ELECTRONICS AND ELECTRICAL ENGINEERING
ISSN 1392 – 1215 2007. No. 5(77)
ELEKTRONIKA IR ELEKTROTECHNIKA

AUTOMATION, ROBOTICS
T125
AUTOMATIZAVIMAS, ROBOTECHNIKA

Application of Fuzzy-Sets Integral in Expert Systems

A. Dervinienė, V. Bagdonas, J. Daunoras
Department of Control Technology, Kaunas University of Technology,
Studentų st. 48-324, LT-51367 Kaunas, Lithuania, tel.: +370 37 300294, e-mail: alma.derviniene@ktu.lt

Introduction

For the process control, for the consummation level estimation of any aim, for the prognosis of receivable solution results and etc., not always we can find and use the analytical methods and optimisation methods based on its. Often methods based on expert estimations are used, that is, human intellect is being used as a measuring instrument. But the results of following methods can be far-off optimal. For error estimation the methods of mathematical statistics are suggested in literature [1]. But for this, usually, we haven't enough data.

The more complicated problem is available than to appeal to a single expert opinion could be too risky (for example, deficit of qualified specialist) and we need to use methods based on collection and processing of group experts' estimations. The two typical cases are possible:

1. Experts group is *large enough*. The criterion "*large enough*" depends on the analyst (the person recipient final decision) available information and demands: the reliance probability and the reliance interval of result and so on. The experts group is enough large or not also depends on the experts' data distribution.

2. The experts group consists of 2-7 persons (common situation in practise).

The robust algorithms for data processing based on defiance of "marginal" estimations are applicable in the first case (when number of experts is large enough).

To apply the robust algorithms for data processing is too risky in the second case (the "marginal" estimations is not necessarily are less reliable). To know the single expert estimations risk level (the expert's weight or competence [2] coefficient) it is desirable in this case. The harmonious methodology of experts' competence rates finding is undefined at this moment. We try to fill the existing gap in this paper.

Estimation's criterions of the single (unique) expert decisions risk level

For estimation of expert facilities the following criterions are offered [3]:

1) *consistency* of estimation;
2) *transitivity* of estimation;
3) *availability of complex strategies.*

These criterions applicable if the *excess* of expert's data is obtainable. The first (consistency) criterion is usable if the expert gives the repeated estimation of the same object. Very important, that the second estimation shouldn't be under influence of the first estimation (this is requirement to methodology of the experts' data collection).

Estimation of expert decisions consistency

These rates (subject to usable data) can estimate the **consistency** (v):

1. When the expert refers tolerant intervals in $[a,b]$: first (x_{j1}, x_{j2}) and repeated (y_{j1}, y_{j2}) (see Fig. 1).

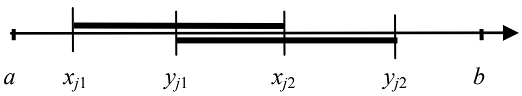

Fig. 1. Expert's referred intervals

In that case the consistency rate is presented

$$v_j = \beta_j \gamma_j = \frac{2\left|(x_{j1}, x_{j2}) \cap (y_{j1}, y_{j2})\right|}{(x_{j2} - x_{j1}) + (y_{j2} - y_{j1})} \times \left[1 - \frac{(x_{j2} - x_{j1}) + (y_{j2} - y_{j1})}{2(b-a)}\right], \quad (1)$$

where $\beta_j = \frac{2\left|(x_{j1}, x_{j2}) \cap (y_{j1}, y_{j2})\right|}{(x_{j2} - x_{j1}) + (y_{j2} - y_{j1})}$ is estimations conjunction rate (it is equals 1 if (x_{j1}, x_{j2}) and (y_{j1}, y_{j2}) are fully conjunct, or 0 - if not crossed);

$\gamma_j = \left[1 - \frac{(x_{j2} - x_{j1}) + (y_{j2} - y_{j1})}{2(b-a)}\right]$ show the relative precision (distinctness) of j-th expert's estimations;

$|(x,y)|$ – the potency of set (*x*, *y*) (when Lebego criterion is usable, $|(x,y)| = y - x$); $\cap$ - symbol of sets crossing.

2. When the expert refers “pessimistic”, “most reliable” and “optimistic” estimations from interval $[a,b]$.

In fact, the expert gives first and repeated estimations submitted in a “triangle” membership function: $\mu_{j1}(x)\big|_{x\in[a,b]}$ and $\mu_{j2}(x)\big|_{x\in[a,b]}$. In that case the consistency rate is presented

$$v_j = \beta_j\gamma_j = \max_{\alpha\in[0,1]}[\alpha\|(H_\alpha(\mu_{j1}(x))\cap H_\alpha(\mu_{j2}(x))\neq\varnothing]\times$$
$$\times\left[1-\frac{|H_0(\mu_{j1}(x))|+|H_0(\mu_{j2}(x))|}{4(b-a)}\right], \qquad (2)$$

where $H_\alpha(\mu_j(x))$ - α-level set of fuzzy number $\mu_j(x)$; $H_0(\mu_j(x))$ - 0-level set of fuzzy number $\mu_j(x)$.

3. When the expert gives fuzzy numbers submitted in a “trapezoid” membership function.

Sometimes a try to indicate the “pessimistic”, “most reliable” and “optimistic” estimations can make some troubles to expert. From psychological point of view a more convenient way is where an expert is asked to point out: interval (x_1, x_2), where the most typical value of the criterion is, by the opinion of an expert; and subjective probability *p*, under which the expert is right. In that case the estimation is presented

$$X_i = ((x_1, x_2), p). \qquad (3)$$

Estimation equation (3) can be changed with a trapezoid type membership function. In that case the consistency rate is presented

$$v_j = \beta_j\gamma_j = \left(\frac{2\int_X(\mu_{j1}(x)\wedge\mu_{j2}(x))dx}{\int_X\mu_{j1}(x)dx+\int_X\mu_{j2}(x)dx}\right)^{0,5}\times$$
$$\times\left[1-\frac{\int_X\mu_{j1}(x)dx+\int_X\mu_{j2}(x)dx}{2(b-a)}\right]. \qquad (4)$$

„$\wedge$“ means „minimum“.

Estimation of expert competence

For estimation of expert facilities the criterion of **availability of complex strategies** is very important. This criterion express expert’s ability to generalize and aggregate available information.

This criterion is in use with **transitivity** criterion [2]. The value of complex expert’s **competence** criterion (weight coefficient) is founded according to scheme (algorithm):

The set $S_i | i = \overline{1,n}$ of an object, action or process *A* goals is formed. The set *S* has to be full, but minimal; it is desirable that the goals shouldn’t be overlap.

After interviewing of each expert such information is obtainable:

a) for estimation of each *A* (there *A* - object, action or process) goal S_i its subjective ***importance rate*** $g_i \in [0,1]$ is suggested;

b) the ***consummation rate*** h_{s_i} (further h_i) of each goal S_i is prognosticated;

c) the system (heuristics) of complex rate (criterion) e_A for aggregation of each *A* goal importance’s and consummation’s estimations is offered;

d) the e_A value by intuition (but not calculated), what further is called h_A, is proposed.

Rates $h_i\big|_{i=\overline{1,n}}$ and e_A have to be quantitative, that is, measurable, calculable or subjectively rateable. They mostly are presented in normal numbers: $e_A,\ h_i \in [0,1]$.

Estimation of expert competence (availability of complex strategies and transitivity of offered estimations) is called α_j (*j* - it is index of the expert). The expert’s competence rate α_j (subject to usable data) is founded the same as consistency rate (according formulas (1) - (2), (4)) where h_A is usable instead primary estimation, e_A - instead repeated estimation. The values h_A and e_A is available operating with a), b) and c) points data.

The estimation of goals importance (g_i) and consummation (h_i) rates don’t cause the some problems, while the formalization of expert heuristics (as the way of decision search) is especially complicated task. The solution of this task in any general way should matter the formalization of human thought.

The experimentation to formalize heuristics is founded in literature [4]. That is usually associated with solution of concrete problem. This also involves in the expert researches.

Such schemes of g_i and h_i estimations aggregation are mostly met [2]:

1) the expert try to estimate average e_A of goals $S_i | i = \overline{1,n}$ consummations h_i;

2) for data aggregation (the evaluation of e_A value) the expert underline the goals with high consummation rate h_i;

3) for data aggregation the expert underline the goals with high product of importance and consummation rates $g_i h_i$;

4) for data aggregation the expert underline the goals with low consummation rate h_i;

5) for data aggregation the expert underline the goals with low product of importance and consummation rates $g_i h_i$.

Sometimes an expert himself can suggest the scheme (heuristics type) of importance and consummation rates aggregation into one complex rate (criterion) e_A. The additional test of the heuristics type identification is suggested.

The first scheme of estimations aggregation is mostly met and best researched. In fact - it is result of formula calculating h_i medium value:

$$e_A = \sum_{i=1}^{n} g_i h_i \;//\; \sum_{i=1}^{n} g_i \,. \qquad (5)$$

Symbol "//" is sign of division. If fuzzy numbers are dividend, then instead ordinary division its adjective division is executed:

$$(GH)_\alpha \;//\; G_\alpha = (a_1, a_2) // (b_1, b_2) =$$
$$= ((a_1 : a_2), (b_1 : b_2)),\; a_1 > 0,\, b_1 > 0, \qquad (6)$$

where $(GH)_\alpha$ and G_α - α-level sets of fuzzy numbers (*GH*) and *G* that is expressed by $\sum_{i=1}^{n} g_i h_i$ and $\sum_{i=1}^{n} g_i$.

Formula (5) is applied when (by the opinion of an expert) the goals $S_i \big| i = \overline{1, n}$ of *A* (there *A* - object, action or process) are independent and its consummations h_i - additive. Applying to this low indistinct Choquet integral, the more universal form (tolerant to interdependence of goals) can be got:

$$C = \sum_{i=1}^{n} (h_i \times [f(S_i, \ldots, S_n) - f(S_{i+1}, \ldots S_n)])\,; \qquad (7)$$

$$f(S_i, \ldots S_n) = 1\,; \qquad (8a)$$

$$f(S_i, \ldots S_n) \geq f(S_j, \ldots S_n),\ \text{jei}\ j > i\,; \qquad (8b)$$

$$f(0) = f(S_{n+1}) = 0\,, \qquad (8c)$$

where $f(S_i, \ldots S_n)$ – indistinct measure of goals $(S_i, S_{i+1}, \ldots, S_n)$ importance, when conditions (8) is supplied.

The expert has to form estimations $f(S_i, \ldots S_n)$ according to test values g_i. Formula (7) is other notation of formula (5).

For description (formalisation) all others schemes of an expert estimations g_i and h_i aggregation into one complex rate e_A is recommendable to apply parameter g_λ and Sugeno integral based on it.

The point of method is such when fuzzy set is given:

$$B = h_1 / g_1 + h_2 / g_2 + \ldots + h_n / g_n\,, \qquad (9)$$

where $0 \leq g_i \leq 1$.

It is possible to define for it (for set *B*) parameter g_λ, whose rating parameter λ must supply condition

$$-1 < \lambda < \infty \qquad (10)$$

and can be found as follows:

$$\frac{1}{\lambda}\left[\prod_{i=1}^{n}(1 + \lambda g_i) - 1\right] = 1 \text{ or } \frac{1}{\lambda}\left[\prod_{i=1}^{n}\left(1 + \lambda g_i h_i\right) - 1\right] = 1\,. \quad (11)$$

The Sugeno integral of fuzzy set (9) is expressed:

$$S = \sup_{\alpha \in [0,1]} \min\{\alpha; g_\alpha)\,. \qquad (12)$$

Using parameter g_λ, g_α in (12) is as follows:

$$g_\alpha = \frac{1}{\lambda}\left[\prod_{i | h_i \geq \alpha}^{n}\left(1 + \lambda g_i\right) - 1\right], \qquad (13a)$$

$$g_\alpha = \frac{1}{\lambda}\left[\prod_{i | g_i h_i \geq \alpha}^{n}\left(1 + \lambda g_i h_i\right) - 1\right]. \qquad (13b)$$

In all cases the complex rate e_A is expressed

$$e_A \to S\,. \qquad (14)$$

If goals with low consummation rate h are meaning (fourth and fifth g_i and h_i aggregation schemes), then it is accented not goal consummation but its *unconsummation level*. So instead estimations $h_i \big|_{i=\overline{1,n}}$ in formulas (13a) and (13b) we need to use theirs inversion: $1 - h_i \big|_{i=\overline{1,n}}$. g_α is as follows:

$$g_\alpha = \frac{1}{\lambda}\left[\prod_{i | (1-h_i) \geq \alpha}^{n}\left(1 + \lambda g_i\right) - 1\right], \qquad (15a)$$

$$g_\alpha = \frac{1}{\lambda}\left[\prod_{i | (1-g_i h_i) \geq \alpha}^{n}\left(1 + \lambda g_i h_i\right) - 1\right]. \qquad (15b)$$

Applying formulas (15), integral (12) express the complex right estimation of goals *unconsummation* (e_A inversion). To get the estimation of complex rate (criterion) e_A, it is enough to have inversion value of formulas (15) and (12):

$$e_A \to S = 1 - \sup_{\alpha \in [0,1]} \min\{\alpha; g_\alpha)\,. \qquad (16)$$

As we have already mentioned, the expert's competence rate α_j (subject to usable data) is founded the same as consistency rate (according formulas (1) - (2), (4)), where h_A is usable instead primary estimation, e_A - instead repeated estimation.

When the expert refers h_A ir e_A as tolerant intervals in $[a, b]$, so the competence rate is presented

$$\alpha_j = \beta_j \gamma_j = \frac{2\left|(h_{j1}, h_{j2}) \cap (e_{j1}, e_{j2})\right|}{(x_{j2} - x_{j1}) + (y_{j2} - y_{j1})} \times$$
$$\times \left[1 - \frac{(h_{j2} - h_{j1}) + (e_{j2} - e_{j1})}{2(b - a)}\right]. \qquad (17)$$

When the expert refers h_A and e_A "pessimistic", "most reliable" and "optimistic" estimations from interval $[a, b]$, so the competence rate is presented

$$\alpha_j = \beta_j \gamma_j = \max_{\alpha \in [0,1]} [\alpha \| (H_\alpha(h_j(x)) \cap H_\alpha(e_j(x)) \neq \varnothing] \times$$

$$\times \left[1 - \frac{|H_0(h_j(x))| + |H_0(e_j(x))|}{4(b-a)} \right]. \quad (18)$$

When the expert gives h_A and e_A submitted in a "trapezoid" membership function, so the competence rate is presented

$$\alpha_j = \beta_j \gamma_j = \left(\frac{2 \int_X (h_j(x) \wedge e_j(x)) dx}{\int_X h_j(x) dx + \int_X e_j(x) dx} \right)^{0,5} \times$$

$$\times \left[1 - \frac{\int_X h_j(x) dx + \int_X e_j(x) dx}{2(b-a)} \right]. \quad (19)$$

The formulas (17–19) are „compatible", i.e. estimations from different experts can be comparable or otherwise processed (for example, searching of its average) even if experts operate with different types of data (estimations submitted in the real numbers intervals, in a "triangle" membership function or in a "trapezoid" membership function). Because estimation in the real numbers intervals can be interpreted as a "rectangular" membership function (as an instance of estimations submitted in a "trapezoid" membership function). The "triangle" membership function is the partial case off "trapezoid" membership

Conclusions

1. For the decision-making based on expert researches the risk level dependent on expert's facilities (defined with the consistency and competence criterions) must be known.

2. To get the upper reliability of operation prognosis it is desirable to use the group expert researches based on single expert facilities evaluation.

3. The classical method of robust decisions search based on elimination of "marginal" estimations is too risky when number of experts is not enough big.

4. For estimation of an expert competence rate, identification of the right scheme (heuristics) of this expert subjective partial estimations aggregation into one complex rate (criterion) is very important.

5. The mostly met schemes of partial estimations aggregation into one complex rate (criterion) can be described with fuzzy integrals.

References

1. **Puškorius S.** Risk Assessment of Decisions Making // Public policy and Administration. – Kaunas: Technologija, 2006. – No. 16. – P. 46–52.
2. **Bagdonas V., Daunoras J., Dervinienė A.** Methodology of Operation Efficiency Prognosis // Electronics and Electrical Engineering. – Kaunas: Technologija, 2006. – No. 5(69). – P. 95–98.
3. **Larichev O. I.** Decision Making as a Research Branch: Methodological Problems. – 7 P. Access on the web: http://www.raai.org/library/papers/Larichev/Larichev.pdf.
4. **Tsarkov D., Horrocks I.** Ordering Heuristics for Description Logic Reasoning // Proceedings of International Joint Conference of Artificial Intelligence 2005, Edinburg, Scotland, 30 July – 5 August 2005. – P. 609–614. Access on the web: http://www.ijcai.org/ papers/1352.pdf.

Submitted for publication 2007 03 21

A. Dervinienė, V. Bagdonas, J. Daunoras. Application of Fuzzy-Sets Integral in Expert Systems // Electronics and Electrical Engineering. – Kaunas: Technologija, 2007. – Nr. 5(77). – P. 45–48.

In more situations for decision-making the methods based on expert estimations are used, that is, human intellect is being used as a measuring instrument. The decision risk level dependent on expert's facilities (defined with the consistency and competence criterions) must be known in the decision-making process. The methodology of experts decisions risk level estimation's is discussed. The methodology is based on search and use of experts' consistency and competence rates, what can be get applying excess of the experts offered information. The application of Sugeno and Choquet fuzzy integrals for processing group expert estimations are offered. Ill. 1, bibl. 4 (in English; summaries in English, Russian and Lithuanian).

А. Дервинене, В. Багдонас, И. Даунорас. Применение нечетких интегралов в экспертных системах // Электроника и электротехика . – Каунас: Технология, 2007. – № 5(77). – С. 45–48.

На практике для принятия решений часто используются методы экспертных оценок, т. е. интеллект человека-специалиста используется как устройство измерения. В процессе принятия решений желательно учитывать показатель уровня их риска, который зависит от возможностей эксперта, характеризируемых критериями устойчивости и компетенции. Анализируется методика оценки уровня риска экспертных решений, предусматривающая определение и учет индивидуальных показателей устойчивости и компетенции экспертов, получаемых на основе избыточности экспертных данных. Методика основана на применении нечетких интегралов Sugeno и Choquet. Ил. 1, библ. 4 (на английском языке; рефераты на английском, русском и литовском яз.).

A. Dervinienė, V. Bagdonas, J. Daunoras. Neraiškiųjų aibių integralų naudojimas ekspertinėse sistemose // Elektronika ir elektrotechnika. – Kaunas: Technologija, 2007. – Nr. 5(77). – P. 45–48.

Praktikoje sprendimai gana dažnai priimami taikant ekspertiniais įvertinimais grindžiamus metodus, kai žmogaus specialisto intelektas naudojamas kaip matavimo prietaisas. Priimant ekspertiniais tyrimais grindžiamus sprendimus, pageidautina žinoti jų rizikos lygmenį, kuris priklauso nuo eksperto galimybių, nusakomų pastovumo bei kompetencijos kriterijais. Nagrinėjama ekspertinių sprendimų rizikos lygmens įvertinimo metodika, grindžiama ekspertų sprendimų pastovumo ir kompetencijos rodiklių paieška ir panaudojimu. Ekspertų sprendimų pastovumo ir kompetencijos rodikliai gali būti gauti naudojant jų (ekspertų) teikiamos informacijos

perteklių. Grupiniams ekspertiniams įvertinimams apdoroti siūloma naudoti Sugeno ir Choquet neraiškiuosius integralus. Il. 1, bibl. 4 (anglų kalba; santraukos anglų, rusų ir lietuvių k.).