Tools for Communication of Bayesian Agents

Vaclav Šmídl, Jan Kracík

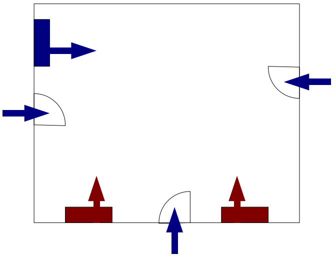
Institute of Information Theory and Automation, Czech Academy of Sciences, Prague, Czech Republic

12th October 2005

Outline

Example: temperature control

Fictious room:



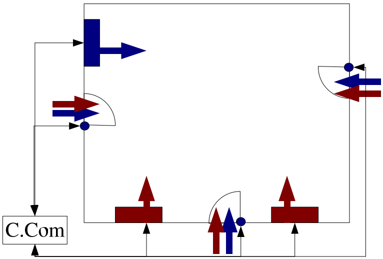
Task:
control the room temperature

reliably: failures,

adaptively: changes in the enviroment

Example: temperature control

Fictious room:

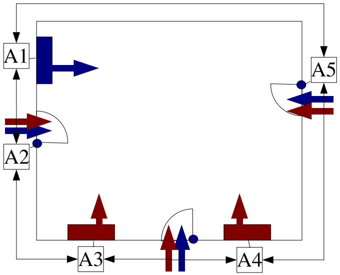


Centralized control:

- optimization
- many possible scenarios
- poor scalability
- error sensitive
- poor reconfiguration

Example: temperature control

Fictious room:



Agent control:

- scalable: agents can be added
- simple: few rules
- cheap: agents in devices
- expensive: communication

Industrial standard: Rockwell automation

Centralized vs. Decentralized Control

Centralized approach:

- Has a *consistent* theory of decision-making under uncertainty (Bayeisan theory),
- Faces the “curse of dimensionality”, solution for complex problems is prohibitive,
- Re-design is not flexible enough and requires a lot of manpower,

Distributed Approach (Multi-agent):

- Complex problem is decomposed into local areas which are governed by autonomous *agents,*
- The agents communicate to each other to achieve overall coordination,
- It is difficult to assess the overall behaviour of the MAS, (game theory),

Centralized vs. Decentralized Control

Centralized approach:

- Has a *consistent* theory of decision-making under uncertainty (Bayeisan theory),
- Faces the “curse of dimensionality”, solution for complex problems is prohibitive,
- Re-design is not flexible enough and requires a lot of manpower,

Distributed Approach (Multi-agent):

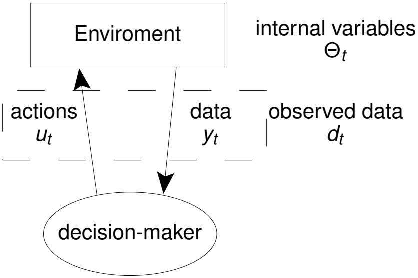
- Complex problem is decomposed into local areas which are governed by autonomous *agents,*
- The agents communicate to each other to achieve overall coordination,
- It is difficult to assess the overall behaviour of the MAS, (game theory),

Proposal: take best of those worlds.

Multiple Participant Decision Making = Bayesian Agents

Adaptive Bayesian Decision-Maker (controller)

Solid *consistent* theory of making decisions under **uncertainty**.



Decision-maker is using probability calculus:

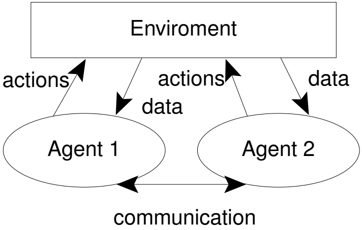
Model: $f(d(t), \Theta(t))$, relation of data and parameters.

Aim: ${}^{\lfloor I}f(d(t), \Theta(t))$, ideal distribution,

Decision: $f(u_t | d(t)) \rightarrow u_t$, optimal decisions.

How to make a Bayesian Agent?

Making Baeysian decision-maker aware of each other

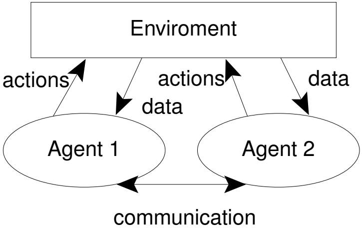


Communication exchange of information $\Rightarrow$ better learning,

Cooperation exchange of aims (pdfs) $\Rightarrow$ avoiding conflicts.

How to make a Bayesian Agent?

Making Baeysian decision-maker aware of each other



Communication exchange of information $\Rightarrow$ better learning,

Cooperation exchange of aims (pdfs) $\Rightarrow$ avoiding conflicts.

The task:
Formalization in terms of ***probability calculus*** and algorithmic solution.

Key technologies: FPD

Fully probabilistic design:

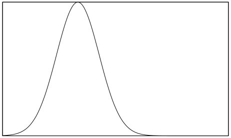
the aim of decision making is formalized in the form of ideal distribution,

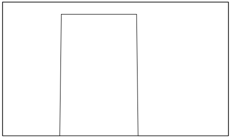
$$^{\lfloor I}f\left(d\left(t\right),\Theta\left(t\right)\right).$$

- the loss function of divergence between the ideal and the true pdf.
  Advantage: **no need to exchange loss functions!**
- Optimal strategy is known: $f\left(u_t|d(t)\right)=\int\int\int\ldots$
- Allows for multi-criteria decision-making
- Solvable for Marcov chains and Gaussian pdf, otherwise approximations.

Key technologies: Merging

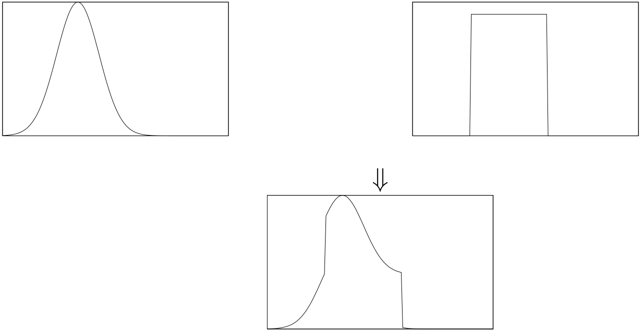
Merging of probability distributions: (information fusion)





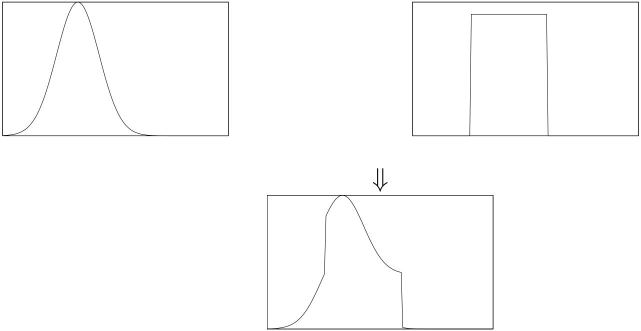
Key technologies: Merging

Merging of probability distributions: (information fusion)



Key technologies: Merging

Merging of probability distributions: (information fusion)

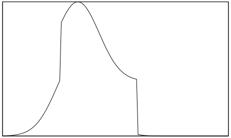


- various types of pdfs (Gauss, discrete, etc.),
- on different variables, of different type (marginalized, conditioned)

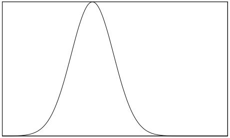
Key technologies: projection

Projection:

- finding 'nicer' distribution, loosing as little information as possible



$\longrightarrow$



Example: Negotiation of temperature

Classical agents:

A1 (cooling): goal 15 °C

A2 (heating): goal 20 °C

scenarios:

1. non-cooperating agents:
   18 °C, both are working on full steam,
2. fully cooperating agents:
   18 °C, lower energy load.

Negotiation: mostly ad hoc methods

Example: Negotiation of temperature

Classical agents:

A1 (cooling): goal 15 °C

A2 (heating): goal 20 °C

scenarios:

1. non-cooperating agents: 18 °C, both are working on full steam,
2. fully cooperating agents: 18 °C, lower energy load.

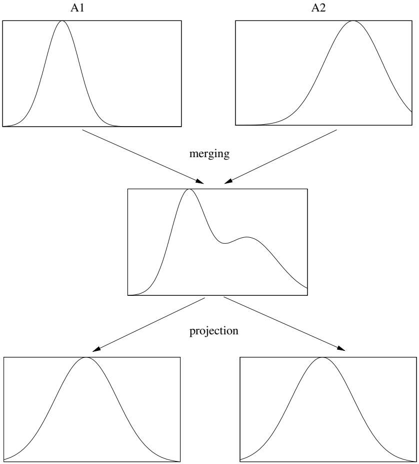
Negotiation: mostly ad hoc methods

Bayesian agents:

A1 (cooling): $^{\lfloor I}f(T) = \mathcal{N}(15, 2)$

A2 (heating): $^{\lfloor I}f(T) = \mathcal{N}(20, 6)$

Example: Negotiation of temperature



Bayesian agents:

A1 (cooling): $^{\llcorner I}f(T) = \mathcal{N}(15, 2)$

A2 (heating): $^{\llcorner I}f(T) = \mathcal{N}(20, 6)$

scenarios:

1. non-cooperating agents: same
2. fully cooperating agents:

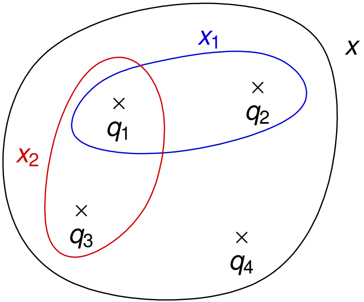
$$^{\llcorner I}f(T) = \mathcal{N}(17, 7),$$

result of ***optimization***.

Negotiation: faster convergence, lower communication load.

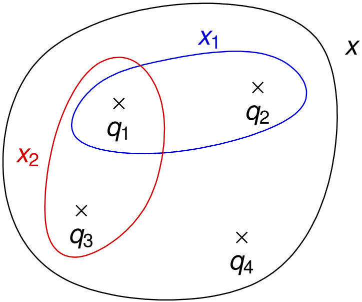
Problem Formulation

- vector random quantity $x = (q_1, \ldots, q_N)$
- n agents, ideal pdfs $f_p(x_p)$
- x_p – random vectors, entries from $\{q_1, \ldots, q_N\}$
- weights $\alpha_p > 0$, $\sum_p \alpha_p = 1$



Problem Formulation

- vector random quantity $x = (q_1, \dots, q_N)$
- n agents, ideal pdfs $f_p(x_p)$
- x_p – random vectors, entries from $\{q_1, \dots, q_N\}$
- weights $\alpha_p > 0$, $\sum_p \alpha_p = 1$



Common ideal pdf $f(x)$?

- How to define $f(x)$?
- How to find it?
- Practical issues

Solution

Common ideal pdf

$$f(x) \in \arg\min_{\tilde{f}} \sum_p \alpha_p D(f_p(x_p)||\tilde{f}(x_p))$$

$D(\cdot||\cdot)$ - Kullback-Leibler divergence

Solution

Common ideal pdf

$$f(x) \in \arg\min_{\tilde{f}} \sum_p \alpha_p D(f_p(x_p)||\tilde{f}(x_p))$$

$$f(x) = \sum_p \alpha_p \frac{f(x)}{f(x_p)} f_p(x_p)$$

$D(\cdot||\cdot)$ - Kullback-Leibler divergence

Approximation of Common Ideal Pdf

$$\mathcal{D}(h) = \sum_p \alpha_p D(f_p(x_p)||h(x_p))$$

$$A(h) = \sum_p \alpha_p \frac{f(x)}{f(x_p)} f_p(x_p)$$

Approximation of Common Ideal Pdf

$$\mathcal{D}(h) = \sum_p \alpha_p D(f_p(x_p)||h(x_p))$$

$$A(h) = \sum_p \alpha_p \frac{f(x)}{f(x_p)} f_p(x_p)$$

$$\begin{aligned} \mathcal{D}(h) &\geq \mathcal{D}(Ah) \ \forall h \\ \mathcal{D}(h) &= \mathcal{D}(Ah) \ \text{iff } h \text{ is optimal} \\ \mathcal{D}(A^k h) &\to \mathcal{D}(f) \end{aligned}$$

Practical Issues

- discrete quantities
  - directly usable
  - marginalization computationally expensive

Practical Issues

- discrete quantities
  - directly usable
  - marginalization computationally expensive
- continuous quantities
  - approximations not in any reasonable class!
  - find optimal pdf f in a predefined class $\mathcal{F}$
  - we have an algorithm for $\mathcal{F}$ being class of mixtures

Conclusions

The proposed method

- fulfills our requirements on ideal pdf fusion
  - independence on the ordering of sources
  - feasible for both discrete and continuous quantities
- fits well into other technologies in our framework
- will be implemented in Matlab toolbox MIXTOOLS 3000