



## Bayesian Approach to Intelligent Control and Its Relation to Fuzzy Control

Kongliang Zhu<sup>†</sup>, Olga Kosheleva<sup>‡</sup> and Vladik Kreinovich<sup>‡,1</sup>

<sup>†</sup>Faculty of Economics, Chiang Mai University  
Chiang Mai, Thailand  
e-mail : 258zk1@gmail.com

<sup>‡</sup>University of Texas at El Paso, El Paso, TX 79968, USA  
e-mail : olgak@utep.edu (O. Kosheleva)  
vladik@utep.edu (V. Kreinovich)

**Abstract :** In many application areas including economics, experts describe their knowledge by using imprecise (“fuzzy”) words from natural language. To design an automatic control system, it is therefore necessary to translate this knowledge into precise computer-understandable terms. To perform such a translation, a special semi-heuristic fuzzy methodology was designed. This methodology has been successfully applied to many practical problem, but its semi-heuristic character is a big obstacle to its use: without a theoretical justification, we are never 100% sure that this methodology will be successful in other applications as well. It is therefore desirable to come up with either a theoretical justification of exactly this methodology, or with a theoretically justified modification of this methodology. In this paper, we apply the Bayesian techniques to the above translation problem, and we analyze when the resulting methodology is identical to fuzzy techniques – and when it is different.

**Keywords :** intelligent control; Bayesian approach; fuzzy control.

**2010 Mathematics Subject Classification :** 62P20; 62P30; 91B70; 93B99; 93C42.

---

This work was supported in part by the National Science Foundation grants HRD-0734825 and HRD-1242122 (Cyber-ShARE Center of Excellence) and DUE-0926721, and by an award “UTEP and Prudential Actuarial Science Academy and Pipeline Initiative” from Prudential Foundation.

<sup>1</sup>Corresponding author.

**Copyright © 2016 by the Mathematical Association of Thailand.  
All rights reserved.**

## 1 Formulation of the Problem

**Need for expert systems and intelligent control.** In many application areas ranging from medicine to economics, there are experts who consistently make very good decision. In the ideal world, we should all be helped by these experts:

- every patient should be seen by the world’s best specialist in treating the corresponding disease,
- everyone’s investment strategy should be advised by the world’s best financial advisors, etc.

Realistically, however, this is not possible: there are very few best experts, and it is not possible to use them for every single decision making process.

It is therefore desirable to incorporate the knowledge of the best experts into a computer-based system so that it will be possible for many people to utilize this knowledge. The resulting computer-based *expert systems* should be able to provide us with intelligent decision-making and intelligent control.

**Challenge: expert knowledge is imprecise (fuzzy).** Sometimes, the experts are able to describe their knowledge in precise (and thus, computer-understandable) terms. For example, a medical expert can say that in the case of common cold, Advil should be taken only when the body temperature is  $38^{\circ}$  C or above. Such knowledge is easy to implement in a computer-based system.

In many practical situations, however, experts cannot describe their knowledge in precise terms, they can only describe their knowledge by using imprecise (“fuzzy”) natural-language words such as “small”, “young”, etc. For example, a financial expert can say that:

- if the price of a financial instrument start decreasing a little bit, it is better to stick to it, while
- if this price starts decreasing rapidly, with little hope of this price climbing back soon, it may be better to sell this instrument as soon as possible – to recover at least some of the original investment.

This reasonable piece of advice contains many imprecise words: “a little bit”, “little” (hope), “rapidly”, “may be better”, etc.

Let us denote the inputs based on which we need to make a decision by  $x_1, \dots, x_n$ , and let us denote the variable that describes the resulting decision by  $u$ .

In general, for each of the variables  $x_i$  and  $u$ , we have several imprecise properties like “small” used in the expert rules:

- for  $x_1$ , we have imprecise properties  $A_{1,1}, A_{1,2}, \dots, A_{1,R_1}$ ;
- for  $x_2$ , we have imprecise properties  $A_{2,1}, A_{2,2}, \dots, A_{2,R_2}$ ;
- ...
- for  $x_n$ , we have imprecise properties  $A_{n,1}, A_{n,2}, \dots, A_{n,R_n}$ ; and

- for  $u$ , we have imprecise properties  $B_1, B_2, \dots, B_{R_U}$ .

An expert knowledge consists of rules that describe what to do under all possible combinations  $r = (r_1, \dots, r_n)$  of conditions on the inputs:

$$\text{If } A_{1,r_1}(x_1) \text{ and } \dots \text{ and } A_{i,r_i}(x_i) \text{ and } \dots \text{ and } A_{n,r_n}(x_n) \text{ then } B_{f(r)}(u),$$

for an appropriate  $u$ -property  $f(r)$ .

How can we translate such imprecise knowledge into precise computer-understandable terms?

**Fuzzy techniques: a brief reminder.** Fuzzy logic was specifically designed to solve this translation problem – i.e., to translate imprecise expert knowledge into a precise control strategy; see, e.g., [1–3].

Fuzzy control methodology starts with an observation that in the above rules, a control  $u$  is reasonable for the input tuple  $x = (x_1, \dots, x_n)$  if and only if one of the expert rules is applicable, i.e.,

$$\text{either } (A_{1,1}(x_1) \text{ and } \dots \text{ and } A_{i,1}(x_i) \text{ and } \dots \text{ and } A_{n,1}(x_n) \text{ and } B_{f(1,\dots,1)}(u))$$

or ...

$$\text{or } (A_{1,r_1}(x_1) \text{ and } \dots \text{ and } A_{i,r_i}(x_i) \text{ and } \dots \text{ and } A_{n,r_n}(x_n) \text{ and } B_{f(r_1,\dots,r_n)}(u))$$

or ...

$$\text{or } (A_{1,R_1}(x_1) \text{ and } \dots \text{ and } A_{i,R_i}(x_i) \text{ and } \dots \text{ and } A_{n,R_n}(x_n) \text{ and } B_{f(R_1,\dots,R_n)}(u)).$$

In line with this representation, fuzzy methodology starts by assigning, to each imprecise property  $A$  to each real number  $x$ , a “degree”  $\mu_A(x) \in [0, 1]$  to which the number  $x$  satisfies this property (e.g., a degree to which a certain number is small). This degree can be obtained:

- either by simply asking an expert to mark this degree on a scale from 0 to 1,
- or in a more probability-like way, e.g., by asking several experts and computing the proportion of the experts who believe that  $x$  satisfies the property  $A$ .

Once we have established the degrees  $\mu_{i,r}(x_i)$  to which  $x_i$  satisfies each property  $A_{i,r}$  and the degrees  $\mu_r(u)$  to which  $u$  satisfies each property  $B_r$ , we need to use these degrees to estimate the degree to which each rule is satisfied. Ideally, we should ask the expert’s opinion about all possible combinations of inputs  $(x_1, \dots, x_n)$ , but this is usually not realistic, so we must use the original degrees to find the corresponding combinations.

In fuzzy methodology, we select two functions  $f_{\&}(a, b)$  and  $f_{\vee}(a, b)$  that correspond to “and” and “or”. These functions should satisfy natural properties coming from the fact that, e.g.,  $A \& B$  and  $B \& A$  mean the same, and that  $A \& (B \& C)$  and  $(A \& B) \& C$  mean the same – thus, both operations should be commutative and associative. Such operations are known as “and”-operations (a.k.a. *t-norms*) and “or”-operations (*t-conorms*).

By applying these operations, we can estimate, for each possible control value  $u$  and for each rule  $r$ , the degree  $d_r(u)$  to which this control value is consistent, for the given input  $x$ , with this rule, as

$$d_r(u) = f_{\&}(\mu_{1,r_1}(x_1), \dots, \mu_{i,r_i}(x_i), \dots, \mu_{n,r_n}(x_n), \mu_{f(r)}(u)),$$

and then we can compute the degree  $\mu(u)$  to which the control value is reasonable as

$$\mu(u) = f_V(d_{(1,\dots,1)}(u), \dots, d_{(r_1,\dots,r_n)}(u), \dots, d_{(R_1,\dots,R_n)}(u)).$$

In situations when the purpose of the expert system is to advise the user, then it is sufficient to describe, for each possible control value  $u$ , to what extent this value is reasonable. In this case, a user can make a decision based on this information. However, in many practical situations, we cannot afford to always have a human decision maker in the loop. In such situations, we want the system to make an automatic decision. For that, we need to select a single value  $\bar{u}$ . Usually, as such a value, we select the “centroid” value

$$\bar{u} = \frac{\int u \cdot \mu(u) du}{\int \mu(u) du}.$$

**Takagi-Sugeno fuzzy techniques.** The above techniques – known as *Mamdani* techniques after the first researcher who successfully applied these techniques to a practical problem – assume that the consequences of all the rules are fuzzy. In practice, sometimes, the experts describe consequences of their rule in precise terms, as an explicit function describing the control value  $u$  in terms of the inputs  $x_i$ . In other words, we have rules of the following type:

if  $A_{1,r_1}(x_1)$  and ... and  $A_{i,r_i}(x_i)$  and ... and  $A_{n,r_n}(x_n)$  then  $u = f_r(x_1, \dots, x_n)$ .

In such situations – first analyzed by Takagi and Sugeno – we first compute the degree  $d_r$  to which each rule is satisfied, as

$$d_r = f_{\&}(\mu_{1,r_1}(x_1), \dots, \mu_{i,r_i}(x_i), \dots, \mu_{n,r_n}(x_n))$$

and then the desired control is produced as the weighted average of the values  $f_r(x_1, \dots, x_n)$  corresponding to different rules:

$$\bar{u} = \frac{\sum_{r=(1,\dots,1)}^{(R_1,\dots,R_n)} d_r \cdot f_r(x_1, \dots, x_n)}{\sum_{r=(1,\dots,1)}^{(R_1,\dots,R_n)} d_r}.$$

**Fuzzy control: successes and limitations.** Fuzzy control methodology has lead to many successful applications; see, e.g., [1,2]. However, a big problem is that the above techniques are largely heuristic, they do not have a precise justification. Because of this lack of justification, there is no guarantee that this empirically successful method will be successful in all future applications as well.

It is therefore desirable to theoretically analyze fuzzy control techniques – and:

- either to justify the existing methodology,

- or, if necessary, come up with a theoretically justified alternative to this methodology.

**What we do in this paper.** In this paper:

- we apply the Bayesian approach to the original problem, and then
- we compare the result of applying this approach to the original fuzzy methodology.

## 2 Let Us Apply Bayesian Methodology to Our Problem

**What is Bayesian methodology: a brief reminder.** In the general Bayesian approach, the original expert knowledge is described as a *prior distribution*.

In situations when we only have partial information about this distribution, we use the *Maximum Entropy approach* to select the corresponding distribution – i.e., we select a probability distribution  $\rho(x)$  for which the entropy

$$S = - \int \rho(x) \cdot \ln(\rho(x)) dx$$

is the largest possible; see, e.g., [4].

Then, we use the *Bayes rule* to update the corresponding probabilities – i.e., to come up with the *posterior* distribution.

**How we will apply the Bayesian approach.** To make this application as clear as possible, we will not start by applying the Bayesian approach to the most general form of the above problem. Instead, we start with the simplest possible case, and then we will gradually add complexity – and at the end, we will have an application to the most general form of intelligent control.

## 3 The Simplest Case: Single Variable and Disjoint Rules

**Analysis of the problem.** The simplest case is when we have only one input  $x_1$ . In this case, the rules have the form “if  $A_{1,r_1}(x_1)$  then  $B_{f(r_1)}(u)$ .”

To apply the Bayesian approach, we need to describe the corresponding uncertainty in probabilistic terms. For each value of the input  $x_1$ , we have  $R_1$  possible properties:  $A_{1,1}, \dots, A_{1,R_1}$ . Uncertainty means that for each value  $x_1$ , we are not 100% sure which of the properties is satisfied. In probabilistic terms, this uncertainty can be naturally described as the conditional probability  $P_{I,1}(r_1 | x_1)$  that the  $r_1$ -st property is satisfied under the condition that the input is equal to  $x_1$  (here,  $I$  stands for “input”).

The above rules describe all possible situations. It makes sense to assume that we cannot have two different properties  $A_{1,r_1} \neq A_{1,r'_1}$  satisfied at the same time: otherwise, two different expert rules with the same condition would lead to two different recommendations about control. Thus, it makes sense to assume that for each  $x_1$ , the corresponding conditional probabilities add up to 1:

$$\sum_{r_1=1}^{R_1} P_{I,1}(r_1 | x_1) = 1.$$

(In the next section, we describe what will happen if this sum is different from 1).

Similarly, the expert's uncertainty about the control values  $u$  can be described by assigning the corresponding probabilities  $P_C(r | u)$ , for which, for every  $u$ , we have  $\sum_{r=1}^{R_U} P_C(r | u) = 1$ ; here,  $C$  stands for "control".

Our goal is to find, for each input  $x_1$  and for each possible control value  $u$ , to what extent this value  $u$  is reasonable. In probabilistic terms, this means that we are interested to find the corresponding conditional probabilities  $P(u | x_1)$ .

Since we assume that different rules  $r_1$  form a whole set of disjoint events, this conditional probability can be computed by using the formula for the complete probability:

$$P(u | x_1) = \sum_{r_1=1}^{R_1} P(u, r_1 | x_1).$$

Here, the rule  $r_1$  determines the control value  $u$ , so each probability  $P(u, r_1 | x_1)$  has the form

$$P(u, r_1 | x_1) = P(u | f(r_1)) \cdot P_{I,1}(r_1 | x_1).$$

In this formula, we know the probabilities  $P_{I,1}(r_1 | x_1)$ . To find the probabilities  $P(u | r)$ , we need to use the Bayes formula:

$$P(u | r) = \frac{P_C(r | u) \cdot P_0(u)}{\int P_C(r | u') \cdot P_0(u') du'},$$

where  $P_0(u)$  is a prior probability of different control values.

In our expert situation, we do not have any prior information about the control. Thus, we have no reason to believe that some values of control are more probable than others. Therefore, it makes sense to assume that the value  $P_0(u)$  is a constant not depending on  $u$ . This assumption is in perfect accordance with the Maximum Entropy approach: if we have no information about the probability distribution on an interval, then this approach leads to a uniform distribution.

When  $P_0(u) \equiv \text{const}$ , we can divide both sides of the probability distribution by this constant, and get the following simplified formula:

$$P(u | r) = \frac{P_C(r | u)}{\int P_C(r | u') du'}.$$

Thus, the desired degree of reasonableness takes the following form.

**The resulting formula for the degree of reasonableness.** For each input value  $x_1$  and each possible control value  $u$ , the degree  $d(u, x_1)$  to which  $u$  is reasonable is described by the following formula:

$$d(u, x_1) = P(u | x_1) = \sum_{r_1=1}^{R_1} \frac{P_C(f(r_1) | u)}{\int P_C(f(r_1) | u') du'} \cdot P_{I,1}(r_1 | x_1).$$

**Comparing with the corresponding fuzzy formula.** In fuzzy terms, the degree  $P_{I,1}(r | x_1)$  corresponds to  $\mu_{1,r_1}(x_1)$ , and the degree  $P_C(r | u)$  corresponds to  $\mu_r(u)$ . In these terms, the above formula has the form

$$d(u, x_1) = \sum_{r_1=1}^{R_1} \frac{\mu_{f(r_1)}(u)}{\int \mu_{f(r_1)}(u') du'} \cdot \mu_{1,r_1}(x_1).$$

Let us compare this formula with the Mamdani formula:

$$d(u, x_1) = f_{\vee}(f_{\&}(\mu_{1,1}(x_1), \mu_{f(1)}(u)), \dots, f_{\&}(\mu_{1,R_1}(x_1), \mu_{f(R_1)}(u))).$$

The Bayesian formula can be obtained from the general Mamdani formula when:

- the “or”-operation is the sum:  $f_{\vee}(a, b) = a + b$ ;
- the “and”-operation is the product  $f_{\&}(a, b) = a \cdot b$ , and
- all “membership functions”  $\mu_r(u)$  have the same value of the integral  $\int \mu_r(u) du$ .

To be more precise, in this case, the Bayesian formula leads to the values not equal but *proportional* to the Mamdani formula – divided by the common integral  $\int \mu_r(u) du$ .

The condition that the integral  $\int \mu_r(u) du$  is the same for all  $r$  is usually satisfied for most applications of fuzzy techniques. But what about the situations when this condition is not satisfied? In this case, as we will show, the Bayesian formula is more appropriate. Indeed, suppose that we have 2 rules:

- that when  $x_1 = 1$ , we should use  $u = 2$ , and
- that when  $x_1$  is small,  $u$  should be small.

In this case, we should expect that when  $x_1 = 1$ , the resulting control value will be exactly 2.

- This is exactly what we observe in the Bayesian case.
- However, in the Mamdani approach, as one can easily check, many other values will also be considered reasonable.

This was a clear advantage of the Bayesian approach. On the other hand, the fuzzy approach has its own advantages. For example, in the fuzzy approach, the possibility to have different “and”- and “or”-operations enables us to adjust

the control depending on whether, e.g., we want it to be more stable or more smooth [5, 6]. The Bayesian approach lacks this flexibility.

**Resulting control values: Bayesian approach.** As we have mentioned earlier, in some cases, in addition to describing which control values  $u$  are more reasonable and which are less reasonable, it is desirable to select a single control value  $\bar{u}$ . It makes sense to select the value for which the expected loss is the smallest possible.

The loss is caused by the difference  $u - \bar{u}$  between the (unknown) actually optimal value  $u$  and the selected value  $\bar{u}$ , so the loss has the form  $J(u - \bar{u})$  for an appropriate loss function  $J(x)$ . For good expert knowledge, the control  $\bar{u}$  is close to the optimal control  $u$ . Thus, the difference  $\Delta u \stackrel{\text{def}}{=} u - \bar{u}$  is small, so we can expand the objective function  $J(u - \bar{u})$  in Taylor series

$$J(\Delta u) = J_0 + J_1 \cdot \Delta u + J_2 \cdot (\Delta u)^2 + \dots,$$

and keep only the main term in the resulting expansion.

When  $\Delta u = 0$ , i.e., when we apply the optimal control value  $u = \bar{u}$ , the loss is 0, so  $J_0 = 0$ . The loss is minimal for  $\Delta u = 0$ , so the derivative  $J_1$  of  $J$  with respect to  $\Delta u$  is equal to 0, hence  $J(\Delta u) = J_2 \cdot (\Delta u)^2 + \dots$ . Thus, the quadratic term is the first non-zero term and therefore, the main term in the Taylor expansion:  $J(\Delta u) = J_2 \cdot (\Delta u)^2$ .

When  $\Delta u \neq 0$ , there is a positive loss, so  $J_2 > 0$ . Minimization does not change if we divide the objective function by a positive constant, so we can assume that  $J(\Delta u) = (\Delta u)^2$ . Thus, the value  $\bar{u}$  should be selected based on the condition that the expected value of the square  $(\Delta u)^2$  of the difference  $\Delta u$  is the smallest possible:

$$\int P(u | x_1) \cdot (u - \bar{u})^2 du \rightarrow \min_{\bar{u}}.$$

Differentiating this objective function with respect to  $\bar{u}$ , equating the derivative to 0, and taking into account that  $\int P(u | x_1) du = 1$ , we conclude that

$$\bar{u} = \int P(u | x_1) \cdot u du.$$

In other words, as  $\bar{u}$ , we should select the expected value of the control  $u$ .

**Resulting control value: conclusion.** As the desired control value, we should select the expected value of the control  $\bar{u} = \int P(u | x_1) \cdot u du$ . In view of the above formula for the probabilities  $P(u | x_1)$ , we get

$$\bar{u} = \sum_{r_1=1}^{R_1} \bar{u}_{f(r_1)} \cdot P_{I,1}(r_1 | x_1),$$

where

$$\bar{u}_r = \frac{\int P(r | u) \cdot u du}{\int P(r | u) du}.$$

**Comparison with the fuzzy formula.** In fuzzy terms, the above formula has the form

$$\bar{u} = \sum_{r_1=1}^{R_1} \bar{u}_{f(r_1)} \cdot \mu_{1,r_1}(x_1),$$

where

$$\bar{u}_r = \frac{\int \mu_r(u) \cdot u \, du}{\int \mu_r(u) \, du}.$$

This formula coincides with the Mamdani formula under the same three conditions as before – that “or” is sum, that “and” is product and that all functions  $\mu_r(u)$  have the same integral.

Interestingly, in general, the Bayesian formula is more similar to the Takagi-Sugeno formula than to the general Mamdani formula.

## 4 What If the Probabilities Do Not Add Up to One?

**Description of the case.** In the above formulas, we assumed:

- that for every input  $x_1$ , the corresponding probabilities  $p_{r_1} = P_{I,1}(r_1 | x_1)$  add up to 1 and
- that for the control  $u$ , the probabilities  $P_C(r | u)$  also add up to 1.

But what if we estimate these probabilities and they do not add up to 1?

**How to deal with this situation: an idea and its formalization.** Estimates are never exact. So, if instead of the actual (unknown) probabilities  $p_r$  for which  $\sum_{r=1}^R p_r = 1$ , we use approximate estimates  $\tilde{p}_r \approx p_r$ , then the sum of the approximate estimates is not necessarily equal to 1.

It is reasonable to assume that the approximation errors  $\tilde{p}_r - p_r$  are independent normally distributed with 0 means and the same standard deviation. In this case, to find the best estimates for the actual probabilities, we must find the values  $p_r$  for which  $\sum_{r=1}^R (\tilde{p}_r - p_r)^2$  under the constraint  $\sum_{r=1}^R p_r = 1$ .

**From the idea to explicit formulas.** For the above constraint optimization problem, the Lagrange multiplier approach leads to the following unconstrained optimization problem:

$$\sum_{r=1}^R (\tilde{p}_r - p_r)^2 + \lambda \cdot \left( \sum_{r=1}^R p_r - 1 \right) \rightarrow \min_{p_r}.$$

When  $p_r \neq 0$ , then the derivative of the above objective function with respect to  $p_r$  should be equal to 0, so we have  $2(p_r - \tilde{p}_r) + \lambda = 0$  and thus,  $p_r = \tilde{p}_r - c$ , where  $c \stackrel{\text{def}}{=} \frac{\lambda}{2}$  is a constant.

In particular, when the original sum  $\sum_{r=1}^R \tilde{p}_r$  is smaller than 1, then we should have  $c < 0$ , and this constant can be computed from the condition that

$$\sum_{r=1}^R p_r = \sum_{r=1}^R (\tilde{p}_r - c) = 1,$$

$$\text{thus } c = -\frac{1 - \sum_{r=1}^R \tilde{p}_r}{R}.$$

Thus, we arrive at the following solution.

**Resulting formula.** When  $\sum_{r=1}^R \tilde{p}_r = 1$ , we take  $p_r = \tilde{p}_r$ .

When  $\sum_{r=1}^R \tilde{p}_r < 1$ , we take  $p_r = \tilde{p}_r - c$ , where

$$c = -\frac{1 - \sum_{r=1}^R \tilde{p}_r}{R}.$$

When  $\sum_{r=1}^R \tilde{p}_r > 1$ , we first try the same formula, but if it leads to negative probabilities, then instead, after sorting the probabilities in the increasing order  $\tilde{p}_1 \leq \dots \leq \tilde{p}_R$ , we take  $p_1 = \dots = p_{r_0} = 0$  and  $p_r = \tilde{p}_r - c$  for  $r > r_0$ , where

$$c = \frac{\sum_{r=r_0+1}^R \tilde{p}_r - 1}{R - r_0}$$

and  $\tilde{p}_{r_0} \leq c$ .

For each  $r_0$ , we have an explicit formula for  $p_r$ . Thus, the desired values  $p_r$  can be computed, e.g., by trying all possible  $r_0 = 1, 2, \dots$ .

**Resulting formulas for intelligent control.** We start with the empirical values  $\tilde{P}_{I,1}(r_1 | x_1)$  and  $\tilde{P}_C(r | u)$ . Based on these values, for each  $x_1$  and each  $u$ , we apply the above procedure and get the updated values  $P_{I,1}(r_1 | x_1)$  and  $P_C(r | u)$  that add up to 1.

Based on these values, we then compute the degrees of reasonableness  $P(u | x_1)$  and, if needed, the resulting single control value  $\bar{u}$ .

## 5 Case of Several Inputs

**Formulation of the problem.** For several inputs, we can use similar Bayesian formulas, but for that, we need to know the probabilities

$$P_I(r | x) = P_I(r_1, \dots, r_n | x_1, \dots, x_n),$$

while what we get from experts are the probabilities  $P_{I,i}(r_i | x_i)$  corresponding to individual inputs  $i$ .

**Let us solve this problem.** We have no information about the dependence between different inputs. To find a joint distribution, we can therefore use the Maximum Entropy approach, which in this case, as is well known, results in independence. Thus,

$$P_I(r | x) = \prod_{i=1}^n P_{I,i}(r_i | x_i).$$

Once we get these values, we can take

$$P(u | x) = \sum_{r=(1,\dots,1)}^{(R_1,\dots,R_n)} \frac{P_C(r | u)}{\int P_C(r | u') du'} \cdot P_I(r | x).$$

**Comparison with fuzzy formulas.** Similar to the case of a single input, the Bayesian case corresponds to situation when the “and”-operation is the product.

However, in this case, we can go beyond the product and still stay within the probabilistic approach: namely, instead of using the independence assumption, we can use *copulas* to combine the probabilities  $P(x_i | r)$  corresponding to different inputs  $x_i$ ; see, e.g., [7–9].

## 6 Takagi-Sugeno Case

**General description.** Let us now consider the case when the expert rules have the form “if  $A_r(x)$ , then  $u = f_r(x_1, \dots, x_n)$ ”, where  $A_r(x)$  stands for

$$“A_{1,r_1}(x_1) \text{ and } \dots \text{ and } A_{n,r_n}(x_n).”$$

**Analysis of the problem.** In this case, we first use the formulas from the previous section to compute the probabilities  $P_I(r | x)$ .

The actual control is thus equal to  $f_r(x)$  with probability  $P(r | x)$ . Similarly to the simplest case, the smallest possible expected loss is when  $\bar{u}$  is equal to the mean value of control, i.e., in this case, to

$$\bar{u} = \sum_{r=(1,\dots,1)}^{(R_1,\dots,R_n)} P_I(r | x) \cdot f_r(x).$$

**Comparison with the fuzzy case.** In the fuzzy case, we have a similar formula, the main difference is that the values  $P(r | x)$  are now computed differently.

## 7 General Conclusions

In this paper, we have shown that Bayesian approach can be efficiently used for solving intelligent control problems.

- In some cases, we get formulas which are very similar to the corresponding fuzzy formulas – and thus, provide the desired probabilistic justifications for the semi-heuristic formulas of fuzzy control.
- In some other cases, we have slightly different formulas – and, e.g., when different rules correspond to different levels of uncertainty, these different formulas more adequately represent expert knowledge.
- In yet other cases, the formulas of fuzzy control are more flexible and thus allow us to describe a higher quality control (e.g., more stable or more smooth).

Each approach – fuzzy and Bayesian – has its own advantages and limitations. It would be great to combine the advantages of fuzzy and Bayesian approaches.

## References

- [1] G. Klir, B. Yuan, *Fuzzy Sets and Fuzzy Logic*, Prentice Hall, Upper Saddle River, New Jersey, 1995.
- [2] H.T. Nguyen, E.A. Walker, *A First Course in Fuzzy Logic*, Chapman and Hall/CRC, Boca Raton, Florida, 2006.
- [3] L.A. Zadeh, *Fuzzy sets, Information and Control* 8 (1965) 338-353.
- [4] E.T. Jaynes, G.L. Bretthorst, *Probability Theory: The Logic of Science*, Cambridge University Press, Cambridge, UK, 2003.
- [5] H.T. Nguyen, V. Kreinovich, *Methodology of fuzzy control: an introduction*, In: H.T. Nguyen, M. Sugeno (eds.), *Fuzzy Systems: Modeling and Control*, Kluwer, Boston, Massachusetts (1998), 19-62.
- [6] M.H. Smith, V. Kreinovich. *Optimal strategy of switching reasoning methods in fuzzy control*, In: H.T. Nguyen, M. Sugeno, R. Tong, R. Yager (eds.), *Theoretical Aspects of Fuzzy Control*, J. Wiley, New York (1995), 117-146.
- [7] P. Jaworski, F. Durante, W.K. Härdle, T. Ruchlik (eds.), *Copula Theory and Its Applications*, Springer Verlag, Berlin, Heidelberg, New York, 2010.
- [8] A.J. McNeil, R. Frey, P. Embrechts, *Quantitative Risk Management: Concepts, Techniques, Tools*, Princeton University Press, Princeton, New Jersey, 2005.
- [9] R.B. Nelsen, *An Introduction to Copulas*, Springer Verlag, Berlin, Heidelberg, New York, 1999.

(Received 30 August 2016)

(Accepted 29 October 2016)