On Dynamic Decision-Making Scenarios with Multiple Participants

Miroslav Kárný, Tatiana V. Guy

ÚTIA AV ČR, POBox 18, 18208 Prague, Czech Republic
tel. +420-266062274, fax. +420-266052068
http://utia.cas.cz/AS, E-mail: school@utia.cas.cz

Abstract. Multiple-participant (MP) dynamic decision making (DDM) is encountered both in societal and technical systems. Unlike in single-participant (SP) DDM, no commonly accepted normative theory exists. The paper contributes to numerous attempts to overcome this state by classifying possible MP scenarios and connecting them with the normative SP DDM. This classification helps us to formulate open problems whose solutions will lead to feasible MP DDM.

Keywords: Bayesian dynamic decision making (DM), multiple participant DM, fully probabilistic design

1 Introduction

In SP scenario, Bayesian paradigm represents a normative theory of DDM that has – probably conceptually – no internally consistent competitors as it guarantees that its strategies are not dominated (DeGroot 1970, Berger 1985). This implies that a normative theory of MP DDM should be searched as an extension of the Bayesian DDM.

Participant's environment $\mathcal{E}$ is a limited sphere of its interests that the participant can influence. The environment is characterized by its possible behaviors defined on the considered data and action spaces as well as on the space of considered but directly unobservable quantities.

Fully probabilistic version of DDM (Kárný 1996, Kárný, et al. 2003), we focus on, models *environment* $\mathcal{E}$ of the *participant* $\mathcal{P}$ as well as its aims in probabilistic terms. Essentially, in SP DDM scenario, $\mathcal{P}$'s *strategy* is chosen so that its interconnection with *environment model* is as close as possible to the *ideal distribution* of environment behavior reflecting these aims. The paper inspects an extension of this methodology to MP DDM. The corresponding normative MP DDM requires specification of admissible strategies of participants, models of their environments $\mathcal{E}$ and relevant ideal distributions.

The paper starts with preliminaries containing a generalized version of the fully probabilistic design of decision strategies and a classification of its ingredients and cooperation stages. Then, a relevant part of Bayesian estimation is recalled. Conjugate prior distributions and the choice of the ideal *probability density function, (pdf)*, forming set-point in the fully probabilistic design, are also briefly discussed. Then, the following MP DDM scenarios are presented: selfish participants, participants cooperating at design stage with and without a priori given hierarchy. Concluding remarks point to the important open problems.

2 Preliminaries

The section defines basic notions adopted; recalls the fully probabilistic SP DDM, introduces basic ingredients of DM and classifies possible cooperation stages.

2.1 Notation

The following agreements are used throughout the paper: $\equiv$ is equality by definition; x^* denotes a set of x-values; $\mathring{x}$ means cardinality of a finite set x^*; $f(\cdot|\cdot)$ is reserved for pdfs; E denotes expectation; t is discrete-time; $d = (\Delta, a)$ is data record consisting of (observable) innovation Δ and action a; d_t is data item at time t; $d(t)$ denotes sequence $(d_1, \ldots, d_t)$; $\propto$ is equality up to a normalizing factor; left-upper superscript distinguishes object categories, e.g. ${}^{\mathcal{P}}f$ means pdf related to the participant $\mathcal{P}$.

The pdfs are distinguished by the identifiers in their arguments. No formal distinction is made between random variable, its realization and an argument of a pdf.

The SP DDM exploits *Kullback-Leibler (KL) divergence* $\mathcal{D}\left(f||\tilde{f}\right)$ that measures of the proximity of a pair of pdfs $f, \tilde{f}$ acting on a common set x^*. It is defined

$$\mathcal{D}\left(f||\tilde{f}\right) \equiv \int f(x) \ln\left(\frac{f(x)}{\tilde{f}(x)}\right) dx. \tag{1}$$

The basic property of the KL divergence (Kullback & Leibler 1951)

$$\mathcal{D}(f||\tilde{f}) \geq 0,\ \mathcal{D}(f||\tilde{f}) = 0 \text{ iff } f = \tilde{f} \text{ almost everywhere}, \tag{2}$$

is used.

2.2 Fully probabilistic SP DDM

To make a decision, a participant $\mathcal{P}$ operates with *considered behaviors* $\mathcal{Q}$ of its environment $\mathcal{E}$. A behavior consists of time sequences of potential participant's actions $a(\mathring{t})$ and of observed innovations $\Delta(\mathring{t})$, informing about environment's state, as well as of considered but directly unobserved $x(\mathring{t})$ quantities. Thus,

$$\mathcal{Q} \equiv (a(\mathring{t}), \Delta(\mathring{t}), x(\mathring{t})) \equiv (d(\mathring{t}), x(\mathring{t})),$$

where $\mathring{t} < \infty$ is decision-making horizon and $d(\mathring{t})$ denotes *observable* data. The joint pdf

$$^{\mathcal{P}}f(\mathcal{Q}) = {}^{\mathcal{P}}f(X|d(\mathring{t})) \underbrace{\prod_{t \in t^*} {}^{\mathcal{P}}f(\Delta_t|a_t, d(t-1))}_{\mathcal{M} \equiv \text{ model of environment}} \times \underbrace{\prod_{t \in t^*} {}^{\mathcal{P}}f(a_t|d(t-1))}_{\mathcal{R} \equiv \text{ decision strategy}} \tag{3}$$

describes possible behaviors of the closed loop formed by interconnection of $\mathcal{P}$ with its $\mathcal{E}$. It is product of the Bayesian estimate $^{\mathcal{P}}f(X|d(\mathring{t}))$ of the unobservable quantity $X \equiv x(\mathring{t})$, the model $\mathcal{M}$ of participant's environment $\mathcal{E}$ (more precisely *outer* model) and the decision strategy $\mathcal{R}$, selected by the participant. The Bayesian estimate of $X \equiv x(\mathring{t})$ and the model of $\mathcal{E}$ are gained via Bayesian filtering (Jazwinski 1970, Peterka 1981) and are fixed when strategy is designed. The decision strategy $\mathcal{R}$ is to be chosen so that aims of $\mathcal{P}$ are met. In the fully probabilistic setting, aims are expressed by the ideal counterpart of the joint pdf (3)

$$^{I}f(\mathcal{Q}) = {}^{I}f(X|d(\mathring{t})) \prod_{t \in t^*} {}^{I}f(\Delta_t|a_t, d(t-1))\, {}^{I}f(a_t|d(t-1)). \tag{4}$$

Ideally, the optimized strategy $\mathcal{R}$ should move the influenced behavior $\mathcal{Q}$ to a desirable area. The ideal pdf $^{I}f(\mathcal{Q})$ expresses this wish. The fully probabilistic design simply searches for the strategy $^{o\mathcal{P}}\mathcal{R}$ that minimizes the KL divergence of the joint pdf $^{\mathcal{P}}f(\mathcal{Q})$ to its ideal counterpart $^{I}f(\mathcal{Q})$.

Proposition 1 (Fully probabilistic design) *The optimal decision strategy $^{o\mathcal{P}}\mathcal{R}$ minimizing $\mathcal{D}\left({}^{\mathcal{P}}f||\,{}^{I}f\right)$ has the form*

$$^{o\mathcal{P}}f(a_t|d(t-1)) \propto {}^{I}f(a_t|d(t-1)) \exp[-\omega(a_t, d(t-1))]. \tag{5}$$

The functions $\omega(a_t, d(t-1))$ are generated recursively in backward manner for $t = \mathring{t}, \mathring{t}-1, \ldots, 1$

$$\begin{aligned} \omega(a_t, d(t-1)) &= \int {}^{\mathcal{P}}f(\Delta_t|a_t, d(t-1)) \ln\left[\frac{^{\mathcal{P}}f(\Delta_t|a_t, d(t-1))}{\gamma(d(t))\, {}^{I}f(\Delta_t|a_t, d(t-1))}\right] d\Delta_t \\ \gamma(d(t-1)) &= \int {}^{I}f(a_t|d(t-1)) \exp[\omega(a_t, d(t-1))]\, da_t. \end{aligned} \tag{6}$$

The recursion starts with

$$\ln(\gamma(d(\mathring{t}))) = -\int {}^{\mathcal{P}}f(X|d(\mathring{t})) \ln\left[\frac{^{\mathcal{P}}f(X|d(\mathring{t}))}{^{I}f(X|d(\mathring{t}))}\right] dX. \tag{7}$$

Proof: Definition of the KL divergence and Fubini theorem on multiple integration imply the following identity

$$\mathcal{D}\left({}^{\mathcal{P}}f||\,{}^{I}f\right) = \int {}^{\mathcal{P}}f(d(\mathring{t})) \left\{ \ln\left(\frac{^{\mathcal{P}}f(d(\mathring{t}))}{^{I}f(d(\mathring{t}))}\right) + \underbrace{\int {}^{\mathcal{P}}f(X|d(\mathring{t})) \ln\left(\frac{^{\mathcal{P}}f(X|d(\mathring{t}))}{^{I}f(X|d(\mathring{t}))}\right) dX}_{\equiv -\ln[\gamma(d(\mathring{t}))]} \right\} dd(\mathring{t}).$$

Using the chain rule for pdfs (Peterka 1981), we can write the part of the KL divergence that depends on ${}^{\mathcal{P}}f(a_{\mathring{t}}|d(\mathring{t}-1))$ in the form

$$E\left\{\int {}^{\mathcal{P}}f(a_{\mathring{t}}|d(\mathring{t}-1))\left[\ln\left(\frac{{}^{\mathcal{P}}f(a_{\mathring{t}}|d(\mathring{t}-1))}{{}^{I}f(a_{\mathring{t}}|d(\mathring{t}-1))}\right)+\omega(a_{\mathring{t}},d(\mathring{t}-1))\right]\,da_{\mathring{t}}\right\}$$

$$\text{with }\ \omega(a_{\mathring{t}},d(\mathring{t}-1))\equiv\int {}^{\mathcal{P}}f(\Delta_{\mathring{t}}|a_{\mathring{t}},d(\mathring{t}-1))\ln\left(\frac{{}^{\mathcal{P}}f(\Delta_{\mathring{t}}|a_{\mathring{t}},d(\mathring{t}-1))}{\gamma(d(\mathring{t}))\,{}^{I}f(\Delta_{\mathring{t}}|a_{\mathring{t}},d(\mathring{t}-1))}\right)\,d\Delta_{\mathring{t}}.$$

The above expected value can be re-written as follows

$$E\left\{\int {}^{\mathcal{P}}f(a_{\mathring{t}}|d(\mathring{t}-1))\ln\left(\frac{{}^{\mathcal{P}}f(a_{\mathring{t}}|d(\mathring{t}-1))}{{}^{o\mathcal{P}}f(a_{\mathring{t}}|d(\mathring{t}-1))}\right)\,da_{\mathring{t}}-\ln(\gamma(d(\mathring{t}-1)))\right\},\quad\text{where}$$

$$\begin{aligned}{}^{o\mathcal{P}}f(a_{\mathring{t}}|d(\mathring{t}-1)) &= \frac{{}^{I}f(a_{\mathring{t}}|d(\mathring{t}-1))\exp[-\omega(a_{\mathring{t}},d(\mathring{t}-1))]}{\gamma(d(\mathring{t}-1))}\\ \gamma(d(\mathring{t}-1)) &\equiv \int {}^{I}f(a_{\mathring{t}}|d(\mathring{t}-1))\exp[-\omega(a_{\mathring{t}},d(\mathring{t}-1))]\,da_{\mathring{t}}.\end{aligned}$$

The last optimized pdf ${}^{\mathcal{P}}f(a_{\mathring{t}}|d(\mathring{t}-1))$ enters the first term having the form of (conditional) KL divergence. Due to (2), it is minimized by the claimed pdf (5). The normalizing factor $\gamma(d(\mathring{t}-1))$ represents remainder to be influenced by pdfs ${}^{\mathcal{P}}f(a_t|d(t-1))$ with $t<\mathring{t}$. Their optimization is, however, formally identical with the demonstrated last step. □

Remarks

1. The explicit solution depends on evaluation of respective integrals. Monte Carlo evaluation can be used in low dimensional case, whenever it is possible to store functions ω, γ.
2. The optimal strategy (5) respects constraints on actions expressed by the support of the ideal pdf ${}^{I}f(a_t|d(t-1))$ due to the inclusion of supports

$$\text{supp}\left[\ {}^{o\mathcal{P}}f(a_t|d(t-1))\right]\subset\text{supp}\left[\ {}^{I}f(a_t|d(t-1))\right].$$

3. The above version of the fully probabilistic design considers the unobservable quantity X, which is expressed in the initial condition (7). This extends the previous solution (Kárný et al. 2003). The extension reduces to it if the quantity X need not or cannot be influenced by the optimized strategy. In these cases, it sufficient to leave X to its fate and to choose ${}^{I}f(X|d(\mathring{t}))={}^{\mathcal{P}}f(X|d(\mathring{t}))$. This reduces the initial condition (7) to $\gamma(d(\mathring{t}))=1$. This start corresponds with the former design version in which $f(X|d(\mathring{t}))$ has no direct influence on the design. The estimate enters the (outer) model $\mathcal{M}$ only, cf. (Peterka 1981) and Section 2.3.2.
4. The design, similarly as general stochastic dynamic programming, relies on the assumption that admissible strategies work with non-decreasing information given by $d(t)=(d_t,d(t-1))$. This implies generally permanent increase of complexity of the resulting strategy with increasing horizon $\mathring{t}$.

2.3 Ingredients of Dynamic Decision Making

Specific instances of DDM are determined by ingredients occurring in Proposition 1. They are briefly discussed here.

2.3.1 Classification of ingredients and cooperation stages

Ingredients of the fully probabilistic design can be classified according to their origin. They can be *internal* with respect to $\mathcal{P}$ or *external*, i.e. received from $\mathcal{E}$ of this $\mathcal{P}$ or both. Variants are listed in the following table.

Ingredient	Source type	Comment
action a_t	internal	no alternative source
innovation Δ_t	external	no alternative source
inner quantity X	internal	$\mathcal{P}$ decides whether X will be included into its model of $\mathcal{E}$
model ${}^{\mathcal{P}}f(\mathcal{Q})$	both types	the model can be supplied externally in MP DDM
ideal ${}^{I}f(\mathcal{Q})$	both types	the ideal pdf can be modified externally in MP DDM; it reflects also constraints

Cooperation stages with other $\mathcal{P}$s can be classified according to the interaction of $\mathcal{P}$ in the design of the decision strategy (*design stage*) and during its application (*acting stage*).

External ingredient	Classification	Comment
behavior $\mathcal{Q}$	cooperation at the acting stage	data shared by $\mathcal{P}$s are specified
factors of ${}^{\mathcal{P}}f(\mathcal{Q})$	cooperation at the design stage	strategies of other $\mathcal{P}$s and/or their models help in building of the model of $\mathcal{E}$ for the considered $\mathcal{P}$
ideal pdf ${}^{I}f(\mathcal{Q})$	cooperation at the design and/or acting stages	knowledge of aims and restrictions of other $\mathcal{P}$s help in modifying $\mathcal{P}$ ideal pdf

Generally, design and acting stages can intertwine and the cooperation extent may vary. Numerous variants may arise in this way up to the powerful adaptive cooperation. These possibilities are too complex for this introductory inspection of MP DDM and their discussion is left aside. Thus, in the rest of the paper, *we take the definition of the external supply of data and models to the inspected $\mathcal{P}$ as given and fixed*. It leaves a few basic ways of cooperation in the design stage.

In the acting stage, the cooperation possible via (partial) sharing of data and (or) the ideal pdf. In the design stage, models, strategies and ideal pdfs can be (partially) shared.

Sharing of ingredients can be either compulsory or facultative. The former option introduces hierarchy into MP DDM. The latter one uses just hierarchy implied by differing "abilities" of individual $\mathcal{P}$s.

Before elaborating the corresponding scenarios, let us discuss freedom in *feasible constructions of ingredients*.

2.3.2 Bayesian Estimate and Model of Environment

For presentation simplicity, an unknown parameter is assumed to be the only unobserved quantity X. Then, the model ${}^{\mathcal{P}}f(\Delta_t|a_t, d(t-1))$ of $\mathcal{E}$ and the Bayesian parameter estimate ${}^{\mathcal{P}}f(X|d(t))$ are described by formulas (Peterka 1981)

$$ {}^{\mathcal{P}}f(\Delta_t|a_t, d(t-1)) = \int {}^{\mathcal{P}}f(\Delta_t|a_t, d(t-1), X)\, {}^{\mathcal{P}}f(X|d(t-1))\, dX \tag{8} $$
$$ {}^{\mathcal{P}}f(X|d(t)) \propto {}^{\mathcal{P}}f(\Delta_t|a_t, d(t-1), X)\, {}^{\mathcal{P}}f(X|d(t-1)). $$

$\mathcal{P}$ has to supply *observation model* ${}^{\mathcal{P}}f(\Delta_t|a_t, d(t-1), X)$ relating the current innovation to the current action a_t, observed data $d(t-1)$ and the unknown parameter X. Moreover, it has to provide prior pdf ${}^{\mathcal{P}}f(X)$ that expresses its prior knowledge about the unknown X. The formulae (8) are valid when X cannot be used for selecting a_t, i.e. when natural conditions of control (Peterka 1981) are met.

For DDM, pdfs as *functions* have to be available. Their dependence on data, whose amount is growing with time, can be feasibly represented only through a finite dimensional statistic. For existence of such a statistics, the observation model (8) has to depend at most on the *finite-dimensional regression vector* $\psi_t = [a_t, \ldots, a_{t-\partial_a}, \Delta_{t-1}, \ldots, \Delta_{t-\partial_\Delta}]$, ∂_a, $\partial_\Delta < \infty$, via a time-invariant function $M(\Psi, X)$ of the *data vector* $\Psi \equiv [\Delta, \psi]$ and the parameter X. Thus, feasible observation models have to have the form

$$ {}^{\mathcal{P}}f(\Delta_t|a_t, d(t-1), X) = M_t([\Delta_t, \psi_t], X) \equiv M(\Psi_t, X) $$

with computationally feasible evaluation of M-values.

Existence of the finite-dimensional *sufficient* statistic V_t, that guarantees feasibility without information loss, is (practically) guaranteed only if $M(\Psi, X)$ belongs to the exponential family (Barndorff-Nielsen 1978)

$$ M(\Psi, X) = \exp \langle B(\Psi), C(X) \rangle \tag{9} $$

with $\langle \cdot, \cdot \rangle$ being a scalar product of compatible, finite-dimensional, practically manageable, functions $B(\cdot)$, $C(\cdot)$ of respective arguments.

A sort of approximation is needed whenever this family is not rich enough. Modelling by finite mixtures with components in the exponential family and development of specialized estimation procedures (Kárný et al. 2003, Titterington, et al. 1985) is one of a few generic approaches to this problem.

2.3.3 Conjugate Prior Pdf

Form of the prior pdf ${}^{\mathcal{P}}f(X)$ is determined by the need to describe Bayesian estimate ${}^{\mathcal{P}}f(X|d(t))$ and the model ${}^{\mathcal{P}}f(\Delta_t|a_t, d(t-1))$ of $\mathcal{E}$ by the finite-dimensional statistic V_t. For the exponential family (9), the prior pdf has to be conjugate prior pdf

$$ {}^{\mathcal{P}}f(X) \equiv {}^{\mathcal{P}}f(X|d(0)) \propto \exp\langle V_0, C(X)\rangle\, \Xi_{X^*}(X), \tag{10} $$

where $\Xi_{X^*}(X)$ is indicator of the set X^*. The optional prior statistic V_0 has to guarantee that right-hand side of (10) can be normalized to pdf. The posterior pdf ${}^{\mathcal{P}}f(X|d(t))$ preserves this form and the functional recursion (8) reduces to the simple algebraic updating

$$ V_t = V_{t-1} + B(\Psi_t),\ V_0 \text{ given.} \tag{11} $$

Specific prior knowledge is to be reflected in the statistic V_0. In (Kárný, et al. 2001a), an algorithmic translation of usual technical knowledge into V_0 is presented. Essentially, Bayesian estimation is performed on simulated or though-of data.

2.3.4 Ideal Pdf

The ideal pdf ${}^{I}f(d_t|d(t-1)) = {}^{I}f(\Delta_t|a_t, d(t-1))\, {}^{I}f(a_t|d(t-1))$ (4) can be constructed through a specification of desirable ranges of respective arguments and their changes. It is done by defining locations and widths of pdfs taken from the exponential family. Restriction of the supports of these pdfs to desired ranges respects hard constraints on the solved DDM.

Often, it happens that the $\mathcal{P}$ is interested only in a few entries $d_{\mathcal{P};t}$ of $d_t \equiv (d_{\mathcal{P};t}, d_{\mathcal{E};t})$. Then, it leaves other variables, $d_{\mathcal{E};t}$, to their fate. It corresponds with the special choice of the ideal pdf

$$ {}^{I}f(d_t|d(t-1)) = {}^{\mathcal{P}}f(d_{\mathcal{E};t}|d_{\mathcal{P};t}, d(t-1))\, {}^{I}f(d_{\mathcal{P};t}|d(t-1)), \tag{12} $$

where ${}^{\mathcal{P}}f(d_{\mathcal{E};t}|d_{\mathcal{P};t}, d(t-1))$ is conditional pdf derived from ${}^{\mathcal{P}}f(d_t|d(t-1))$, describing the estimated model of $\mathcal{E}$ connected with the designed decision strategy $\mathcal{R}$. It is straightforward to modify Proposition 1 to this special case.

It should be mentioned that *the use of finite mixtures as ideal pdfs allows us to express multiple aims*.

3 Decision-Making Scenarios with Multiple Participants

Participant's location (level of hierarchy) and type of interrelation with other participants (selfish, cooperative, hierarchic), determined by the communication structure, define its operation domain and type of information shared. As it has been mentioned before, the *acting stage* realizes the DM strategy determined at the design stage. From the cooperation viewpoint, the acting stages for different types of scenarios are distinguishes by sharing the data about the common part of considered behaviors only. The highest cooperation level of cooperation at this stage can be based on sharing the statistics on that part of the behavior space, which is not "reachable" by another participant.

The more rich and influential is the cooperation in the *design stage* where the cooperation relies on sharing the models and ideal pdfs, possibly creating a common group model and (or) ideal, etc.

This section describes cooperations for different scenarios, outlines possible technical solutions and emphasizes open problems to be solved.

The design stage of all types of scenarios consists of two phases: *communication phase*, when information exchange takes place and *updating phase*, when particular participant updates its models and (or) ideals. The strength of the mutual influence makes the main difference between individual variants. Otherwise, the scenarios differ only by the order of alternating and by the frequency of repetitions of these stages.

Selfish scenario supposes that a communicating participant either send to or receive from another participant information in communication phase. Then, the *facultative* updating phase for the information-receiving participant starts. The receiving participant may be selfish, it may respect the information provided partially only. In extreme it does not make any change of its model or ideal pdf. The same sequence is repeated for another participant of the communicating pair. Thus the whole design stage of selfish scenario results in exchange of models and possible model updating. No common global model or ideal pdf for the both communicating participants is expected to be chosen. It means that each participant will use its own model of $\mathcal{E}$ at the next acting stage.

Even selfish scenario requires at least minimum cooperation during communication phase: participant should provide its model or/and ideal at least partially. Providing correct models or real ideal is not however compulsory, i.e. possibility to lie consciously has been preserved for participant.

In contrast to the selfish scenario, the *cooperative scenario* assumes that the mutual exchange of models and ideals at communication phase is followed by *obligatory* updating phase. The result of the negotiations at design stage must be *common model and ideal* of communicating participants concerning to their common part of data space.

The last considered *hierarchical scenario* assumes existence of a participant coordinating whole group of participants. Exceptional role of the coordinator is by its power to influence the members of the group either via their models or ideals or even via data supplied in acting stage.

The following subsections describe technical details of different scenarios.

3.1 Selfish Scenario and Its Estimation Aspects

In a non-trivial MP DDM, the environment of an individual $\mathcal{P}$ overlaps at least partially with that of other $\mathcal{P}$s.

The simplest MP DDM scenario discussed here takes individuals as selfish (egoistic) $\mathcal{P}$s. They simply act as if they were alone in their $\mathcal{E}$s, i.e. they act according to the theory recalled in Section 2.2. When modelling their environments, each of them relies on its observation model ${}^{\mathcal{P}}f(\Delta_t|, a_t, d(t{-}1), X)$ and prior pdf ${}^{\mathcal{P}}f(X)$. Presence of other acting participants makes the modelling and subsequent design are sensitive to modelling errors. Thus, inspection of achievable quality of estimation and consequently of resulting strategy becomes dominating aspect of this MP DDM scheme. The following proposition indicates how this problem can be addressed. It represents a version of large-deviations theorems (Sanov 1957, Kulhavý 1994).

Proposition 2 (Asymptotic properties of Bayesian estimation) *Let the observed data be generated by so called "objective" model (Berec & Kárný 1997) ${}^{o}f(\Delta_t|a_t, d(t-1))$, which is the best description of reality for the inspected DDM. Let also the used strategy meet natural conditions of control and bounds $\bar{B}(X)$, $\underline{B}(X)$ exist such that*

$$\infty > \bar{B}(X) \geq \frac{{}^{o}f(\Delta_t|a_t, d(t-1))}{{}^{\mathcal{P}}f(\Delta_t|\psi_t, X)} \geq \underline{B}(X) > 0 \tag{13}$$

for almost all $X \in \operatorname{supp}\left[\,{}^{\mathcal{P}}f(X)\right]$ and $d(\mathring{t}) \in d^*(\mathring{t})$.

Then, for $\mathring{t} \to \infty$,

$$\operatorname{supp}\left[\,{}^{\mathcal{P}}f(X|d(\mathring{t}))\right] \equiv \operatorname{Arg} \inf_{X \in \operatorname{supp}[\,{}^{\mathcal{P}}f(X)]} \frac{1}{\mathring{t}} \sum_{t \in t^*} \ln\left(\frac{{}^{o}f(\Delta_t|a_t, d(t-1))}{M(\Psi_t, X)}\right) \text{ almost surely.} \tag{14}$$

Proof: Using the Bayes rule and the definition of proportionality sign $\propto$, the posterior pdf can be given the form

$${}^{\mathcal{P}}f(X|d(\mathring{t})) \propto {}^{\mathcal{P}}f(X) \exp\left\{-\mathring{t}\left[\underbrace{\frac{1}{\mathring{t}} \sum_{t \in t^*} \ln\left(\frac{{}^{o}f(\Delta_t|a_t, d(t-1))}{M(\Psi_t, X)}\right)}_{\equiv \mathcal{H}(d(\mathring{t}), X)} - \inf_{X \in \operatorname{supp}[\,{}^{\mathcal{P}}f(X)]} \mathcal{H}(d(\mathring{t}), X)\right]\right\}$$

Due to the factor $-\mathring{t}$, for $\mathring{t} \to \infty$,

$$\operatorname{supp}\left[f(X|d(\mathring{t}))\right] = \operatorname{Arg} \inf_{X \in \operatorname{supp}[f(X)]} \mathcal{H}(d(\mathring{t}), X)$$

if the infimum is finite. To verify the finiteness, let us introduce, for a fixed X and the considered strategy fulfilling natural conditions of control, the deviations

$$e(d(t-1), X) \equiv \ln\left(\frac{{}^{o}f(\Delta_t|a_t, d(t-1))}{M(\Psi_t, X)}\right) - E\left[\ln\left(\frac{{}^{o}f(\Delta_t|a_t, d(t-1))}{M(\Psi_t, X)}\right)|d(t-1)\right].$$

With them,

$$\mathcal{H}(d(\mathring{t}), X) = \frac{1}{\mathring{t}} \sum_{t \in t^*} E\left[\ln\left(\frac{{}^{o}f(\Delta_t|a_t, d(t-1))}{M(\Psi_t, X)}\right)|d(t-1)\right] + \frac{1}{\mathring{t}} \sum_{t \in t^*} e(d(t-1), X).$$

The first term is average of (conditional) KL divergences and as such it is non-negative. The second term is average of uncorrelated zero-mean variables that are almost surely bounded due to (13). Consequently, (Loeve 1962), the second term converges almost surely to zero. Thus, $\mathcal{H}(d(\mathring{t}), X)$ is almost surely bounded from below that guarantees finiteness of the infimum. $\square$

The fact that the posterior pdf concentrates on points in $\mathrm{supp}\left[\, {}^{\mathcal{P}}f(X)\right]$ minimizing asymptotically entropy rate $\mathcal{H}(d(\infty), X) \equiv$

$$\equiv \limsup_{t \to \infty} \frac{1}{\mathring{t}} \sum_{t \in t^*} \ln\left(\frac{{}^{o}f(\Delta_t|a_t, d(t-1))}{M(\Psi_t, X)}\right) = \limsup_{t \to \infty} \frac{1}{\mathring{t}} \sum_{t \in t^*} E\left[\ln\left(\frac{{}^{o}f(\Delta_t|a_t, d(t-1))}{M(\Psi_t, X)}\right)|d(t-1)\right]$$

is the known message of Proposition 2. For the inspected MP DDM, it is, however, important to notice that this divergence may be so large that the obtained model is useless. To see it more clearly, let us discuss influence of modelling error on DDM. The optimization has to deal with the estimated model and to minimize the KL divergence $\mathcal{D}\left({}^{\mathcal{P}}f||\,{}^{I}f\right)$.

Let us consider the ideal situation that a unique model parameterized ${}^{\mathcal{P}}X \in X^*$ is asymptotically obtained, i.e. ${}^{\mathcal{P}}f(\Delta_t|a_t, d(t-1)) =$

$$= \int {}^{\mathcal{P}}f(\Delta_t|a_t, d(t-1), X)\, {}^{\mathcal{P}}f(X|d(t-1))\, dX \rightarrow_{t \to \infty} {}^{\mathcal{P}}f(\Delta_t|a_t, d(t-1), {}^{\mathcal{P}}X) = M\left(\Psi_t, {}^{\mathcal{P}}X\right).$$

Let also assume that, with this model $M\left(\Psi_t, {}^{\mathcal{P}}X\right)$, the perfect matching of ${}^{\mathcal{P}}f(\mathcal{Q})$ with the ideal pdf ${}^{I}f(\mathcal{Q})$ is achieved, i.e. ${}^{\mathcal{P}}f(\cdot) = {}^{I}f(\cdot)$. The loss per time step, corresponding to the designed strategy ${}^{o\mathcal{P}}f(a_t|d(t-1))$ connected to the objective model ${}^{o}f(\Delta_t|a_t, d(t-1))$, becomes

$$\frac{1}{\mathring{t}}\mathcal{D}\left({}^{o}f||\,{}^{I}f\right) = \frac{1}{\mathring{t}}\mathcal{D}\left({}^{o}f||\,{}^{\mathcal{P}}f\right) = \frac{1}{\mathring{t}} E\left[\ln\left(\frac{\prod_{t \in t^*} {}^{o}f(\Delta_t|a_t, d(t-1))}{\prod_{t \in t^*} M(\Psi_t, {}^{\mathcal{P}}X)}\right)\right] \rightarrow_{\mathring{t} \to \infty} E\left[\mathcal{H}\left(d(\infty), {}^{\mathcal{P}}X\right)\right] \quad (15)$$

Thus, the loss is the higher the greater is asymptotically achieved minimum of the entropy rate. Continuity arguments extend this property to ${}^{\mathcal{P}}f$s in neighborhood of ${}^{I}f$. A finer robustness analysis is generally needed (Perez 1965, Martin, et al. 2003).

Remarks

1. The condition (13) given by bounds $\underline{B}(X)$, $\overline{B}(X)$ excludes significant differences in tails of the objective ${}^{o}f$ and ${}^{\mathcal{P}}f$ pdfs. It can be definitely modified and refined so that $\mathring{t}^{-1} \sum_{t \in t^*} e(d(t-1), X)$ still converges to zero.
2. The infimum is taken over the support of the prior pdf ${}^{\mathcal{P}}f(X)$ that is included in X^*, cf. (10). The proposition can be simply re-stated to other mutual relationships of $\mathrm{supp}\left[\, f(X)\right]$ and X^*.

3.2 Cooperation at Design Stage of Selfish Scenario

Even in selfish scenario, the participant is given a chance to cooperate. In the communication phase of the design stage, it gets information on knowledge aims and restrictions of its neighbors and it *may* exploit them for updating of its model and ideal in accordance with its fidelity into the information obtained.

The exploitation of models of $\mathcal{E}$s as well as of prior, ideal and strategy-describing pdfs have a common formal structure as all these objects are pdfs. The common exploitation way is discussed here.

At abstract level, $\mathcal{P}$ deals with the pdf ${}^{\mathcal{P}}f\left({}^{\mathcal{P}}Q\right)$ and gets the pdf ${}^{\tilde{\mathcal{P}}}f\left({}^{\tilde{\mathcal{P}}}Q\right)$ from other participant $\tilde{\mathcal{P}}$. The cooperation makes sense if the intersection of sets of possible behaviors ${}^{\cap}Q^* \equiv {}^{\mathcal{P}}Q^* \cap {}^{\tilde{\mathcal{P}}}Q$ is non-empty. Let us decompose behavior ${}^{\mathcal{P}}Q \equiv (G, H)$ and ${}^{\tilde{\mathcal{P}}}Q \equiv (H, I)$ with $G \in {}^{\mathcal{P}}Q^* \setminus {}^{\cap}Q^*$, $H \in {}^{\cap}Q^*$, $I \in {}^{\tilde{\mathcal{P}}}Q^* \setminus {}^{\cap}Q^*$. Thus, we deal with a pair of pdfs ${}^{\mathcal{P}}f(G, H)$, ${}^{\tilde{\mathcal{P}}}f(H, I)$. Generally, there is a discrepancy between these pdfs, i.e. the marginal pdfs ${}^{\mathcal{P}}f(H)$, ${}^{\tilde{\mathcal{P}}}f(H)$ differ.

Let ${}^{\mathcal{P}}\alpha \in [0, 1]$ be belief of the participant into correctness (relevance) of its pdf ${}^{\mathcal{P}}f(G, H)$ and $1 - {}^{\mathcal{P}}\alpha$ to that provided by the participant $\tilde{\mathcal{P}}$.

The addressed problem can be now formulated as a construction of estimate ${}^{\mathcal{P}}\hat{f}(G,H)$ of the unknown objective pdf (Berec & Kárný 1997) ${}^{o\mathcal{P}}f(G,H)$ using the available knowledge ${}^{\mathcal{P}}f(G,H),\ {}^{\tilde{\mathcal{P}}}f(H,I)$.

The advocated solution is based on several assumptions formulated and justified now.

The conditional expectation ${}^{\mathcal{P}}\hat{f}(G,H) \equiv E\left[{}^{o}f|\,{}^{\mathcal{P}}f,\,{}^{\tilde{\mathcal{P}}}f\right]$ is the best estimate of ${}^{o\mathcal{P}}f(G,H)$ minimizing the expected KL divergence $E\mathcal{D}\left({}^{o}f||\,{}^{\mathcal{P}}\hat{f}\right)$ that should serve as the adequate expected loss of the problem, as it follows from Proposition 2.

The expectation is taken over the set of unknown objective pdfs ${}^{o}f^{*}$. Its conditional version will be constructed using the observations listed further on.

The participant $\mathcal{P}$ is not aware of the space ${}^{\tilde{\mathcal{P}}}Q^{*} \setminus {}^{cap}Q^{*}$ and thus it may exploit at most the marginal pdf ${}^{\tilde{\mathcal{P}}}f(H)$. Similarly, the participant $\tilde{\mathcal{P}}$ is not aware of the space ${}^{\mathcal{P}}Q^{*} \setminus {}^{cap}Q^{*}$ thus it brings no information on ${}^{o\mathcal{P}}f(G,H)$.

The only available information about relationship between G and H is in ${}^{\mathcal{P}}f(G|H)$ which should thus coincide with the corresponding conditional expectation. These observations motivate assumptions

$$E\left[{}^{o\mathcal{P}}f(G,H)|\,{}^{\mathcal{P}}f(G,H),\,{}^{\tilde{\mathcal{P}}}f(H,I)\right] \equiv E\left[{}^{o\mathcal{P}}f(G,H)|\,{}^{\mathcal{P}}f(G,H),\,{}^{\tilde{\mathcal{P}}}f(H)\right]$$

$$E\left[{}^{o\mathcal{P}}f(G|H)|\,{}^{\mathcal{P}}f(G,H),\,{}^{\tilde{\mathcal{P}}}f(H)\right] \equiv E\left[{}^{o\mathcal{P}}f(G|H)|\,{}^{\mathcal{P}}f(G|H),\,{}^{\mathcal{P}}f(H)\right] \equiv {}^{\mathcal{P}}f(G|H). \tag{16}$$

No information on correlation of ${}^{o\mathcal{P}}f(G|H)$ and ${}^{o\mathcal{P}}f(H)$ is available. Thus, we can assume furthermore that

$$E\left[{}^{o\mathcal{P}}f(G|H)\,{}^{o\mathcal{P}}f(H)|\,{}^{\mathcal{P}}f(G|H),\,{}^{\mathcal{P}}f(H),\,{}^{\tilde{\mathcal{P}}}f(H)\right] \underbrace{=}_{(16)} {}^{\mathcal{P}}f(G|H)E\left[{}^{o\mathcal{P}}f(H)|\,{}^{\mathcal{P}}f(G|H),\,{}^{\mathcal{P}}f(H),\,{}^{\tilde{\mathcal{P}}}f(H)\right]. \tag{17}$$

Thus it remains to specify the expectation of ${}^{o\mathcal{P}}f(H)$ given (essentially) by a pair of alternatives with given degrees of belief. Thus, it makes sense to adopt the last assumption

$$E\left[{}^{o\mathcal{P}}f(H)|\,{}^{\mathcal{P}}f(G|H),\,{}^{\mathcal{P}}f(H),\,{}^{\tilde{\mathcal{P}}}f(H)\right] \equiv E\left[{}^{o\mathcal{P}}f(H)|\,{}^{\mathcal{P}}f(H),\,{}^{\tilde{\mathcal{P}}}f(H)\right] \equiv {}^{\mathcal{P}}\alpha\,{}^{\mathcal{P}}f(H)+\left(1-{}^{\mathcal{P}}\alpha\right)\,{}^{\tilde{\mathcal{P}}}f(H). \tag{18}$$

This implies the final composition rule defining the new pdf $\mathcal{P}\hat{f}$ of the participant P corrected by information provided by the participant $\tilde{\mathcal{P}}$. It is summarized in the following algorithmic form.

Algorithm 1 (Correction of pdf ${}^{\mathcal{P}}f$ by the pdf ${}^{\tilde{\mathcal{P}}}f$) *1. Participant $\mathcal{P}$ selects belief ${}^{\mathcal{P}}\alpha \in [0,1]$ in correctness of its pdf*

2. *Participant $\mathcal{P}$ gets marginal pdf ${}^{\tilde{\mathcal{P}}}f(H)$ of its neighbor $\tilde{\mathcal{P}}$ describing the common part H of behaviors ${}^{\mathcal{P}}Q \equiv (G,H),\ {}^{\tilde{\mathcal{P}}}Q \equiv (H,I)$ considered by the involved participants $\mathcal{P}$, $\tilde{\mathcal{P}}$*
3. *Participant $\mathcal{P}$ defines the corrected pdf ${}^{\mathcal{P}}\hat{f}(G,H)$ by the formula*

$${}^{\mathcal{P}}\hat{f}(G,H) = {}^{\mathcal{P}}f(G|H)\left[{}^{\mathcal{P}}\alpha\,{}^{\mathcal{P}}f(H) + \left(1-{}^{\mathcal{P}}\alpha\right)\,{}^{\tilde{\mathcal{P}}}f(H)\right]. \tag{19}$$

Remarks

1. Application of Algorithm 1 to prior pdfs, models of environments as well as to ideal pdfs is generally non-trivial as the recognition of the common variable H and evaluation of the marginal pdf need not be straightforward. Examples of detailed evaluations are out of scope of this paper.
2. Section 3.4 that deals with hierarchical scenario combines individual aims indirectly via estimation. Even in this case, Algorithm 1 is useful as it shows that the combined model should be searched as the mixture model (19.
3. The improved model ${}^{\mathcal{P}}\hat{f}$ is supposed to serve to $\mathcal{P}$. Thus, $\mathcal{P}$ has a freedom to assign belief ${}^{\mathcal{P}}\alpha$ to itself. Obviously, it is wise to make this probability proportional to its predictive ability on the common observable part of H in comparison with this ability of information provider $\mathcal{P}$.
4. Extension of the proposed combination to a larger number of cooperating participants is nontrivial. Often, however, pair-wise comparison (possibly iterative) can be sufficient.

3.3 Cooperative Scenario

In the selfish scenario, communicating partners have right to decide on degree of exploitation of the information provided by their neighbors. In the cooperative scenario, they are obliged to come to a common model and ideal concerning to the common part of their behavior. It means that have to negotiate on the belief values assigned to the partner. They have to reach the situation with ${}^{\mathcal{P}}\alpha = 1 - {}^{\tilde{\mathcal{P}}}\alpha$, they have to *negotiate*.

The fair negotiation requires to judge final loss on union of both considered behaviors, i.e. to define the common model ${}^{\mathcal{P}\tilde{\mathcal{P}}}f(G,H,I)$ and common ideal ${}^{I\mathcal{P}\tilde{\mathcal{P}}}f(G,H.I)$ as a function of the common belief(s) α and to minimize the smallest value of the KL divergence $\mathcal{D}\left({}^{\mathcal{P}\tilde{\mathcal{P}}}f||\,{}^{I\mathcal{P}\tilde{\mathcal{P}}}f\right)$ with respect to these beliefs.

Practically, the minimization can be performed by comparing values on a relatively sparse grid in belief-values space. Thus, the extension of the involved pdfs on the union of behaviors is the key problem to be addressed. There are indicators that the pdf

$$ {}^{\mathcal{P}\tilde{\mathcal{P}}}f(G,H,I) = {}^{\mathcal{P}}f(G|H)\,{}^{\tilde{\mathcal{P}}}f(I|H)\left[\alpha\,{}^{\mathcal{P}}f(H) + (1-\alpha)\,{}^{\tilde{\mathcal{P}}}f(H)\right] \tag{20} $$

is the adequate extension. No rigorous justification is, however, available.

3.4 Hierarchical Scenario

Hierarchical scenario supposes that actions of individual participants in the group $\mathcal{P}^*$ are coordinated by a single member of the group, referred to as *coordinator* $\mathcal{C} \in \mathcal{P}^*$. *The coordinator has the power to influence other participants in the group* by: i) enforcing ideal pdfs and models of their environments (at design stage); ii) data supplied (at acting stage). Although individual participants may behave as selfish or cooperative, they must accept some ingredients enforced to them by the coordinator $\mathcal{C}$. The degree of obedience is predefined by strength of hierarchy.

Type of coordination is determined by the way how $\mathcal{C}$ constructs its own ideal according to which it coordinates the group $\mathcal{P}^*$. A selfish coordinator respecting aims neither group nor particular participants behaves as a dictator. The coordinator that constructs its ideal pdf so that group aim as well as aims of individual members are respected as much as possible can be labelled as democrat. The construction of such a "democratic ideal" ${}^{I\mathcal{C}}f(d_t|d(t-1))$ is the major new task connected with the hierarchical scenario. The task solution is discussed below. For a complementary information see (Kárný & Kracík 2004).

In the task formalization, the coordinator $\mathcal{C}$ is assumed:

- To have at disposal data that inform it about the particular participants behavior and behavior of the group as a whole.
- To share a part of its data d_t with all participants of the group. These data may include actions of the coordinator as well as its innovations serving for coordination.
- To have information about externally supplied group aim (if any).
- To have at disposal information about aims of individual participants. It serves for harmonizing these individual aims both mutually and with the group aim.
- To obtain regularly data with respect to which members of the group have expressed their aims.

The construction of the coordinator's ideal ${}^{I\mathcal{C}}f$ is straightforward if the coordinated group has $\mathring{\mathcal{P}}$ members with a small $\mathring{\mathcal{P}}$. It has to describe the common data available to each participant and thus it suffices to create the ideal pdf from the corresponding marginal pdfs of respective ideals. Their convex combination solves this problem. The weights, selected by the coordinator, reflects the significance it assigns to participants. The constant weights $\mathring{\mathcal{P}}^{-1}$ can be taken as prototype of the fair weights. This simple weighting is possible due to the common expression of losses in terms of the KL divergences.

Real problem arises when the number of coordinated participants is large. It this case, however, it can be expected that participant aims can be clustered in relatively small amount $\mathring{\mathcal{P}} < \infty$ of clusters. A cluster contains an ${}^{\mathcal{P}}\alpha > 0$ part of the group participants $\left(\sum_{\mathcal{P}=1}^{\mathring{\mathcal{P}}} {}^{\mathcal{P}}\alpha = 1\right)$. Each cluster has data ${}^{\mathcal{P}}d_t$ on which cluster participants express their aims through a cluster ideal ${}^{I\mathcal{P}}f\left({}^{\mathcal{P}}d_t|\,{}^{\mathcal{P}}d(t-1)\right)$.

The coordinator processes the provided cluster ideals by employing data ${}^{\bar{\mathcal{P}}}d_t$ not available to the participants from this cluster. Actually, the extension (12) of the cluster ideal pdf with respect to the full data record $d_t = \left({}^{\mathcal{P}}d_t,\, {}^{\bar{\mathcal{P}}}d_t\right)$ available to the coordinator, is defined as

$$ {}^{I\mathcal{P}}f(d_t|d(t-1)) = {}^{I\mathcal{C}}f\left({}^{\bar{\mathcal{P}}}d_t|d_t, d(t-1)\right)\,{}^{I\mathcal{P}}f\left({}^{\mathcal{P}}d_t|\,{}^{\mathcal{P}}d(t-1)\right) \text{ for any } \mathcal{P} = 1,\ldots,\mathring{\mathcal{P}} \tag{21} $$

where the marginal pdf of the constructed coordinator's ideal ${}^{I\mathcal{C}}f(d_t|d(t-1))$ is used. With the extension (21), the coordinator's democratic ideal ${}^{I\mathcal{C}}f(d_t|d(t-1))$, respecting ideals of all clusters within a group, is searched for. The pdf defined on the coordinator's data d_t and closest to the extended ideals of particular clusters is selected, i.e.:

$$^{IC}f(d_t|d(t-1)) = \mathrm{Arg}\min_{f(d_t|d(t-1))} \sum_{\mathcal{P}\in\mathcal{P}^*} {}^{\mathcal{P}}\alpha \mathcal{D}\left({}^{I\mathcal{P}}f(d_t|d(t-1)) \,||\, f(d_t|d(t-1))\right). \quad (22)$$

This formulation can be interpreted as minimization of the expected KL divergence of the objective but unknown coordinator's ideal $^{oIC}f(d_t|d(t-1))$ and its approximation $^{IC}f(d_t|d(t-1))$. For it, we have to assume that $^{oIC}f(d_t|d(t-1)) \in {}^{oIC}f^*(d_t|d(t-1)) \equiv \left\{ {}^{I\mathcal{P}}f(d_t|d(t-1))\right\}_{\mathcal{P}\in\mathcal{P}^*}$ and

$$\mathsf{Probability}\left[{}^{oIC}f(d_t|d(t-1)) = {}^{I\mathcal{P}}f(d_t|d(t-1)) | {}^{oIC}f^*(d_t|d(t-1))\right] = {}^{\mathcal{P}}\alpha.$$

Unrestricted minimization gives implicit equation for the ideal search for

$$^{IC}f(d_t|d(t-1)) = \sum_{\mathcal{P}\in\mathcal{P}^*} {}^{\mathcal{P}}\alpha\, {}^{I\mathcal{P}}f(d_t|d(t-1)) \underbrace{=}_{(21)} \sum_{\mathcal{P}\in\mathcal{P}^*} {}^{\mathcal{P}}\alpha\, {}^{IC}f\left({}^{\bar{\mathcal{P}}}d_t|d_t, d(t-1)\right) {}^{I\mathcal{P}}f\left({}^{\mathcal{P}}d_t| {}^{\mathcal{P}}d(t-1)\right). \quad (23)$$

Equation (23) has no analytic solution for generic cases with $\mathcal{P}$. This problem can be practically avoided by replacing the extension (21). Alternatively, the cluster ideal can be extended on the full data record $d = \left({}^{\mathcal{P}}d, {}^{\bar{\mathcal{P}}}d\right)$ by multiplying the marginal pdf $^{I\mathcal{P}}f\left({}^{\mathcal{P}}d_t| {}^{\mathcal{P}}d(t-1)\right)$ by a flat pdf $^{I\bar{\mathcal{P}}}\bar{f}\left({}^{\bar{\mathcal{P}}}d_t|d(t-1)\right)$ defined on the data $^{\bar{\mathcal{P}}}d_t$ out of the interests of the $\mathcal{P}$th cluster. Then, the *optimal joint ideal pdf of the whole group* equals

$$^{I}f(d_t|d(t-1)) = \sum_{\mathcal{P}\in\mathcal{P}^*} {}^{\mathcal{P}}\alpha\, {}^{I\mathcal{P}}f\left({}^{\mathcal{P}}d_t| {}^{\mathcal{P}}d(t-1)\right) {}^{I\bar{\mathcal{P}}}\bar{f}\left({}^{\bar{\mathcal{P}}}d_t|d(t-1)\right). \quad (24)$$

The obtained explicit group ideal (24), representing compromise between ideals of all clusters entering the group, can be taken as the coordinator's ideal $^{IC}f(d_t|d(t-1)) \equiv {}^{I}f(d_t|d(t-1))$. However even in this simple case, the direct use of the explicit group ideal, driving the fully probabilistic design of coordinator's strategy, is inhibited by the difficulty to cluster the participants properly. For it, we have to define which ideal pdfs are taken as practically identical and the results may be quite sensitive to this definition. Moreover, the well-established clustering techniques are not elaborated for such objects as pdfs are. Thus, it makes sense to search for an alternative way of constructing of the coordinator's ideal.

Constructing of the group ideal, addressed as *estimation problem*, seems to offer a remedy of all mentioned troubles. The above considerations indicate the ideal should be search in the form of finite mixture. The solution is described by the following algorithm.

Algorithm 2 (Data-based construction of the group ideal) *1. The coordinator's ideal is supposed to have a form of a finite mixture (Titterington et al. 1985) composed of $\mathring{\mathcal{P}}$ components $^{I}f\left(d_t|d(t-1), {}^{\mathcal{P}}\Theta\right)$ parameterized by finite collection of unknown parameters $^{\mathcal{P}}\Theta$ and weighted by unknown probabilistic weights $^{\mathcal{P}}\alpha$*

$$^{I}f(d_t|d(t-1),\Theta) = \sum_{\mathcal{P}=1}^{\mathring{\mathcal{P}}} {}^{\mathcal{P}}\alpha\, {}^{I}f\left(d_t|d(t-1), {}^{\mathcal{P}}\Theta\right).$$

2. *Each participant is required to provide data records characterizing its ideal. The data format and number of data records is previously defined and fixed for all participants of the group. All entries of the data record available to the coordinator has to be filled in.*
3. *The data, provided by participants, are processed to estimate a group ideal in the form of mixture model suggested. The standard Bayesian methodology, see e.g. (Kárný et al. 2003) and references there, is used for estimation of unknown mixture parameters $\Theta = \{\alpha_c, \mathring{c}, \Theta_c\}$. Resulting mixture model approximates the whole group data with particular components modelling individual clusters of the group.*

Remarks

1. In all variants, the degree of compromise between contradictory ideals is defined by the cluster (component) weights $^{\mathcal{P}}\alpha$.
2. As a rule, $\mathcal{P}$ expresses its wishes about sub-selection of data entries only, i.e. $d = \left({}^{\mathcal{P}}d, {}^{\bar{\mathcal{P}}}d\right)$. This introduces the hard problem of missing data into the mixture estimation. One possibility is to ask the participant

to provide lower and upper bound on entries of its interest and complement it by the lowest and highest bounds on the variables it is not interested in.
Alternatively, the fact that often whole groups of entries are out of the participant interest can be exploited. Considering for presentation simplicity the static parameterized ideal ${}^{I}f(d_t|d(t-1),\Theta) = {}^{I}f(d_t|\Theta)$, we get

$$f(\Theta|d_{\mathcal{P}}) \propto f(\Theta)\,{}^{I\mathcal{P}}f\left({}^{\mathcal{P}}d|\Theta\right) = f(\Theta)\int {}^{I\mathcal{P}}f(d|\Theta)\,d\,{}^{\bar{\mathcal{P}}}d.$$

To get a practical algorithmic solution of this formally correct treatment is, however, difficult and generally un-elaborated problem.

3. The data-based presentation of ideals is "natural" in societal problems (Kárný & Kracík 2004). It may be advantageous even in technical problems. Let ${}^{I\mathcal{P}}f({}^{\mathcal{P}}d_t|{}^{\mathcal{P}}d(t-1))$ be the ideal expressing the decision aims and restrictions. As a rule, it is rather complex pdf and the fully probabilistic design with it or with mixtures having it as a component is quite hard. Then, it makes sense to take samples from this pdf and use it for estimation of a mixture with numerically tractable components. Then, the estimation-based fitting of the ideal pdf performs simultaneously projection of the original "wild" ideal onto a feasible set.
4. Standard mixture estimation procedures can be used (Titterington et al. 1985). An advanced initialization (Kárný, et al. 2001b) combined with structure estimation becomes inevitable in high-dimensional cases. It guarantees practically appropriate clustering of the group.
5. The participant may have at disposal data not available to the coordinator. Then, their existence does not influence $\mathcal{C}$ who can assume that such surplus data (taken from its perspective) do not exist. Availability of surplus data of $\mathcal{C}$ that are strongly related to aims of the coordinated $\mathcal{P}$s is the main source of its ability to construct the proper ideal.
6. If power of enforcing of the ideal and models is weaken for the coordinator, the *scenario of horizontally cooperating group* can be obtained. Practically it means that some participant (called *probabilistic advisor*) is assumed to have enough modelling and evaluation power allowing to support other participants in their environment (Kárný et al. 2003, Quinn, et al. 2003). Instead of enforcing, probabilistic advisor offers projections of the optimized models as advices to respective participants. Supported participants have freedom to take the ideal pdfs optimized by advisor as the ideal pdfs for the data shared with advisor. In extreme case, they can follow blindly up to copying the advised actions. They are, however, responsible for the achieved quality and thus they should harmonize these advices with the information and actions related to their surplus data that are not at disposal to advisor.

4 Concluding Remarks

MP DDM discussed in this paper has many facets and problems. Besides the problems unsolved for SP DDM and thus transferred to MP DDM, new specific problems arise. Let us list some of them, we are aware of.

- It is not clear whether classification introduced within this paper is general enough and that no significant aspect has been left aside.
- Communication ways are extremely important in MP DDM. The net of participants should be structured – both in time and space – in order to get a high overall performance or even just stable behavior.
- Efficient storing of function of many variables (both pdfs and functions appearing in the fully probabilistic design) becomes even more important in MP DDM than in SP DDM.
- Estimation with incomplete data becomes especially urgent in MP DDM.
- Sharing of probabilistic models leads directly to finite mixtures whose set is not closed with respect to learning and conditioning. These general problems still wait for their solutions.
- Extension of pdfs (a solution of general marginal problems) is to be elaborated as the outlined solution is neither unique nor sufficiently justified.

Each item of the above list represents difficult and fundamental research problems whose quick solution can hardly be expected. In spite of this, partial results reflected in the paper improve understanding of MP DDM and can be used in solution of particular problems like normative fair governing (Kárný & Kracík 2004), advising to other $\mathcal{P}$s or sensor fusion.

Concerning to relationships of discussed scenarios, it is clear that

- The hierarchical scenario has a significant potential to provide the highest performance if the cooperating group is small enough.
- The horizontal cooperation has potential to be better than selfish MP DDM. Comparing to groups manage-

able in hierarchical scenario, it is expected to be worse. It has, however, an extreme potential in *scalability*. Horizontal cooperation of participants with a sufficiently dense subset of coordinators can be successfully applied to extremely large MP DDM where hierarchical scenario cannot be applied.

- Combination of small-scale hierarchies with horizontal cooperations, enhanced potentially by the adaptive shifting of boundaries between participant environments, seems to be the most powerful way of solving large scale MP DDM. It also seems that Nature came to the same conclusion.

Acknowledgments

This work was supported by ESF, project TED, GAČR 102/03/P010 and GAČR 102/03/0049.

References

O. Barndorff-Nielsen (1978). *Information and exponential families in statistical theory*. Wiley, New York.

L. Berec & M. Kárný (1997). 'Identification of reality in Bayesian context'. In K. Warwick & M. Kárný (eds.), *Computer-Intensive Methods in Control and Signal Processing: Curse of Dimensionality*, pp. 181–193. Birkhäuser.

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

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

A. Jazwinski (1970). *Stochastic Processes and Filtering Theory*. Academic Press, New York.

M. Kárný (1996). 'Towards fully probabilistic control design'. *Automatica* **32**(12):1719–1722.

M. Kárný, et al. (2003). 'Mixture-based adaptive probabilistic control'. *International Journal of Adaptive Control and Signal Processing* **17**(2):119–132.

M. Kárný, et al. (2001a). 'Quantification of prior information revised'. *International Journal of Adaptive Control and Signal Processing* **15**(1):65–84.

M. Kárný & J. Kracík (2004). 'A Normative Probabilistic Design of a Fair Governmental Decision Strategy'. *Journal of Multi-Criteria Decision Analysis* **10**:1–15.

M. Kárný, et al. (2001b). 'Initial description of multi-modal dynamic models'. In V. Kůrková, R. Neruda, M. Kárný, & N. C. Steele (eds.), *Artificial Neural Nets and Genetic Algorithms. Proceedings*, pp. 398–401, Wien. Springer.

R. Kulhavý (1994). 'Bayesian estimation, large deviations, and incomplete data'. In *IEEE Conference on Decision and Control*, vol. 1, pp. 755–756. Orlando, FL.

S. Kullback & R. Leibler (1951). 'On information and sufficiency'. *Annals of Mathematical Statistics* **22**:79–87.

M. Loeve (1962). *Probability theory*. D. van Nostrand Company Inc., Princeton, New Jersey. Russian translation, Moscow 1962.

J. Martin, et al. (2003). 'Joint sensitivity in Bayesian decision theory'. *TEST* **12**(1):173–194.

A. Perez (1965). 'Information, ε-sufficiency and data reduction problem'. *Kybernetika* **1**:299–310.

V. Peterka (1981). 'Bayesian system identification'. In P. Eykhoff (ed.), *Trends and Progress in System Identification*, pp. 239–304. Pergamon Press, Oxford.

A. Quinn, et al. (2003). 'Probabilistic advisory systems for data-intensive applications'. *International Journal of Adaptive Control and Signal Processing* **17**(2):133–148.

I. Sanov (1957). 'On probability of large deviations of random variables'. *Matematičeskij Sbornik* **42**:11–44. in Russian, translation in Selected Translations mathematical statistics and probability, I, 1961, 213–244.

D. Titterington, et al. (1985). *Statistical Analysis of Finite Mixtures*. John Wiley & Sons, Chichester, New York, Brisbane, Toronto, Singapore. ISBN 0 471 90763 4.