mathcal{V}}$

$$\hat{\mathcal{V}} \in \text{Arg} \min_{\mathcal{V} \in \mathcal{V}^*} - \int \mathsf{E}\left[g(b)|g_\iota(b), \iota \in \iota^*\right] \ln(g(b|\mathcal{V}))\, \mathrm{d}b, \tag{7}$$

where the expectation is taken over uncertainty about values of the global pdf $g(b)$.

Thus, the construction of $g(b|\hat{\mathcal{V}})$ relies on the evaluation of the conditional expectation

$$\hat{g}(b) \equiv \mathsf{E}\left[g(b)|g_\iota(b), \iota \in \iota^*\right]. \tag{8}$$

Note that Assumption 3 allows us to use "shorthand" condition $g(b)$, instead of the more correct and complex $g(\cdot)$, $b$.

### *3.3. Knowledge on modelled relation*

In order to obtain the conditional expectation (8), we have to model the relation of the data to the unknown parameter. Definitely, the model can hardly be unique. Here, we present two unsatisfactory and one promising variants. Their choice was driven by the wish to make the estimation feasible, see Assumption 2 (b). We temporarily assume that the treated pdfs have the common domain $b^* = b^*_\iota$, $\forall \iota \in \iota^*$, (4). This assumption is relaxed in Section 4.

All meaningful models have to respect the nature of the processed data (the pdfs $g_\iota(b)$, $\iota \in \iota^*$,) and of the estimated parameter (the global pdf $g(b)$). They are non-negative functions having the integral equal to unity. We ignore the latter condition in modelling as it can be simply respected by normalization of the final merger. Various models are obtained by complementing prior knowledge by various additional modelling assumptions. Let us list first the common:

ASSUMPTION 4 (CONDITIONS ON PDF $G(g_\iota(b)|g(b), \cdot)$ RELATING $\{g_\iota\}_{\iota \in \iota^*}$ TO $g$).

(a) *Finite-dimensional functions $\Psi$, $\Omega$ exist such that*

$$\mathsf{E}[\Psi(g_\iota(b))|g(b), \Psi, \Omega] \equiv \int \Psi(v) G(v|g(b), \Psi, \Omega)\, \mathrm{d}v = \Omega(g(b)), \ \forall b \in b^*, \ \iota \in \iota^*, \tag{9}$$

*i.e., the expectation $\mathsf{E}[\cdot|g(b), \Psi, \Omega]$ is made over the random mechanism generating the values $v = g_\iota(b)$, $\iota \in \iota^*$.*

(b) *Pdfs $g_\iota$, $\iota \in \iota^*$, provided by different participants are conditionally independent for the global pdf $g$ given.*

(c) *Values $v_\iota$, $\tilde{v}_\iota$, of the pdf $g_\iota$, $v_\iota \equiv g_\iota(b)$, $\tilde{v}_\iota = g_\iota(\tilde{b})$ are independent for $b \neq \tilde{b}$ when $g, \Psi, \Omega$ are given.*

(d) *The constructed distribution of the pdfs' values* $v_\iota = g_\iota(b)$, $\iota \in \iota^*$, $b \in b^*$, *has the highest entropy among those meeting (a), (b), (c).*

REMARK 3. *Assumption (a) expresses our hope that the individual pdfs are related to the global pdf. Respective models, discussed below, differ in the choice of functions* $\Psi$ *and* $\Omega$ *in (9). Their finite-dimensional ranges are inevitable for meeting Assumption 2 (b). Independence of those functions of* $\iota$ *is a sort of common scaling, which is consistent with the assumed lack of preferences among participants, cf. Assumption 2 (c).*

*Assumption (b) postulates independence of personal deviations and it is predominantly acceptable. The independence assumption (c) and the assumption (d) offer an extreme freedom in forms of the merged pdfs. This extreme attitude is consistent with the attempt to create a universal interface.*

Both data and estimated parameter are generally infinite dimensional and should be treated as random processes. This machinery can be avoided as we search for an estimate $\hat{g}(b)$ of $g(b)$ specified for each $b \in b^*$ (Assumption 3). Indeed, if we take any finite collection of different behaviors $(b, {}^1b, \ldots, {}^nb)$, $n < \infty$, then Assumption 4 (b), (c) restrict the search to finite-dimensional distributions of the random vector

$$[\underbrace{g_1(b), \ldots, g_{\hat{\iota}}(b)}_{\alpha}, \underbrace{g_1({}^1b), \ldots, g_{\hat{\iota}}({}^1b), \ldots, g_1({}^nb), \ldots, g_{\hat{\iota}}({}^nb)}_{\beta}] \tag{10}$$

with conditionally independent entries. Thus, the distribution of the part of interest $\alpha$ in (10) is uninfluenced by the distribution of the extendable part $\beta$. Moreover, the constructed parametric model is fully determined by the marginal pdf $G(v|g, \Psi, \Omega)$ on values $v = g_\iota(b)$ of the modelled pdfs.

Assumption 4 (d) defines $G(v|g(b), \Psi, \Omega)$ as an entropy maximizer restricted by linear-in-$G$ constraint implied by Assumption 4 (a). This convex optimization problem has the solution

$$G(v|g(b), \Psi, \Omega) = \frac{\exp[-\lambda'(b)\Psi(v)]}{\int_0^\infty \exp[-\lambda'(b)\Psi(v)]\,\mathrm{d}v}, \ ' \text{ denotes transposition,} \tag{11}$$

with the vector function of Lagrangian coefficients $\lambda(b)$ solving the equation

$$\int \Psi(v) G(v|g(b), \Psi, \Omega)\,\mathrm{d}v = \Omega(g(b)), \ b \in b^*. \tag{12}$$

The choice of the functions $\Psi$, $\Omega$ has to guarantee that integral in (11) is finite and (12) (simply) solvable. The subsequent sections discuss candidates of this type.

### *3.4. Exponential model*

In the cooperative context, it is "natural" to assume that values of the pdfs offered by participants are unbiased guesses of the values of the unknown global pdf $g$, i.e.,

$$\mathsf{E}[g_\iota(b)|g(b)] = g(b). \tag{13}$$

When considering just this, i.e., choosing $\Psi$, $\Omega$ in (9) as one-dimensional identities, the constructed parametric model (11), (12) is the exponential one with $\lambda(b) = 1/g(b)$

$$G(g_\iota(b)|g(b), (\Psi, \Omega) \equiv \textit{identities}) = G(g_\iota(b)|g(b)) = \chi(g_\iota(b) \geq 0)\frac{1}{g(b)} \exp\left[-\frac{g_\iota(b)}{g(b)}\right], \tag{14}$$

with $\chi(\cdot)$ denoting indicator of the set in its argument.

The Bayesian parameter estimation with the model (14) is easy. For the completely flat improper prior distribution on $g(b)$, it has posterior Gamma distribution with $\mathring{\iota}$ degrees of freedom and expectation equal to the arithmetic mean of the measurements

$$\bar{g}_a(b) = \mathsf{E}[g(b)|g_\iota(b), \iota \in \iota^*] \equiv \frac{1}{\mathring{\iota}} \sum_{\iota \in \iota^*} g_\iota(b). \tag{15}$$

This merging is known as arithmetic pooling of opinions, O'Hagan et al. (2006). It is widely used but it is adequately criticized for sensitivity to outlying opinions, which make it spread too much. More importantly, the exponential model does not respect the fact that relative errors of small values are usually larger than that of large values. The exponential model is unable to describe this situation. Indeed, the exponential model has the property

$$\frac{\text{standard deviation}}{\text{expectation}} = 1. \tag{16}$$

### 3.5. Truncated normal model

In order to avoid the unpleasant property (16), we add to the unbiasedness (13) the requirement on the second moment

$$\mathsf{E}[g_\iota^2(b)|g(b), \rho(b)] = \rho(b) + g^2(b), \;\; \rho(b) > 0,$$

i.e., $\Psi(v) = [v, v^2],\ \Omega(g(b)) = [g(b), \rho(b) + g^2(b)]$.

In this case, the parametric distribution (11) becomes a normal one, truncated on non-negative domain of the modelled values

$$G(g_\iota(b)|g(b), \rho(b)) = \chi(g_\iota(b) \geq 0) \frac{\mathcal{N}_{g_\iota(b)}(\mu(b), r(b))}{\int_0^\infty \mathcal{N}_v(\mu(b), r(b))\,\mathrm{d}b}, \;\; \mathcal{N}_v(\mu, r) \equiv \frac{\exp\left[-\frac{(v-\mu)^2}{2r}\right]}{\sqrt{2\pi r}}.$$

Evaluation of the posterior pdf is relatively complex due to the dependence of the denominator on unknown parameters $\mu(b)$, $r(b)$. In principle, it is possible to developed an adequate, relatively fast, numerical procedure for doing it. The real trouble is that parameters $\mu(b)$, $r(b)$ are related to moments in relatively complex way. For instance, if we rely on unbiased measurements (13), we have to respect the highly non-linear relation

$$g(b) \;\; = \;\; \mu(b) + \sqrt{r(b)} \frac{\mathcal{N}_{v=0}(\mu(b), r(b))}{\int_0^\infty \mathcal{N}_v(\mu(b), r(b))\,\mathrm{d}v}.$$

Relation of the parameter $r(b)$ to the variance $\rho(b)$ is even more complex Greene (2003). The overall complexity makes us to drop the development in this direction.

### 3.6. Log-normal model

Here, we again require unbiasedness and allow variance of deviations $g_\iota(b) - g(b)$ be tuned independently. The non-negativity of processed data is, however, respected by relating their logarithms. We have to adopt:

Assumption 5 (Partial compatibility of pdfs). *The pdfs $g_\iota(b)$, $\iota \in \iota^*$, and $g(b)$ are strictly positive on $b^*$.*

Under Assumption 5 and modelling Assumptions 4 with

$$\mathsf{E}[\ln(g_\iota(b))|g,\rho] = \ln(\mu(b)),\ \mathsf{E}[\ln^2(g_\iota(b))|g,\rho] \quad = \quad \rho(b) + \mu^2(b),$$

we get $G(v|\rho,\rho)$ (11) in log-normal form. The unbiasedness requirement $\mathsf{E}[g_\iota(b)|g(b),\rho(b)] = g(b)$ and the formula for the mean of the log-normal distribution lead to the definition

$$\log(\mu(b)) \equiv \log(g(b)) - 0.5\rho(b).$$

We take completely flat prior on $\rho(b)$. To get finite $\hat{g}(b)$ (8), the prior pdf has to fall to infinity faster than $\exp\left(-\frac{2\rho}{\mathring{\iota}}\right)$. This makes us to select

$$G(g(b),\rho(b)) \propto \exp\left(-\frac{(2+\varepsilon)\rho(b)}{\mathring{\iota}}\right),\ \varepsilon > 0 \text{ arbitrarily small.}$$

It gives the posterior pdf $G(g(b),\rho(b)|g_\iota(b),\iota\in\iota^*)$ of the form (the argument $b$ is temporarily dropped)

$$G(g,\rho|g_\iota,\iota\in\iota^*) \quad \propto \quad \rho^{-\frac{\mathring{\iota}}{2}} \exp\left(-\frac{(2+\varepsilon)\rho}{\mathring{\iota}}\right) \exp\left\{-\frac{\mathring{\iota}}{2\rho}\left[\left(\ln\left(\frac{g}{\bar{g}}\right) - 0.5\rho\right)^2 + \hat{\rho}\right]\right\} \quad (17)$$

$$\bar{g} \quad \equiv \quad \left[\prod_{\iota\in\iota^*} g_\iota\right]^{\frac{1}{\mathring{\iota}}} \quad \text{(geometric mean)} \quad (18)$$

$$\hat{\rho} \quad \equiv \quad \frac{1}{\mathring{\iota}} \sum_{\iota\in\iota^*} \ln^2(g_\iota) - \bar{g}^2 \quad \text{(normalized least-squares remainder).} \quad (19)$$

Subsequent evaluation of integrals $E_k,\ k = 0,1$ is needed to find normalization of the posterior pdf (17) and to compute its first moment. {text} written in the equality chain indicates the used substitution.

$$\begin{aligned} E_k \quad &\equiv \quad \int_0^\infty g^k \exp\left\{-\frac{\mathring{\iota}}{2\rho}\left[\left(\ln\left(\frac{g}{\bar{g}}\right) - 0.5\rho\right)^2\right]\right\} \mathrm{d}g \\ &= \quad \left\{x = \ln\left(\frac{g}{\bar{g}}\right) - 0.5\rho,\ g = \bar{g}\exp(x+0.5\rho),\ \mathrm{d}g = g\,\mathrm{d}x\right\} \\ &= \quad \bar{g}^{k+1}\exp(0.5(k+1)\rho)\int_{-\infty}^{\infty} \exp\left\{(k+1)x - \frac{\mathring{\iota}x^2}{2\rho}\right\} \mathrm{d}x \\ &= \quad (\rho 2\pi/\mathring{\iota})^{0.5}\bar{g}^{k+1}\exp\left[\frac{(k+1)^2\rho}{2\mathring{\iota}}\right]. \end{aligned}$$

This leads to the factorized form of the posterior pdf in which the last factor is proportional to the normalized posterior pdf of $g$ conditioned, moreover, on $\rho$

$$G(g,\rho|g_\iota,\iota\in\iota^*) \propto \rho^{-\frac{\mathring{\iota}-1}{2}} \exp\left\{-\frac{\mathring{\iota}\hat{\rho}}{2\rho} - \frac{(3+\varepsilon)\rho}{2\mathring{\iota}}\right\} \frac{\exp\left\{-\frac{\mathring{\iota}}{2\rho}\left(\ln\left(\frac{g}{\bar{g}}\right) - 0.5\rho\right)^2\right\}}{\rho^{0.5}\bar{g}\exp\left(\frac{\rho}{2\mathring{\iota}}\right)}.$$

It gives the posterior expectation of $g$ conditioned moreover on $\rho$

$$\mathsf{E}[g|\rho,g_\iota,\iota\in\iota^*] = E_1/E_0 = \bar{g}\exp\left(\frac{3}{2\mathring{\iota}}\rho\right).$$

The marginal posterior pdf on $\rho$ reads

$$G(\rho|g_\iota, \iota \in \iota^*) \quad \propto \quad \rho^{-\frac{\mathring{\iota}-1}{2}} \exp\left\{-\frac{\mathring{\iota}\hat{\rho}}{2\rho} - \frac{(3+\varepsilon)\rho}{2\mathring{\iota}}\right\}.$$

This is a proper pdf and provides the finite expectation of $\exp\left(\frac{3\rho}{\mathring{\iota}}\right)$ needed for the evaluation of $\hat{g}(b) \equiv \mathsf{E}[g|g_\iota, \iota \in \iota^*] = \mathsf{E}\left\{\mathsf{E}[g|\rho, g_\iota, \iota \in \iota^*]|g_\iota, \iota \in \iota^*\right\}$. The evaluation of the "outer" expectation has no closed form. This makes us to substitute it by inserting maximum a posteriori probability estimate of $\rho$ into the "inner" expectation while letting $\varepsilon \to 0$. It gives

$$\hat{g}(b) \equiv \mathsf{E}\left[g(b)|g_\iota(b), \iota \in \iota^*\right] \approx \bar{g}(b) \exp\left[\frac{\mathring{\iota}}{\mathring{\iota}-1} \frac{2\hat{\rho}(b)}{1+\sqrt{1+\frac{12\hat{\rho}(b)}{(\mathring{\iota}-1)^2}}}\right]. \tag{20}$$

Note that the approximate equality has to be taken as approximate proportionality.

The formula (20) increases the geometric mean (18) whenever the normalized least squares remainder $\hat{\rho}(b)$ (19) is large. It is intuitively plausible property illustrated by Figure 1 presenting the merging results according to the formula (20). Details of the experiment are in the script testsingle generating it, see Section 7.

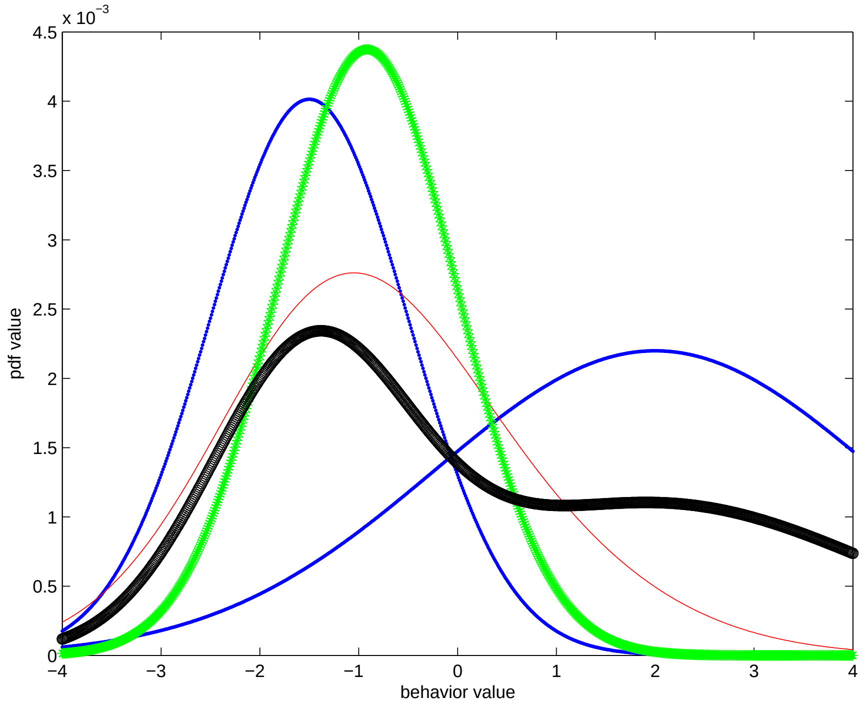


**Fig. 1.** Comparison of various merging possibilities (the script testsingle in Section 7): processed pdfs blue dots; the proposed merger red line; geometric mean (18) green stars; arithmetic mean (15) black circles.

## 4. Processing fragmental information pieces

The simplicity of the merging described in previous section stems from the fact that the inherent complexity of the projecting operators (6) is avoided. This makes us to *extend the processed marginal and conditional pdfs to the joint space of behaviors $b^*$ and then combine these extensions according to (20).* This idea is elaborated here.

The $\iota$th participant deals with, cf. (5),

$$\begin{aligned} b_\iota &\equiv (f_\iota, c_\iota) \\ &\equiv \text{(modelled, conditioning) variables for } \iota\text{th participant}, \end{aligned}$$

where at least $f_\iota$ is non-void.

Let us construct an extension ${}^e g_\iota(b)$ of the given pdf $g_\iota(f_\iota|c_\iota)$ to the global space $b^*$ of behaviors $b$. This extension should approximate the global pdf as tightly as possible, i.e., Bernardo (1979), to minimize the expected Kerridge distance $-\int \hat{g}(b)\ln({}^e g_\iota(b))\,\mathrm{d}b$. At the same time, it should preserve the given pdf $g_\iota(f_\iota|c_\iota)$. These requirements determine the extension *uniquely*

$${}^e g_\iota(b) = \hat{g}(u_\iota|f_\iota, c_\iota) g_\iota(f_\iota|c_\iota)\hat{g}(c_\iota), \tag{21}$$

where $\hat{g}(\cdot|\cdot)$ are projections of the constructed estimate $\hat{g}(b)$ (8) of the global pdf $g(b)$.

Inserting the extensions ${}^e g_\iota(b)$ into the formula (20) for the estimate of the global pdf, we get the *final* non-linear equation for it

$$\hat{g}(b) \quad \propto \quad {}^e\bar{g}(b) \exp\left[\frac{\mathring{\iota}}{\mathring{\iota}-1} \frac{2\,{}^e\hat{\rho}(b)}{1+\sqrt{1+\frac{12\,{}^e\hat{\rho}(b)}{(\iota-1)^2}}}\right]. \tag{22}$$

The left superscript ${}^e$ stresses that the geometric mean (18) and normalized least-squares remainder (19) are evaluated from the pdfs ${}^e g_\iota(b),\ \iota \in \iota^*$ (21).

The explicit solvability of (22) can hardly be expected but the equation is "naturally" prepared for solution by successive approximations.

Note that after finding $\hat{g}(b)$, it is necessary to compute its $\iota$-specific projections $\hat{g}_\iota(f_\iota|c_\iota) \equiv \mathsf{O}_\iota(\hat{g}(b))$ (6) as mergers provided to respective participants.

## 5. Illustrative examples

### *5.1. Processing of fragmental information*

The example demonstrates behavior of the proposed merging (22) in a simple case with two-dimensional behavior $b$. The 1st participant provides only marginal pdf on the 1st entry of $b$, the 2nd one offers the joint pdf. The proposed methodology is implemented in the script testfragment, Section 7, which generated self-explaining illustrative Figures 2, 3, 4 and 5.

The shown results were obtained after 8 iterations when the norm between two successive iterations dropped below the threshold 1e-5.

Note that the results are strongly influenced by the correlation between $b$-entries considered by the 2nd participant: for higher correlations, the shifts of the marginal pdf describing the 2nd behavior entry are more pronounced.

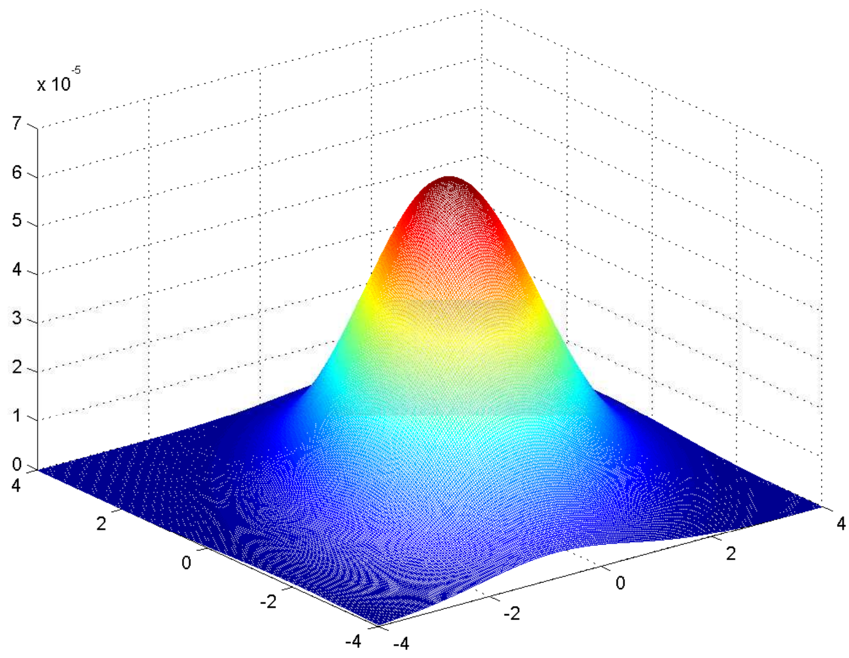


**Fig. 2.** Two-dimensional merging result. The 1st participant offered marginal pdf on the axis heading to the right (the script testfragment in Section 7).

### *5.2. Elicitation of prior knowledge*

This example indicates the use of the developed methodology to knowledge elicitation.

Let us assume that the participant $\iota = 1$ has a prior pdf $g_1(b)$ on behavior $b$ and the participant $\iota = 2$ offers him partial information in the form of a (generalized) moment

$$\mu = \int \phi(b) g_2(b)\, \mathrm{d}b, \tag{23}$$

where the finite-dimensional vector $\mu$ and the vector function $\phi(b)$ are the offered probabilistic elements. In harmony with the proposed methodology, the offered information (23) is extended to the pdf ${}^e g_2(b)$. The extension ${}^e g_2(b)$ is merged with $g_1(b)$ according to (22). The extension ${}^e g_2(b)$ is again taken as the pdf nearest to the constructed merger $\hat{g}(b)$ in Kerridge-inaccuracy sense under the constraint (23). It has the form

$${}^e g_2(b) \propto \hat{g}(b) \exp[-\zeta' \phi(b)] \tag{24}$$

with the vector $\zeta$ chosen so that $g_2(b) = {}^e g_2(b)$ in (24) meets (23).

Figure 6 presents the result of such merging with scalar $b$, $\mu$ and $\phi(b)$, obtained after 11 iterations of successive approximations, for $\mu = 0.8$ and $\phi(b) = b$, i.e., for the offered information on expected value of $b$. Figure 7 shows the result of such merging, obtained also after 11 iterations, for $\mu = 0.5$ and $\phi(b) = \chi(b > -1)$, i.e., for the offered information on the median of $b$. The results were obtained with the script telicit, Section 7.

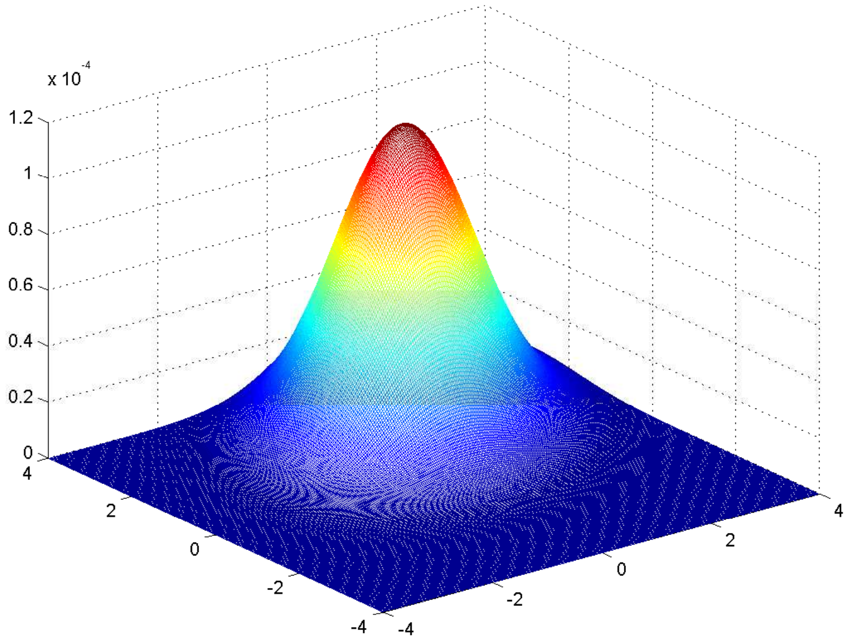


**Fig. 3.** Two-dimensional pdf offered by the 2nd participant (the script testfragment in Section 7).

The examples illustrate plausible properties of the proposed methodology. To get practically useful result, the approximation task (7) is to be solved within an appropriate class $g(b|\mathcal{V}),\ \mathcal{V} \in \mathcal{V}^*$, containing typically more regular functions than the obtained $\hat{g}(b)$.

## 6. Concluding remarks

In this paper, merging of participants' information and aims is treated as Bayesian estimation problem. This idea was already proposed and tested but the presented solution is more complete and leads to feasible algorithms covering infinite-dimensional cases. The case with different observation spaces was not treated in this way before.

The important methodological shift consists in a clear separation of non-linear deterministic mapping of joint distributions to the supplied conditional or marginal pdfs from handling the influence of "personal noise". The proposed modelling and processing order: combine noisy joint distributions and then map the results on observed lower dimensional pdfs is (probably) new and computationally simpler. The application of this way to fragmental information pieces, given by noisy projections of the underlying global pdf is enabled by extending them to a full space in a unique, well-justified, way.

The paper proposes a unified methodology. It does not mean that there is no ambiguity in solving a specific problem. However, the modelling of the relations of the unknown merger (parameter) to respective processed pdfs (data) is the only optional step. Thus, modelling can only be blamed if the results are not satisfactory enough.

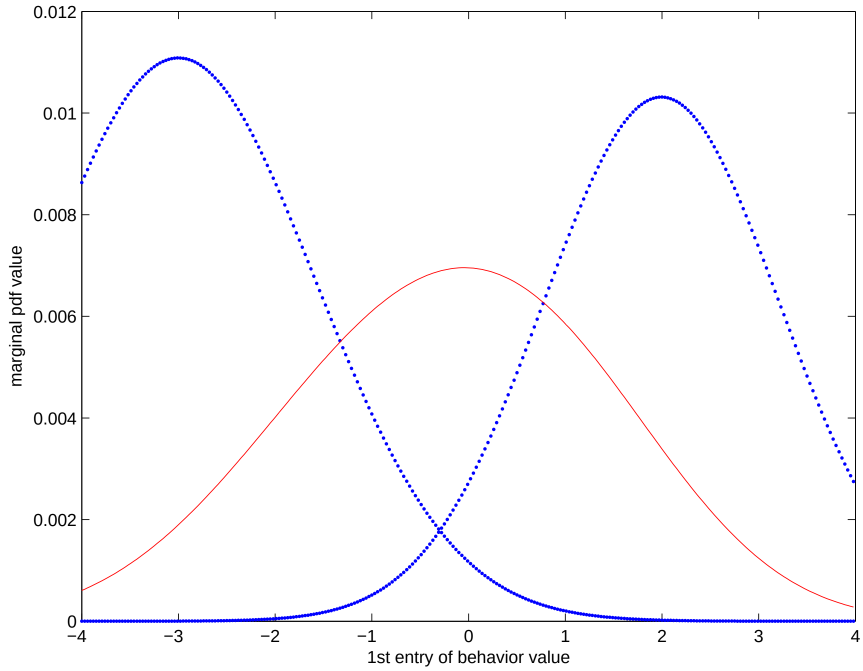


**Fig. 4.** Marginal pdfs in the 1st behavior entry characterizing merging of fragmental information (the script testfragment in Section 7). Processed pdfs blue dots; merger red line.

The algorithm based on log-normal model provides practically applicable algorithm whose results lie between usual, often too spread, arithmetic pooling and geometric one, which is often discarded due to the over-optimistic concentration of the results.

Naturally, a lot remains to be solved. The following problems should be surely addressed.

- Uniqueness of the participant-specific projections of the merger (22) is conjectured but not proved.
- General applicability of successive approximations for solving (22) is not verified.
- The important case with offered sample pdfs is not covered due to the required non-singularity of processed pdfs (the only restriction on compatibility of the processed information pieces). Ideally, the merging should reduce to the Bayes rule when the sample pdf is processed.
- Conversion of offered parametric distributions and generalized moments to non-parametric pdfs has to be solved systematically. It will represent a sort of communication protocol studied and used in connection with closely related multi-agent systems (MAS, Vlassis (2003)).
- Algorithms projecting non-parametric estimates to a priori given families specified by finite dimensional statistic need to be developed.

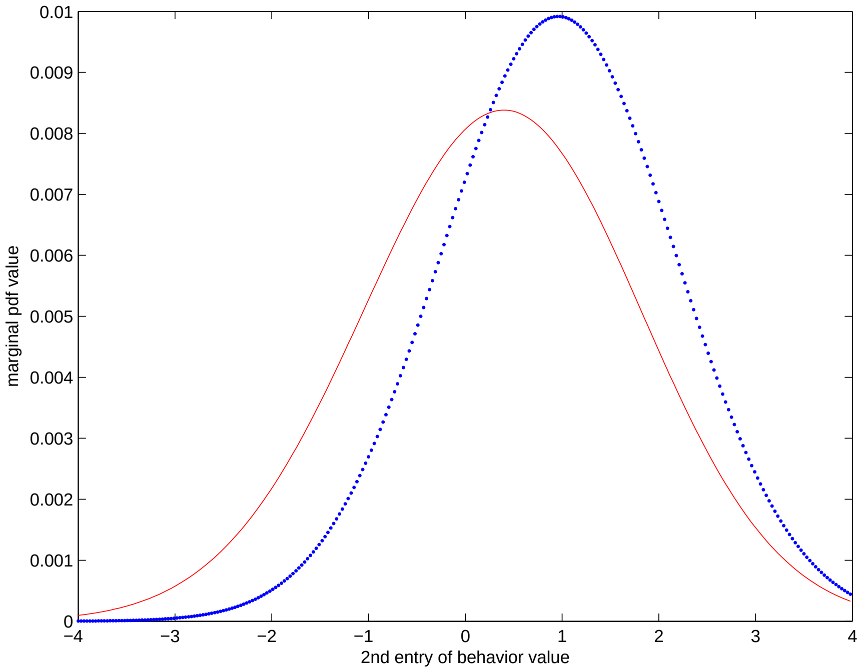


**Fig. 5.** Marginal pdfs in the 2nd behavior entry characterizing merging of fragmental information (the script testfragment in Section 7). Processed pdf blue dots; merger red line.

- Actions controlling communications have to be explicitly considered and their optimization design. These extensions have to respect the considered flat structure and will create counterparts of negotiation, bargaining and conflict resolution strategies studied in MAS.
- Study and exploitation of the extensive overlap with a range of existing techniques originating in knowledge elicitation, probabilistic expert systems, MAS, etc. should be performed.

Even this demanding list is by no means complete but the importance of the aim and the results obtained make their solution worth trying.

*Acknowledgements*
This research was supported by a common project IMATI CNR Milano and ÚTIA AV ČR, GAČR 102/08/0567 as well as by MŠMT ČR 2C06001.

## References

Andrýsek, J., A. Bodini, M. Kárný, J. Kracík, and F. Ruggeri (2007). On combining partial and incompatible information in e-negotiation and e-arbitration. *Group Decision and Negotiation 10 (2007)*, 1–10.

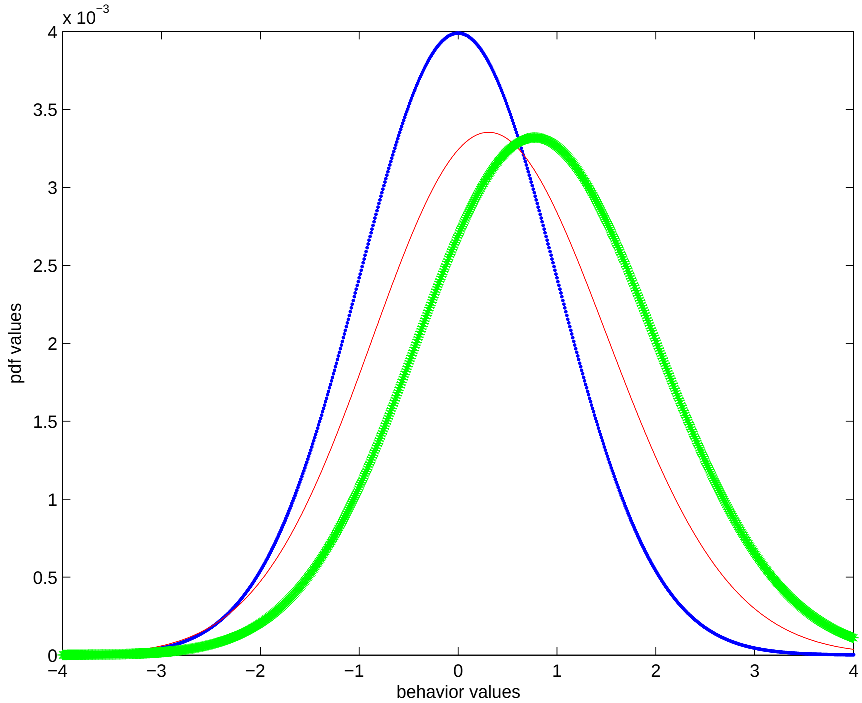


**Fig. 6.** Original pdf (blue dots) is modified to the merger (red line) by information about mean value, which was extended to the pdf ${}^{e}g_2(b)$ marked by green stars; the script telicit, case cas = 1


Berger, J. (1985). *Statistical Decision Theory and Bayesian Analysis*. New York: Springer-Verlag.

Bernardo, J. M. (1979). Expected information as expected utility. *The Annals of Statistics 7*(3), 686–690.

Cowell, R., A. Dawid, S.L.Lauritzen, and D. Spiegelhalter (2003). *Probabilistic networks and expert systems*. Springer. 2nd edition.

DeGroot, M. (1970). *Optimal Statistical Decisions*. New York: McGraw-Hill.

Greene, W. (2003). *Econometric Analysis*. Prentice Hall. 5th edition.

Harsanyi, J. (2004). Games with incomplete information played by Bayesian players, I–III. *Management Science 50*(12). Supplement.

J.Oakley and A. O'Hagan (2005). Uncertainty in prior elicitations: a non-parametric approach. Technical report, Department of Probability and Statistics, University of Sheffield, Sheffield. Revised version of research report No. 521/02.

Kárný, M. (2008, July). Bayesian paradigm and fully probabilistic design. In *Preprints of the 17th World Congress of the International Federation of Automatic Control*, Soul. IFAC.

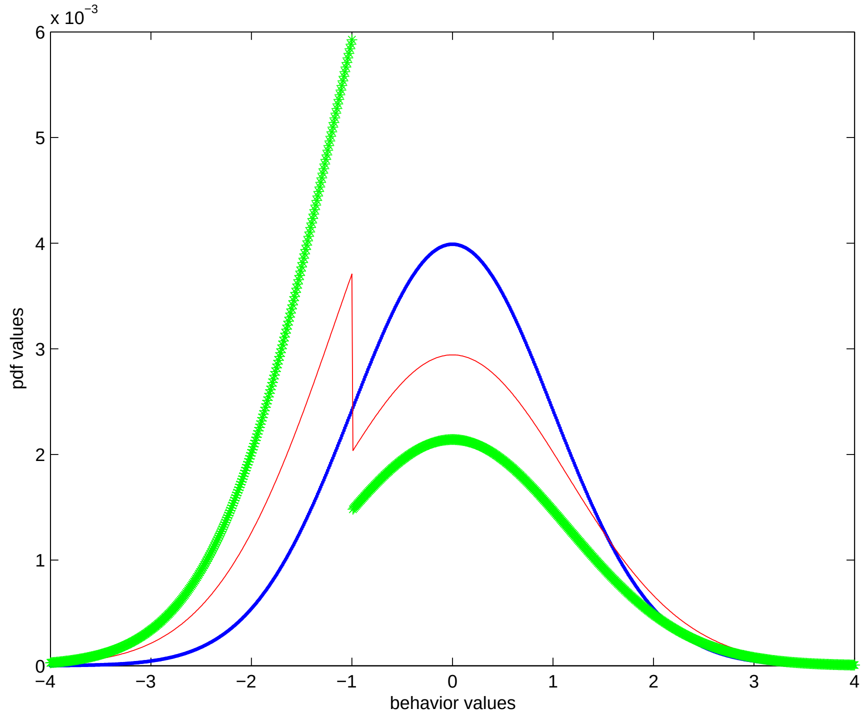


**Fig. 7.** Original pdf (blue dots) is modified to the merger (red line) by information about median, which was extended to the pdf ${}^{e}g_2(b)$ marked by green stars; the script telicit, case cas = 2

Kárný, M. and T. V. Guy (2006). Fully probabilistic control design. *Systems & Control Letters 55*(4), 259–265.

Kárný, M., J. Kracík, and T. Guy (2007). Cooperative decision making without facilitator. In F. A. Andrievsky B.R. (Ed.), *IFAC Workshop "Adaptation and Learning in Control and Signal Processing" /9./*. IFAC.

Kerridge, D. (1961). Inaccuracy and inference. *Journal of Royal Statistical Society B 23*, 284–294.

O'Hagan, A., C.E.Buck, A. Daneshkhah, J. Eiser, P. Garthwaite, D.J.Jenkinson, J. Oakley, and T. Rakow (2006). *Uncertin judgement: eliciting experts' probabilities*. John Wiley & Sons.

Stirling, W. (2004). Satisficing games for multiple-participant coordinated decision-making. In J. Andrýsek, J. Kracík, and M. Kárný (Eds.), *Mutliple Participant Decision-making*, pp. 3–16. Advanced Knowledge International, Maggil Adelaide.

Vajda, I. (1989). *Theory of statistical inference and information*. Dordrecht: Kluwer Academic Publishers.

Vlassis, N. (2003). A concise introduction to multi-agent systems and distributed AI. nformatics Institute, University of Amsterdam, 2003, http://www.sciece.uva.nl/ vlassis/cimasdai.

Wolper, D. and K. Tumer (1999). Collective intelligence for optimization. In J. Boyan (Ed.), *Sixteenth International Joint Conference on Artificial Intelligence, Workshop ML-1, Statistical Machine Learning for Large-Scale Optimization*, pp. 121–128. Stockholm.

Wolper, D. and K. Tumer (2001). Optimal payoff functions for members of collectives. *Advances in Complex Systems 4*(2/3), 265–279.

Yiming, Y., Y. Xun, and J. Lee (2003). Performance and attention in multiagent object search team. *IEEE Transactions on Systems Man and Cybernetics, Part A – Systems and Humans 33*(2), 257–263.

## 7. Matlab scripts

Here, Matlab scripts are attached for anybody to play with the simple examples presented in Sections 3.2, 4. The function geo implements the key merging formula (20).

```
%  testsingle: test of the merging of continuous pdfs of a single variable
%
% designed = MK
% updated  = 8-08-08
% project  = Puzzle, UTIA-IMATI
%
% call: geo

 close all
 clear all

% Generating of the processed pdfs
%
 s = [-4:0.01:4];              % behavior's grid on which pdfs are merged
 g = [exp(-0.5/1*(s+1.5).^(2));% the merged pair of un-normalized pdfs
      exp(-0.5/5*(s-2).^(2))];
 [po,so]=size(g);              % the number of participants, grid cardinality
 g = eye(po)/diag(sum(g,2))*g; % normalization of the sampled merged pdfs
 % merging
 [ghat,ge,ar,hatrho]=geo(g);   % merger, geometric and arithmetic mean,
                               % least-squares remainder
% graphical outputs
 figure(1)
 plot(s,g,'b.',s,ghat,'r-',s,ge,'g*',s,ar,'ko')
 xlabel('behavior value')
 ylabel('pdf value')
 print   -depsc 't3-fig1'
```

```
=========================================================================
function [hatg,ge,ar,hatrho]=geo(g)
% merging several pdfs
% [hatg,ge,ar,hatrho]=geo(g)
% hatg   = merged pdf
% ge     = geometric mean;
% ar     = arithmetic mean
% g      = array of pdfs evaluated on the same argument grid:
%          pth row corresponds with pth participant
% hatrho = normalized least-squares reminder
%
% designed = MK
% updated  = 11-8-08
% project  = Puzzle, UTIA-IMATI
%

% problem dimensions
 [po,no]=size(g);           % [the number of participants,grid cardinality]
 g      = eye(po)/diag(sum(g,2))*g;% normalization of  participants' pdfs
 ar     = zeros(1,no);            % arithmetic mean
 ge     = zeros(1,no);            % geometric mean
 sum(g,1)/po;
 ge     = prod(g).^(1/po);        % geometric mean
 ge     = ge/sum(ge);             % geometric mean
 hatrho = zeros(1,no);            % normalized least squares remainder
 for p=1:po                       % point-wise summing over participants
     z  = g(p,:);
     y  = log(z);
     ar = ar + z;
     ge = ge + y;
     hatrho =hatrho+y.^2;
 end
 ar     = ar/po;                        % arithmetic mean
 ge     = exp(ge-max(ge));              % geometric mean
 ge     = ge/sum(ge);                   % normalization
 hatrho = max(hatrho/po - ge.^(2),1e-20);% normalized least squares remainder
                                        % with numerical pre-cation
 hatg  = log(ge)+po/(po-1)*2*hatrho./(1+(1+12*hatrho/(po-1)^2).^(0.5));
                                        % scaling of the result
 hatg = exp(hatg-max(hatg));            % normalization
 hatg = hatg/sum(hatg);

=========================================================================
% testfragment: test of the behavior of merging of continuous pdfs
% on different observation spaces
%
% designed = MK
% updated  = 14-08-08
```

```
% project  = Puzzle, UTIA-IMATI
%
% call: geo, norm

 close all
 clear all

 % Generating of the processed pdfs
 %
 s1 = [-4:0.03:4];   % grid of the first variable
 s2 = s1;            % grid of the second variable
 po = 2;             % the number of participants
 so = length(s1);    % the number of discrete points

 mu1 = -3;           % mean of the first participant data
 r1  = 2;            % variance of the first participant data
 g1  = exp(-0.5/r1*(s1-mu1).^(2));
                     % probabilities provided by the 1st participant
 g1  = g1/sum(g1);   % normalization
 mu2 = [2,1];        % mean of the second participant data
 r2  = 1.5*[1 0.3;0.3 1];% covariance of the second participant data
 g2 = zeros(so,so);
 om = r2^(-1);
  for i=1:so;
      for j=1:so
          g2(i,j)= exp(-0.5*([s1(i),s2(j)]-mu2)*om*([s1(i),s2(j)]-mu2)');
      end
  end
 g2=g2/sum(sum(g2));   % normalization
 % tuning knobs of the merging algorithm
 no   = 1000;         % upper bound on the number of iterations
 eps1 = 1e-5;         % threshold for stopping of iterations
 % auxiliary arrays
 f1e  = zeros(size(g2));  % extension of the first marginal
 nor  = 0;                % norm for stopping
 so2  =so*so;
 ghato= ones(1,so2);      % array for the old estimate of the merger
 ghat = ones(1,so2)/so2;  % row ordered merger
 % successive approximations for solving equation determining the merger
 fg   =zeros(size(g2));   % auxiliary array
 for n=1:no
     f=[];
     % extension of the fragmental information
     for i=1:so
         il=(i-1)*so+1;iu=i*so;
         fg(i,:) = ghat(il:iu);
         g1e(i,:) = g1(i)*fg(i,:)/sum(fg(i,:)); % extension of the 1st pdf
         f=[f,[g1e(i,:);g2(i,:)]];% row ordering of data
```

```
    end
    [ghat,ge,ar,rhato]=geo(f);  % merging of the extended processed pdfs
    nor = norm(ghat-ghato);
    if n ==1, initial_error = nor,end
    if nor< eps1
        n,break
    else
        ghato=ghat;
    end
end
% characterization of the iterative process run
terminal_norm = nor, iteration_termined = n
% graphical outputs
for i=1:so                        % back to matrix form
    il=(i-1)*so+1;iu=i*so;
    fg(i,:) = ghat(il:iu);
end
  g11  = g1;           % first marginal of the first participant
  g21  = sum(g2,2);    % first marginal of the second participant
  g22  = sum(g2,1);    % second marginal of the second participant
  hg1  = sum(fg,2);    % estimate of the first marginal
  hg2  = sum(fg,1);    % estimate of the second marginal
% title('estimate of the global pdf')
figure(1)
mesh(s1,s2,fg)
print   -depsc 't4-fig1'
 % title('joint pdf offered by the second participant')
figure(2)
mesh(s1,s2,g2)
print   -depsc 't4-fig2'
% 1st marginal
figure(3)
plot(s1,g11,'b.',s1,g21,'b.',s1,hg1','r-')
xlabel('1st entry of behavior value')
ylabel('marginal pdf value')
% 'processed data blue dots, merger red line'
print   -depsc 't4-fig3'
% 2nd marginal
figure(4)
plot(s1,hg2','r-',s1,g22','b.')
xlabel('2nd entry of behavior value')
ylabel('marginal pdf value')
print   -depsc 't4-fig4'
% 'processed data blue dots, merger red -'

============================================================

% telicit: test of the merging applied to elicitation tasks
```

```
%
% designed = MK
% updated  = 8-08-08
% project  = Puzzle, UTIA-IMATI
%
% call: geo, fzero, norm, extend

 close all
 clear all

 % Generating of the processed pdfs
 %
 s1  = [-4:0.01:4];     % grid of the variable
 po  = 2;               % the number of participants
 so  = length(s1);      % the number of discrete points
 g1  = exp(-0.5*s1.^2); % distribution of the 1st participant
 g1  = g1/sum(g1);      % normalization 1st distribution

 % mean value processed
 cas  = 1;            % mean value offered by the 2nd participant
 par  = 0.8;          % the value offered
 name ='mean';        % label  for figures
 % quantile processed
  cas = 2;         % quantile value offered by the 2nd participant
  par = [-1,0.5]   % [quantile,its probability]
  name='quant';    % label  for figures

 % tuning knobs of the merging algorithm
 no   = 1000;          % upper bound on the number of iterations
 eps1 = 1e-5;          % threshold for stopping of iterations
 ghat = g1.^(0.001);   % initial guess on the global distribution
 ghat = ghat/sum(ghat);

 % auxiliary arrays
 g2e  = zeros(size(ghat));% extension of the first marginal
 nor  = 0;                % norm for stopping
 ghato= ones(1,so);       % array for the old estimate of the merger
 % successive approximations for equation determining the merger
 zeto = -0.3;              % old guess of Lagrangian
 for n=1:no
     zet = fzero(@(zet) extend(zet,ghat,s1,cas,par),zeto);
     [dev,g2e] = extend(zet,ghat,s1,cas,par); % the final extension
     g=[g1;                 % merged data
        g2e];
     [ghat,ge,ar,hatrho]=geo(g);  % merging of the extended pdfs
      nor = norm(ghat-ghato);
     if n ==1, initial_error = nor,end
     if nor< eps1
```

```
        n,break
    else
        zeto=zet;
        ghato=ghat;
    end
end
% characterization of the iterative process run
terminal_norm = nor, iteration_termined = n
% graphical outputs
% title('data blue*, merger red-, extended pdf green.')
figure(1)
plot(s1,g1,'b.',s1,ghat,'r-',s1,g2e,'g*')
xlabel('behavior values')
ylabel('pdf values')
print('-depsc',name)

==============================================================

function  [dev,fe] = extend(zet,fhat,s1,cas,par)
% for a Lagrangian-factor guess extends pdf fhat
% to pdf fe on grid s1 and evaluate constraints deviation dev
% supported cases distinguished by cas
%
% [dev,fe] = extend(zet,fhat,s1,cas,par)
%
% dev  = deviation in the constrains
% fe   = guess of the extended distribution corresponding to the zet
% zet  = guess of the Lagrangian factor
% fhat = guess of the merged pdf
% cas  = [1,2]  given [mean, quantile]
% par  = mean value                                    if cas = 1
%        [quantile value,probability of the quantile]     cas = 2
%
% designed = MK
% updated  = 2-08-08
% project  = Puzzle, UTIA-IMATI

if cas == 1                      % mean value
   fe = log(fhat)-zet*s1;
   fe = exp(fe-max(fe));
   fe = fe/sum(fe);             % extended pdf
   dev=sum(s1.*fe)-par;         % deviation in mean
else
   if cas == 2
      val = s1> par(1);
      fe  = log(fhat)- zet*val;
      fe = exp(fe-max(fe));
      fe = fe/sum(fe);             % extended pdf
```

```
      dev=sum(val.*fe)-par(2);     % deviation in probability
   else
     error('case unsupported')
   end
end
```