DISTRIBUTED BAYESIAN DECISION MAKING: EARLY EXPERIMENTS

Václav Šmídl, Josef Andrýsek

Institute of Information Theory and Automation,
Prague, Czech Republic,
`[smidl,andrysek]@utia.cas.cz`

1. INTRODUCTION

Decision-making under uncertainty is a natural part of everyday life of every human being. In societal science, various aspects of decision-making were studied, mostly in the area of psychology. In technical science, the process was formalized using probability theory yielding so called Bayesian theory of decision making (Berger, 1985). However, one of the key assumptions of this theory is that the decision-maker is the only entity that intentionally influences the system. This assumption is certainly violated in more complicated systems, such as human society or distributed control. Recently, a serie of papers attempts to offer an extension of the Bayesian theory for many decision-makers (Kárný and Guy, 2004), i.e. decentralized stochastic control. Since there are no proofs of optimality of the proposed Bayesian distributed decision making in the literature, we study this approach via experimental simulation studies. In this paper we present the first experimental results of the approach.

2. SUMMARY OF DISTRIBUTED BAYESIAN DECISION MAKING

Bayesian decision making (DM) is based on the following principle (Berger, 1985): *Incomplete knowledge and randomness have the same operational consequences for decision making*. Therefore, all unknown quantities are treated as random variables and formulation of the problem and its solution are firmly based within the framework of probability calculus.

This task of designing of a decision-maker consists of the following sub-tasks:

Model Parametrization: Each agent must have its own model of its *neighbourhood*, i.e. part of the overall environment. Uncertainty in the model is described by parametrized probability density functions.

Learning: Is an operation that reduces uncertainty in the neighbourhood model, using the observed data. In Bayesian paradigm, this task is reduced to estimation of the model parameters.

DM Strategy Design: Is an operation that produce a rule for generating decisions based on the observations. The goal of this task is to design the best possible strategy in order to reach presribed aims of the decision-maker.

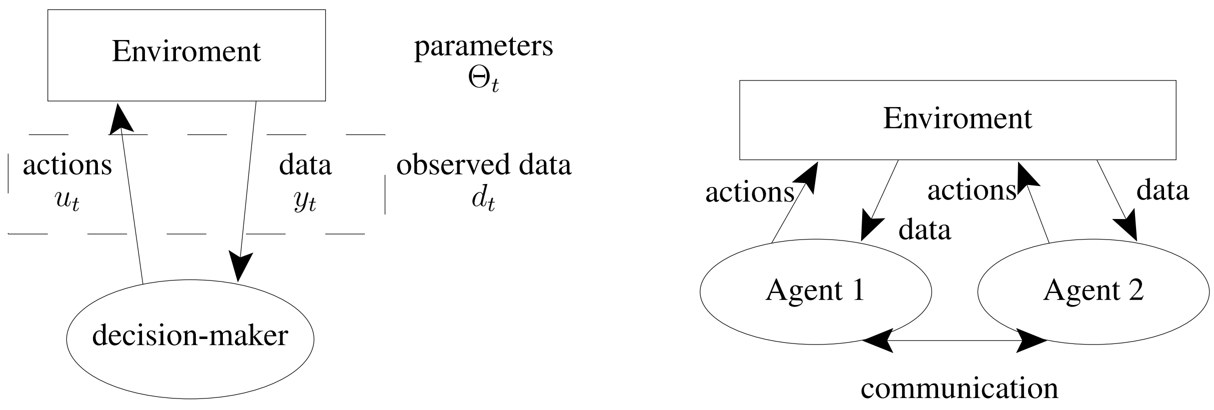

Figure 1: Illustration of Bayesian (left) and distributed Bayesian decision-making (right).

The previous operations are defined in standard Bayesian decision-making. In order extend the paradigm for distributed decision-making, extra sub-task must be addressed:

Communication & Negotiation: Decision-makers echange messages with information about models and aims. The goal of this sub-task is to design proper handling of this information which ensures that individual decision-makers do not act against each other, but cooperate whenever possible.

With this extra operation, the decision-makers behave like intelligent agents in multi-agent systems (Weiss, 2000). Hence, we will refer to the extended decision-makers as *agents*. The above sub-tasks will be now described in detail.

2.1 Model Parametrization

All quantities observable on the environment at time t will be denoted d_t. These quantities can be either data, y_t, or actions, u_t, $d_t = [y_t', u_t']'$. Θ_t is an unknown parameter of the model of the environment. In Bayesian framework, the *closed loop*—i.e. the environment *and* the agent—is described by the following probability density function:

$$f\left(d^{1:t}, \Theta^{1:t}\right) = \prod_{\tau=1}^{t} f\left(y_\tau | u_\tau, \Theta_\tau\right) f\left(\Theta_\tau | u_\tau, \Theta_{\tau-1}\right) f\left(u_\tau | d^{1:\tau-1}\right). \tag{1}$$

Here, $f(\cdot)$ denotes probability density function (pdf) of its argument. $d^{1:t}$ denotes the observation history $d^{1:t} = [d_1, \ldots, d_t]$. The model represents the whole trajectory of the system, including inputs u_t which can be influenced by the agent. The chosen order of conditioning in (1) distinguishes the following important pdfs: (i) observation model $f(y_t|u_t, \Theta_t)$, (ii) internal model $f(\Theta_t|u_t, \Theta_{t-1})$, and (iii) decision-making strategy $f(u_t|d^{1:t-1})$.

We suppose that the observation model and the internal model are given (or chosen from given alternatives), while the decision-making strategy is being designed.

2.2 Learning via Bayesian filtering

In Bayesian paradigm, the task of learning is equivalent to evaluation of posterior distribution of unknown parameters conditioned on the observed data, $f(\Theta_t|d^{1:t})$. This pdf can be computed recursively as follows:

$$f\left(\Theta_t|u_t,d^{1:t-1}\right) = \int f\left(\Theta_t|u_t,d^{1:t-1},\Theta_{t-1}\right) f\left(\Theta_{t-1}|d^{1:t-1}\right) d\Theta_{t-1}, \tag{2}$$

$$f\left(\Theta_t|d^{1:t}\right) \propto \frac{f\left(y_t|u_t,d^{1:t-1},\Theta_t\right) f\left(\Theta_t|u_t,d^{1:t-1}\right)}{f\left(y_t|u_t,d^{1:t-1}\right)}, \tag{3}$$

$$f\left(y_t|u_t,d^{1:t-1}\right) = \int f\left(y_t|u_t,d^{1:t-1},\Theta_t\right) f\left(\Theta_t|u_t,d^{1:t-1}\right) d\Theta_t. \tag{4}$$

In general, evaluation of the above pdfs is a complicated task, which is often intractable and many approximate techniques must be used (Chen, 2003). In this text, we are concerned with conceptual issues and we assume that all operations (2)–(4) are tractable.

2.3 Design of DM strategy

In this Section, we review *fully probabilistic design (FPD)* of the DM strategy (Kárný, 1996). This approach is an alternative to the standard stochastic control design, which is formulated as minimization of an expected loss function with respect to decision making strategies (Astrom, 1970; Bertsekas, 2001). The FPD starts with specification of the decision making aim in the form of *ideal pdf* of the closed loop. This ideal pdf—which is the key object of this approach—is constructed in the same form as (1), from which it is distinguished by superscript ^{I}:

$$f\left(d^{1:t},\Theta^{1:t}\right) \rightarrow {}^I\! f\left(d^{1:t},\Theta^{1:t}\right). \tag{5}$$

Similarly to (1), the ideal distribution is decomposed into ideal observation model, internal model, and ideal DM strategy. Recall, from Section 2.1, that model (1) contains the DM strategy, which can be freely chosen. Therefore, the optimal DM strategy can be found by functional optimization of the following loss function

$$\mathcal{L}\left(f\left(u_t|d^{1:t-1}\right),\mathring{t}\right) = \mathsf{D}\left(f\left(d^{1:\mathring{t}},\Theta^{1:\mathring{t}}\right) || {}^I\! f\left(d^{1:\mathring{t}},\Theta^{1:\mathring{t}}\right)\right), \tag{6}$$

where $\mathsf{D}(\cdot,\cdot)$ denotes the Kullback-Leibler divergence between the current (learnt) and the ideal pdf (Kullback and Leibler, 1951), and $\mathring{t} > t$ is the decision making horizon.

Minimum of the loss (6)—i.e. the optimal DM strategy—is found in *closed form*:

$$f\left(u_t|d^{1:t-1}\right) = {}^I\! f(u_t|d^{1:t-1}))\frac{\exp[-\omega(u_t,d^{1:t-1})]}{\gamma(d^{1:t-1})}, \tag{7}$$

$$\omega(u_t,d^{1:t-1}) = \int f(y_t|u_t,d^{1:t-1}) \ln\left(\frac{f(y_t|u_t,d^{1:t-1})}{\gamma(d^{1:t})\, {}^I\! f(y_t|u_t,d^{1:t-1})}\right) dy_t. \tag{8}$$

$$\gamma(d^{1:t-1}) = \int {}^I\! f(u_t|d^{1:t-1}) \exp[-\omega(u_t,d^{1:t-1})]\, du_t. \tag{9}$$

The decisions are then generated backward in time starting at horizon $\mathring{t}$ with initial value $\gamma(\mathring{t}) = 1$.

2.4 Communication & Negotiation

Cooperation between autonomously acting agents can arise if their aims an models are mutually compatible. This can be assured at the beginning of an experiment by design. However, the models and aims can diverge, since each agent is continually updating them using the observed data. The process of synchronizing of aims is known as negotiation in multi-agent systems (Weiss, 2000). In Bayesian paradigm, this operation must also be formalized by means of probabilistic calculus. No such operation exists in standard Bayesian theory. In (Kárný and Guy, 2004), one possible formalization of negotiation was proposed using probabilistic merging. In general, merging refers to a process of creating one *target* pdf from two *source* pdfs.

When we want to synchronize the aims defined on the same variable, e.g. via ideal distribution on observation y_t, we seek a distribution defined on the same variable. For this purpose, we will use *direct* merging operation (Kracík, 2004). Synchronization of models is more demanding, since each decision maker may have different model, moreover its parametrization is unknown to the neighbour. The agents may exchange only distributions on commonly known variables. The task is then to use the neighbours distribution on the data $d^{1:t}$ to modify the posterior distribution $f\left(\Theta_t|d^{1:t}\right)$. This operation is known as *indirect* merging (Kracík, 2005).

Direct merging In this Section, we restrict our attention to cases where the sources, and thus the target, are defined on the same multivariate variable. For merging of partially overlapping multivariate variables see (Kracík, 2004). Merging is defined as optimization problem, where the target distribution is found as follows:

$$\tilde{f}\left(y_t\right) = \arg\min_{f}\left(\alpha \mathsf{D}\left(f_1\left(y_t\right)||f\left(y_t\right)\right) + \left(1-\alpha\right)\mathsf{D}\left(f_2\left(y_t\right)||f\left(y_t\right)\right)\right). \tag{10}$$

Here, D denotes the Kullback-Leibler divergence, and scalar parameter $\alpha \in \langle 0, 1\rangle$ is used to tune the importance of each source. Functional minimization of (10), yields the solution in the form of probabilistic mixture:

$$\tilde{f}\left(y_t\right) = \alpha f_1\left(y_t\right) + \left(1-\alpha\right) f_2\left(y_t\right). \tag{11}$$

Note that complexity of this solution grows with repeated use of this rule, and soon it may become computationally prohibitive. Therefore, an alternative approach is to minimize (10) under the restriction of $f()$ being from a class parametrized by multivariate parameter Φ. Then, the necessary condition for an optimum is equality of all moments of $\tilde{f}()$ that depends on Φ, e.g.

$$\mathsf{E}_{\tilde{f}(y_t|\Phi)}(y_t) = \mathsf{E}_{\alpha f_1(y_t)+(1-\alpha)f_2(y_t)}(y_t), \tag{12}$$

for the first moment and similarly for higher moments. Here, $\mathsf{E}\left[\,\right]$ denotes expected value of the argument with respect to pdf defined in the subscript. The subscript can be omitted in situations when it is clear which pdf is being used.

Alternatively, the problem can be formulated in the reverse KL divergence

$$\tilde{f}\left(y_t\right) = \arg\min_{f}\left(\alpha \mathsf{D}\left(f\left(y_t\right)||f_1\left(y_t\right)\right) + \left(1-\alpha\right)\mathsf{D}\left(f\left(y_t\right)||f_2\left(y_t\right)\right)\right). \tag{13}$$

Optimum of (13), is found in the form of a geometric mean of the source pdfs:

$$\tilde{f}\left(y_t\right) = \left(f_1\left(y_t\right)\right)^{\alpha}\left(f_2\left(y_t\right)\right)^{(1-\alpha)}. \tag{14}$$

Indirect merging This operation is closely related to Bayesian estimation, and thus it can be formalized using formula for evaluation of sufficient statistics (Kracík and Kárný, 2005) as follows:

$$\tilde{f}\left(\Theta_t|d^{1:t}\right) = f_1\left(\Theta_t|d^{1:t}\right)\exp\left[\beta\int f_2(\Psi_t)\ln f_1\left(\Psi_t,\Theta_t\right)d\Psi_t\right]. \quad (15)$$

Here, Ψ_t denotes the regressor, i.e. a vector of delayed observation $\Psi_t = [d_t', d_{t-1}', \ldots]'$, scalar parameter β governs the importance of the communicated information.

Negotiation Is a process of finding such an aim (or model) that would be acceptable for both agents. In distributed Bayesian DM, the task of negotiation is reduced to selection of weights α and β for direct and indirect merging respectively. It is possible to formulate several strategies how to choose these parameters (Kárný and Guy, 2004), however in this paper, we will assume that they are chosen a priori.

3. ILLUSTRATIVE EXAMPLE: ROOM TEMPERATURE CONTROL

Consider the following autoregressive two-input one-output model of the environment:

$$y_t = ay_{t-1} + by_{t-2} + cu_{1,t} + du_{2,t} + e_t, \quad (16)$$

where a, b, c, d are scalar parameters and $e_t \sim \mathcal{N}(0,\sigma)$ is a realization of normally distributed noise with zero-mean and variance σ. The environment is influenced by two agents, each controlling one of the inputs u in (16), i.e. A_1 decides on strategy of $u_{1,t}$ and A_2 on strategy of $u_{2,t}$. However, the agents have incomplete model of the environment:

$$A_1: \quad y_t = a_1y_{t-1} + b_1y_{t-2} + c_1u_{1,t} + 0u_{2,t} + \sigma_1e_{1,t}, \quad (17)$$

$$A_2: \quad y_t = a_2y_{t-1} + b_2y_{t-2} + 0u_{1,t} + c_2u_{2,t} + \sigma_2e_{2,t}, \quad (18)$$

Incompleteness of the model is hard-coded by zero coefficients of input actions of the other agent, which means that the agents are, by design, unaware of the actions of the neighbour. The task is to compensate for this ignorance via communication of knowledge.

3.1 Learning

Since both agents are using the same model structure, we derive the learning algorithm only for one of them. The unknown parameters, a, b, c, and σ, are aggregated into a vector $\theta_t = [a, b, c]$, $\Theta_t = [\theta_t, \sigma]$. The parameters are assumed to be stationary, i.e. the parameter evolution model is

$$f\left(\Theta_t|\Theta_{t-1}\right) = \delta\left(\Theta_t - \Theta_{t-1}\right),$$

where $\delta(\cdot)$ denotes the Dirac delta function. Under the assumption of Gaussian noise, the observation model is Normal distribution

$$f\left(y_t|u_t,\Theta_t,d^{1:t-1}\right) = \mathcal{N}\left(\theta_t'\psi_t, r_t\right), \quad (19)$$

where $\psi_t = [y_{t-1}, y_{t-2}, u_t]$'. Since (19) is from the dynamic exponential family (Kárný *et al.*, 2005), it is reasonable to choose the prior distribution as conjugate to it, i.e. Normal-inverse-gamma distribution:

$$f\left(\Theta_t\right) = \mathcal{N}i\mathcal{G}\left(V_0,\nu_0\right). \quad (20)$$

Then, the solution of one step update (3) has also the form of (20) with statistics V_t, ν_t updated as follows:

$$V_t = V_{t-1} + \varphi_t \varphi_t', \quad \nu_t = \nu_{t-1} + 1.$$

Here, $\varphi_t = [y_t, \psi_t']'$. Using assignment $\lambda_t = V_{yy,t} - V_{\psi y,t} V_{\psi\psi,t}^{-1} V_{y\psi,t}$ and decomposition $V_t = \begin{bmatrix} V_{yy,t} & V_{y\psi,t} \\ V_{\psi y,t} & V_{\psi\psi,t} \end{bmatrix}$, moments of the posterior distribution are:

$$\mathsf{E}\left[\theta_t\right] = \hat{\theta}_t = V_{\psi\psi,t}^{-1} V_{y\psi,t}, \quad \mathsf{E}\left[r_t\right] = \hat{r}_t = \frac{1}{\nu - 7}\lambda_t.$$

The predictive distribution (4), is of Student-t type with $\nu_t - 1$ degrees of freedom

$$f\left(y_{t+1}|d^{1:t}\right) = \mathcal{S}t\left(\hat{\theta}_t \psi_{t+1}, \frac{1}{1+\zeta_t}\lambda_t, \nu_t - 1\right), \tag{21}$$

where $\zeta_t = \psi_t' V_{\psi\psi,t}^{-1} \psi$. Moments of (21) are

$$\begin{aligned} \hat{y}_{t+1} &= \hat{\theta}_t \psi_{t+1}, \\ \mathsf{E}\left[(y_{t+1} - \hat{y}_{t+1})^2\right] &= \hat{r}_t (1 + \zeta_t). \end{aligned}$$

With growing t, (21) rapidly converges to a Gaussian distribution with the same mean and variance, yielding a computationally tractable approximation.

3.2 Design of DM strategy

Ideal distributions Each agent assigns its own aims of decision making in the form of Gaussian pdfs:

$${}^I f\left(y_t\right) = \mathcal{N}\left({}^I\hat{y}_t, {}^I r_t\right), \tag{22}$$

$${}^I f\left(u_t\right) = \mathcal{N}\left({}^I\hat{u}_t, {}^I \rho_t\right), \tag{23}$$

where constants ${}^I\hat{y}_t$, ${}^I\hat{u}_t$, and ${}^I\rho_t$ can be chosen arbitrarily, however, ${}^I r_t$ is chosen as ${}^I r_t = \hat{r}_t$ for each participant. This choice reveals important computational simplifications in solution of (7)–(9).

Fully probabilistic design Note that equation (7) involves predictive distribution $f\left(y_t|u_t, d^{1:t-1}\right)$ which is the Student-t pdf (21) for the chosen model. Evaluation of (7) with for Student-t distribution is computationally intractable, however, it is tractable for the following approximation:

$$f\left(y_t|u_t, d^{1:t-1}\right) \approx \mathcal{N}\left(\hat{\theta}_t' \psi_t, \mathsf{E}\left[\sigma_t^2\right]\right).$$

Under this simplification and the chosen ideal distributions, the functional recursions (7)–(9) have the following algebraic form:

$$\begin{aligned} -\log \gamma(y_t, y_{t-1}) &= [y_t, y_{t-1}, 1] J_t [y_t, y_{t-1}, 1]', \\ \omega(u_t, y_{t-1}, y_{t-2}) &= [u_t, y_{t-1}, y_{t-2}, 1] K_t [u_t, y_{t-1}, y_{t-2}, 1]'. \end{aligned}$$

Kernels K_t and J_t can be evaluated recursively:

$$
\begin{aligned}
K_t &= \frac{1}{2\,{}^I r_t}\left[\hat{c},\hat{a},\hat{b},-{}^I\hat{y}_t\right]'\left[\hat{c},\hat{a},\hat{b},-{}^I\hat{y}_t\right] + \left[\hat{c},\hat{a},\hat{b},-{}^I\hat{y}_t\right] J_t \begin{bmatrix} \hat{c} & \hat{a} & \hat{b} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&+ \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.5\ln(\frac{{}^I r_t}{\hat{r}_t}) - 1 + \frac{\hat{r}_t}{{}^I r_t} + J_{1,1;t}\hat{r}_t \end{bmatrix}, \\
X_t &= K_t + \begin{bmatrix} \frac{1}{2\rho_t} & 0 & 0 & -\frac{{}^I\hat{u}_t}{2\rho_t} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -\frac{{}^I\hat{u}_t}{2\rho_t} & 0 & 0 & \frac{{}^I\hat{u}_t}{2\rho_t} + 0.5\log(2\pi\rho_t) \end{bmatrix} = \begin{bmatrix} X_{uu} & X_{u\psi} \\ X_{\psi u} & X_{\psi\psi} \end{bmatrix}, \\
J_{t-1} &= X_{\psi\psi} - \frac{X_{\psi u}X_{u\psi}}{X_{uu}X_{uu}}.
\end{aligned}
$$

The stochastic controller is then:

$$
\begin{aligned}
f\left(u_t|d^{1:t-1}\right) &= \mathcal{N}\left(\hat{u}_t, \eta_t\right), \\
\hat{u}_t &= -\frac{X_{u\psi}}{X_{uu}}, \eta_t = \frac{1}{sX_{uu}}.
\end{aligned}
\tag{24}
$$

3.3 Direct Merging

In this application, direct merging will be applied for merging of ideal distributions on the output y_t,

$$
{}^I\tilde{f}_i(y_t) \longleftarrow {}^I f_i(y_t),\ {}^I f_j(y_t).
$$

where j is used as an index of the neighbouring pdf, in our case $j = 3 - i$. We will distinguish two types of merging, each is optimal in one sense of Kullback-Leibler divergence.

Merging via linear combination Since the previous results heavily depends on the assumption of Gaussian ideal distributions, we seek the best possible target distribution from this class. For the case of linear combination, the optimal solution is found using moment matching (12) as follows:

$$
\begin{aligned}
{}^I\tilde{f}_i(y_t) &= \mathcal{N}\left({}^I\tilde{y}_{i,t},\ {}^I\tilde{\sigma}^2_{i,t}\right), \\
{}^I\tilde{y}_{i,t} &= \alpha\, {}^I y_{i,t} + (1-\alpha)\, {}^I y_{j,t}, \\
{}^I\tilde{\sigma}^2_{i,t} &= \alpha\left({}^I\sigma^2_{i,t} + {}^I y^2_{i,t}\right) + (1-\alpha)\left({}^I\sigma^2_{j,t} + {}^I y^2_{j,t}\right) - {}^I\tilde{y}^2_{i,t}.
\end{aligned}
$$

Merging via geometric combination Geometric combination of Gaussian distributions in (14) is again a Gaussian distribution, hence, the optimal distribution is:

$$
\begin{aligned}
{}^I\tilde{f}_i(y_t) &= \mathcal{N}\left({}^I\tilde{y}_{i,t},\ {}^I\tilde{\sigma}^2_{i,t}\right), \\
{}^I\tilde{y}_{i,t} &= {}^I\tilde{\sigma}^2_{i,t}\left(\alpha\, {}^I\sigma^{-2}_{i,t}\, {}^I y_{i,t} + (1-\alpha)\, {}^I\sigma^{-2}_{j,t}\, {}^I y_{j,t},\right), \\
{}^I\tilde{\sigma}^2_{i,t} &= \left(\alpha\, {}^I\sigma^{-2}_{i,t} + (1-\alpha)\, {}^I\sigma^{-2}_{j,t}\right)^{-1}.
\end{aligned}
$$

3.4 Indirect merging

Indirect merging will be used for merging of information about the environment from one agent to the other. An agent that receives knowledge from its neighbour will update its posterior distribution using (15). Note however, that we can not use this formula directly, since the agents can exchange information only about mutually known variables. In our case, $\varphi_{1,t} = [y_t, y_{t-1}, y_{t-2}, u_{1,t}]$ and $\varphi_{2,t} = [y_t, y_{t-1}, y_{t-2}, u_{2,t}]$, which have an intersection on $\tilde{\varphi}_t = [y_t, y_{t-1}, y_{t-2}]$. Therefore, in our experiment, we will use the following approximation.

1. ith agent send to its neighbour the following marginal:

$$f_i(\tilde{\varphi}_t) = \int f(y_t, y_{t-1}, y_{t-2}, u_{i,t})\, du_t = \int f(y_t|u_{1,t}, y_{t-1}, y_{t-2})\, f(u_{1,t}|y_{t-1}, y_{t-2}) \times$$
$$f(y_{t-1}|y_{t-2}, \boldsymbol{y}_{t-3})\, f(y_{t-2}|\boldsymbol{y}_{t-3}, \boldsymbol{y}_{t-4})\, du_t.$$

   Here, $\boldsymbol{y_{t-3}}$ denotes the observed value of y_t at time $t-3$. Note that the result depends not only on the estimated parameters $\hat{\theta}_t$, $\mathsf{E}[\sigma_t^2]$ via (21), but also on the designed strategy (24). Hence, even if the agents observe the same data, their predictors $f(\tilde{\varphi}_t)$ will be different if they follow different aims.

2. The recipient (jth agent) complements the obtained $f_i(\tilde{\varphi}_t)$ by *its own* DM strategy

$$\overline{f}_i(\varphi_t) = f(u_{2,t}|y_{t-1}, y_{t-2})\, f_i(\tilde{\varphi}_t),$$

   which can be substituted into (15).

Under the adopted approximation, the result of (15) for the jth agent is again in the Normal-inverse-gamma form (20) with statistics:

$$\tilde{V}_{j,t} = V_{j,t} + \beta \overline{V}_{i,t}, \quad \tilde{\nu}_{j,t} = \nu_{j,t} + \beta.$$

Here, $\overline{V}_{i,t}$ denotes matrix of expected values

$$\overline{V}_{i,t} = \mathsf{E}_{\overline{f}_i(\varphi_t)}[y_t, y_{t-1}, y_{t-2}, u_t]'\,[y_t, y_{t-1}, y_{t-2}, u_t],$$

which is easy to evaluate, since $\overline{f}_i(\varphi_t)$ is a Gaussian pdf.

4. SIMULATION EXPERIMENTS

The example from Section was studied in simulation. The environment parameters were chosen as follows: $a = 0.8$, $b = 0.2$, $c = 1$, $d = -1$ and $\sigma = 0.1$. The agents were initialized with flat non-informative prior knowledge on the model, and prior aims ${}^I\hat{y}_{1,t} = 20$, ${}^I\hat{y}_{2,t} = 10$, and ${}^I\sigma_{1,t} = {}^I\sigma_{2,t} = 1$. These parameters were chosen intentionally to yield conflict when both agents act in autonomous mode. The devices are chosen to have the same power, which is restricted by ideal pdf on input signal (24) with assignments ${}^I\hat{u}_{1,t} = {}^I\hat{u}_{2,t} = 0$, and ${}^I\hat{\rho}_{1,t} = {}^I\hat{\rho}_{2,t} = 1$.

In the first experiment, the first 40 steps of the operation are used as training period when the agents do not apply their DM strategy, but generate their actions randomly. The agents start to apply their strategy at times $t = 50$ and $t = 80$. The resulting temperatures, $y^{1:t}$, and control

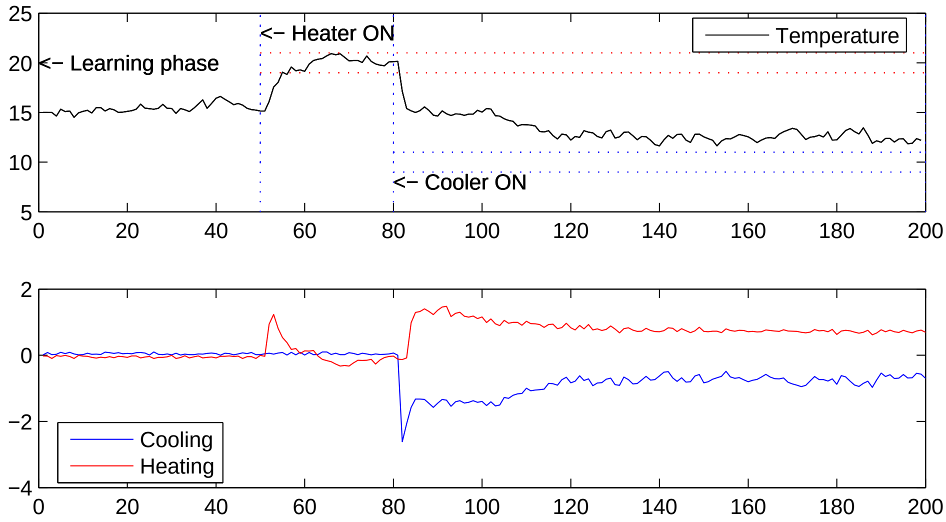

Figure 2: Simulation results for two agents with conflicting aims.

actions, $u_1^{1:t}$ and $u_2^{1:t}$, are displayed in Figure 2. Note that since the devices have the same power, shortly after the second agent starts applying its strategy, both agents are running at (almost) full power, but none of them is able to reach its aim. This conflicting situation results in waste of energy.

The second experiment was designed to remedy this situation by merging of agents aims at time $t = 90$. The agents fully cooperate, i.e. they both of them choose the negotiation weight $\alpha = 0.5$. Two merging techniques were studied: linear and geometric combination of aims. Results are displayed in Figures 3 and 4 respectively. Note that in this case, after $t = 90$ the power generated by both agents drops significantly, and they both cooperate in order to maintain the negotiated temperature which is $15°$ C. The difference between linear and geometric combination is in the variance of the aim. Linear combination yields wider margins on the ideal distribution and thus the temperature varies within this range. Geometric combination yields tighter margins and the agents must apply more aggressive control to keep the temperature within the limits, Figure 4.

Even with synchronised aims, the agents may still be in conflict since they may predict different behaviour of the observed system. Therefore, the third experiment was designed to merge agents models at $t = 140$, see Figure 5. This operation has less dramatic effect, however, it also results in decrease of generated power while preserving the ability to control the temperature.

5. CONCLUSION

The theory of distributed Bayesian decision-making was applied to the task of control of an ARX model with two inputs using two autonomous agents. It was shown that in this case, the proposed techniques of probabilistic merging are able to resolve possible conflicts of the two agents. This result is encouraging for further development of the methodology.

Acknowledgement: This work was supported by grants MŠMT 1M0572, AVČR 1ET 100 750 401.

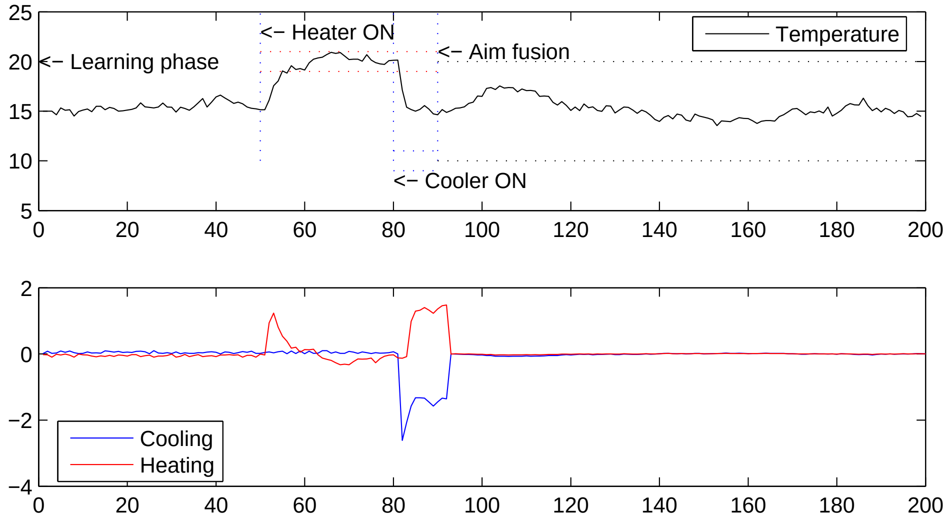

Figure 3: Conflict resolution using linear combination of aims.

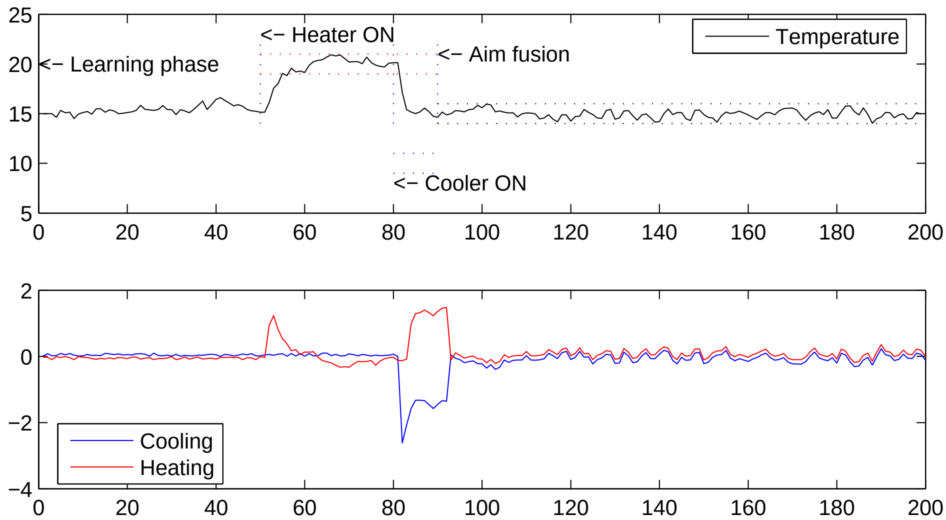

Figure 4: Conflict resolution using geometric combination of aims.

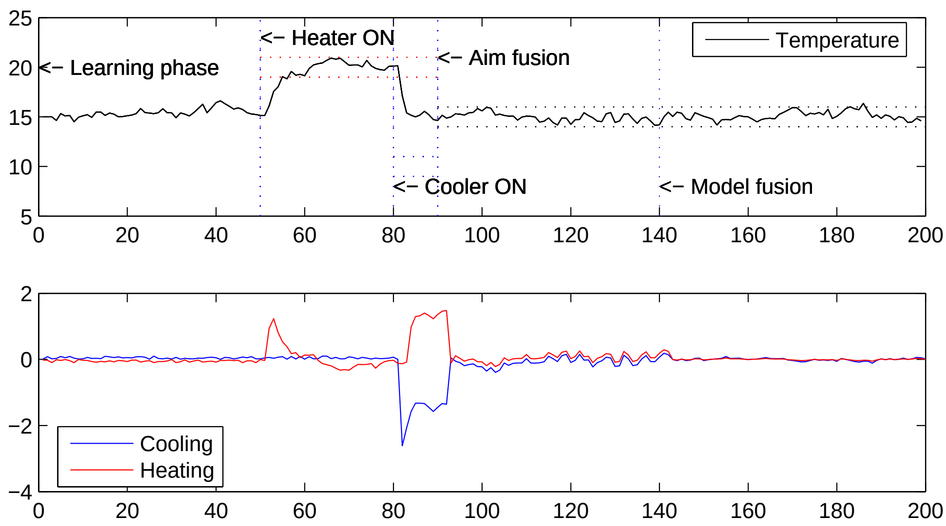

Figure 5: Conflict resolution using geometric combination of aims and indirect merging of models.

REFERENCES

Astrom, K.J. (1970). *Introduction to Stochastic Control*. Academic Press. New York.

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

Bertsekas, D.P. (2001). *Dynamic Programming and Optimal Control*. Athena Scientific. Nashua, US. 2nd edition.

Chen, Z. (2003). Bayesian filtering: From Kalman filters to particle filters, and beyond. Technical report. Adaptive Syst. Lab., McMaster University, Hamilton, ON, Canada.

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

Kárný, M. and T.V. Guy (2004). On dynamic decision-making scenarios with multiple participants. In: *Multiple Participant Decision Making* (J. Andrýsek, M. Kárný and J. Kracík, Eds.). Advanced Knowledge International. Adelaide. pp. 17–28.

Kárný, M., J. Böhm, T. V. Guy, L. Jirsa, I. Nagy, P. Nedoma and L. Tesař (2005). *Optimized Bayesian Dynamic Advising: Theory and Algorithms*. Springer. London.

Kracík, J. (2004). Composition of Probability Density Functions - Optimizing Approach. Technical Report 2099. ÚTIA AV ČR. Praha.

Kracík, J. (2005). Processing of expert information in bayesian parameter estimation. In: *Proceedings of the 6th International PhD Workshop on Systems and Control - a Young Generation Viewpoint*. Josef Stefan Institute. Ljubljana. p. 4. accepted.

Kracík, J. and M. Kárný (2005). Merging of data knowledge in Bayesian estimation. In: *Proceedings of the Second International Conference on Informatics in Control, Automation and Robotics* (J. Filipe, J. A. Cetto and J. L. Ferrier, Eds.). INSTICC. Barcelona. pp. 229–232.

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

Weiss, Gerhard, Ed. (2000). *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*. The MIT Press.