Z. GÓRNY*, S. KLUSKA-NAWARECKA*, D. WILK-KOŁODZIEJCZYK**

ATTRIBUTE-BASED KNOWLEDGE REPRESENTATION IN THE PROCESS OF DEFECT DIAGNOSIS

ATRYBUTOWA REPREZENTACJA WIEDZY W PROCESIE DIAGNOSTYKI WAD

The problem of casting defects diagnosis includes several distinct areas in which only the relevant global perspective allows for a satisfactory solution of the diagnostic task. One of these areas relates to knowledge about the parameters of the possible occurrence of defects. This knowledge is the more valuable, the more extensive it is, and the more numerous are the sources it has been acquired from. It should be combined into a coherent whole and given the form allowing for component processing. The paper proposes a solution in which the methods of knowledge engineering and artificial intelligence have been used. In this solution, a very important role is played by formalisms and inference algorithm design as they have a decisive impact on the effectiveness of the diagnostic process. A new solution proposed in this paper is to use attribute table as, a tool supporting the identification of defects. To this purpose the table has been developed (as on open system) in which descriptions of defects from various sources (international literature, expert's knowledge, production data) was collected. The entire system creates a consistent methodological approach, enabling more comprehensive treatment of the diagnostic process, what should be noted as a new solution of the problem. All this results in increased efficiency and reliability of the diagnostic process.
Keywords: konwledge representation, attributa table, identification of defects, casting defects

Problem diagnostyki wad wyrobów odlewniczych obejmuje kilka odrębnych dziedzin, dla których dopiero odpowiednie globalne spojrzenie pozwala na satysfakcjonujące rozwiązanie zadań diagnostycznych. Jedna z tych dziedzin dotyczy wiedzy o parametrach możliwości wystąpienia wad. Wiedza ta jest tym bardziej cenna im jest bardziej rozległa, pozyskana z większej ilości źródeł. Taką wiedzę należy połączyć w spójna całość i nadać jej formę umożliwiającą przetwarzanie komponentowe. W artykule zaproponowano rozwiązanie, w którym wykorzystane zostały inżynieria wiedzy oraz metody sztucznej inteligencji. W omawianym rozwiązaniu bardzo ważną rolę odgrywają formalizmy oraz konstrukcja algorytmu wnioskowania gdyż mają one decydujący wpływ na efektywność procesu diagnostycznego. Całość systemu tworzy spójne podejście metodyczne, umożliwiające bardziej kompleksowe traktowanie procesu diagnostycznego. Równocześnie, zastosowanie alternatywnych rozwiązań modułu określającego przyczyny wady (logika rozmyta, LPR, logika klasyczna), pozwala na dostosowanie procedur decyzyjnych do aktualnych kompetencji użytkownika, umożliwiając działania w warunkach niepewności. Wszystko to składa się na zwiększenie efektywności i wiarygodności procesu diagnostycznego.

1. Introduction

The problem of casting defects diagnosis includes several distinct areas in which only the relevant global perspective allows for a satisfactory solution of the diagnostic task.

One of these areas relates to knowledge about the parameters of the possible occurrence of defects. This knowledge is the more valuable, the more extensive it is, and the more numerous are the sources it has been acquired from. It should be combined into a coherent whole and given the form allowing for component processing.

The paper proposes a solution in which the methods of knowledge engineering and artificial intelligence have been used. In this solution, a very important role is played by formalisms and inference algorithm design as they have a decisive impact on the effectiveness of the diagnostic process.

A new solution proposed in this paper is to use attribute table as, a tool supporting the identification of defects. To this purpose the table has been developed (as on open system) in which descriptions of defects from various sources (international literature, expert's knowledge, production data) was collected. Each entry in the attribute table was

* FOUNDRY RESERCH INSTITUTE, 30-418 KRAKÓW, 73 ZAKOPIAŃSKA STR., POLAND
** ANDRZEJ FRYCZ KRAKOW UNIVERSITY, 30-090 KRAKÓW, 1 HERLINGA GRUDZIŃSKIEGO STR., POLAND

consulted with an expert in the field. Another innovative information presented in this paper is to create procedures identifying defects described in chapter 4. Information contained in chapter 5 and 6 are supplementary, which are described in the previously presented publications, but it can not be ignore because of the whole defects identification process which is described in Fig. 1 (which figure is also innovative elaboration). The various stages shown in this diagram have been partially described already in published by the team positions, nevertheless such comprehensive approach to defects diagnostics hasn't been presented yet.

2. The structure of diagnostic processes

The problem of the design and implementation of diagnostic systems is still open, because for each application area, different solutions are needed, particularly as regards the choice of formalisms of knowledge representation and decision algorithms, tailored to the user's specific needs.

In the approach proposed here, the following stages can be distinguished:

– identification of the type (name) of defect based on knowledge representation in the form of attribute table;
– identification of the causes of defect using knowledge representation and allowing for its incomplete and uncertain character;
– identification of the methods to prevent defects using knowledge acquired from diffuse and heterogeneous sources.

The breakdown of the diagnostic process into the above mentioned steps enables serialisation of the asked queries in a logical sequence, which directs the user's way of thinking, helping him to find adequate answers. Fig. 1 illustrates these steps.

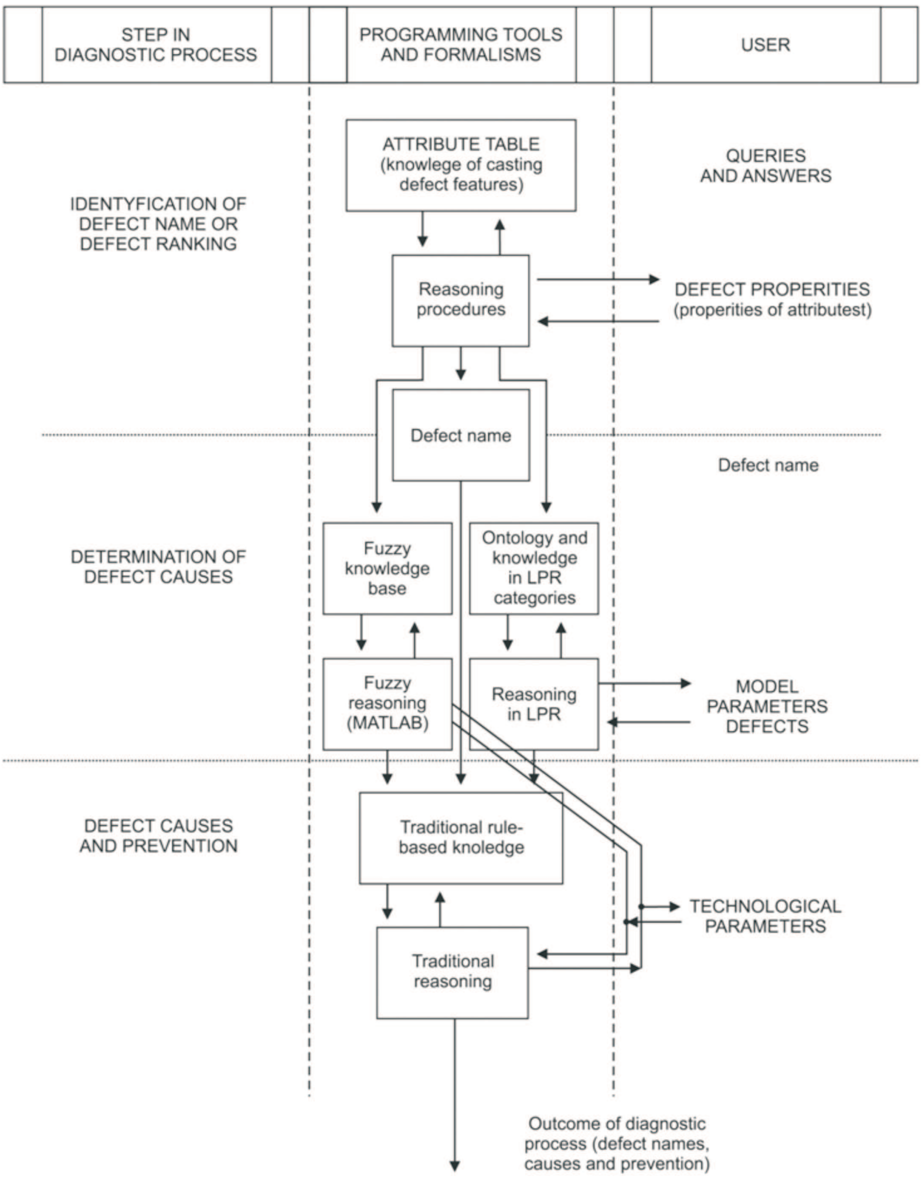

Fig. 1. Schematic diagram of the process of defect diagnosis

In any case, the starting point for the process of diagnosis is identification of the type (name) of defect. In this study, in the solution proposed and described in the present paper, this step has been implemented based on the attribute Table.

If user declares that he knows the name of the defect, direct transition to the second step follows, i.e. determination of the causes of defect occurrence.

There are three possibilities (variants) here:

1) User has sufficiently precise information about the parameters of the technological process, which allows direct transition to step 3, i.e. to a combined determination of the causes of defects and means to prevent their occurrence.
2) User's knowledge of the technological parameters is of approximate and uncertain character. Then it is recommended to use the solutions based on fuzzy logic, which result in a determination of the likely causes of defects or in a ranking of possible causes.
3) User's knowledge is of an intuitive character and relates to certain relationships and linkages that exist in the process rather than to the specific values of various parameters. In this situation, it is proposed to apply the logic of plausible reasoning (LPR).
4) The last step in the diagnostic procedure is indication of the diagnostic measures to prevent the occurrence of a given type of defect. This action can be preceded by an indication of the cause of defect.

3. Attribute Table

The concept that was adopted in creation of this module resulted from several important tasks that it should satisfy:

1) provide the ability to integrate knowledge from different (possibly heterogeneous) sources;
2) provide the ability to create dialogue procedures, tailored to the user's needs and competencies;
3) provide the ability to use incomplete and uncertain knowledge;
4) provide options for application of the reasoning procedures forming a flexible interpretation of the existing knowledge, at different levels of generality.

It should be noted that in the proposed solution the starting point for the creation of an attribute table are casting defect classification systems, described in standards, directories, and national and international handbooks. The need for an ability to integrate knowledge from different sources results from the fact that in various systems of classification there are different divisions into classes, groups and names of the defects. Below, systems used in this study have been depicted graphically and briefly characterised.

Figure 2 shows the system based on a French method of classification published in the atlas of defects. [1, 2] This division is of a very complex nature. The classes of defects are designated with capital letters from A to G. Each class comprises several groups of defects (e.g. A100, A200, A300), these, in turn, comprise subgroups (e.g. A110, 120…), subgroups include defects designated with a letter that denotes the class of defect they belong to, and a three-digit number that denotes their membership in group and subgroup. (e.g. A111, A112, A113….).

Fig. 2. The division of defects into classes, groups and names according to a French system

Fig. 3. The division of defects into classes and names according to a Polish system

The classification in Fig. 3 is a system described in Polish Standard PN 85/H-83105 [3]. According to this system, the defects are divided into 4 classes designated with the following symbols: W-100, W-200, W-300, W400. To each class belong the defects designated with letter W and a three-digit number denoting the class number (e.g. W 206, W-107, W 301);

Fig. 4. The division of defects into classes, groups and names according to a Czech system

The system shown in Fig. 4 was developed in the Czech Republic [4]. According to this system, the defects are divided in 7 classes designated with three-digit numbers (e.g. 100, 200....), each class is divided further into groups (110, 120, 130...), each group may include specific names of defects (111, 112, 113....), or sometimes the name of the group is also the name of the defect. As mentioned previously, conditions from 1 to 4 must be satisfied. The knowledge representation with attribute table (decision table) was used here, i.e. the method which has not been used until now in solutions of this type [5, 6].

Figure 5 shows a general form of an attribute table, where A ={A_1, A_2,... A_n}, {D_1, D_2......, D_n} are domains, and D_i is the domain of attribute A_i for (=1,2,.....,n). The table is a finite set of names.

In the description of object properties (O_i), the values of all attributes or conditions that these objects should satisfy are given. An elementary notation of the fact which says that the value of attribute A_i for object O_j is t_i has the form $A_i = t_{ij}$, where $t_{ij} \subseteq D_i$ while $O_j \in \Omega$ (where: Ω-the set of all described objects). Owing to such notation, a significant extension compared to the classical relational data model is obtained. Attribute values need not be explicit, so the representation of the conditions to be met by individual attributes allows for an interval specification, or for a specification in the form of a set (of value or names). Figure 5 shows schematic representation of information. The columns of this table are the values of relevant attributes, while rows correspond to the descriptions of successive objects O_j (j=1,....,k), on which the rules are based.

A_1	A_2	A_3	...	A_1	O
$t_{1,1}$	$t_{1,2}$	$t_{1,3}$	...	$t_{1,5}$	O_1
$t_{2,1}$	$t_{2,2}$	$t_{2,3}$	...	$t_{2,5}$	O_2
...	...	...	...	...	...
$t_{1,1}$	$t_{1,2}$	$t_{1,3}$	...	$t_{1,5}$	O_1
...	...	...	...	...	...
$t_{k,1}$	$t_{k,2}$	$t_{k,3}$	...	$t_{k,5}$	O_k.

Fig. 5. The structure of attribute table

Decect name	Damage type	Visibility	Damage size	Material loss/gain	Range	Location	Moulding material	Inclusions	Occur-rence rate	Shape
121 Hot casting break off	breaking off	well visible	distinct	loss	local	surface	any	absent	single cases	insignificant
						edge				
		invisible to naked eye			spread	protruding elements				
					clusters	part				
122 Hot Cold casting break off Cz	breaking off	well visible	distinct	loss	local	surface	any	absent	single cases	insignificant
						edge				
		invisible to naked eye			spread	protruding elements				
					clusters	part				
320 Cracks Cz	crack	well visible	distinct	loss	local	wall	any	oxides	single cases	straight
						surface				zizagged
										curved
331 Hot fracture Cz	crack	invisible to naked eye	small	insigni-ficant	clusters	interior	any	absent		idefinite
						surface				

Fig. 6. Fragment of an attribute table with names of defects regardless of their classification system

It is assumed that objects are homogeneous, i.e. each of them is described by specifying the value of the same set of attributes (assuming that an attribute can have the zero value). The contents of the table can represent data (attribute values are independent), data patterns (attribute values are then subsets of the fields), and rules of inference.

The attribute table can be used in defect identification, based on attribute values specified in the table. Analysing the descriptions of defects included in the source documents [1,2,3,4], a list of the attributes of defects which occurred in any of the considered systems (sources) has been made. In this way, an attribute set A has been defined.

For each of these attributes, the respective domains (D_i) were defined. A complete attribute table includes 90 names of defects; to each of these defects, the attributes were ascribed in 18 columns (a fragment is shown in Fig. 6). These sets are not equinumerous, which means that the size of set D_i is not necessarily equal to the size of set D_j (e.g. for the attribute damage type 42 values were specified, while for the attribute size only 3 values were stated).

The essential difference between the traditional decision tables and attribute tables constructed in this study consists in this that the table proposed here has empty places.

It is worth noting that, directly from the table, one can obtain logical expressions defining the name of the defect as a conjunction of attribute values assigned to this defect.

$$\rho_i = [A_1 = t_{i,1}] \wedge [A_2 = t_{i,2}] \wedge ... \wedge [A_n = t_{i,n}] \Rightarrow [O = o_i] \quad (1)$$

Thus, the rule specifies the name of the defect the attributes of which have the value $t_{i1} \ldots\ldots\ldots t_{in}$. In practical interpretation, this rule in linguistic notation can, e.g., express the following statement:

Ri=[damage type=break off] ∧ [visibility=well visible] ∧ [range=local] ∧ [location=surface]∧ [moulding material=any] ∧ [inclusions=absent]∧ [occurrence rate=single forms] ∧ [defect shape=insignificant] ∧ [cast material=any]∧ [penetration=surface] ∧ [surface colour=metal colour] ∧ [orientation = insignificant]∧ [surface oxidation=oxidised] ∧ [defect surface=data not available] ∧ [moment of defect formation (process stage)= mechanical fettling of casting] ⇒ hot casting break off.

(2)

Rules of the structure corresponding to expression (1) can be formed directly from the rows of attribute table.

At the same time, attention deserves the fact that, using the table, one can formulate rules of more complex structure regarding, e.g., indication of a group of defects or equivalence with defects described by other standards.

Figure 7 shows block diagram of the application implementing the defect identification procedure based on the knowledge contained in an attribute table.

Below an algorithm of defect identification is presented. To facilitate the interpretation of this algorithm in terms of technology, a descriptive form was used to allow for comments on the successive stages of the procedure.

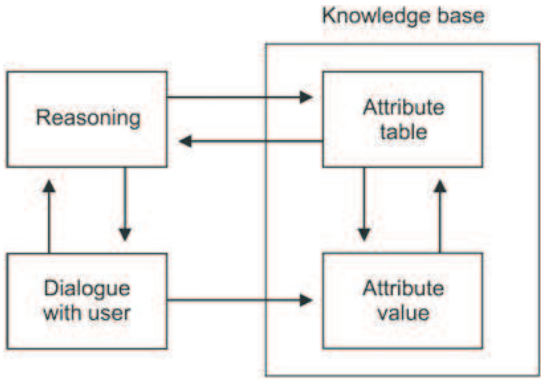

Fig. 7. Block diagram of pilot application

Initial state:
Knowledge base:
Attribute table T^{mxn}
n-number of rows (rules)
m-number of columns (attributes)
Sets of attribute values V_j (j=1...m).

$$\begin{aligned} V_1 &= (a_{11}, a_{12}, \ldots, a_{1k_1}) \\ V_2 &= (a_{21}, a_{22}, \ldots, a_{2k_2}) \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ V_m &= (a_{m1}, a_{m2}, \ldots, a_{mk_m}) \end{aligned} \quad (3)$$

It should be noted that the numerical values of individual attributes (k_1, k_2, ..., k_m) are generally different.

4. Procedure for defect identification

The procedure consists in asking queries about the values of the successive attributes (a_{jk})

Step 1: Query: Give value of attribute a1 from the set V_1

$$Answer : a_1 = v_{1k}$$
$$Reasoning : T^{nm}(X(0), a_1 = v_{1k}) \rightarrow T^{\lambda(1)m}(X(1)) \quad (4)$$

where:

X(1) – the set of objects (defects) for which the value of attribute a_1 is v_{1k}

that is

$$\rho: (x_i, a_1) = v_{1k}, x_i \in X, 1, a_j \in A oraz v_{1k} \in V_1 \quad (5)$$

$\lambda(1)$ – the number of objects satisfying condition (5).

So, as a result of the execution of the first step in an algorithm, the number of objects (defects) examined in further reasoning has been reduced and now is λ(i), and not n as it was at the beginning.

Step 2: Query: Give value of attribute a_2 from the set V_2

$$Answer : a_2 = v_{2l}$$
$$Reasoning : T^{\lambda(1),mm}[X(1), a_2 = v_{2l}] \rightarrow T^{\lambda(2),m}(X(2)) \quad (6)$$

where:

X(2), $\lambda(2)$ – the set of objects and their number, respectively, obtained as a result of operation 6

. . . .

. . . .

Step m: Query: Give value of attribute a_m from the set V_m

$$Answer : a_m = v_{m\tau}$$
$$Reasoning : T^{\lambda(m-1),m}[X(m-1), a_m = v_{m\tau}] \rightarrow T^{\lambda(m),m}(X(m),) \quad (7)$$

where:

X(m) – the set of objects (defects) determined as a result of ascribing the values to all attributes $a_j \in A$.

At the end of the procedure, three cases can be distinguished:

1) If X(m) is a singleton,

that is λ(m)=1, then the name of the defect is clearly indicated and the procedure is completed. However, compliance with this condition is the case rather theoretical. In practical situations, the following possibilities may occur:

2) Indication of specific defect may occur earlier, i.e. after the number of steps k<m, if by this time the n-1 objects (defect types) have already been eliminated. In other words, even stating the value of k attributes is enough to distinguish given object (defect) from all other attributes described in the attribute table.

3) Having ascribed values to all attributes, i.e. after m steps of the algorithm, the set X(m) still holds more than one object. So, λ(m)>1.

This corresponds to the situation when certain objects (defects) are described with the same values of attributes. This situation may occur when the same defect has different names (e.g. when knowledge comprised in standards or catalogues edited by different nations is used). This happens in the case of defect A in Polish classification, and defect B in French classification, as well as defect C in Czech classification.

5. Diagnosis of the causes of defects

To determine the causes of defects on the basis of incomplete and uncertain data, two types of logic have been used, viz. logic of plausible reasoning and fuzzy logic. The logic of plausible reasoning is used in those areas where knowledge about the causes of defect formation is of an intuitive character, most often "man-associated". This logic is also used to connect the data on causes of defects originating from different sources. The problem has been described in reference literature [7].

For indefiniteness considered in two aspects, viz.:

- at the level of knowledge about the defect, which does not explicitly attribute to the defect the reasons of its occurrence;
- referred to the parameters of technological process (in the course of which the defect has occurred), usually determined in an approximate way;

a formalism was used which takes into account both the lack of precise knowledge about the defects, as well as an approximate nature of information about the parameters of the technological process.

The formalism, which by definition is dedicated to the description of such situations, is fuzzy logic [8, 9].

6. Determination of the method to prevent defect occurrence

This module has been based on bivalent logic. The formulation of rules in terms of bivalent logic consists in identification of causal-resultant relationship between the variables of values taken from the set {0,1}, with respective interpretation {false, true}.

A typical for expert systems procedure (applied, among others, in CastExpert system) consists in sequential application of rules. The course of the final diagnostic procedure can be represented as a schematic diagram shown in

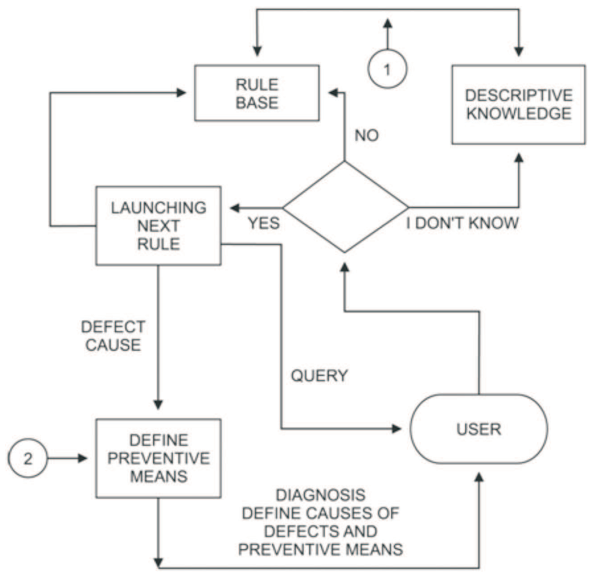

Fig. 8. Schematic diagram of final diagnostic procedure

Indicating the name of defect results in one of the rules being selected from the base. An answer to this query is meant to confirm the truth of the premises in the rule under consideration.

7. Summary

The entire system creates a consistent methodological approach, enabling more comprehensive treatment of the diagnostic process, what should be noted as a new solution of the problem. So far, discussed in the literature solutions were presented individually, while at the same time, the use of alternative solutions for module defining the causes of defects (fuzzy logic, LPR, classical logic) allows adaptation of decision-making procedures to current user's competences, and hence acting under the conditions of uncertainty.

All this results in increased efficiency and reliability of the diagnostic process. At the same time it must be acknowledged that the lack of actual data on the numerical characteristics of a technological process prevents the acquisition of assessments characterising the diagnostic process in a quantitative manner (in order to carry out such studies, data from multiple cycles of casting are necessary). Therefore, out of necessity, the developed solutions could be verified basing only on the qualitative assessments made by experts and process engineers. The said assessments issued by the cooperating foundry plants were definitely favourable.

Received: 10 July 2010.

REFERENCES

[1] Casting Defects Hand-Book; American Foundrymen Society INC., 1984.
[2] G. Horon, C. Mascar, G. Blanc, Recherche de la qualite des pieces de fonderie; Editions Techniques des Industries de la Fonderie, Paris, 1986.
[3] Odlewy: Podział i terminologia wad, Polski Komitet Normalizacji Miar i Jakości, PN-85/H-83105, Dz. Norm i Miar, 1 pozycja 2, 1986.
[4] T. Elbel, F. Havliček, P. Jelinek, P. Leviček, J. Rous, K. Stransky, Vady odlitkú ze slitin železa (klasifikace, přičiny a prevence), MATECS, Brno, 1992.
[5] A. Ligięza, M. Szpyrka, R. Klimek, T. Szmuc, Weryfikacja wybranych własności jakościowych systemów tablicowych z bazą wiedzy, Oficyna wydawnicza Politechniki Wrocławskiej, Wrocław, 2000.
[6] Z. Pawlak, Rough sets, International Journal of Information and Computer Science **11**, 5, 341-356 (1982).
[7] S. Kluska-Nawarecka, Z. Górny, D. Wilk-Kołodziejczyk, A. Smolarek-Grzyb, The logic of plausible reasoning in the diagnosis of castings, Archives of Metallurgy and Materials, Polish Academy of Sciences. Committee of Metallurgy. Institute of Metallurgy and Materials Science; ISSN 1733-3490 **52**, 3 375-380 (2007).
[8] A. Łachwa, Rozmyty świat zbiorów, liczb, relacji, faktów, reguł i decyzji, wyd.: AOW Exit, Kraków 2001.
[9] J. Talar, Metody eksploracji danych – zastosowanie w metalurgii; Archives of Metallurgy and Materials **2** (2009). Wydawca: Instytut Metalurgii i Inżynierii Materiałowej PAN Rok wydania: 2009 ISSN: 1733-3490.