BAYESIAN APPROACH TO MULTI-AGENT SYSTEMS

Václav Šmídl

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

Abstract: Multi-agent systems are increasingly popular approach to control of complex industrial processes. The idea of distribution of a complex task into many semi-autonomous cooperating units has been formalized using many frameworks. In this paper, we review the close relation of distributed Bayesian decision making and multi-agent systems. The Bayesian methodology was primarily designed for systems with uncertainty. Therefore, a distinctive feature of a Bayesian agent is that all information is represented by probability density functions. Many algorithms derived for a single Bayesian decision maker are suitable for use in multiagent scenarios, however, more work is required to resolve issues related to Bayesian approach to communication and cooperation. The challenges for future research will be outlined. It is concluded that the Bayesian paradigm provides a solid, consistent framework for formalization of the task.

Keywords: Bayesian decision making, Multi-agent systems, information fusion

1. INTRODUCTION

In recent years, it becomes obvious that the traditional centralized approach to control of large systems has reached its limits. Decentralization of control and decision making is seen as future direction of research in both academia (Haimes and Li, 1988) and industry (Hall *et al.*, 2005). Many successful applications of so called holonic or multi-agent systems has been published. This paradigm presents a new challenge for designers of these systems, since the traditional methodologies of design became obsolete and no consistent replacement is available (Hall *et al.*, 2005). One possible solution of this problem is to extend the existing methodologies to accommodate the distributed setup.

In control applications, we can see an agent as an entity consisting of two principal parts: (i) *autonomous subsystem*, which is responsible for agents ability to act according to its own aims and needs, and (ii) *communication and negotiation subsystem*, which is responsible for exchanging its knowledge with other agents and adjustment of its aims in order to cooperate and thus achieve better overall performance. The autonomous subsystem can be seen as a controller in the traditional sense, hence a number of methodologies for its design is readily available (Tamer, 2001). From this range of theories, we seek a methodology which is able to embrace not only the autonomous but also the communication and negotiation subsystem. The most promising candidate is the Bayesian theory of decision making, since (i) it is a consistent theory for dealing with uncertainty which is ever present in real environments (Berger, 1985), (ii) the task of agent communication and negotiation can be formalized as decision making problem, and (iii) it is successfully applied in controller design and in design of advanced applications such as advising systems (Kárný *et al.*, 2005).

Traditionally, the decision-maker is assumed to be the only entity that intentionally influences the environment. It consists of a model of its environment, its individual aims, and a pre-determined strategy of decision making. On the contrary, an agent in multi-agent systems is known to influence only a part of the environment, i.e. its *neighbourhood*, while the rest of the environment is modelled by other agents. In order to obtain relevant information from distant parts of the environment, an agent relies on communication with other agents in its neighbourhood. If the agents are able to exchange their aims and take them into account, they can cooperate and improve the overall performance of the system. The challenge for Bayesian decision making theory is to formalize communication and negotiation as operations on probability distributions. It was shown that the technique of *fully probabilistic design (FPD)* (Kárný, 1996) reduces the task of agent cooperation into *reporting and merging of probability density functions* (Andrýsek *et al.*, 2004).

2. 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 decision making can be decomposed into the following principal sub-tasks: (i) model parametrization, (ii) learning of model parameters, and (iii) design of the control strategy. These tasks will be now described in detail.

2.1 Model Parametrization

In this text, d_t denotes all observable quantities on the environment, i.e. data, y_t, and 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 decision-maker—is described by the following probability density function:

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

Here, $f\left(\cdot\right)$ denotes probability density function (pdf) of its argument. $d\left(t\right)$ denotes the observation history $d\left(t\right) = [d_1, \ldots, d_t]$. The model represents the whole trajectory of the system, including inputs u_τ which can be influenced by the decision-maker. The chosen order of conditioning distinguishes the following important pdfs; (i) observation model, $f\left(y_t | u_t, d(t-1), \Theta_t\right)$, (ii) internal model, $f\left(\Theta_t | u_t, d(t-1), \Theta_{t-1}\right)$, and (iii) DM strategy, $f\left(u_t | d\left(t-1\right)\right)$. The first two models are considered to be known, while the DM strategy is to be found.

2.2 Learning via Bayesian filtering

The task of learning is to infer posterior distribution of unknown parameters from the observed data, $f(\Theta_t|d(t))$. This pdf can be computed recursively as follows:

$$
\begin{aligned}
f(\Theta_t|u_t, d(t-1)) &= \int f(\Theta_t|u_t, d(t-1), \Theta_{t-1}) f(\Theta_{t-1}|d(t-1))\, d\Theta_{t-1}, &\quad (2)\\
f(\Theta_t|d(t)) &\propto f(y_t|u_t, d(t-1), \Theta_t) f(\Theta_t|u_t, d(t-1)), &\quad (3)
\end{aligned}
$$

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 operation (2)–(3) 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 (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) distinguished by superscript $^{\lfloor I}$:

$$
f(d(t), \Theta(t)) \rightarrow {}^{\lfloor I}f(d(t), \Theta(t)). \quad (4)
$$

Similarly to (1), the ideal distribution is decomposed into ideal observation model, internal model, and ideal DM strategy. The loss function of the decision making has the form of Kullback-Leibler divergence between the model and the ideal. This has the following consequences: (i) the loss function is very well interpretable and it can be simply tailored to practical problems, and (ii) minimum of the KL divergence—i.e. the optimal DM strategy—is found in *closed form*.

3. BAYESIAN DECISION-MAKER

In practise, the task of adaptive decision making is typically solved in two stages (Kárný *et al.*, 2005): (i) off-line, and (ii) on-line. The off-line stage is dedicated to design of the structure and fixed parameters (such as initial conditions) of the decision-maker. When the structure and fixed parameters are determined, the decision-maker operates autonomously in on-line mode, where it is able to adapt (by adjusting model parameters) to changes in the environment and improve its DM-strategy. Operation needed in both stages are described in this Section.

3.1 Off-line stage

In this stage, it is necessary to determine structure of the model (1) and prior distribution of model parameters. If there is no physically justified model of the environment, the technique of *model selection* test many possible parametrization of the model, and selects one, which is best suited for the observed data. The technique of *elicitation of prior* distributions converts the expert knowledge which is not available in the form of pdfs into probabilistic terms. When the model and ideal distributions are chosen, the optimal DM strategy is computed using FPD (Section 2).

All these tasks are computationally demanding and thus they are traditionally solved off-line, i.e. only once for all available data. This is acceptable, since all expert information is available a priori, and model of the environment is assumed to be constant.

3.2 On-line stage

A typical adaptive decision-maker operates by recursive repetition of the following steps:

read: the observed data are read from the environment and pre-processed.

learn: the observed data are used to increase the knowledge about the environment.

adapt: the decision-maker use the improved knowledge of the system to improve its DM strategy.

decide: the adapted DM strategy is used to choose an appropriate action.

write: the chosen action is written into the environment.

Note that due to computational constraints, all operations in this stage are defined on finite dimensional parameters or statistics.

3.3 Merging of pdfs

For the task of prior elicitation, we need to define a new probabilistic operation for merging of information from many sources. The merging operation is defined as a mapping of two pdfs into one:

$$f_1\left(\Theta_t|d\left(t\right)\right), f_2\left(\Theta_t|d\left(t\right)\right) \stackrel{\text{merge}}{\longrightarrow} \tilde{f}\left(\Theta_t|d\left(t\right)\right), \tag{5}$$

where f_1 and f_2 are the *source pdfs*, and the $\tilde{f}$ is the *merged pdf*. Many approaches are available, e.g. (Jiroušek, 2003), with different assumptions and properties. However, in the sequel, we rely on the results of (Kracík, 2004) since these have the following properties: (i) defined as optimization problems, with a reasonable loss function, (ii) their results are intuitively appealing and well interpretable, (iii) the optimum is reached for a class of pdfs which is uniquely defined, (iv) is applicable to both discrete and continuous distributions, and (v) algorithmic solutions are available.

We distinguish two kinds of merging: (i) direct, when the source and the merged pdfs are defined on the same variables, and (ii) indirect, when the source distributions are defined on the variable in condition of the merged pdf.

4. BAYESIAN AGENTS

The Bayesian agent is an extended Bayesian decision-maker described in previous Section. The additional features are the ability and need of agents to communicate and cooperate. In the Bayesian framework, all knowledge is stored in pdfs. The challenge is to formalize communication and cooperation within the framework of probability calculus. In this Section, we propose a simple probabilistic model of negotiation. For clarity of explanation, we consider only two agents, $A_{[1]}$ and $A_{[2]}$, where agent number is always in subscript in square brackets.

Each agent has the following quantities:

Observed data d_t: Naturally, each agent can observe different subset of variables, i.e. $d_{t,[1]}$ and $d_{t,[2]}$, for A_1 and A_2, respectively.

Internal quantities Θ_t: We do not impose any structure of the environment model for the agents, hence, internal quantities $\Theta_{t,[1]}$ and $\Theta_{t,[2]}$ are in general disjoint sets.

Environment Model: e.g. $f_{[1]} = f\left(d_{[1]}(t), \Theta_{[1]}(t)\right)$ for agent $A_{[1]}$.

Ideal distributions: e.g. ${}^{\lfloor I}f_{[1]} = {}^{\lfloor I}f\left(d_{[1]}(t), \Theta_{[1]}(t)\right)$ for agent $A_{[1]}$.

Negotiation weights: For the purpose of negotiation, we define a scalar quantity $\alpha_{2,[1]} \in \langle 0, 1\rangle$ denoting the level of belief of agent A_1 in information received from A_2. Analogically, $\alpha_{1,[2]}$ is defined in A_2.

4.1 Communication

The agents can communicate two kinds of information: (i) about the environment, and (ii) about their individual aims. The easiest way how to exchange the information about the environment is to exchange the observed data. The observed data can be seen as a special case of pdf, namely empirical pdf $f\left(d_{[2]}(t)\right)$. Then, the task is formally identical with the task of indirect merging of pdf, Section 3.3. The observed data from A_2 are merged with the existing model of A_1 using (5). When the operation is finished, the merged pdf $\tilde{f}_{[1]}$ is then used as the new model of the environment. The ideal distributions can be communicated and merged in the same way, using direct merging, Section 3.3.

4.2 On-line algorithm of Bayesian agents

On-line operation of each Bayesian agent is an extension of the on-line steps of Bayesian decision-maker (Section 3).

read: the observed data are read from the system (environment). Possible communication (via pdf) from the neighbour is also received in this step.

learn: the observed data are used to increase the knowledge about the system (environment).

merge: if the communication from the neighbour contains information about the environment, the merge operation is called in order to merge it with the current knowledge.

adapt: the decision-maker use the improved knowledge of the environment to improve its DM strategy.

decide: the adapted DM strategy is used to choose an appropriate action. In multi-agent scenario, the tasks of communication and negotiation are also part of the decision making process and are done in this step.

write: the chosen action is written into the system (environment). If the decision to communicate was made, a message to the neighbour is also written in this step.

Note that acquisition of the observed data is synchronized with communication. In each time step, only one message from the neighbour is received, processed and answered. This allows seamless merging of knowledge from direct observations and from communication. If the periods of data sampling and communication differ, the smaller one is chosen as the period of one step of an agent.

5. CONCLUSION

We have presented a Bayesian framework for extension of decision-makers for cooperation and cooperation. We have shown that this aim can be achieved using available theory and methods. Detailed implementation of this approach and analysis of its properties are subject of ongoing research.

Acknowledgements

This work was supported by grants MŠMT 1M0572 (DAR), AVČR 1ET 100 750 401 (BADDYR), and MŠMT 8-2006-06.

REFERENCES

Andrýsek, J., Kárný, M. and Kracík, J., Eds. (2004). *Multiple Participant Decision Making*. Advanced Knowledge International. Adelaide.

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.

Haimes, Y.Y. and D. Li (1988). Hierarchical multiobjective analysis for large scale systems: Review and current status. *Automatica* **24**(1), 53–69.

Hall, K. H., R. J. Staron and P. Vrba (2005). Holonic and agent-based control. In: *Proceedings of the 16th IFAC Congress*.

Jiroušek, R. (2003). On experimental system for multidimensional model development MUDIN. *Neural Network World* (5), 513–520.

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

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). On composition of probability density functions. In: *Multiple Participant Decision Making* (J. Andrýsek, M. Kárný and J. Kracík, Eds.). Vol. 9 of *International Series on Advanced Intelligence*. pp. 113–121. Advanced Knowledge International. Adelaide, Australia.

Tamer, B., Ed. (2001). *Control Theory*. IEEE Press. New York.