Decision Theoretic Approaches for Focussed Bayesian Fusion

Jennifer Sander[1], Jürgen Beyerer[1,2]

[1] Lehrstuhl für Interaktive Echtzeitsysteme, Institut für Anthropomatik (IFA), Karlsruher Institut für Technologie (KIT), Adenauerring 4, 76131 Karlsruhe
[2]Fraunhofer-Institut für Optronik, Systemtechnik und Bildauswertung IOSB, Fraunhoferstraße 1, 76131 Karlsruhe

jennifer.sander@kit.edu, juergen.beyerer@iosb.fraunhofer.de

Abstract: Focussed Bayesian fusion is a local Bayesian fusion technique by that high costs caused by Bayesian fusion can get circumvented. This publication addresses globally optimal decision making on the basis of a focussed Bayesian model. Therefore, common decision criteria under linear partial information and in particular principles of lazy decision making are applied. We also present an interval scheme for global posterior probabilities whose informativeness is notably high.

1 Introduction

Bayesian theory delivers a powerful methodology for the fusion of homogeneous and heterogeneous information sources [BHSG08]. By the Bayesian fusion methodology, a lot of problems within the context of information fusion can be solved [Koc10]. To reduce the computational costs of Bayesian fusion, we developed local Bayesian fusion approaches, at which the actual fusion task gets concentrated on a local context $\boldsymbol{U}$. See for example [BHS06, SB08]. A local context is an adequately chosen subset of the space $\boldsymbol{Z}$ that is spanned by the range of the Properties of Interest (PoI). Ignoring $\boldsymbol{Z} \setminus \boldsymbol{U}$ completely delivers a straightforward local Bayesian fusion scheme, which we termed focussed Bayesian fusion. See for example [SHGB09, SHGB10].

On the basis of the resulting focussed posterior distribution[1], it is not possible to reconstruct the global posterior distribution, which would result if Bayesian fusion was performed with respect to whole $\boldsymbol{Z}$ in a unique manner. However, combining facts about the connection between the focussed and the global posterior distribution with facts about construction rules that lead to an adequately chosen local context $\boldsymbol{U}$, the unknown global posterior distribution can be bounded from both above and below within $\boldsymbol{U}$ [San09, SHGB10]. Due to additional constraints, the informativeness of the resulting interval scheme for

[1]According to the nature of the involved quantities, the term distribution has a mixed meaning as discrete probability function and probability density.

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4
weitere Artikel online:
http://informatik2011.de/519.html

global posterior probabilities is notably high. Statements with respect to global posterior probabilities that are surely valid may be obtained by its analysis.

At Bayesian fusion also subsequent decisions could be aimed [BS04, BHS06]. Thereby, a rational decision maker should maximize the expected utility with respect to the global posterior distribution. We trace the problem of making globally optimal decisions on the basis of a focussed Bayesian model back to a decision problem under linear partial information (LPI) [KM76], which is manageable effectively. We adapt common decision criteria under LPI and in particular principles of lazy decision making [Pre02] for focussed Bayesian fusion. If it is not possible to identify globally optimal decisions due to the focussing, there are different ways to improve the basis for decision making. Such an improvement corresponds to a shrinking of the intervals for global posterior probabilities.

The next section of this publication constitutes a short introduction into both, Bayesian fusion and focussed Bayesian fusion. Section 3 addresses interval schemes for the global posterior distribution and for global posterior probabilities. Section 4 deals with the use of the theory of LPI and the principles of lazy decision making for globally optimal decision making in the context of focussed Bayesian fusion. After providing illustrative examples in section 5, we finish with a short conclusion.

The intervals for global posterior probabilities represent the uncertainty that results additionally from the focussing explicitly in a non-probabilistic manner. Provided that such a distinction of uncertainties is not used effectively, facts and corresponding uncertainties are represented in an adequate manner by probability in the sense of the Degree of Belief (DoB) interpretation [BS04]. Hence, if the aim of the fusion task is to obtain a comprehensive representation of the posterior state of knowledge, a reduction of the intervals to a unique posterior distribution should be done. The choice of such a unique global posterior distribution is also a decision [BS04]. For details with regard to the choice of a unique global posterior distribution, the reader is referred to the technical report [San10], which is the basis for the current paper. There, the maximum entropy principle is applied and analyzed with regard to this task.

2 Introduction to (Focussed) Bayesian Fusion

In the following, $x_s \in X_s$ denotes the information contribution of source number s, $s \in \{1, \dots, S\}$, $S \in \mathbb{N}$, and $\boldsymbol{x} := (x_1, \dots, x_S)$ embodies all information from the information sources. Let $\boldsymbol{z} \in \boldsymbol{Z}$ denote the PoI, which specify the desired information. As usual in Bayesian fusion, it is assumed that $\boldsymbol{z}$ adopts a "true" value which is not directly observable. Prior knowledge and $\boldsymbol{x}$ are used to infer about $\boldsymbol{z}$.

In the Bayesian theory, all quantities are assumed to be random and all available information is represented probabilistically in the sense of the DoB interpretation. The prior distribution $p(\boldsymbol{z})$ and the Likelihood function $l(\boldsymbol{x}|\boldsymbol{z})$ are combined via the Bayesian theorem to the posterior distribution $p(\boldsymbol{z}|\boldsymbol{x}) \propto l(\boldsymbol{x}|\boldsymbol{z})\, p(\boldsymbol{z})$. Since the computational complexity for the necessary operations to obtain this quantity is $\mathrm{O}(|\boldsymbol{Z}|)$, Bayesian fusion is prohibitive in many real world tasks [BHS06].

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4
weitere Artikel online:
http://informatik2011.de/519.html

At focussed Bayesian fusion, the actual Bayesian fusion is performed only with respect to the local context $\boldsymbol{U} \subset \boldsymbol{Z}$ which lowers the computational complexity to $\mathrm{O}(|\boldsymbol{U}|)$. As explained in [SHGB09], the corresponding fusion scheme is

$$p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U}) \propto l(\boldsymbol{x}|\boldsymbol{z})\, p(\boldsymbol{z}|\boldsymbol{U}) \ , \quad \boldsymbol{z} \in \boldsymbol{U} \ . \tag{1}$$

In the focussed Bayesian model, $\boldsymbol{Z} \setminus \boldsymbol{U}$ is ignored completely and the posterior probability of events[2] $\boldsymbol{E} \subseteq \boldsymbol{U}$ gets distorted according to the global posterior probability[3] $P(\boldsymbol{U}|\boldsymbol{x}) = \int_{\boldsymbol{U}} p(\boldsymbol{z}|\boldsymbol{x}) \,\mathrm{d}\boldsymbol{z}$ of the local context $\boldsymbol{U}$. More precisely, it holds

$$P(\boldsymbol{E}|\boldsymbol{x},\boldsymbol{U}) = \int_{\boldsymbol{E}} p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U}) \,\mathrm{d}\boldsymbol{z} \quad \text{with} \quad p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U}) = \frac{p(\boldsymbol{z}|\boldsymbol{x})}{P(\boldsymbol{U}|\boldsymbol{x})} \ , \quad \boldsymbol{z} \in \boldsymbol{U} \ . \tag{2}$$

A local context $\boldsymbol{U}$ is specified adequately if it contains at least these values of the PoI for which the standardized Likelihood function $l_{\mathrm{st}}(\boldsymbol{x}|\boldsymbol{z}) := \frac{l(\boldsymbol{x}|\boldsymbol{z})}{\max_{\boldsymbol{z}^*} l(\boldsymbol{x}|\boldsymbol{z}^*)}$ is larger than a suitable threshold. We addressed the rationale behind this specification, its extensions, and the threshold determination in previous publications, see [SB08, SHGB09, SKB10].

If the information contributions $x_1, \ldots, x_S$ are conditionally independent given $\boldsymbol{z}$, we have $l(\boldsymbol{x}|\boldsymbol{z}) = \prod_{s=1}^{S} l(x_s|\boldsymbol{z})$ and Bayesian fusion is realizable sequentially in an uncomplicated manner [BHS06]. In this case, for the determination of the local context $\boldsymbol{U}$, the information contributions can be also evaluated individually. Here, it is reasonable to chose $\boldsymbol{U}$ such that it contains at least these values of the PoI for that there exists at least one $s \in \{1, \ldots, S\}$ such that $l_{\mathrm{st}}(x_s|\boldsymbol{z})$ is larger than a suitable threshold. Compare also [SB08, SHGB09, SKB10].

3 Probability Intervals at Focussed Bayesian Fusion

Assume that $l(\boldsymbol{x}|\boldsymbol{z}) \leq \delta$ holds for all $\boldsymbol{z} \notin \boldsymbol{U}$ with $\delta := \tilde{\delta} \cdot \max_{\boldsymbol{z}^*} l(\boldsymbol{x}|\boldsymbol{z}^*)$ and a suitable threshold $\tilde{\delta} \in (0, 1)$. Then, the following lower bound for the global posterior probability $P(\boldsymbol{U}|\boldsymbol{x})$ of the local context $\boldsymbol{U}$ results [San09, SHGB10]:

$$P(\boldsymbol{U}|\boldsymbol{x}) \geq \frac{\int_{\boldsymbol{U}} l(\boldsymbol{x}|\boldsymbol{z})\, p(\boldsymbol{z}|\boldsymbol{U}) \,\mathrm{d}\boldsymbol{z}}{\int_{\boldsymbol{U}} l(\boldsymbol{x}|\boldsymbol{z})\, p(\boldsymbol{z}|\boldsymbol{U}) \,\mathrm{d}\boldsymbol{z} + (\frac{1}{P(\boldsymbol{U})} - 1)\, \delta} =: \beta \ . \tag{3}$$

β is computable in the focussed Bayesian model provided that the prior probability $P(\boldsymbol{U}) = \int_{\boldsymbol{U}} p(\boldsymbol{z}) \,\mathrm{d}\boldsymbol{z}$ of the local context $\boldsymbol{U}$ is ratable.

Assume that $x_1, \ldots, x_S$ are conditionally independent given $\boldsymbol{z}$ and assume that, for all $s \in \{1, \ldots, S\}$ and for all $\boldsymbol{z} \notin \boldsymbol{U}$, $l(x_s|\boldsymbol{z}) \leq \delta_s$ holds with $\delta_s = \tilde{\delta} \cdot \max_{\boldsymbol{z}^*} l(x_s|\boldsymbol{z}^*)$ and $\tilde{\delta} \in (0, 1)$. Then, (3) is valid with $l(\boldsymbol{x}|\boldsymbol{z}) = \prod_{s=1}^{S} l(x_s|\boldsymbol{z})$ and $\delta = \prod_{s=1}^{S} \delta_s$.

Combining (2) and (3), one obtains for $\boldsymbol{z} \in \boldsymbol{U}$ that it holds

$$p(\boldsymbol{z}|\boldsymbol{x}) \in [\mathrm{a}(\boldsymbol{z}), \mathrm{b}(\boldsymbol{z})] := [\beta\, p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U}), p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})] \ . \tag{4}$$

[2] Events are sets to that a probability is assigned.

[3] We use an integral notation is for both, the summation of discrete and the integration of continuous quantities. A summation sign is used only if the respective formula is to refer exclusively to the discrete case.

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4 weitere Artikel online:
http://informatik2011.de/519.html

Hence, an interval scheme for global posterior probabilities results:

$$P(\boldsymbol{E}|\boldsymbol{x}) \in [\mathrm{a}(\boldsymbol{E}), \mathrm{b}(\boldsymbol{E})] := \begin{cases} [\beta\, P(\boldsymbol{E}|\boldsymbol{x}, \boldsymbol{U}), P(\boldsymbol{E}|\boldsymbol{x}, \boldsymbol{U})]\ , & \boldsymbol{E} \subseteq \boldsymbol{U}\ , \\ [0, 1-\beta]\ , & \emptyset \neq \boldsymbol{E} \subseteq \boldsymbol{Z} \setminus \boldsymbol{U}\ . \end{cases} \tag{5}$$

Thereby, it is not possible that the posterior probabilities of events that are contained in the local context $\boldsymbol{U}$ vary arbitrarily within the corresponding intervals [SHGB10]. From (2), one obtains additionally for events $\boldsymbol{E}^*, \boldsymbol{E}^{**} \subseteq \boldsymbol{U}$ the connection

$$\frac{P(\boldsymbol{E}^*|\boldsymbol{x}, \boldsymbol{U})}{P(\boldsymbol{E}^{**}|\boldsymbol{x}, \boldsymbol{U})} = \frac{P(\boldsymbol{E}^*|\boldsymbol{x})}{P(\boldsymbol{E}^{**}|\boldsymbol{x})} =: o(\boldsymbol{E}^*, \boldsymbol{E}^{**})\ . \tag{6}$$

Let $\boldsymbol{E}^{**} \subseteq \boldsymbol{U}$ be arbitrary but fixed. If we assume that $P(\boldsymbol{E}^{**}|\boldsymbol{x})$ is equal to a certain value in the interval $[\mathrm{a}(\boldsymbol{E}^{**}), \mathrm{b}(\boldsymbol{E}^{**})]$, the posterior probabilities of all other events $\boldsymbol{E} \subseteq \boldsymbol{U}$ are uniquely determined because, according to (6), we have

$$P(\boldsymbol{E}|\boldsymbol{x}) = o(\boldsymbol{E}, \boldsymbol{E}^{**})\, P(\boldsymbol{E}^{**}|\boldsymbol{x})\ . \tag{7}$$

Thereby, $o(\boldsymbol{E}, \boldsymbol{E}^{**})$ is computable within the focussed Bayesian model.

4 A Framework Based on the Theory of LPI

4.1 Decisions under Risk, Partial Information, and LPI

Let $\boldsymbol{A}$ be a set of available actions. The utility of action $\boldsymbol{a} \in \boldsymbol{A}$ provided that $\boldsymbol{z} \in \boldsymbol{Z}$ is the "true" value of the PoI is denoted by $u(\boldsymbol{a}, \boldsymbol{z})$. If the global posterior distribution $p(\boldsymbol{z}|\boldsymbol{x})$ was known completely, the decision making was done under risk [Rüg99]. According to the principle of expected utility [BS04], a rational decision maker should chose an action $\boldsymbol{a}^* \in \boldsymbol{A}$ that maximizes the global posterior expected utility, i.e.,

$$\boldsymbol{a}^* = \arg\max_{\boldsymbol{a} \in \boldsymbol{A}} \mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{a}, \boldsymbol{z})] \quad \text{with} \quad \mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{a}, \boldsymbol{z})] = \int_{\boldsymbol{Z}} u(\boldsymbol{a}, \boldsymbol{z})\, p(\boldsymbol{z}|\boldsymbol{x})\ \mathrm{d}\boldsymbol{z}\ . \tag{8}$$

In this publication, we assume all used utility functions to be bounded and an act that is optimal in the sense of (8) to exist.

Global decision making on the basis of a focussed Bayesian model is decision making under partial information [Pre02]. Here, it is only known that $p(\boldsymbol{z}|\boldsymbol{x})$ is contained in the set $\mathcal{P}_{\mathrm{F}}$ of all probability distributions on $\boldsymbol{Z}$ that are consistent with the constraints (2) and (3). As consequence, a set of possible values for the global posterior expected utility of an action has to be considered explicitly at decision making–instead of one unique value as at decision making under risk.

In the following, $\boldsymbol{Z} \setminus \boldsymbol{U}$ is regarded as a (possibly large) finite set[4]: $\boldsymbol{Z} \setminus \boldsymbol{U} = \{\boldsymbol{z}^1, \ldots, \boldsymbol{z}^M\}$, $M \in \mathbb{N}$. Then, the global decision problem which has to be solved on the basis of a

[4]To our mind, this assumption is justifiable in many tasks. However, in [San10], also a generalization of the derived concept for the solution of arbitrary decision problems is sketched additionally.

ISBN 978-3-88579-286-4

focussed Bayesian model is traceable back to a decision problem under LPI. The following definition of LPI is based on [KM76]:

Definition 1. *Partial information about a probability distribution over a finite set of cardinality k is LPI if the respective subarea $\boldsymbol{W}$ of the k-dimensional probability simplex $[0,1]^k$ can be specified by a system of inequalities such that it holds*

$$\boldsymbol{W} = \left\{ \boldsymbol{p} = (p^1, \dots, p^k)^{\mathrm{T}} \in \mathbb{R}^k \,\middle|\, \sum_{i=1}^{k} p^i = 1, 0 \leq p^i \text{ for } i \leq k, \boldsymbol{B}\boldsymbol{p} \geq \boldsymbol{c} \right\} . \tag{9}$$

with $\boldsymbol{B} \in \mathbb{R}^{l \times k}$ and $\boldsymbol{c} \in \mathbb{R}^l$, $k, l \in \mathbb{N}$.

Lemma 1. *Assume that it is only known that $p(\boldsymbol{z}|\boldsymbol{x}) \in \mathcal{P}_{\mathrm{F}}$ holds. For each action $\boldsymbol{a} \in \boldsymbol{A}$, the set of possible values for the global posterior expected utility $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{a}, \boldsymbol{z})]$ of $\boldsymbol{a}$ is identical to the set of possible values for the expected utility of $\boldsymbol{a}$ in a decision problem under LPI over $\boldsymbol{Z}_{\mathrm{F}} := \{\boldsymbol{z}^1, \dots, \boldsymbol{z}^M, \boldsymbol{z}^{M+1}\}$ with $\boldsymbol{z}^{M+1} := \boldsymbol{U}$ if the respective subarea of $[0,1]^{M+1}$ is given by*

$$\boldsymbol{W}_{\mathrm{F}} := \left\{ \boldsymbol{p} = (p^1, \dots, p^{M+1})^{\mathrm{T}} \,\middle|\, \sum_{i=1}^{M+1} p^i = 1, 0 \leq p^i \text{ for } i \leq M, \beta \leq p^{M+1} \right\} \tag{10}$$

and if the respective utility function is

$$u_{\mathrm{F}}(\boldsymbol{a}, \boldsymbol{z}^i) := \begin{cases} u(\boldsymbol{a}, \boldsymbol{z}^i) \ , & i \in \{1, \dots, M\} \ , \\ \mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x}, \boldsymbol{U})}[u(\boldsymbol{a}, \boldsymbol{z})] \ , & i = M+1 \ . \end{cases} \tag{11}$$

Proof. We set

$$p^i := p(\boldsymbol{z}^i|\boldsymbol{x}) \ , i \in \{1, \dots, M\} \ , \quad \text{and} \quad p^{M+1} := P(\boldsymbol{U}|\boldsymbol{x}) \ . \tag{12}$$

$\boldsymbol{W}_{\mathrm{F}}$ contains exactly all global posterior distributions $p(\boldsymbol{z}|\boldsymbol{x})$ on $\boldsymbol{Z}$ that satisfy the condition (3). According to the structure of $\boldsymbol{W}_{\mathrm{F}}$, (3) specifies LPI about $\boldsymbol{p}$.
Because of (2), for each action $\boldsymbol{a} \in \boldsymbol{A}$, it holds:

$$\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{a}, \boldsymbol{z})] = \int_{\boldsymbol{Z} \setminus \boldsymbol{U}} u(\boldsymbol{a}, \boldsymbol{z})\, p(\boldsymbol{z}|\boldsymbol{x}) \; \mathrm{d}\boldsymbol{z} \; + P(\boldsymbol{U}|\boldsymbol{x})\, \mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x}, \boldsymbol{U})}[u(\boldsymbol{a}, \boldsymbol{z})] \ . \tag{13}$$

Because $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{a}, \boldsymbol{z})]$ is identical to

$$\mathrm{E}_{\boldsymbol{p}}[u_{\mathrm{F}}(\boldsymbol{a}, \boldsymbol{z}^i)] = \sum_{i=1}^{M} u(\boldsymbol{a}, \boldsymbol{z}^i)\, p^i + p^{M+1}\, \mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x}, \boldsymbol{U})}[u(\boldsymbol{a}, \boldsymbol{z})] \ , \tag{14}$$

the utility function $u_{\mathrm{F}}(\boldsymbol{a}, \boldsymbol{z}^i)$ on $\boldsymbol{A} \times \boldsymbol{Z}_{\mathrm{F}}$ that has been introduced in (11) is compatible to the original utility function $u(\boldsymbol{a}, \boldsymbol{z})$ on $\boldsymbol{A} \times \boldsymbol{Z}$. □

A set $\boldsymbol{W}$ of probability functions that is specified by LPI constitutes a convex polyhedron [Fis01, KM76]. The next known lemma connects this geometric consideration and linear optimization. For a polyhedron $\boldsymbol{W} \subseteq \mathbb{R}^k$, let $\boldsymbol{V}(\boldsymbol{W})$ denote the set of edges of $\boldsymbol{W}$.

ISBN 978-3-88579-286-4

Lemma 2 (see [NM04] or [Fis01])**.** *Let* $\mathrm{f} : \mathbb{R}^k \to \mathbb{R}$ *be a linear function and* $\boldsymbol{W} \subseteq \mathbb{R}^k$ *a convex polyhedron,* $k \in \mathbb{N}$*. Then, there exist points* $\boldsymbol{w}_{\min} \in \boldsymbol{V}(\boldsymbol{W})$ *and* $\boldsymbol{w}_{\max} \in \boldsymbol{V}(\boldsymbol{W})$ *such that* $\mathrm{f}(\boldsymbol{w}_{\min}) = \min_{\boldsymbol{w}\in\boldsymbol{W}} \mathrm{f}(\boldsymbol{w})$ *and* $\mathrm{f}(\boldsymbol{w}_{\max}) = \max_{\boldsymbol{w}\in\boldsymbol{W}} \mathrm{f}(\boldsymbol{w})$*.*

The next lemma shows that $\boldsymbol{V}(\boldsymbol{W}_{\mathrm{F}})$ has a convenient structure:

Lemma 3. $\boldsymbol{p}^{\mathrm{T}} = (p^1, \ldots, p^{M+1})^{\mathrm{T}} \in \mathbb{R}^{M+1}$ *is an edge of* $\boldsymbol{W}_{\mathrm{F}}$ *iff it holds*

$$p^i = \begin{cases} 0\,, & i \leq M\,, \\ 1\,, & i = M+1\,, \end{cases} \qquad \textit{or} \quad p^i = \begin{cases} 1-\beta\,, & \textit{for one } i_0 \in \{1, \ldots, M\}\,, \\ 0\,, & i \leq M \wedge i \neq i_0\,, \\ \beta\,, & i = M+1\,. \end{cases} \tag{15}$$

Proof. It is known [KM76, Pre02] that a point of $\boldsymbol{W}_{\mathrm{F}}$ is an edge iff $M+1$ of the $M+2$ conditions in (10) are satisfied as equations. Here, this means that the values of M components of $\boldsymbol{p}$ must be equal to the respective lower bounds. □

4.2 Decision Criteria under LPI

Because of lemma 1, criteria for global decision making on the basis of a focussed Bayesian model result if we consider the respective decision problem under LPI.

Theorem 1 (**Expected Utility Intervals**)**.** *If it is only known that* $p(\boldsymbol{z}|\boldsymbol{x}) \in \mathcal{P}_{\mathrm{F}}$ *holds, for each action* $\boldsymbol{a} \in \boldsymbol{A}$*, the set of possible values for the global posterior expected utility* $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{a}, \boldsymbol{z})]$ *is identical to the interval*

$$\mathbf{I}_{\boldsymbol{W}_{\mathrm{F}}}[u_{\mathrm{F}}(\boldsymbol{a}, \boldsymbol{z}^i)] := \left[\underline{\mathrm{E}}_{\boldsymbol{W}_{\mathrm{F}}}[u_{\mathrm{F}}(\boldsymbol{a}, \boldsymbol{z}^i)], \overline{\mathrm{E}}_{\boldsymbol{W}_{\mathrm{F}}}[u_{\mathrm{F}}(\boldsymbol{a}, \boldsymbol{z}^i)]\right] \tag{16}$$

whereby

$$\begin{aligned} \underline{\mathrm{E}}_{\boldsymbol{W}_{\mathrm{F}}}[u_{\mathrm{F}}(\boldsymbol{a}, \boldsymbol{z}^i)] = \min \Bigg\{ &(1-\beta) \min_{1\leq i\leq M} u(\boldsymbol{a}, \boldsymbol{z}^i) + \beta\, \mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a}, \boldsymbol{z})], \\ &\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a}, \boldsymbol{z})] \Bigg\}\,, \end{aligned} \tag{17}$$

$$\begin{aligned} \overline{\mathrm{E}}_{\boldsymbol{W}_{\mathrm{F}}}[u_{\mathrm{F}}(\boldsymbol{a}, \boldsymbol{z}^i)] = \max \Bigg\{ &(1-\beta) \max_{1\leq i\leq M} u(\boldsymbol{a}, \boldsymbol{z}^i) + \beta\, \mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a}, \boldsymbol{z})], \\ &\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a}, \boldsymbol{z})] \Bigg\}\,. \end{aligned} \tag{18}$$

Proof. $\mathrm{E}_{\boldsymbol{p}}[u_{\mathrm{F}}(\boldsymbol{a}, \boldsymbol{z}^i)]$ is a linear function with respect to $\boldsymbol{p}$, compare (14). Because $\boldsymbol{W}_{\mathrm{F}}$ is a convex polyhedron and lemma 2 holds, the set of values of the global posterior expected utility of $\boldsymbol{a}$ that results if $\boldsymbol{p}$ varies within $\boldsymbol{W}_{\mathrm{F}}$ is an interval and we have

$$\mathbf{I}_{\boldsymbol{W}_{\mathrm{F}}}[u_{\mathrm{F}}(\boldsymbol{a}, \boldsymbol{z}^i)] = \left[\min_{\boldsymbol{p}\in V(\boldsymbol{W}_{\mathrm{F}})} \mathrm{E}_{\boldsymbol{p}}[u_{\mathrm{F}}(\boldsymbol{a}, \boldsymbol{z}^i)], \max_{\boldsymbol{p}\in V(\boldsymbol{W}_{\mathrm{F}})} \mathrm{E}_{\boldsymbol{p}}[u_{\mathrm{F}}(\boldsymbol{a}, \boldsymbol{z}^i)]\right]\,. \tag{19}$$

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4
weitere Artikel online:
http://informatik2011.de/519.html

$\boldsymbol{V}(\boldsymbol{W}_{\mathrm{F}})$ has been identified in lemma 3 and from (14), we obtain

$$\left\{\mathrm{E}_{\boldsymbol{p}}[u_{\mathrm{F}}(\boldsymbol{a},\boldsymbol{z}^i)]\middle|\boldsymbol{p}\in\boldsymbol{V}(\boldsymbol{W}_{\mathrm{F}})\right\} = \left\{\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a},\boldsymbol{z})]\right\}$$
$$\bigcup\left\{(1-\beta)\,u(\boldsymbol{a},\boldsymbol{z}^i)+\beta\,\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a},\boldsymbol{z})]\middle|i=1,\ldots,M\right\}. \quad (20)$$

$\boldsymbol{z}^{i_0} \in \{\boldsymbol{z}^1,\ldots,\boldsymbol{z}^M\}$ minimizes the term contained in the second set on the right side if it holds that $u(\boldsymbol{a},\boldsymbol{z}^{i_0}) = \min_{1\le i\le M} u(\boldsymbol{a},\boldsymbol{z}^i)$. This proves (17). An analog consideration with respect to maximization delivers (18). □

An action $\boldsymbol{a}^* \in \boldsymbol{A}$ dominates another action $\boldsymbol{a}^{**} \in \boldsymbol{A}\setminus\{\boldsymbol{a}^*\}$ if, in terms of expected utility, $\boldsymbol{a}^*$ is surely at least as good as $\boldsymbol{a}^{**}$ is. The condition

$$\overline{\mathrm{E}}_{\boldsymbol{W}_{\mathrm{F}}}[u_{\mathrm{F}}(\boldsymbol{a}^{**},\boldsymbol{z}^i)] \le \underline{\mathrm{E}}_{\boldsymbol{W}_{\mathrm{F}}}[u_{\mathrm{F}}(\boldsymbol{a}^{*},\boldsymbol{z}^i)] \quad (21)$$

is sufficient to guarantee that $\boldsymbol{a}^*$ dominates $\boldsymbol{a}^{**}$. The next theorem provides an additional criterion by which dominance relations may be identified also if (21) does not hold.

Theorem 2 (**Additional Dominance Criterion**). *Provided that it is known that $p(\boldsymbol{z}|\boldsymbol{x}) \in \mathcal{P}_{\mathrm{F}}$, an action $\boldsymbol{a}^{**} \in \boldsymbol{A}$ is dominated by another action $\boldsymbol{a}^* \in \boldsymbol{A}\setminus\{\boldsymbol{a}^{**}\}$ if the following criterion is satisfied:*

$$\max\left\{(1-\beta)\max_{1\le i\le M}\left\{u(\boldsymbol{a}^{**},\boldsymbol{z}^i)-u(\boldsymbol{a}^{*},\boldsymbol{z}^i)\right\}\right.$$
$$\left.+\beta\,\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a}^{**},\boldsymbol{z})-u(\boldsymbol{a}^{*},\boldsymbol{z})],\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a}^{**},\boldsymbol{z})-u(\boldsymbol{a}^{*},\boldsymbol{z})]\right\}\le 0\,. \quad (22)$$

Condition (22) *can be only satisfied if $\boldsymbol{a}^*$ dominates $\boldsymbol{a}^{**}$ in the focussed model, i.e., if $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a}^{**},\boldsymbol{z})] \le \mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a}^{*},\boldsymbol{z})]$ holds.*

Proof. $\boldsymbol{a}^{**}$ is dominated by $\boldsymbol{a}^*$ if, for all $\boldsymbol{p}\in W_{\mathrm{F}}$, it holds that

$$\mathrm{E}_{\boldsymbol{p}}[u_{\mathrm{F}}(\boldsymbol{a}^{**},\boldsymbol{z}^i)]-\mathrm{E}_{\boldsymbol{p}}[u_{\mathrm{F}}(\boldsymbol{a}^{*},\boldsymbol{z}^i)] = \mathrm{E}_{\boldsymbol{p}}[u_{\mathrm{F}}(\boldsymbol{a}^{**},\boldsymbol{z}^i)-u_{\mathrm{F}}(\boldsymbol{a}^{*},\boldsymbol{z}^i)]\le 0\,. \quad (23)$$

This condition is satisfied if we have

$$\max_{\boldsymbol{p}\in W_{\mathrm{F}}}\mathrm{E}_{\boldsymbol{p}}[u_{\mathrm{F}}(\boldsymbol{a}^{**},\boldsymbol{z}^i)-u_{\mathrm{F}}(\boldsymbol{a}^{*},\boldsymbol{z}^i)]\le 0\,. \quad (24)$$

Because $\mathrm{E}_{\boldsymbol{p}}[u_{\mathrm{F}}(\boldsymbol{a}^{**},\boldsymbol{z}^i)-u_{\mathrm{F}}(\boldsymbol{a}^{*},\boldsymbol{z}^i)]$ is also a linear function with respect to $\boldsymbol{p}$, it adopts its maximum at least in one edge of $\boldsymbol{W}_{\mathrm{F}}$. Compare lemma 2. Introducing the notation $v(\boldsymbol{a}^*,\boldsymbol{a}^{**},\boldsymbol{z}) := u(\boldsymbol{a}^{**},\boldsymbol{z}^i)-u(\boldsymbol{a}^{*},\boldsymbol{z}^i)$, we obtain

$$\left\{\mathrm{E}_{\boldsymbol{p}}[u_{\mathrm{F}}(\boldsymbol{a}^{**},\boldsymbol{z}^i)-u_{\mathrm{F}}(\boldsymbol{a}^{*},\boldsymbol{z}^i)]\middle|\boldsymbol{p}\in\boldsymbol{V}(\boldsymbol{W}_{\mathbf{F}})\right\} = \left\{\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[v(\boldsymbol{a}^*,\boldsymbol{a}^{**},\boldsymbol{z})]\right\}$$
$$\bigcup\left\{(1-\beta)\,v(\boldsymbol{a}^*,\boldsymbol{a}^{**},\boldsymbol{z})+\beta\,\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[v(\boldsymbol{a}^*,\boldsymbol{a}^{**},\boldsymbol{z})]\middle|i=1,\ldots,M\right\}. \quad (25)$$

From this, it becomes clear that condition (22) is a sufficient dominance criterion: taking the maximum over (25), we just have to eliminate from the second set of the right side these elements for that $\mathrm{E}_{\boldsymbol{p}}[v(\boldsymbol{a}^*,\boldsymbol{a}^{**},\boldsymbol{z})]$ is surely not maximal.

Trivially, (22) can only hold if $\boldsymbol{a}^*$ dominates $\boldsymbol{a}^{**}$ in the focussed Bayesian model . □

ISBN 978-3-88579-286-4

Also if no action which maximizes the expected utility with respect to all $p(\boldsymbol{z}|\boldsymbol{x}) \in \mathcal{P}_\mathrm{F}$ is identifiable, the decision maker may be able to make an adequate decision: he may be able to chose an action which is guaranteed to be good enough with respect to the task at hand. Therefore, he may consider also regret values. The regret of an action is defined to be the maximal deficit in terms of expected utility that can arise from the choice of this action.

Theorem 3 (Regret). *Provided that it is known that $p(\boldsymbol{z}|\boldsymbol{x}) \in \mathcal{P}_\mathrm{F}$, for each action $\boldsymbol{a}^* \in \boldsymbol{A}$, the regret $\mathrm{R}_{\boldsymbol{W}_\mathrm{F}}(\boldsymbol{a}^*)$ of $\boldsymbol{a}^*$ is bounded from above:*

$$\mathrm{R}_{\boldsymbol{W}_\mathrm{F}}(\boldsymbol{a}^*) \leq \max \Big\{ \mathrm{R}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}(\boldsymbol{a}^*), \\ (1-\beta) \max_{\boldsymbol{a}\in\boldsymbol{A}} \max_{1\leq i\leq M} \left\{ u(\boldsymbol{a},\boldsymbol{z}^i) - u(\boldsymbol{a}^*,\boldsymbol{z}^i) \right\} + \beta \, \mathrm{R}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}(\boldsymbol{a}^*) \Big\} . \quad (26)$$

Thereby, $\mathrm{R}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}(\boldsymbol{a}^*)$ is the regret in the focussed Bayesian model: $\mathrm{R}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}(\boldsymbol{a}^*) := \max_{\boldsymbol{a}\in\boldsymbol{A}} \mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a},\boldsymbol{z})] - \mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a}^*,\boldsymbol{z})]$.

Proof. It holds

$$\mathrm{R}_{\boldsymbol{W}_\mathrm{F}}(\boldsymbol{a}^*) := \max_{\boldsymbol{a}\in\boldsymbol{A}} \max_{\boldsymbol{p}\in\boldsymbol{W}_\mathrm{F}} \left\{ \mathrm{E}_{\boldsymbol{p}}[u_\mathrm{F}(\boldsymbol{a},\boldsymbol{z}^i)] - \mathrm{E}_{\boldsymbol{p}}[u_\mathrm{F}(\boldsymbol{a}^*,\boldsymbol{z}^i)] \right\} . \quad (27)$$

Basically, the set in (27) is equal to the set in (24). Therefore, theorem 3 follows from a nearly analogous proceeding as the one applied at the maximization of the set in (24): performing an additional maximization with respect to $\boldsymbol{a}$, noting that this maximization is subadditive, and respecting the definition of $\mathrm{R}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}(\boldsymbol{a}^*)$ directly leads to (26). □

4.3 Improvement of the Basis for Decision Making

The decision maker may render the original LPI from lemma 1 more precisely if it is not possible to him to chose an adequate action. For this, he may expand the focussed Bayesian fusion to a superset of the local context $\boldsymbol{U}$. For simplicity, it is assumed without loss of generality that $\boldsymbol{U}$ is enlarged to $\boldsymbol{U}_\mathrm{L} := \boldsymbol{U} \cup \{\boldsymbol{z}^M\}$ and that it holds $p(\boldsymbol{x}|\boldsymbol{z}) \leq \delta_\mathrm{L}$ for all $\boldsymbol{z} \in \boldsymbol{U}_\mathrm{L}$ with a threshold δ_L such that $\delta_\mathrm{L} \leq \delta$ holds. Compare the beginning of section 3. By this, in (10), the inequality $0 \leq p^M$ gets sharpened to the equality $p^M = p^{M+1} \frac{p(\boldsymbol{z}^M|\boldsymbol{x},\boldsymbol{U}_\mathrm{L})}{1-p(\boldsymbol{z}^M|\boldsymbol{x},\boldsymbol{U}_\mathrm{L})}$ and the inequality $p^{M+1} \geq \beta$ gets sharpened to

$$p^{M+1} \geq \frac{\int_{\boldsymbol{U}_\mathrm{L}} l(\boldsymbol{x}|\boldsymbol{z})\, p(\boldsymbol{z}|\boldsymbol{U}_\mathrm{L}) \,\mathrm{d}\boldsymbol{z}}{\int_{\boldsymbol{U}_\mathrm{L}} l(\boldsymbol{x}|\boldsymbol{z})\, p(\boldsymbol{z}|\boldsymbol{U}_\mathrm{L}) \,\mathrm{d}\boldsymbol{z} + (\frac{1}{P(\boldsymbol{U}_\mathrm{L})} - 1)\, \delta_\mathrm{L}} \; (\geq \beta) . \quad (28)$$

Hence, enlarging $\boldsymbol{U}$ to $\boldsymbol{U}_\mathrm{L}$ results in LPI which is specified by a subset $\boldsymbol{W}_\mathrm{L}$ of the set $\boldsymbol{W}_\mathrm{F}$ in (10). Generally, this leads to a shrinking of the expected utility intervals, a larger set of dominated actions and smaller regret values. For the numerical evaluation of the decision criteria, the new LPI can get redrafted: setting $M_\mathrm{L} := M-1$, $Z_\mathrm{L} := \{\boldsymbol{z}^1, \ldots, \boldsymbol{z}^{M_\mathrm{L}+1}\}$, $\boldsymbol{p} := (p^1, \ldots, p^{M_\mathrm{L}+1})^\mathrm{T}$ with $\boldsymbol{z}^{M_\mathrm{L}+1} := \boldsymbol{U}_\mathrm{L}$, $p^i = p(\boldsymbol{z}^i|\boldsymbol{x})$ for $i \in \{1, \ldots, M_\mathrm{L}\}$, and

ISBN 978-3-88579-286-4

$p^{M_{\mathrm{L}+1}} := P(\boldsymbol{U}_{\mathrm{L}}|\boldsymbol{x})$, the results from the sections 4.1 and 4.2 are directly applicable by replacing M by M_{L}.

Alternatively, by making the additional assumption that $P(\boldsymbol{Z} \setminus \boldsymbol{U}|\boldsymbol{x})$ is not concentrated on small parts of $\boldsymbol{Z} \setminus \boldsymbol{U}$, the decision maker may also precise the LPI by the inclusion of non-trivial upper bounds for p^i, $i \in \{1, \ldots, M\}$. This also leads to a subset of $\boldsymbol{W}_{\mathrm{F}}$. If he assumes $p^i \leq \frac{1-\beta}{k}$ with a $k \in \{2, \ldots, M\}$, the edge structure of the resulting polygon gets changed and formulas (17), (18), (22), and (26) must be modified: the minimal and maximal values of the utility and utility differences of actions with respect to $\boldsymbol{Z} \setminus \boldsymbol{U}$ get substituted by the arithmetic means of the respective k-th lowest and largest values. The case $k = M$ corresponds to the assumption that either $P(\boldsymbol{Z} \setminus \boldsymbol{U}|\boldsymbol{x}) = 0$ or $p^i = \frac{1-\beta}{M}$, $i \in \{1, \ldots, M\}$, holds.

5 Examples

5.1 Example 1

As first example, we will prove[5] the following theorem by the use of theorem 2:

Theorem 4. *For events $\boldsymbol{E}^* \subseteq \boldsymbol{U}$, $\boldsymbol{E}^{**} \subseteq \boldsymbol{Z}$, we have $P(\boldsymbol{E}^*|\boldsymbol{x}) \geq P(\boldsymbol{E}^{**}|\boldsymbol{x})$ if the corresponding interval bounds satisfy* $\mathrm{b}(\boldsymbol{E}^{**} \cap (\boldsymbol{Z} \setminus \boldsymbol{U})) \leq \mathrm{a}(\boldsymbol{E}^*) - \mathrm{a}(\boldsymbol{E}^{**} \cap \boldsymbol{U})$.

Theorem 4 can be formulated as decision problem by defining $\boldsymbol{A} = \{\boldsymbol{E}^*, \boldsymbol{E}^{**}\}$, $u(\boldsymbol{E}, \boldsymbol{z}) = \mathbf{1}_{\boldsymbol{E}}(\boldsymbol{z})$ such that $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{E}, \boldsymbol{z})] = P(\boldsymbol{E}|\boldsymbol{x})$, $\boldsymbol{E} \in \boldsymbol{A}$. Thereby, $\mathbf{1}_{\boldsymbol{E}}(\boldsymbol{z})$ has the value 1 if $\boldsymbol{z} \in \boldsymbol{E}$ and the value 0 if $\boldsymbol{z} \notin \boldsymbol{E}$.

Proof. Without loss of generality, it can be assumed that we have $\boldsymbol{E}^* \cap \boldsymbol{E}^{**} = \emptyset$: if this does not hold, theorem 4 follows from a comparison of the posterior probabilities of $\boldsymbol{E}^* \setminus (\boldsymbol{E}^* \cap \boldsymbol{E}^{**})$ and $\boldsymbol{E}^{**} \setminus (\boldsymbol{E}^* \cap \boldsymbol{E}^{**})$.
We will show that $\boldsymbol{E}^*$ dominates $\boldsymbol{E}^{**}$ if the conditions in theorem 4 hold. It is assumed, also without loss of generality, that we have $\boldsymbol{E}^{**} = \boldsymbol{U}^{**} \cup \left(\bigcup_{i \in I^{**}} \boldsymbol{z}^i\right)$ with $\boldsymbol{U}^{**} \subseteq \boldsymbol{U}$ and $I^{**} \subseteq \{1, \ldots, M\}$. The dominance criterion (22) delivers the condition

$$\max\left\{\mathrm{b}(\boldsymbol{E}^{**} \cap \boldsymbol{U}) - \mathrm{b}(\boldsymbol{E}^*), \mathrm{b}(\boldsymbol{E}^{**} \cap (\boldsymbol{Z} \setminus \boldsymbol{U})) + \mathrm{a}(\boldsymbol{E}^{**} \cap \boldsymbol{U}) - \mathrm{a}(\boldsymbol{E}^*)\right\} \leq 0\ . \quad (29)$$

The second element of the set in (29) is not larger than zero if it holds

$$\mathrm{b}(\boldsymbol{E}^{**} \cap (\boldsymbol{Z} \setminus \boldsymbol{U})) \leq \mathrm{a}(\boldsymbol{E}^*) - \mathrm{a}(\boldsymbol{E}^{**} \cap \boldsymbol{U})\ . \quad (30)$$

The condition $\mathrm{b}(\boldsymbol{E}^{**} \cap \boldsymbol{U}) - \mathrm{b}(\boldsymbol{E}^*) \leq 0$ holds if $\mathrm{a}(\boldsymbol{E}^{**} \cap \boldsymbol{U}) - \mathrm{a}(\boldsymbol{E}^*) \leq 0$. Because $\mathrm{b}(\boldsymbol{E}^{**} \cap (\boldsymbol{Z} \setminus \boldsymbol{U})) \geq 0$, the (30) guarantees that $\boldsymbol{E}^*$ dominates $\boldsymbol{E}^{**}$. □

The case $\boldsymbol{E}^{**} \subseteq \boldsymbol{U} \setminus \boldsymbol{E}^*$ shows that (21) is a sufficient but not necessary condition for dominance: here, it holds $\mathbf{I}_{\boldsymbol{W}_{\mathrm{F}}}[u_{\mathrm{F}}(\boldsymbol{E}, \boldsymbol{z}^i)] = [\mathrm{a}[\boldsymbol{E}], \mathrm{b}[\boldsymbol{E}]]$, $\boldsymbol{E} \in \{\boldsymbol{E}^*, \boldsymbol{E}^{**}\}$. According to theorem 4, $\boldsymbol{E}^*$ dominates $\boldsymbol{E}^{**}$ if $\mathrm{a}[\boldsymbol{E}^{**}] \leq \mathrm{a}[\boldsymbol{E}^*]$, which is a weaker condition than $\mathrm{b}[\boldsymbol{E}^{**}] \leq \mathrm{a}[\boldsymbol{E}^*]$ is.

[5] An alternative proof, which is completely based on the results of the sections 2 and 3, is given in [San10].

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4

	$\boldsymbol{z}^1$	$\boldsymbol{z}^2$	$\boldsymbol{z}^3$	$\boldsymbol{z}^4$	$\boldsymbol{z}^5$	$\boldsymbol{z}^6$
$p(\boldsymbol{z})$	$\frac{4}{15}$	$\frac{2}{15}$	$\frac{2}{15}$	$\frac{3}{20}$	$\frac{4}{15}$	$\frac{1}{20}$
$l(x_1\|\boldsymbol{z})$	$\frac{1}{20}$	$\frac{1}{40}$	$\frac{1}{28}$	$\frac{1}{36}$	$\frac{1}{40}$	$\frac{1}{2}$
$l(x_2\|\boldsymbol{z})$	$\frac{1}{80}$	$\frac{1}{40}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{4}$	$\frac{1}{12}$

Table 1: Prior distribution and Likelihood functions at example 2.

	$\boldsymbol{z}^1$	$\boldsymbol{z}^2$	$\boldsymbol{z}^3$	$\boldsymbol{z}^4$	$\boldsymbol{z}^5$	$\boldsymbol{z}^6$
$u(\boldsymbol{a}_1, \boldsymbol{z})$	1	4	3	4	4	9
$u(\boldsymbol{a}_2, \boldsymbol{z})$	3	8	4	3	4	4
$u(\boldsymbol{a}_3, \boldsymbol{z})$	2	3	3	3	1	10

Table 2: Utility function at example 2.

5.2 Example 2

To hold the results verifiable easily, we assume in this example that $|\boldsymbol{Z}|$ is rather small: $\boldsymbol{Z} := \{\boldsymbol{z}^1, \ldots, \boldsymbol{z}^6\}$. Further, we assume that two information sources are available and that the respective information contributions x_1, x_2 are conditionally independent given $\boldsymbol{z}$. Let $p(\boldsymbol{z})$, $l(x_1|\boldsymbol{z})$, and $l(x_2|\boldsymbol{z})$ be as in table 1.

We apply a threshold $\tilde{\delta}$ on the standardized Likelihood functions $l(x_1|\boldsymbol{z})$ and $l(x_2|\boldsymbol{z})$. Compare section 2 and section 3. If we chose $\tilde{\delta} = \frac{1}{8}$, an element of $\boldsymbol{Z}$ is contained in the local context $\boldsymbol{U}$ if it holds $l(x_1|\boldsymbol{z}) > \delta_1$ or if it holds $l(x_1|\boldsymbol{z}) > \delta_2$ with $\delta_1 = \frac{1}{16}$ and $\delta_1 = \frac{1}{32}$. Hence, we obtain $\boldsymbol{U} = \{\boldsymbol{z}^5, \boldsymbol{z}^6\}$.

With $\delta = \delta_1 \cdot \delta_2 = \frac{1}{512}$, (3) delivers the lower bound $\beta \approx 0.7375$ for $P(\boldsymbol{U}|\boldsymbol{x})$ and according to (5), one obtains

$$p(\boldsymbol{z}|\boldsymbol{x}) \in [\mathrm{a}(\boldsymbol{z}), \mathrm{b}(\boldsymbol{z})] \approx \begin{cases} [0, 0.2625] & \text{if } \boldsymbol{z} \in \boldsymbol{Z} \setminus \boldsymbol{U} \ , \\ [0.3278, 0.4444] & \text{if } \boldsymbol{z} = \boldsymbol{z}^5 \ , \\ [0.4097, 0.5556] & \text{if } \boldsymbol{z} = \boldsymbol{z}^6 \ . \end{cases} \tag{31}$$

We have $\mathrm{b}(\boldsymbol{z}) \leq \mathrm{a}(\boldsymbol{z}^6)$ for $\boldsymbol{z} \in \boldsymbol{Z} \setminus \boldsymbol{U}$ and $\mathrm{a}(\boldsymbol{z}^5) \leq \mathrm{a}(\boldsymbol{z}^6)$. From this, we can conclude that $p(\boldsymbol{z}^6|\boldsymbol{x}) \geq p(\boldsymbol{z}|\boldsymbol{x})$ holds for all $\boldsymbol{z} \in \boldsymbol{Z} \setminus \{\boldsymbol{z}^6\}$. Compare theorem 4.

Now, let $\boldsymbol{A} = \{\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3\}$ be a set of available actions and let the utility function $u(\boldsymbol{a}, \boldsymbol{z})$ be as in table 2. Then, it holds $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a}_1, \boldsymbol{z})] = \frac{61}{9}$, $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a}_2, \boldsymbol{z})] = 4$, and $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a}_3, \boldsymbol{z})] = 6$. After the determination of the minimum and maximum values of $u(\boldsymbol{a}_i, \boldsymbol{z})$ for $\boldsymbol{z} \in \boldsymbol{Z} \setminus \boldsymbol{U}$, theorem 1 approximately delivers

$$\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{a}|\boldsymbol{z})] \in \begin{cases} [5.2611, 6.7778] & \text{for } \boldsymbol{a} = \boldsymbol{a}_1 \ , \\ [3.7375, 5.0500] & \text{for } \boldsymbol{a} = \boldsymbol{a}_2 \ , \\ [4.9501, 6.0000] & \text{for } \boldsymbol{a} = \boldsymbol{a}_3 \ . \end{cases} \tag{32}$$

From (32), one directly obtains that action $\boldsymbol{a}_1$ is always better than action $\boldsymbol{a}_2$. Compare (21). After the determination of the maximum value of $u(\boldsymbol{a}_3, \boldsymbol{z}) - u(\boldsymbol{a}_1, \boldsymbol{z})$ for $\boldsymbol{z} \in \boldsymbol{Z} \setminus \boldsymbol{U}$, by the application of the dominance criterion in theorem 2, we can additionally conclude that $\boldsymbol{a}_1$ dominates $\boldsymbol{a}_3$. Hence, it is sure that action $\boldsymbol{a}_1$ has the maximum posterior expected utility within the global Bayesian model.

Now, we assume that only action $\boldsymbol{a}_2$ and action $\boldsymbol{a}_3$ are available, i.e., we assume that it holds $\boldsymbol{A} = \{\boldsymbol{a}_2, \boldsymbol{a}_3\}$. Also in this case, we can determine the action that has the maximum posterior expected utility within the global Bayesian model: the application of the dominance criterion in theorem 2 yields that action $\boldsymbol{a}_3$ dominates action $\boldsymbol{a}_2$.

To demonstrate that it is not always possible to identify the action that is optimal which respect to the global model, we assume again that $\boldsymbol{A} = \{\boldsymbol{a}_2, \boldsymbol{a}_3\}$ holds. However, we modify the value of $u(\boldsymbol{a}_2, \boldsymbol{z}^1)$ to 8. This modification has no effect on the expected utility intervals for $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{a}_2, \boldsymbol{z})]$ and $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{a}_3, \boldsymbol{z})]$–although the value of $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{a}_2, \boldsymbol{z})]$ has changed by the modification of $u(\boldsymbol{a}, \boldsymbol{z})$. If we assume that, globally, it holds that $p(\boldsymbol{z}^1) = 1 - \beta$, we obtain $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{a}_2, \boldsymbol{z})] \approx 5.0500$ and $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{a}_3, \boldsymbol{z})] \approx 4.9501$. Hence, in this case, action $\boldsymbol{a}_2$ is better than action $\boldsymbol{a}_3$. However, if we assume that, globally, it holds that $p(\boldsymbol{z}^4) = 1 - \beta$, we obtain $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{a}_2, \boldsymbol{z})] \approx 3.7375$ and $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{a}_3, \boldsymbol{z})] \approx 5.2125$. Here, action $\boldsymbol{a}_3$ is better than action $\boldsymbol{a}_2$.

In this situation, theorem 3 approximately delivers $\mathrm{R}_{\boldsymbol{W}_\mathrm{F}}(\boldsymbol{a}_3) \leq 0.0999$. Depending on the actual task, this upper bound for the regret of action $\boldsymbol{a}_3$ may be low enough therefore that the decision maker is able to choose this action. If this it not possible, he may expand the local context $\boldsymbol{U}$. For example, if he lowers the threshold which is applied to the standardized Likelihood functions to $\tilde{\delta} = \frac{1}{10}$, the local context contains additionally $\boldsymbol{z}^3$ and $\boldsymbol{z}^4$. With $\delta = \frac{1}{800}$, (3) delivers the lower bound $\beta \approx 0.8888$ for $P(\boldsymbol{U}|\boldsymbol{x})$. We have $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a}_2, \boldsymbol{z})] \approx 3.9711$, and $\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x},\boldsymbol{U})}[u(\boldsymbol{a}_3, \boldsymbol{z})] \approx 5.8139$. Here, theorem 1 approximately delivers

$$\mathrm{E}_{p(\boldsymbol{z}|\boldsymbol{x})}[u(\boldsymbol{a}|\boldsymbol{z})] \in \begin{cases} [3.9711, 4.4191] & \text{for } \boldsymbol{a} = \boldsymbol{a}_2 \ , \\ [5.3898, 5.8139] & \text{for } \boldsymbol{a} = \boldsymbol{a}_3 \ . \end{cases} \tag{33}$$

Hence, it becomes directly clear that, within the global Bayesian model, the posterior expected utility of action $\boldsymbol{a}_3$ is larger than the posterior expected utility of action $\boldsymbol{a}_2$.

Of course, by enlarging $\boldsymbol{U}$, the lower and upper bounds for the global posterior probabilities get sharpened as well. Now, we have

$$p(\boldsymbol{z}|\boldsymbol{x}) \in [\mathrm{a}(\boldsymbol{z}), \mathrm{b}(\boldsymbol{z})] \approx \begin{cases} [0, 0.1112] & \text{if } \boldsymbol{z} \in \boldsymbol{Z} \setminus \boldsymbol{U} \ , \\ [0.0294, 0.0331] & \text{if } \boldsymbol{z} = \boldsymbol{z}^3 \ , \\ [0.0257, 0.0289] & \text{if } \boldsymbol{z} = \boldsymbol{z}^4 \ . \\ [0.3705, 0.4169] & \text{if } \boldsymbol{z} = \boldsymbol{z}^5 \ , \\ [0.4632, 0.5211] & \text{if } \boldsymbol{z} = \boldsymbol{z}^6 \ . \end{cases} \tag{34}$$

6 Conclusion

Focussed Bayesian fusion is a local Bayesian fusion technique that distinguishes itself by a straightforward fusion scheme. We developed decision theoretic approaches to make globally optimal decision making possible in a consistent and timely manner. For this, an interval scheme for global posterior probabilities and common decision criteria under

LPI together with principles of lazy decision making have been used profitably. Local Bayesian fusion approaches are widely new approaches to circumvent high computational costs of Bayesian fusion. They may also get combined with the concept of conjugate priors or Markov Chain Monte Carlo Methods. Further research may address this topic in detail.

References

[BHS06] Jürgen Beyerer, Michael Heizmann, and Jennifer Sander. Fuselets – An Agent Based Architecture for Fusion of Heterogeneous Information and Data. In Belur V. Dasarathy, editor, *Multisensor, multisource information fusion: architectures, algorithms, and applications 2006*, Proceedings of SPIE 6242, pages 235–243, 2006.

[BHSG08] Jürgen Beyerer, Michael Heizmann, Jennifer Sander, and Ioana Gheta. Bayesian Methods for Image Fusion. In Tania Stathaki, editor, *Image Fusion: Algorithms and Applications*, pages 157–192. Academic Press, 2008.

[BS04] José M. Bernardo and Adrian F. M. Smith. *Bayesian Theory*. Wiley, 2004.

[Fis01] Gerd Fischer. *Analytische Geometrie*. Vieweg, 2001.

[KM76] Eduard Kofler and Günter Menges. *Entscheidungen bei unvollständiger Information*. Springer, 1976.

[Koc10] Wolfgang Koch. On Bayesian Tracking and Data Fusion: a Tutorial Introduction with Examples. *Aerospace and Electronic Systems Magazine, IEEE*, 25(7):29 –52, 2010.

[NM04] Klaus Neumann and Martin Morlock. *Operations Research*. Hanser, 2004.

[Pre02] Gero Presser. *Lazy Decision Making*. Dissertation, Universität Dortmund, 2002.

[Rüg99] Bernhard Rüger. *Test- und Schätztheorie, Band 1, Grundlagen*. Oldenbourg, 1999.

[San09] Jennifer Sander. Further Investigation of Focussed Bayesian Fusion. Technical Report IES-2009-03, 2009.

[San10] Jennifer Sander. Use of a Probability Interval Scheme for Focussed Bayesian Fusion. Technical Report IES-2010-03, 2010.

[SB08] Jennifer Sander and Jürgen Beyerer. Decreased Complexity and Increased Problem Specificity of Bayesian Fusion by Local Approaches. In *Proceedings of 11th Conference on Information Fusion 2008 (Fusion 2008)*, pages 1035–1042. IEEE, 2008.

[SHGB09] Jennifer Sander, Michael Heizmann, Igor Goussev, and Jürgen Beyerer. A Local Approach for Focussed Bayesian Fusion. In Belur Dasarathy, editor, *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications, Proceedings of SPIE Vol. 7345*, 2009.

[SHGB10] Jennifer Sander, Michael Heizmann, Igor Goussev, and Jürgen Beyerer. Global Evaluation of Focussed Bayesian Fusion. In Jerome J. Braun, editor, *Multisensor, Multisource Information Fusion: Architectures, Algorithms, and Applications, Proceedings of SPIE Vol. 7710*, 2010.

[SKB10] Jennifer Sander, Jonas Krieger, and Jürgen Beyerer. The Importance of Statistical Evidence for Focussed Bayesian Fusion. In Rüdiger Dillmann, Jürgen Beyerer, Uwe D. Hanebeck, and Tanja Schultz, editors, *KI 2010: Advances in Artificial Intelligence*, volume 6359, pages 299–308. Springer, 2010.

erschienen im Tagungsband der INFORMATIK 2011
Lecture Notes in Informatics, Band P192
ISBN 978-3-88579-286-4
weitere Artikel online:
http://informatik2011.de/519.html