Development of INFOCAST: Information System for Foundry Industry

Grzegorz Dobrowolski[1], Robert Marcjan[1], Edward Nawarecki[1], Stanisława Kluska-Nawarecka[2], Joanna Dziaduś[3] and Teresa Wójcik[3]

[1] *Department of Computer Science, University of Mining and Metallurgy, Mickiewicza 30, 30-059 Cracow, Poland, grzela@agh.edu.pl*

[2] *Department of Industrial Informatics, University of Mining and Metallurgy, Mickiewicza 30, 30-059 Cracow, Poland, nawar@iod.krakow.pl*

[3] *Foundry Research Institute, Zakopianska 73, 30-418 Cracow, Poland, jdziadus@iod.krakow.pl*

(Received 30 November 2002; revised manuscript received 18 March 2003)

Abstract: The article presents the current state of development of INFOCAST – an information decision system supporting technological problems of the foundry industry. Two aspects of the system are related: enrichment of its information and knowledge resources and changes in architecture, both oriented towards improvement of its applicability. The reported stage of development of INFOCAST is characterized by a close integration of its data and knowledge bases by means of the agent-based technology. Due to decentralization, the knowledge and information can be used not only in research and design activities, but also in the exploitation of technological processes.

Keywords: multi-agent systems, expert systems, databases, foundry industry, diagnosis

1. Introduction

The information resources available in the databases currently operating within the INFOCAST system [1, 2] are of great importance for research work and projects, as well as for production planning and marketing in the field of the foundry industry. The information is the property of the Foundry Research Institute in Cracow, who make it available and keep it up-to-date. The computer implementation of the system has been developed by Department of Computer Science of University of Mining and Metallurgy, also in Cracow. The core of the INFOCAST system consists of four currently operating and continually updated data and knowledge bases: SINTE, NORCAST, CASTSTOP and CAST-EXPERT.

Computerization of the bases of the INFOCAST system was carried out within the framework of a research project sponsored by KBN (State Committee of Scientific Research) in the years 1997–2001, in two stages. Information stored in its bases has been collected since 1977.

During the last stage, in 1999–2001, two kinds of activities were carried out:

- One was oriented towards the implementation of a new version of the INFOCAST system on a supercomputer at ACK CYFRONET AGH Centre, based on DBMS ORACLE. An immediate effect of this activity is that the system is easily available on the net and operates more efficiently.
- The other was related to designing and building a decentralized version of INFOCAST, based on the agent technology. There are programming tools (an agent software platform along with the relevant programmer's environment) designed specifically for the construction of decentralized information decision systems distributed over the net.

All the time, enrichment of INFOCAST's information resources has continued.

The decentralized version of the system (its agents), like CASTEXPERT, takes advantage of rule-based knowledge representation. Its implementation makes use of the JAVA language.

INFOCAST in its decentralized structure is expected to offer direct (network) access to foundry mills along with the possibility to utilize information coming from the currently running production processes.

The article is organised as follows. Section 2 is a recapitulation of the current state of resources of INFOCAST and a brief introduction to the system. Section 3 presents basic information regarding the idea of a decentralized version of the system. An outline of its overall architecture, designed according to the agent-based technology, is given. Information about the general structure of an agent and the inter-agent communication platform completes the picture.

2. Utilitarian characteristic of INFOCAST

2.1. Field of application

The system under consideration has been designed to serve as a multi-purpose tool of technical assistance, an aid in designing new technologies and quality assurance for foundry products. The following can be mentioned as typical tasks solved by the system:

1. assistance in reference literature from the field of foundry practice,
2. analyses in the field of manufacturing technologies,

3. expertise in the field of casting defects diagnosis,
4. analyses for casting products – a module of knowledge to be designed in the future.

It is assumed that the services provided by the system can cover the analyses ordered by:

- research and development centers and specialized laboratories,
- design and trade offices,
- production plants and supervising staff,
- potential users of produced castings.

The system, which is accessible[1] free of charge to the domestic research institutions, serves as a basis for many research projects, including MSc and PhD theses. At present, it is the only Internet source of information on foundry practice in Poland.

2.2. Information and knowledge resources

INFOCAST is composed of four data and knowledge bases (SINTE, NORCAST, CASTSTOP, CASTEXPERT) [1, 2]. They contain a wide range of information of foreign and Polish literature (the SINTE database), European, international, Polish and EU standards (the NORCAST database), standard grades of cast alloys (the CASTSTOP database) and casting defects with reasons of their occurrence (the CASTEXPERT system with its knowledge base). The databases are relational.

SINTE thematically covers the following foundry-related problems: metallurgy and metal science of cast alloys; heat treatment of castings; iron castings; steel castings; non-ferrous metals castings; metal matrix composite castings; melting processes and melting installations; technological processes of moulding sand preparation; mould and core making; casting technologies; environmental protection; fettling and finishing of castings; use of computers to aid foundry production; foundry machines and equipment; mechanization and automation; quality control and quality systems; marketing; management; organization and cost of production.

Information included in the database is taken from all the leading Polish and international (American, English, French, German, Czech, Slovenian, Russian, Portuguese, Swedish) journals issued in the years 1977–2001, available in the Library of the Foundry Research Institute.

There are about 32 000 records now, while an average 1 000 is added each year. Two special ways of searching are available: according to descriptors using the thesaurus or using a dedicated classification system.

NORCAST comprises 4 058 records containing up-to-date information on foundry standards: Polish, European, international, and foreign national from selected countries (UK, Germany, France, Norway, Finland, Sweden, Italy, USA), as well as information on PN, ISO and EN standards related to a variety of problems concerning quality assurance systems, environmental management and certification.

Recorded aspects of the standards include: general issues, charge materials, moulding materials and methods of their testing, auxiliary materials, machines and equipment, castings, examination of cast alloys and testing of castings, occupational health and safety, quality assurance systems and certification.

CASTSTOP database gives information on Polish, European and foreign standard grades of cast alloys, including their chemical composition, mechanical properties and heat treatment, augmented whenever possible with some additional information. At present, its 1 098 records contain information taken from the international standards (ISO), European standards (EN), French (NF), German (DIN), Norwegian (NS), Swedish (SS, SIS), British (BS) and Italian standards (UNI).

CASTEXPERT is an expert system for diagnosing casting defects. Its user is aided by a knowledge base in the form of decision rules with an appropriate inference engine and a database of photographs of castings defects, as an exemplification. Based on the knowledge comprised (7 200 rules) and information acquired from process engineers (by way of dialogue), CASTEXPERT supports identification of possible reasons of occurrence of defects in castings with an indication of actions to be undertaken to eliminate these defects in the future. The current version of the system embraces defects of castings made of grey iron, ductile iron and non-ferrous metal alloys, as well as defects in steel castings.

CASTEXPERT data may be also regarded as an important source of information on casting defects that may be used in quality control of

1. All the bases of INFOCAST system are available at the following address: `http://czapla.iod.krakow.pl/infocast/`

castings, or as reference material for training engineers and students.

3. Agent architecture of the decentralized INFOCAST

Apparently, the idea of multi-agent systems [3–5] is an extension or modification of the idea of decentralized systems, so that they may be even regarded as synonymous.

A typical system of this kind is composed of a set of subsystems/nodes (agents) that cooperate with each other, linked to form an organizational structure. The links are dependent on the functional characteristics of subsystems and can be created in a dynamic way in order to attain a global goal (of the whole system). Because agents are assumed to have some autonomy, the process of searching for a solution is decentralized. The potential of this system (its functionality) result from the functions of its individual agents and a certain *added value*, which is created in their cooperation.

The main nodes or subsystems of the decentralized version of INFOCAST are its computerized information or knowledge sources. Assuming the future development of INFOCAST, it is possible to enumerate many possible types of sources:

- Relational databases
  Information is accessed using an appropriate query language (SQL). Handling of information (making a query) requires knowledge of the database structure.
- Knowledge bases of reasoning systems
  Expert systems contain fragments of the domain knowledge. In the case described here, these are systems with rule-based knowledge representation. Communication with systems of this type consists in exchange of facts and rules.
- Experts – humans
  It is assumed that an expert is a source of the domain knowledge. Communication is in an off-line mode. Adequate understanding of questions addressed to an expert should be assured by the system.
- Other sources of information
  It is allowed to use other sources of information (*e.g.* other information systems or control and measuring instruments). Including such a source of information into the system requires the creation of special communication software.

In the solution described here, access to the various types of knowledge and information sources mentioned above is through agents assigned to these sources.

It should be mentioned here that the goal is not to create a decentralized diagnostic system from the very beginning, but to design a technology which would enable utilization of the existing information systems available via the internet and their arrangement into a system capable of solving complex tasks.

3.1. Agents and their functions

Given the above described application field of INFOCAST and the idea of its operation in a decentralized regime, the main functions which are to be performed by agents, can be defined as follows:

- reasoning based on local information resources,
- communication with other agents,
- communication with humans acting within the system (user, expert),
- making available various sources of knowledge created using various technologies,
- creation of an information network through identification of agents necessary to employ to solve a specific problem,
- assuring the conformity of knowledge components used in the diagnostic process,
- integration of information, synthesis and propagation of results, and arbitrage in the case of conflict.

A full set of agents proposed for the system architecture is shown schematically in Figure 1.

AG-DB: an agent for a source of knowledge of the database type – possesses a description of the source of knowledge (definition of the database, characteristic of the tables) and is capable of translating the contents of KQML messages [6] into the corresponding phrases (queries) written in SQL. A majority of the systems of this type enable communication through a JDBC/ODBC interface.

AG-SE: an agent for a source of knowledge of the expert system type – both the expert system and the agent use the same representation of information (facts and rules written in the JESS language [7]). Interactions consist in direct start ups of the reasoning process and appropriate data exchanges.

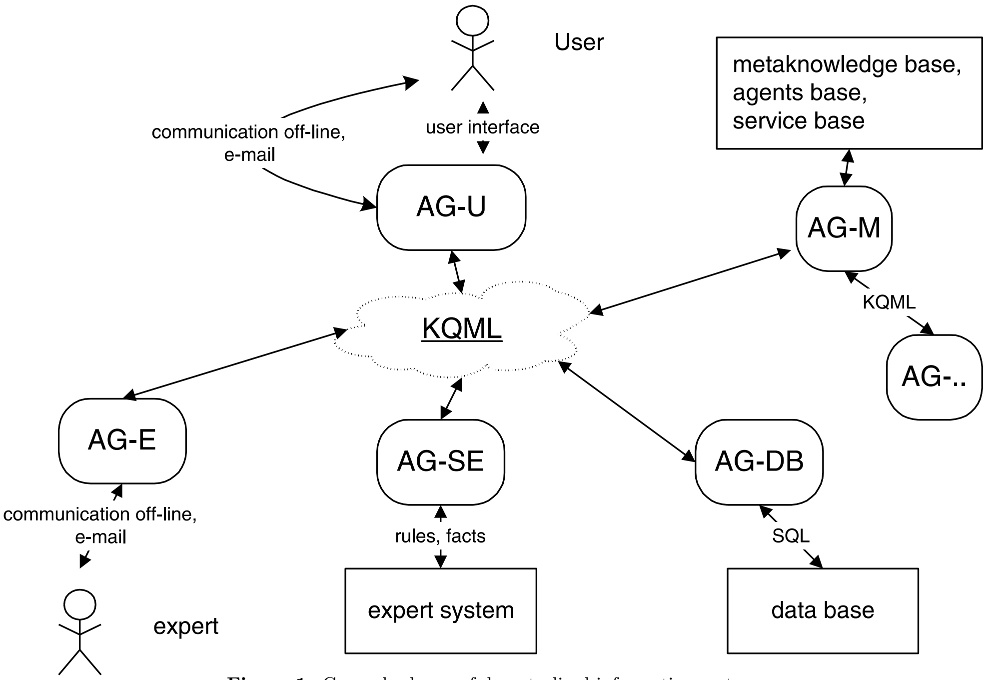

Figure 1. General scheme of decentralised information system

AG-E: an agent for a source of knowledge of the expert type – assists an expert in his communication with the system, receives messages in KQML, generates their HTML forms and sends them by electronic mail to the expert's address. When a replay comes, the contents is translated back to KQML.

AG-U: the user's agent – is responsible for communication with the user from the moment an analysis is initiated, through generation of auxiliary queries, until results are presented. These functions can be executed via the standard windows interface (based on a WWW browser), or by means of electronic mail. In the latter case, communication is carried out in the same way as it is with an expert.

AG-M: an agent managing meta-knowledge. This is also the agent allowing access to the source of knowledge/information, but of special functions in the system. There are:

- a database of the agents – information about the system structure (identification, addresses, *etc.*),
- ontologies – sets of notions (facts) together with relations between them which are used by the agents; reasoning in the system is made according to the relevant ontology (in foundry technology, the ontology is based on a thesaurus developed by the Foundry Research Institute),
- a database of services offered individually by the agents – information which can be provided by the given source of knowledge (the agent that operates it).

The system structure described above and the organization of access to the sources of information enable the process of reasoning (preparing an analysis) to be carried out irrespective of the structure of a given source (a coherent set of notions/facts, uniform methods of communication). They also make it an open system and create possibilities of easy adaptation of new sources of knowledge (extension of the system domain, new architectures of knowledge sources, *etc.*).

3.2. Agent as an inference engine

A sketch of the internal structure of an agent and its functional blocks are shown in Figure 2. This is a general agent architecture in that it involves all the modules (elements) which can appear among the system agents, while in specific cases some of them may be unnecessary (restricted configuration is applied).

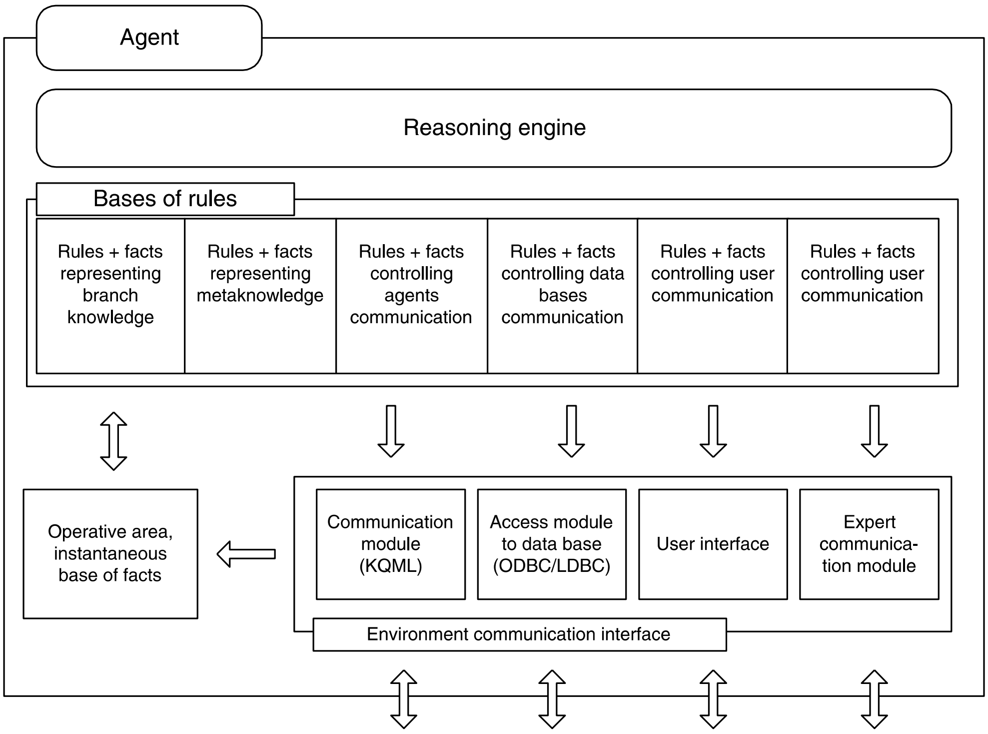

Figure 2. General architecture of an agent

The agent's core is an inference engine that accepts a set of rules in the form of *premise* ⟶ *action* and facts that can be divided into several categories depending on what these facts refer to [8]. This results in differentiation of the functional blocks (modules) of an agent. Each block consists of a sub-base of rules/facts and an appropriate module, which ensures communication with the environment (*i.e.* with other agents or sources of knowledge). Communication among an agent's functional blocks is achieved through an instantaneous base of facts (an operational area or agenda, in the terminology of rule-based systems). Appearance of new facts in this area activates the rules for which the premises become true. Thus, modifications are introduced into the instantaneous base of facts, *i.e.* new facts appear, which in turn activate the rules of other modules. The appropriate communication modules are responsible for the exchange messages with the environment [5, 9].

3.3. Agent communication platform

The main task of the platform is to provide an appropriate communication infrastructure and organize agents into a coherently functioning system.

The platform implementation uses the JATLite (Java Agent Template) [10] package of software. JATLite offers a library, written in Java, comprising classes which enable communication at different levels of abstraction, defined as communication layers and protocols (including communication using the KQML protocol/language).

The JATLite software also serves as a simple communication platform. Its services are represented by the specialized agent called AMR (Agent Message Router). The main tasks of the agent are given below:

- Managing the system structure (*i.e.* the name space and communication addresses). Before an individual agent starts functioning in the system, it registers with AMR. Then, using its name and password, the agent can connect and disconnect with the system, and exchange information with other agents operating on different computers until it registers off the platform. AMR stores and manages physical addresses of agents.

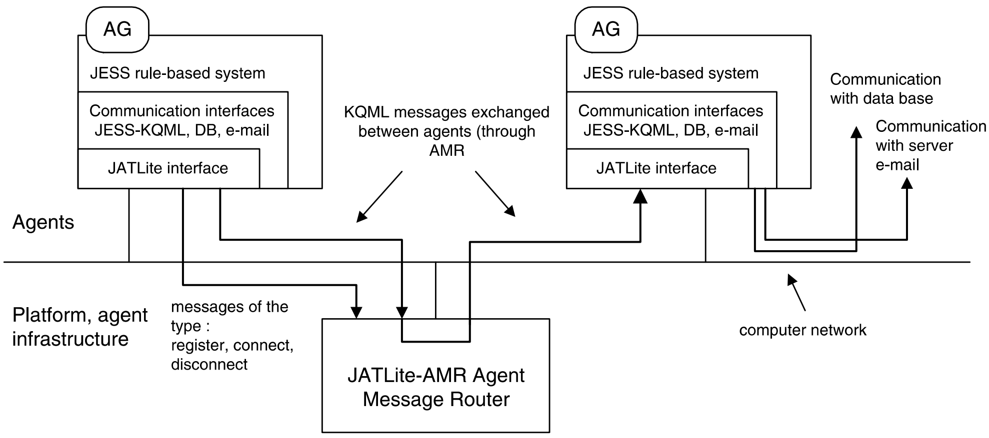

Figure 3. General concept of the agent communication platform

- Acting as an intermediate link between the agents who participate in the system. The agents do not communicate directly with each other, but use AMR to this purpose. Using the JATLite platform, each agent forms a single connection to AMR and uses this connection for exchanging messages with other registered agents. If an agent is disconnected or damaged (stops functioning), the messages sent to it are stored by AMR until being re-connected.

General schematic presentation of the platform functionality as well as types of messages and their means of transportation are shown in Figure 3.

4. Final remarks

The article describes information decision system INFOCAST dedicated to aiding the foundry technologies. The system has been installed on the ACK CYFRONET AGH supercomputer using an ORACLE data base management system and is available through the Internet to users from Poland and abroad. It offers the knowledge and information generated and collected at the Foundry Research Institute in Cracow. The system has proven its usefulness in many research, production planning and marketing projects.

The decentralized version of INFOCAST has been designed to semantically integrate the resources of the Foundry Research Institute with similar resources of other research centers and the knowledge of individual experts and real data collected by industrial enterprises. It is assumed that, based on proper network infrastructure and appropriate computer tools, the system makes it possible even for quite complex analyses to be done automatically without a deeper intervention of the user.

Full capabilities of the system will be gained when direct network connections with industrial enterprises (and their appropriate departments) have been established. Such target situation should enable, on one hand, making local analyses right on the spot, utilizing the general knowledge collected in the INFOCAST system. On the other hand, the communication links should intensify further enrichment of this knowledge through generalization of experience acquired within the industry.

The idea of building a decentralized information system using the agent-based technology is straightforward and effective. Invention and application of a unified architecture of the agents, using rule-based representation of their behavior, has given a similarly positive effect.

References

[1] Kluska-Nawarecka S, Dobrowolski G and Marcjan R 2001 *Acta Metallurgica Slovaca* **7** 441

[2] Kluska-Nawarecka S 1996 *J. Mat. Sci. and Technol.* **4** 11

[3] Ferber J 1999 *Multi-Agent Systems*, Addison-Wesley

[4] Weiss G (Ed.) 1999 *Multiagent Systems: A Modern Approach to Distributed Artificial Intelligence*, The MIT Press

[5] Dobrowolski G 1998 *Management and Control of Production and Logistics*, MCPL'97 IFAC/PERGAMON (Binder Z, Hirsch B, Aguilera L, Eds) **2** 393

[6] Finin T, Fritzon R, McKay D and McEntire R 1994 *Proc. 13th Int. Workshop on Distributed Artificial Intelligence*, Seatle, WA, pp. 126–136

[7] Friedman-Hill E 1999 *JESS – the Java Expert System Shell* http://herzberg.ca.sandia.gov/jess/
[8] Negrete Martinez J and Gonzalez Perez P 1998 *Expert Systems with Applications* **14** 102
[9] Dobrowolski G and Nawarecki E 2001 *Management and Control of Production and Logistics 2000*, PERGAMON (Binder Z, Ed.), pp. 445–450
[10] Jeon H, Petrie C and Cutkosky M R 2000 *IEEE Internet Computing* **4** 89

Diagnosing Skin Melanoma: Current versus Future Directions

Zdzisław S. Hippe[1], Stanisław Bajcar[2], Piotr Blajdo[1], Jan P. Grzymala-Busse[3], Jerzy W. Grzymala-Busse[3], Maksymilian Knap[1], Wiesław Paja[1] and Mariusz Wrzesień[1]

[1] *Department of Expert Systems and Artificial Intelligence, University of Information Technology and Management, Sucharskiego 2, 35-225 Rzeszow, Poland, zhippe@wenus.wsiz.rzeszow.pl*

[2] *Regional Dermatology Center, Warzywna 3, 35-310 Rzeszow, Poland,*

[3] *Department of Electrical Engineering and Computer Science, University of Kansas Lawrence, KS 66045, USA*

(Received 18 January 2003)

Abstract: A new database containing 410 cases of *nevi pigmentosi*, in four categories: *benign nevus*, *blue nevus*, *suspicious nevus* and *melanoma malignant*, carefully verified by histopathology, is described. The database is entirely different from the base presented previously, and can be readily used for research based on the so-called constructive induction in machine learning. To achieve this, the database features a different set of thirteen descriptive attributes, with a fourteenth additional attribute computed by applying values of the remaining thirteen attributes. In addition, a new program environment for the validation of computer-assisted diagnosis of melanoma, is briefly discussed. Finally, results are presented on determining optimal coefficients for the well-known ABCD formula, useful for melanoma diagnosis.

Keywords: melanoma, TDS, machine learning in diagnosis of

1. Introduction

In recent papers [1, 2], we have presented the results of experiments on new samples relating to changes in skin melanoma, using machine learning with the idea of generating a model of learning to help identify and classify cases of skin melanoma. Skin melanoma may be a symptom of serious skin diseases, or even cancer, which has a high mortality rate. The numbers of victims of this type are rising because of the high levels of ultraviolet radiation entering the atmosphere and the increasingly thin ozone layer [3]. Anonymous data sets pertaining to cases of skin cancer have been collected by the Regional Dermatology Center in Rzeszow, Poland [4]. This data set of cases has been analyzed and expert systems based upon it have been implemented at the Department of Expert Systems and Artificial Intelligence, University of Information Technology and Management in Rzeszow, Poland. The first version of the data set has been analyzed in a paper presented at the INFOBAZY'99 conference [1, 2]. The actual production version contains (i) new internal structures and (ii) an increased number of registered cases (from 250 to 410). Regarding (i), the data set information is stored in 13 attributes that are regularly used in dermatology for typical analysis of skin-based melanoma. In the context of these attributes, we calculate the **TDS** (Total Dermatoscopy Score) indicator [5]. The underlying idea of our experiments was to prepare our sample in both Polish and English and use specialized software algorithms in the verification of its accuracy, as well as generate learning models to help diagnose diseases. The data sets were acquired in studies taking place simultaneously in Rzeszow, Poland (University of Information Technology and Management) and the United States (University of Kansas, Lawrence, Kansas). In the following sections, the data sets have its statistical analysis and its machine learning results discussed. An earlier version of this paper was presented at the 3rd National Conference INFOBAZY'2002, Gdansk, Poland, June 24–26, 2002 [6].

2. Statistical analysis of the data sets

The attributes used in deducing diagnoses of melanoma have been broken down into 5 categories: ⟨Asymmetry⟩, ⟨Border⟩, ⟨Color⟩, ⟨Diversity⟩ and ⟨TDS⟩. The ⟨Asymmetry⟩ parameter can have the following values: symmetrical, single-axis asymmetry and dual-axis asymmetry. ⟨Border⟩ is a numerical attribute with discrete values between 0 and 8. The next two categories, ⟨Color⟩ and ⟨Diversity⟩, have symbolic values. ⟨Color⟩ can have six allowed values: black, blue, light brown, dark brown, red, and white. Likewise, ⟨Structure⟩ has five possible values: pigment dots, pigment globules, pigment network, structureless areas and branched