AGH Journal of Mining and Geoengineering • vol. 36 • No. 3 • 2012

*Edyta Brzychczy**

# MODELLING UNCERTAINTY IN AN ADVISORY SYSTEM FOR MINING WORKS PLANNING IN HARD COAL MINES**

## 1. Introduction

The literature presents several definitions of expert systems [7], however, it also consentaneously state, that these are computer software which use knowledge within the problem area for which they were designed. The software, through the adequately assumed notation (representation) contains knowledge regarding the analyzed problem (comparable with that of an expert knowledge) and via reasoning procedures enable the search for its solution.

The advisory systems are the sub-group of the expert systems, the aim of which is the analysis and providing solution for the formulated problem. The power of the advisory software lies in the knowledge encoded therein, and not in the course of the inference process [11]. Hence, in order for such systems to function properly, the quality of knowledge stored inside is important. This knowledge, and most importantly, the expert knowledge is a result of the human thought activity as well as the execution of some perception process. Hence its character is most often uncertain, inaccurate and/or fuzzy.

The classification of information stored in the expert systems has been shown in the Figure 1.

Perfect information is the information which is true without reservations; otherwise it is imperfect information. Imperfect information may be uncertain (resulting from uncertain data or knowledge) or incomplete (incomplete problem description). In the work [9] uncertainty is pointed out as one of the two most important forms of partial ignorance in the knowledge base systems (the second one being inaccuracy).

* AGH University of Science and Technology, Krakow

** The paper is supported by Polish Ministry of Science and Higher Education as research project No. N N524 468939.

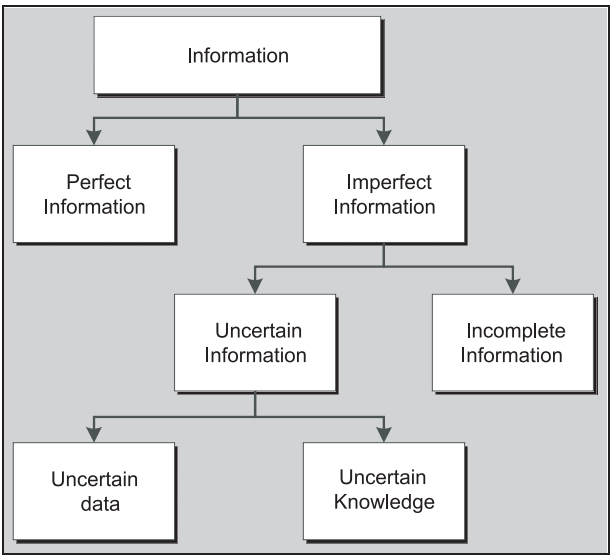


**Fig. 1.** Information in expert systems [6]

Knowledge uncertainty in the expert systems results mainly from [6]:

— random nature of events,
— expert/experts' convictions about the analyzed event,
— imprecise or incomplete information.

Considering the fact, that the uncertainty has essential impact on the deduction process and the results of advisory system operations, the models representing uncertainty should be selected adequately to the character of the considered problem. This article presents the considerations regarding the description of uncertainty in a knowledge base of the advisory system designed for the planning of the preparatory and exploitation works in the hard coal mines, which was mentioned in the work [2].

## 2. Uncertainty models in the expert systems

The literature presents the following models enable modelling the uncertainty in the expert systems [13, 17]:

— probabilistic models (Bayes),
— models using certainty factors,
— models based on the belief functions (Dempster-Shafer theory),
— fuzzy models.

The models enumerated above belong to the quantitive methods of expressing uncertainty. Probabilistic methods and methods using the certainty factors give one measure of uncertainty (successively, probability and certainty factors), while other give a set of certain values.

In the next section, as the most popular models, certainty factors and fuzzy models will be presented.

### 2.1. Models using the certainty factors

Knowledge recorded in a knowledge base is expressed by rules with the following form:

$$A \Rightarrow B \tag{1}$$

which shall be read: IF $A$ THEN $B$. In case of the complex rules the record is the following:

$$A_1 \wedge A_2 \wedge A_3 \ldots \wedge A_n \Rightarrow B \tag{2}$$

The expressions to the left of the sign are referred to as premises (conditions), whereas those to the right are called conclusions (inferences).

Models using the certainty factors constitute simple and very popular description of uncertainty. The certainty factors allow for the expression of subjective assessment of rules and conditions reliability, made by the creator or user of knowledge base. However, in a number of cases these might be verified experimentally [13]. The first meaningful use of this method was MYCIN expert system, described among others in [15], used to diagnose blood diseases.

For a condition $A$, a number called a certainty measure $C(A)$ is associated [6]. The certainty measure is defined as follows: $C(A) = 1$ if $A$ is known to be true, $C(A) = -1$ if $A$ is known to be false, and $C(A) = 0$ if nothing is known about $A$. Every rule has an associated number from the interval [–1, 1], called a certainty factor, denoted $CF$.

Algebra of certainty factors consists in several rules [6, 8, 13]:

1) For each A condition, the initial value $C(A) = 0$. In introducing conditions into the knowledge base their certainty factors are declared by the user prior to inference or during it. For the conditions resulting from inference these are determined by considering the uncertainty propagation through inference.
   Moreover $C(\neg A) = 1 - C(A)$.

2) For the conclusion of the R rule:

$$A \Rightarrow B \text{ z } CF(R) = x \text{ i } C(A) = 1,\ C(B) = x \tag{3}$$

3) Premise certainty factor being a conjunction of several conditions equals the minimum value of certainty grades of these conditions, which can be noted as:

$$C(A_1 \wedge A_2) = \min\,(C(A_1), C(A_2)) \tag{4}$$

4) In case of alternative conditions forming a rule premise, the certainty factor of this premise shall be determined as follows:

$$C(A_1 \vee A_2) = \max\,(C(A_1), C(A_2)) \tag{5}$$

5) The certainty factor of the rule conclusion is the product of the rule certainty factor and certainty factor of the conditions conjunction of this rule:

$$CF(B) = C(A_1 \wedge A_2) \cdot CF(R) \tag{6}$$

In the standard case, the rule certainty factor is a kind of strengthening determining the influence of rule conditions certainty on the rule conclusion certainty.

Advantages of this method for uncertainty description undoubtfully listed can be the following [13]:

— possibility of ex post verification of the assumed certainty factors of conditions and rules, based on historic data describing results of decisions based on the selected certainty factors;
— possibility to use them to express user preferences;
— consistency and intuitive triviality of the inference results using the certainty factors (the results turn out to be consistent with the common-sense, non-formal inference practically always applicable in the uncertainty conditions).

Main disadvantages of models with certainty factors are: lack of deeper theoretical justification and the necessity to take assumptions on the independence of conditions of rules with the same conclusions.

### 2.2. Fuzzy models

Fuzzy models make it possible to model uncertainty related to imprecision by determining, to what extend an object corresponds to the imprecise description. Fuzzy logic is an extension of the classical Aristotle's logic in case of imprecise, ambiguous and indistinct descriptions. In a classic set theory the expression of the $x \in A$ type where $x$ denotes element, whereas $A$ is a set of elements — might be true or false [13]. In the fuzzy sets theory, the affiliation of $x$ element to $A$ set might be partial and as such, might be described thanks to so called membership function, which to each of the $x$ element of the considered space assigns a real number from the [0, 1] range called grade of membership of this element to the $F$ fuzzy set.

Thus the characteristic feature of the fuzzy interference is the replacement of classic bivalent characteristic function of $x$ object membership to $A$ set [8]:

$$f(x) = \begin{cases} 1 & \text{gdy} \quad x \in A \\ 0 & \text{gdy} \quad x \notin A \end{cases} \tag{7}$$

with the „fuzzy" function:

$$\mu_A : A \rightarrow [0,1] \tag{8}$$

A fuzzy set is defined by the pair [7]:

$$\{X, \mu_A\} \tag{9}$$

where:

$X$ — is a space of the considered objects,

$\mu_A$ — is a function determining for each element from $X$ membership grade to $A$ set (as described in formula (8)).

Each set in a classic sense is in the fuzzy formulation the so called acute set, with which the membership function is connected, according to formula (7).

The fuzzy rules shall be recorded in the form:

if $X$ is $A$ then $Y$ is $B$

where:

$X$, $Y$ — variables,

$A$, $B$ — linguistic values.

The fuzzy inference [8] consists of converting the quantitative variables into the linguistic notions (variables) and then modelling system on the basis of rule base and finally converting results again to the quantitative variables.

For the adequately fuzzy inference, there appeared a need to define set operations on fuzzy sets, which are generalization of operations defined for acute sets. Therefore [7, 13]:

— Fuzzy set $\{X, \mu_B\}$is an empty set when and only when, for all x the $\mu_B(x) = 0$,

— Complement of fuzzy set $\{X, \mu_B\}$is a set$\{X, 1 - \mu_B\}$,

— Fuzzy set $\{X, \mu_B\}$is contained in fuzzy set$\{X, \mu_C\}$,if $\mu_B(x) <= \mu_C(x)$,

— Product of fuzzy sets $\{X, \mu_B\}$ and $\{X, \mu_C\}$ is a set $\{X, \mu_W\}$, where $\mu_W(x) = \min(\mu_B(x), \mu_C(x))$,

— Sum of fuzzy sets $\{X, \mu_B\}$ i $\{X, \mu_C\}$is a set $\{X, \mu_W\}$, where $\mu_W(x) = \max(\mu_B(x), \mu_C(x))$.

Fuzzy sets and operations performed on them together with various methods of defuzzification (sharpening) results are presented in detail in literature (ex. [6, 10]) together with practical use [5, 14].

## 3. Advisory system supporting the planning of the preparatory and exploitation works in the hard coal mines

The aim of the designed advisory system is to provide support for the designers of the mining production, within the scope [2]:

— the selection of equipment with respect to geological and mining conditions of the planned excavations,
— combining mining machinery into longwall complexes,
— determining the production results in the planned excavations.

The simplified scheme of the designed system has been presented in Figure 2.
The main component of the system, i.e. the knowledge base, will consist of [1]:

— a database — containing detailed information with regard to first working and mining works in the past,
— a model base — containing models of advances of first working and mining works (stochastic models),
— a rule base — containing the rules of equipment selection with respect to the excavation conditions.

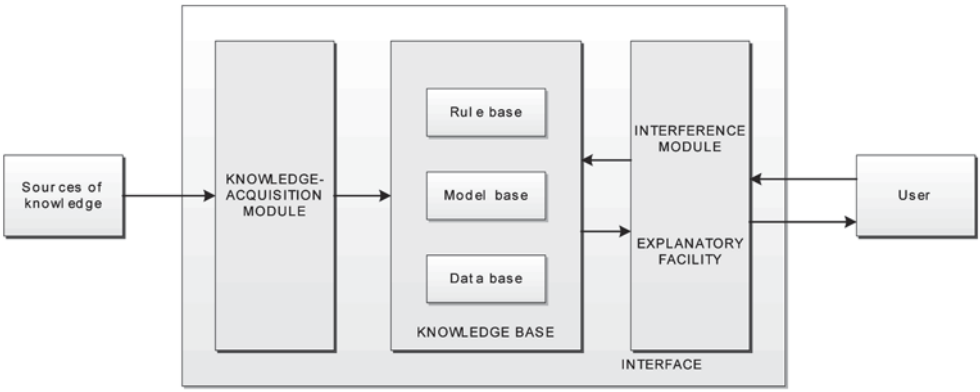


**Fig. 2.** Designed advisory system
(Source: own development)

The information (facts, statements) regarding longwalls and preparatory works performed in the past, are stored in data base. These data pertain to the geological-mining conditions, technical-organizational conditions and equipment, as well as the adequate model regarding the advance of the carried out works, which is stored in the model base. Whereas, in the

rule base the rules are stored (dependences between facts) acquired from experts and determined with the use of machines learning algorithms (data mining). These rules pertain to the selection of equipment for the excavation conditions as well as combining machines and equipment together to make sets.

Considering the uncertainty of information and knowledge in the designed knowledge base, in Table 1 the selected characteristics (quantitive and qualitative) for the longwall were presented.

TABLE 1

**Selected characteristics (quantitative and qualitative) of the longwall**

| Data | Unit | Quality of information |
|---|---|---|
| Thickness of the seam | m | U |
| Seam layer (0, I, II, III) | – | C |
| Fall of the roof | m | U |
| Impurity in the longwall cross-section | m | U |
| Transverse slope of the longwall | ° | U |
| Longitudinal slope of the longwall | ° | U |
| Floor class (I, II, III) | – | CC |
| Roof class (I–IV) | – | CC |
| Coal workability | – | U |
| Methane hazard category (I–IV) | – | CC |
| Degree of water hazard (I, II, III) | – | CC |
| Rockburst hazard degree (I, II, III) | – | CC |
| Temperature of formation | °C | U |
| Group of coal combustion hazard (I–V) | – | CC |
| Height of the longwall | m | U |
| Length of the longwall | m | U |
| Roof maintenance (none, filing) | – | C |
| Direction of the exploitation (to the shaft, from the shaft) | – | C |
| Type of excavation (longitudinal, transverse, other) | – | C |
| Number of maintained entries | pc | C |
| Shearer type | | C |
| Conveyor type | | C |
| Mechanized support type | | C |

(Source: own development)

In the table also comments pertaining to the possible quality of the information were presented, according to the following rules:

— certain information — *C*,
— conditionally certain information — qualitative parameter established in accordance with the certain administrative regulations, still based on variables which might change on the excavation length — *CC*,
— uncertain information — quantitative parameter which might by characterized by the distinct variability on the excavation length or lack of measurement accuracy — *U*.

Data gathered in the database are used to set the rules feeding the rule base and multi-dimensional comparative analysis, it is therefore important to adequately involve the uncertainty thereof.

The important problem related to the quality of knowledge also is the certainty of rules, under which the inference process shall be carried out. The rules originating from the machine learning algorithms (data mining software) are characterized by the quality factors (support, confidence factor). The rules coming from experts might be characterized by different level of certainty, which results from the informal character of human reasoning process and the lack of dependable inference schemes. Expert opinions often vary, both as regards the statements evaluation and the conclusions and require evaluation of their certainty.

In further part of the article the selected models shall be presented, which will be implemented in the designed advisory system to represent uncertainty in the knowledge base.

## 4. Proposal to include uncertainty in the designed expert system

The selection of adequate models in the designed system shall first of all be conditioned by the statements' uncertainty and the rules' uncertainty.

Uncertainty of statements (presented in Table 1) depends on the character of the parameter (qualitative or quantitative variable) and its possible changeability, but mostly the accuracy of the measurement.

Certain statements will be exploited to generate rules using selected data mining techniques (eg. association rules to determine the most common connections between devices in a longwall complex — Table 2).

Rules obtained in this way have specific indicators of quality (e.g., support, confidence factor) — but their value depends on the diversity of the training set — hence they will be further evaluated by domain experts (by certainty factors).

The statements, which are characterized by a degree of uncertainty, will be expressed in terms of fuzzy sets and will be used to determine the fuzzy rules.

TABLE 2

**Selected rules for equipment combinations (shearer – conveyor)**

| Rule ID | IF | ==> | THEN | Support, % | Confidence factor, % |
|---|---|---|---|---|---|
| 1 | KSW 1140 | ==> | Rybnik 1100 | 5,15 | 55,55 |
| 2 | KGS 600 | ==> | Rybnik 750 | 3,09 | 50,00 |
| 3 | KGE 710 | ==> | Glinik 298/800 | 2,06 | 66,66 |
| 4 | KSW 500 | ==> | Longwall-724/AFC | 3,09 | 100,00 |
| 5 | Joy 6LS | ==> | Glinik 298/800 | 2,06 | 100,00 |

(Source: own development)

Fuzzy rules could be generated according to the Wang-Mendel method described in [10, 18]. This method consists of five steps:

1) Divide the input and output spaces into fuzzy regions.
   Domain intervals of $x_1$, $x_2$ and $y$ are [$min\ x_1$, $max\ x_1$], [$min\ x_2$, $max\ x_2$], [$min\ y$, $max\ y$] respectively. In the next stage divide each domain interval into $N$ regions ($N$ can be different for each variable). Assign for each region a fuzzy membership function.

2) Generate fuzzy rules from the given data pairs.
   Determine the degrees of membership of the given factors $x_1(i)$, $x_2(i)$, and $y(i)$ in different regions. Obtain one rule from one pair of data, taking into account the maximum value in one region of each variable.

3) Assign a degree to each rule.
   Degree of the rule: „IF $x_1$ is $A$ and $x_2$ is $B$ THEN $y$ is $C$" can be defined as:

$$D_{(R)} = m_A(x_1)\ m_B(x_2)\ m_C(y) \tag{10}$$

   where:
   $m_i$ — denotes the value of the membership function for each region of the variable.

4) Create a fuzzy rule base.
   Complete the boxes of the base with fuzzy rules. If there is more than one rule in one box, use the rule which has the maximum degree.

5) Determine a mapping based on the combined fuzzy rule base (for numerical outputs).

For example, we would like to find rules for a longwall shearer selection where excavation conditions are expressed by the height of the longwall ($h$) and the workability of the coal ($f$).

The parameters of the excavation can be expressed as fuzzy sets with the following membership functions:

— for coal workability (the f coefficient is used according to [12]:

$$\mu_{(good\ workability)} = \begin{cases} 1 \ \text{dla} \ \ f \le 0 \\ \dfrac{1,2-f}{1,2} \ \text{dla} \ \ f \le 1,2 \\ 0 \ \text{dla} \ \ f > 1,2 \end{cases}$$

$$\mu_{(hard\ workability)} = \begin{cases} 0 \ \text{dla} \ \ f \le 1,2 \\ \dfrac{f-1,2}{0,4} \ \text{dla} \ \ 1,2 \le f \le 1,6 \\ \dfrac{1,6-f}{0,4} \ \text{dla} \ \ 1,6 \le f \le 2 \\ 0 \ \text{dla} \ \ f > 2 \end{cases} \tag{11}$$

$$\mu_{(very\ hard\ workability)} = \begin{cases} 0 \ \text{dla} \ \ f \le 2 \\ \dfrac{f-2}{0,2} \ \text{dla} \ \ 2 \le f \le 2,2 \\ 1 \ \text{dla} \ \ f > 2 \end{cases}$$

— for the height of the longwall:

$$\mu_{(low\ longwall)} = \begin{cases} 1 \ \text{dla} \ \ h \le 1 \\ \dfrac{2,75-f}{1,75} \ \text{dla} \ \ h \le 2,75 \\ 0 \ \text{dla} \ \ h > 2,75 \end{cases}$$

$$\mu_{(middle\ longwall)} = \begin{cases} 0 \ \text{dla} \ \ h \le 1 \\ \dfrac{h-1}{1,75} \ \text{dla} \ \ 1 \le h \le 2,75 \\ \dfrac{4,5-h}{1,75} \ \text{dla} \ \ 2,75 \le h \le 4,5 \\ 0 \ \text{dla} \ \ h > 4,5 \end{cases} \tag{12}$$

$$\mu_{(high\ longwall)} = \begin{cases} 0 \ \text{dla} \ \ h \le 2,75 \\ \dfrac{h-2,75}{1,75} \ \text{dla} \ \ 2,75 \le h \le 4,5 \\ 1 \ \text{dla} \ \ h > 4,5 \end{cases}$$

Designated fuzzy sets are shown in Figures 3 and 4.

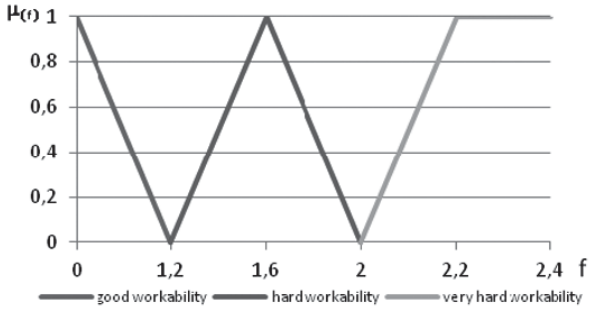


**Fig. 3.** Fuzzy sets for coal workability

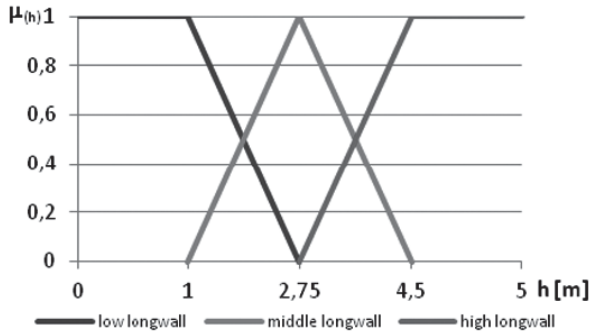


**Fig. 4.** Fuzzy sets for the height of the longwall

In the original Wang-Mendel method, input and output data are represented as numerical values. In the designed system, the output value will be used to describe the type of equipment for the longwall working conditions (discrete set). A sample of a discrete fuzzy set is presented in Figure 5.

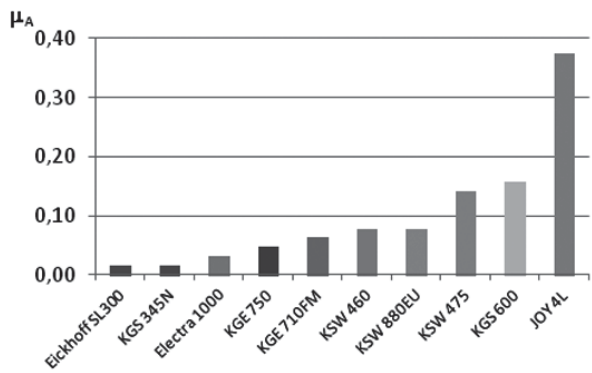


**Fig. 5.** Discrete fuzzy set of longwall shearers for good coal workability and middle longwall

On the basis of a set of 279 examples of longwalls, the designated fuzzy sets (Fig. 3 and 4) and discrete fuzzy sets of longwall shearers, which can be used under certain conditions, result in the obtainment of 279 fuzzy rules with different degrees of rule (10). After selection of the rules with the highest rate $D_{(R)}$, the parameter space has been covered by the selected rule presented in Table 3.

TABLE 3
**Examples of rules for longwall shearer selection**

| | **good workability** | **hard workability** | **very hard workability** |
|---|---|---|---|
| **low longwall** | KSW 460 (0,26)<br>KGS 600 (0,25) | Eickhoff SL300 (0,19)<br>KGE 710FM (0,16) | KSW 460 (0,19) |
| **middle longwall** | JOY 4L (0,31) | Eickhoff SL300 (0,18)<br>KGS 600 (0,16) | KSW 880EU (0,23) |
| **high longwall** | JOY 4L (0,41) | KSW 620EZ (0,36)<br>Electra 1000 (0,27) | KGE 750 (0,36)<br>JOY7LS6 (0,24) |

(Source: own development)

In the table a great number of possible assignments (resulting from other rules) is omitted because of the lower degree of the rule.

Taking into account the advisory function of the designed system and the specific issues for which it is intended, the knowledge base should contain omitted rules.

This will allow for the same premises (the conditions of the excavation), and indicate the various possibilities for the use of appropriate equipment (not just one or two longwall shearers), which is consistent with the practice of the mining design process.

Due to the fact that the quality of the fuzzy rules ($D_{(R)}$) also depends on the diversity of the training set, all designated rules will be evaluated by domain experts using certainty factors.

Additionally, for an adequate description of acquired expert knowledge by the new rules, which could not be automatically extracted from the set of examples, certainty factors will be used.

## 5. Summary

Advisory systems belong to the best information technology solutions in the form of knowledge base systems. They are a result of studies based on artificial intelligence. The main task of these systems is to find a solution for the formulated problem on the basis of facts and rules recorded in the knowledge base. Effectiveness and correctness of advisory system operations depends mainly on the quality of the collated knowledge. The character of the knowledge is most often uncertain, inaccurate and/or fuzzy.

On the basis of the selected models review, a proposal concerning the uncertainty description in the designed advisory system for mining works' planning in hard coal mines was presented. The assumed fuzzy models and certainty factors for particular elements of the knowledge base, shall have an influence on the direction of the inference process and the operational effects of the system. The fuzzy approach enables modelling of the uncertainty of parameters characterizing the excavation, and can also be used for the generalization of rules, which is especially important in the case of high complexity within the modelled objects/processes. The result of works conducted within this area shall constitute the subject of further publications.

REFERENCES


[1] *Brzychczy E., Magda R., Franik T., Kęsek M., Woźny T., Napieraj A.*: An expert system for supporting mine production planning in multi-plant mining enterprises. 22nd World Mining Congress, Red. S.Eskikaya, Istambul 2011.

[2] *Brzychczy E.*: The planning optimization system for underground hard coal mines. Archives of Mining Sciences, vol.56, no 2, 2011.

[3] *Cholewa W., Pedrycz W.*: Systemy doradcze. Skrypt Uczelniany Nr 1447 Politechniki Śląskiej, Gliwice 1987.

[4] *Cichosz P.*: Systemy uczące się. Wyd. Naukowo-Techniczne, Warszawa 2004.

[5] *Grychowski T.*: Hazard Assessment Based on Fuzzy Logic. Archives of Mining Sciences vol.53, no 4, 2008.

[6] *Grzymala-Busse J.W.*: Managing Uncertainty in Expert Systems. Kluwer Academic Publishers, 1991.

[7] Inteligentne systemy w zarządzaniu. Teoria i praktyka. Red. J.S. Zieliński, Wydawnictwo Naukowe PWN, 2000.

[8] *Jagielski J.*: Inżynieria wiedzy. Uniwersytet Zielonogórski, Zielona Góra 2005.

[9] *Kruse R., Schwecke E., Heinsohn J.*: Uncertainty and Vagueness in Knowledge Based Systems, Springer-Verlag, 1991.

[10] *Łęski J.*: Systemy neuronowo-rozmyte. Wyd. Naukowo-Techniczne, Warszawa 2008.

[11] *Mulawka J.*: Systemy ekspertowe. Wyd. Naukowo-Techniczne, Warszawa 1996.

[12] *Napieraj A.*: Warunki geologiczno-górnicze a systemy mechanizacyjne stosowane w przodkach ścianowych kopalń — wybrane zagadnienia W: Materiały Krakowskiej Konferencji Młodych Uczonych 2008 (Sympozja i Konferencje KKMU; nr 3), Kraków 2008.

[13] *Niederliński A.*: Regułowo-modelowe systemy ekspertowe rmse, Wyd. Pracowni Komputerowej Jacka Skalmierskiego, Gliwice 2006.

[14] *Rutkowski L.*: Metody i techniki sztucznej inteligencji. Wydawnictwo Naukowe PWN, Warszawa 2005.

[15] *Shortliffe E.*: Computer Based Medical Consultations: MYCIN. Elsevier, New York 1976.

[16] *Starzyńska W.*: Statystyka praktyczna. Wyd. Naukowe PWN, Warszawa 2002.

[17] *Walley P.*: Measures of uncertainty in expert systems. Artificial Intelligence, 83, 1996.

[18] *Wang L.X., Mendel J.M.*: Generating fuzzy rules by learning from examples. IEEE Transactions in Systems, Man and Cybernetics, vol. 22, no 6, 1992.