[Chemical Technology, Control and Management](https://uzjournals.edu.uz/ijctcm)

Volume 2018 | Issue 4 Article 11

12-26-2018

Integrated intelligent systems for industrial automation: the challenges of Industry 4.0, information granulation and understanding agents .

N.R Yusupbekov
Tashkent State Technical University, Tashkent city, Republic of Uzbekistan, dodabek@mail.ru

V.B Tarassov
Bauman Moscow State Technical University, Moscow, Russia, Vbulbov@yahoo.com

ISSN 1815-4840
CHEMICAL TECHNOLOGY. CONTROL AND MANAGEMENT
2018, №6 (84) pp.62-75
International scientific and technical journal
journal homepage: ijctcm.com
Since 2005

N.R.YUSUPBEKOV (TSTU), V.B.TARASSOV (MSTU)

INTEGRATED INTELLIGENT SYSTEMS FOR INDUSTRIAL AUTOMATION: THE CHALLENGES OF INDUSTRY 4.0, INFORMATION GRANULATION AND UNDERSTANDING AGENTS

Автоматлаштиришнинг Industry 4.0 янги парадигмаси муаммолари ва киберфизик тизимлар, булутли ҳисоблаш, нарсалар Интернети ва Маълумотларнинг катта дунёси каби илгор ахборот ва коммуникация технологиялари соҳасидаги энг янги технологиялар кўриб чиқилган. Саноатда катта маълумотларни кўп ўлчамли, кўп қиррали акс эттириш ва уларни таҳлил қилишнингбир қатор усуллари таклиф этилган. Грануляциялovчи ҳисоблаш усулларидан фойдаланиб, катта маълумотларга ишлов бериш асослари ишлаб чиқилган. интеграллашган интеллектуал корхоналар учун тушунишнинг сунъий агентларини яратишнинг махсус когнитив воситаларини яроатиш муаммолари юзага келмоқда.

***Таянч сўзлар**: саноат автоматлаштириши, агент, интеллектуал ишлаб чиқариш тизими, Саноат 4.0, рақамлаштириш, виртуал корхона, киберфизик тизим, нарсалар интернети, булутли ҳисоблаш, қўшма робототехника, катта маълумотлар, гранулалар, ахборотларни гранулалаш, тушуниш модели.*

Рассмотрении проблем новой парадигмы автоматизации Industry 4.0 и рассмотрении новейших технологий в области ее передовых информационных и коммуникационных технологий, включая киберфизические системы, облачные вычисления, Интернет вещей и большой мир Данные. Предложены некоторые способы многомерного, многогранного промышленного представления и анализа больших данных. Были разработаны основы обработки больших данных с использованием методов гранулярных вычислений. Возникла проблема создания специальных когнитивных инструментов для создания искусственных агентов понимания для интегрированных интеллектуальных предприятий.

***Ключевые слова**: промышленная автоматизация, агент, интеллектуальная производственная система, индустрия 4.0, оцифровка, виртуальное предприятие, киберфизическая система, интернет вещей, облачные вычисления, совместная робототехника, большие данные, гранулы, гранулирование информации, модель понимания.*

The objective of the paper consists in considering the challenges of new automation paradigm Industry 4.0 and reviewing the-state-of-the-art in the field of its enabling information and communication technologies, including Cyberphysical Systems, Cloud Computing, Internet of Things and Big Data. Some ways of multi-dimensional, multi-faceted industrial Big Data representation and analysis are suggested. The fundamentals of Big Data processing with using Granular Computing techniques have been developed. The problem of constructing special cognitive tools to build artificial understanding agents for Integrated Intelligent Enterprises has been faced.

***Keywords:** Industrial Automation, Agent, Intelligent Manufacturing System, Industry 4.0, Digitization, Virtual Enterprise, Cyberphysical System, Internet of Things, Cloud Computing, Collaborative Robotics, Big Data, Granule, Information Granulation, Model of Understanding.*

Introduction

The aim of the paper consists in reviewing the-state-of-the-art in modern communication and information technologies, associated with the advent of the 4th Industrial Revolution (Digital Enterprise, Cyberphysical Systems, Internet of Things, Big Data Mining & Analytics, etc.) and specify appropriate new trends in building new generation intelligent systems of industrial automation. A 5V Big Data representation is discussed, a keynote role of transitions «Big Data – Information – Knowledge – Meta-Knowledge» in raising the efficiency of Big Data Use is highlighted. The perspectives of Granular Computing in Big Data Analytics are shown. Top-down and bottom-up approaches in Data Granulation are considered, non-standard set-based models of granules are constructed. Finally, the problem of developing artificial cognitive agents able to «understand» required behavior patterns in a specific industrial situation is faced.

1. Intelligent Manufacturing Systems: Some Historical Prerequisites

Nowadays a main challenge in CIM and industrial automation is the development of integrated intelligent manufacturing systems and technologies enabling the implementation of Industry 4.0 initiative. These innovative approaches and technologies don't develop from scratch. The first step in making manufacturing intelligent was an international program *«Intelligent Manufacturing Systems»* (IMS) which was started in the mid 1990's (see [1,2]). The objective of this program was the creation of new generation manufacturing systems and technologies by performing global intercontinental joint projects. These big projects, such as GLOBEMAN'21 (the abbreviation of Global Manufacturing in XXI century) [2], Next Generation Manufacturing Systems, GNOSIS, Holonic Manufacturing Systems, faced many aspects of automated, integrated, intelligent manufacturing.

Primarily, the topics of IMS program were divided into five groups: (1) modeling and management of total product lifecycle; (2) analysis and development of production and business processes for enterprises of different industries;(3) enterprise strategy planning and engineering techniques and tools; (4) human, organizational, social factors of production; (5) virtual and extended enterprises.

Let us focus on GLOBEMAN'21 project. It was devoted to the problems of enterprise integration and product lifecycle modeling for global manufacturing of XXI century. Both industrial and university partners from various countries and even continents took part at the project.

So GLOBEMAN'21 was positioned as an international consortium to develop and demonstrate the enterprise integration tools and methods. Its purpose was enabling manufacturing enterprise to form a mission oriented project organization, i.e. a virtual corporation, for networked manufacturing business.

A virtual corporation is loosely coupled enterprise which is formed by many partners to fulfill a difficult mission requiring shared resources or organize world class production. The motivation for building virtual enterprise is to integrate available partners' resources and raise the production efficiency [3,4].

The following GLOBEMAN'21 results are worth mentioning [2]: a) design of direct and inverse product lifecycle management systems; b) creation of new technologies of intelligent simulation, decision support and production management for the enterprise networks; c) development of innovative CIM architectures in the enterprise networks to obtain world class products; d) remote customer service and support by using the information about real customers needs and real products manufactured by plants located in different parts of the world; e) more complete and deep understanding of new trends in manufacturing related to the use of advanced information and communication technologies.

It is worth stressing that within the IMS program some basic knowledge engineering problems were already faced: 1) acquisition of manufacturing knowledge and experience; creation of large and distributed knowledge bases; 2) manufacturing data mining and knowledge discovery; 3) development of heavyweight ontological models for virtual enterprises (see ToVE project [5]); 4) creation of intelligent information technologies for production management; 5) multi-agent representation and modeling of knowledge management in networked enterprises; 6) design of innovative enterprises and manufacturing systems on the basis of A-life approaches and bionic (swarm cognition) algorithms. The first generation of IMS tended to solve these problems.

The emergence of new strategic initiative Industry 4.0 generates the need in new approaches to IMS, including integrated Computational Intelligence [6] and Synergetic Intelligence [7] systems.

2. Industry 4.0 and Its Components: On Integrated Intelligent Manufacturing

Main current trends in using advanced information and communication technologies for industrial automation are related to the initiative called *Industry 4.0* [8-10]. The term «The 4th Industrial Revolution (for short, Industry 4.0) firstly appeared at Hannover Messe in 2011 as an outline of German industrial perspectives. Nowadays the concept of Industry 4.0 has spread far beyond Germany and is widely used all over the world. The similar initiatives are called: *Industrial Internet* in the USA, *High Value Manufacturing Catapult* in the United Kingdom, *Usine du Futur* in France, *Fabbrica del Futuro* in Italy, *Smart Factory* in Netherlands, *Made Different* in Belgium, *Industrial Value Chain Initiative* in Japan, *Made in China 2025*, *National Technology Initiative* in Russia, and so on.

To differ from previous third industrial revolution Industry 3.0 that faced the problem of automating single machines and improve technological processes by using computers and electronic devices, Industry 4.0 focuses on the end-to-end digitization of all physical assets and their integration into digital ecosystems with value chain partners.

Digitization means converting information into a digital form; the result is called digital representation or, more specifically, digital image. Here the basic principle is «Ubiquitous Digitization», i.e. digitization and integration of both vertical and horizontal value chains, digitization of product and service offerings, digitization of business models and customer access, etc. [10]

In 2015, McKinsey [11] defined Industry 4.0 as «the next phase in the digitization of the manufacturing sector, driven by four basic factors: a) an astonishing rise in data volumes, computational power and connectivity, in particular, low-power wide-area networks; b) the emergence of analytics and business-intelligence capabilities; c) new forms of human-machine interaction such as touch interfaces and augmented-reality systems; d) improvements in transferring digital instructions to the physical world, such as advanced robotics and 3D printing».

The most important idea of Industry 4.0 is the fusion of the physical and the virtual worlds [8-10] provided by *cyberphysical systems* (CPhS) [12]. The emergence of CPhS means the inclusion of computational resources into physical-technical processes. In others words, embedded computers and networks monitor and control the physical-technical processes with feedback loops, where physical processes affect computations and vice versa. A virtual copy of physical world is created and well-timed decentralized decisions are made.

Let us note that CPhS are mechatronic systems enhanced by advanced tools of data/knowledge acquisition, communication and control. Their components continuously interact, providing CPhS self-adjustment and adaptation to changes. It is obvious that CPhS are crucial for production digitization and agentification. Here work pieces, devices, equipment, production plant and logistics components with embedded software are all talking to each other. Smart products know how they are made and what they will be used for. Thus, both production machines and equipment and products become cognitive agents involved into manufacturing and logistics processes.

Nowadays, BCG's nine basic technologies for Industry 4.0 are usually considered [13] (see Figure 1).

Figure 1. ***Nine Key Technologies for Industry 4.0.***

The implementation of Industry 4.0 is closely related to enterprise engineering problems [14], in particular, strategic and ontological engineering [15]. With Industry 4.0, companies, departments, functions, and capabilities will become much more cohesive, as cross-company, universal data-integration networks evolve and enable truly automated value chains. Both horizontal and vertical enterprise integration takes place [10].

On the one hand, the initiative Industry 4.0 means digital representation and vertical integration of basic processes across the entire enterprise, from product development and purchasing, through manufacturing, logistics and service.

On the other hand, horizontal integration goes beyond the internal enterprise operations by involving both suppliers and customers together with all other value chain partners. It includes technologies from track and trace devices to real time integrated planning with execution.

The Internet of Things (*IoT*) is a vision where every object in the world has the potential to connect to the Internet and provide their data so as to derive actionable insights on its own or through other connected objects. The Internet of Things allows people and things to be connected anytime, anyplace, with anything and anyone, ideally using any path/network and any service [16]. The appropriate technologies open new, wide opportunities for engineering networked enterprises.

The term «Internet of Things» was first coined by Kevin Ashton, the founder and head of Auto-ID Center in MIT, in 1999. As he stated, «the IoT has the potential to change the world, just as the Internet did – maybe even more so» [17]. It will comprise many billions of Internet-connected objects (ICOs) or «things» that can sense, communicate, compute, evaluate, interpret and potentially actuate, as well as have intelligence, multimodal interfaces and social ability.

Gartner defines IoT as «the network of physical objects that contain embedded technology to communicate and sense or interact with their internal states or the external environment» [18]. According to the International Data Corporation (IDC) [19], IoT is a dynamic «network of networks of uniquely identified objects that communicate without human interaction by using IP». This infrastructure, possessing self-configurating capabilities [16], is based on standard and interpretable communication protocols, where physical and virtual things have identifiers and physical attributes, use intelligent interfaces and are tightly interconnected.

Nowadays, such communication and network technologies as IPv6, web-services, Radio Frequency IDentification (RFID) and high speed mobile 6G Internet networks are employed.

The IoT incorporates basic concepts from pervasive, ubiquitous, cognitive computing, which have been evolving since the late 1990's and have now reached some level of maturity.

The IoT is an enabler to many application domains including intelligent manufacturing, product lifecycle management, smart logistics and transportation, aerospace and automotive industries. Society-oriented applications of IoT include smart cities, smart buildings (both home and office), telecommunications, new generation media, smart grids, medical technology, social robotics. Environment-focused applications include agriculture, breeding, recycling, environment monitoring and disaster alerting.

Cloud Computing [20,21] is a general term that refers to delivering computational services (servers, storage, databases, software, networking, analytics, etc.) over the Internet («the cloud»). It is a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources that can be rapidly provisioned and released with minimal management effort or service provider interaction.

Production resources and capacities can be intelligently sensed and connected into the cloud. The scalability of resources makes cloud computing interesting for business owners, as it allows enterprises to start small projects and invest in more resources only if there are rises in further service demand.

Today most production-related companies require intensive data and knowledge sharing across sites and partnerships. At the same time, the performance of cloud technologies will improve, achieving reaction times of just several milliseconds. As a result, machine data and functionality will increasingly be deployed to the cloud, enabling more data-driven services for production systems.

The advent of the Industry 4.0 initiative has brought serious changes to the *simulation paradigm* [22]. Now it supposes modeling of manufacturing and other systems by using virtual factory and digital twin concepts. The idea of digital twin extends the use of simulation techniques to all phases of the product lifecycle, where the products are developed and thoroughly tested in a virtual environment. Combining the real life data with the simulation results enables the rise of productivity and the improvement of product quality. Thus, simulations will be used more extensively in plant operations to leverage real-time data and mirror the physical world into a virtual model, which can include machines, products, and humans.

The new simulation paradigm is closely related to considerable technological advances of *augmented reality*. The last one brings users the chance to experience an augmented world by overlaying virtual information in the real world. This way the user can be in touch with both the real and virtual manufacturing worlds and receive real-time data or statistics.

For Industry 4.0, this may bring several advantages. It can be the perfect method to represent relevant information for technicians and workers in the enterprise, allowing them to watch real time information from the work they are performing. This is a suitable way to improve decision-making procedures and working operations.

Augmented-reality-based systems support a variety of services, such as selecting parts in a warehouse and sending repair instructions over mobile devices. Another great advantage is the possibility of enhancing industrial training and learning while reducing risks and costs.

Additive manufacturing is a transformative approach to industrial production that enables the development of lighter, stronger parts and systems. To differ from usual manufacturing, additive manufacturing (AdM) adds material to create an object. The processes of 3D printing and rapid prototyping are actually viewed as AdM components. In perspective, a combined approach based on «additive-subtractive» manufacturing will be of special concern [23].

Methods and tools of AdM are widely used in Industry 4.0 to produce small batches of customized products [13]. High-performance, decentralized AdM systems reduce transport distances and stock on hand.

In the framework of Industry 4.0, *robots* become more autonomous, intelligent and cooperative. Here the concepts of collective and collaborative robotics are of special concern. *Collective robotics* considers various groups of robots working collectively to solve a problem. It investigates both teams of cognitive robots and swarms of reactive robots. In its turn, *Collaborative Robotics* (briefly, *Cobotics)* deals with *cobotic systems*. A cobotic system is a «man – robot» system, where the participants collaborate in synergy to perform some tasks.

In other words, a *Collaborative Robot* (cobot) [24] is intended to work safely side by side with humans, i.e. physically interact with them in a shared workspace and learn. This is the difference with respect to conventional industrial robots, that operate autonomously or with a limited guidance.

Any cobot has to be equipped with a powerful onboard computer and complex sensor system, including an advanced computer vision and learning facilities. It allows to avoid the collisions with human partners and obstacles, as well as to operate in case of software crash.

These cobots will cost less and have a greater range of capabilities than those used in Industry 3.0.

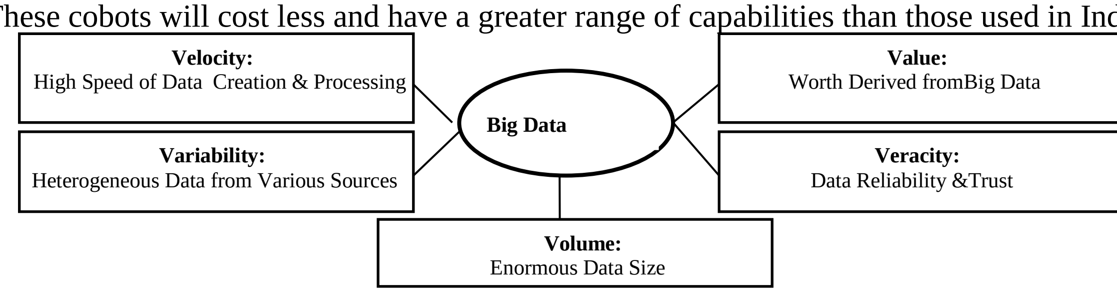

Figure 2. ***5V Representation of Big Data***

The term «*Big Data*» stands for large data sets that may be analyzed computationally to reveal useful patterns, trends and associations. Following [25], we adopt a 5V concept of Big Data (Figure 2) that associates it with data *volume, variability, velocity, veracity, value*. Here data variability is related to different distributed data sources, data velocity means the speed of data generation and processing, data veracity is attributed to a specific data source and data value expresses its utility. Moreover, the concept of Big Data Value Chain was introduced in [26]; it encompasses big data capture, processing, interpretation (visualization) while preparing decision.

In Industry 4.0 context, the collection and comprehensive evaluation of data from many different sources – production equipment and systems, as well as product lifecycle and enterprise management– becomes a necessary step to support real-time decision making.

Now the main technical challenge in the field of Big Data is that the speed of data generation exceeds the capacity of processing.

In the near future, this situation can seriously deteriorate. Indeed, IoT will be a major source of big data, contributing massive amounts of streamed information from billions of ICOs. Here M2M communications will generate enormous Internet traffic.

So decentralized Big Data Analytics and Information Mining are needed to cope with 5V. Specifically, Visual Analytics supports analytical reasoning by interactive user-friendly visual interfaces. In manufacturing, Data Analytics tools allow to optimize production quality, save energy and improve equipment service.

With the increased connectivity and use of standard communications protocols that come with Industry 4.0, the need in *cybersecuirity* to protect critical industrial systems and manufacturing lines from cyberattacks danger becomes crucial. In case of Industrial Internet, the design, deployment and maintenance of communications between heterogeneous and geographically distributed things create grand challenges related to security and privacy. To deal with such hard problems, reliable and safe communications, sophisticated identity and secure access management of machines and users are essential.

In our opinion, there exist at least two different ways in developing integrated intelligent systems and technologies for Industry 4.0. The first one consists in building intelligent counterparts of main Industry 4.0 components by employing conventional AI technologies, such as: Intelligent Simulation [27], Intelligent Cyberphysical Systems [28], Intelligent Cloud Computing and so on. The second way supposes the development of new trends in AI and Cognitive Sciences, such as Decision Theory with Imperfect Information [29,30], Granular Computing [31], Granular Measurements by Cognitive Sensor Networks [32], Context Aware Search, General Understanding Theory with practice-oriented models, etc.

3. Big Data and Information Granules

An exponential growth in industrial data volume, variability, value, originated by IoT and information services, drives various industries to develop innovative models and distributed tools for handling Big Data. Although many useful approaches and technologies, such as MapReduce, Hadoop, Disco (see [33]), have been successfully implemented, the need in new paradigms in Data Science becomes crucial.

Mining Information and Discovering Knowledge from Big Data requires the support of special techniques and new advanced technologies, in particular, data granulation and intelligent clustering methods and tools.

To start, it is worth using for the Big Data Analysis a classical model of transitions «Data-Information-Knowledge-Wisdom» [34] (in terms of M.Zeleny [35], these transitions correspond to various types of knowledge – «Nothing-Knowledge, What-Knowledge, How-Knowledge and Why-Knowledge»); it provides decision-making processes with a suitable information environment. To put in practice these transitions, we have to synthesize and study in detail multi-leveled understanding mechanisms: a) understanding situational links between objects or events; b) understanding behavior norms and patterns; c) understanding situation as a whole via basic principles (see Figure 3).

Granular Computing is a novel, promising approach to Big Data processing and analysis, involving theories, methods and models of Information Granulation [31,36]. Here the objects of computation are granules rather than singletons.

Data Granulation refers to the process of construction, representation and interpretation of granules. This concept is closely related to data abstraction and derivation of knowledge from information. By selecting different levels of granulation one can obtain different models of knowledge.

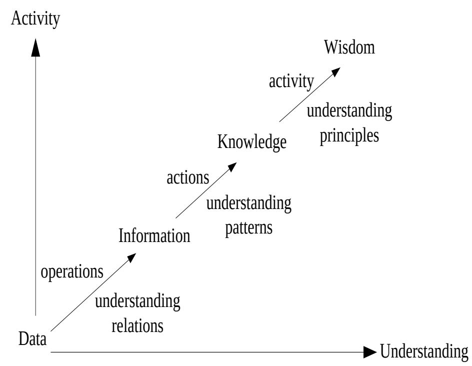

Figure 3. ***Representation of Basic Transitions «Data-Information-Knowledge-Meta-Knowledge» Using the «Understanding – Activities» Co-Ordinates***

The term «*Granule*» is originated from Latin word *granum* that means grain to denote a small particle in the world. Information granules are complex dynamic information entities which are formed to achieve some goal.

In [36] L.Zadeh specifies *granule* as a clump of entities drawn together by indistinguishability, proximity, similarity or functionality relations. Typical interpretations of granules are: part of the whole, sub-problem of the problem, uncertainty regions, knowledge entities, variable constraints.

The granules differ one from another by theirs sizes, complexity and abstraction level. There are various classifications of granules: physical and conceptual, one-dimensional and multidimensional, crisp and fuzzy, time and space granules, etc. In particular, information granules of *higher type* (2^{nd} type and n^{th} type, in general) are worth mentioning, where the granule is specified on the basis of another granule without recourse to numeric data.

Basic problems and components of data granulation theory are presented below:

- ❑ General data granulation principles and criteria;
- ❑ Basic approaches and methods of data granulation;
- ❑ Specification, interpretation, classification of granules;
- ❑ Formal models of granules and granular structures;
- ❑ Mappings between different granulation levels and various granular models;
- ❑ Quantitative characteristics of both granules and granulation process.

The granulation principle by Zadeh [36] is formulated as follows: to exploit the tolerance for imprecision, employ the coarsest level of granulation which is consistent with the allowable level of imprecision. In his turn, Pedrycz [31] proposes to find a compromise between granule specificity and coverage criteria.

The level of detail (which is represented by the size of information granules) becomes an essential part of the hierarchical processing of information, where different levels of the hierarchy correspond to the size of information granules.

Two keynote characteristics of information granulation are specificity and coverage. The specificity sp expresses a level of abstraction conveyed by the information granules. The higher is the specificity, the lower is the level of abstraction. It is clear that any specificity measure sp(•) on the set of granules satisfies the property of anti-monotonicity together with a boundary condition for singletons [37]. In a formal way, if G_1 and G_2 are granules, and $G_1 \subset G_2$, then $\mathrm{sp}(G_1) \geq \mathrm{sp}(G_2)$. In the limit case for a singleton $\boldsymbol{x}_0$ we have $\mathrm{sp}(\{\mathbf{x}_0\})=1$. A dual non-specificity measure nsp is monotonous, $G_1 \subset G_2, \Rightarrow \mathrm{nsp}(G_1) \leq \mathrm{nsp}(G_2)$.

The concept of coverage of information granule, cov(•) is discussed with regard to some experimental data existing in $\mathbf{R}^n$, that is $\{x_1, x_2, \ldots, x_N\}$. It is associated with a capability to represent (cover) these data. In general, the larger number of data is being "covered", the higher the coverage measure. Formally, the coverage is a non-decreasing function of the number of data that are represented by the given information granule G. Depending on the nature of information granule, the definition of $\mathrm{cov}(G)$ can be easily expressed. For instance, when dealing with a multi-dimensional interval (hypercube) H, $\mathrm{cov}(H)$ in its normalized form is related with the cardinality of the data belonging to H, $\mathrm{cov}(H)=(1/\mathrm{N})\,\mathrm{card}\{x_i |\ x_i \in H\}$.

Big Data Granulation can be achieved by integration and detailing operations. In case of integration (abstraction) larger and/or higher level granules are developed based on smaller and/or lower level granules. Inversely, for detailing (dispersion), things are done in the opposite way.

To express «data–information» transitions, opposite granulation–degranulation processes were specified in [31]. *Granulation* is made by representing data with using the prototypes (clusters). Inversely, *degranulation* means a reconstruction of data from information granules.

3.1. Top-Down vs Bottom-Up Approaches to Information Granulation

There exist two generic formal approaches to constructing granules – top-down and bottom-up. The top-down granulation process is based on universal set that is divided into some families of subsets, where each subset is partitioned into smaller subsets. So the role of the granule in data mining is similar to that of subset, class or cluster in a universe U.

In case of top-down approach, two basic models of granules are coverings and partitions. Let U be a covering: $U = G_1 \cup G_2 \cup \ldots \cup G_n$. This covering condition together with $G_i \cap G_j = \varnothing$, $\forall$i,j=1,…,n specifies a partition, i.e. a family of pairwise non-overlapping sets, the union of which is the universe U.

Now let E be an equivalence relation that is reflexive, symmetric and transitive. This relation E can be represented by a mapping from U to 2^U given by $[x_E] = \{y \in U \mid xEy\}$. The subset $[x_E]$, called an equivalence class containing x, is a granule that consists of indistinguishable elements.

A family of all equivalence classes is called a quotient set represented as $U/E = \{[x_E] | x \in U\}$. Let us recall that there is a one-to-one correspondence between the relations of equivalence and the partitions of sets.

Therefore, the equivalence class $[x_E]$ that specifies a granule, can be interpreted as: 1) a subset of the universe U; 2) an element of the quotient set U/E. As a result, we obtain two information-oriented representations of the same universe – a fine-grained U and a coarse-grained U/E respectively.

In case of bottom-up granulation process larger granules are built from smaller ones (from numbers). Here a primary granule can be formed as a selected point neighborhood. Then a pre-topology of neighborhood system can be constructed.

Another natural granulation technique is clustering viewed as a generation of information granules on the basis of experimental data. Clustering means grouping a set of objects in such a way that objects in the same group (cluster) are more similar to each other than to those in other group. Here some similarity or dissimilarity measure is used, for instance, the squared Euclidean distance $D(x, y) = \| x - y \|^2$.

In granular clustering [31] the transition from data to information granules means the shift from numeric prototypes to granular prototypes.

3.2. Big Data Granulation with Non-Standard Sets

Big Data are really riddled by uncertainty, vagueness, inaccuracy, ambiguity factors. For instance, uncertainties not only exist in data themselves, but may occur at each phase of Big Data processing. Ambiguity is caused by inconsistent data from different information sources. Besides, measurement is never exact (see [32]). According to the Guide to the Expression of Uncertainty in Measurement [38], a difference is made between the type A stochastic uncertainty and the type B non-stochastic uncertainty that can have interval or fuzzy nature.

Stochastic or Probability-oriented data granules are expressed in the form of probability density functions or (simple) probability functions; they capture a collection of elements resulting from an experiment. The granularity, based on the concept of probability, reflects an occurrence of different elements. Each element in a set has a probability function belonging to [0,1], which quantifies the degree of membership to the information granule.

Interval analysis is also a good framework for uncertain, imprecise, approximate data representation. The tools for interval systems in controlling technological processes were investigated in [39].

Let us briefly consider some non-classical set theories in which information granules can be formally defined and processed. Basic characteristics of any set are boundaries and measure. Classical sets have crisp boundaries and additive measure. They satisfy two basic postulates: membership postulate and distinguishability postulate. The *membership postulate* is similar to the excluded middle law in classical logic: every element of a set must be uniquely classified as belonging to the set or not. Moreover, according to the *postulate*, a set is viewed as a collection of quite distinguishable elements which can be enumerated, represented by the list. If these postulates are rejected, then we deal with non-standard set theories.

Below we shall consider non-standard sets depending on observer's awareness. Let us denote by U some universal set and by $A \subset U$ its subset. It is easy to define $B = U \setminus A$. Here the procedure of information granulation is reduced to specifying two regions: certainty region A and uncertainty region B (Figure 4).

Now let us denote A by X^+, and B by X^-. A non-standard set may be introduced by a triple $X = \langle X^+, X^-, X^0 \rangle$, where $X^+ = \{x | x \in X\}$, $X^- = \{x | x \notin X\}$, $X^0 = \{x | x ? X\}$, where ? stands for ignorance. Equivalently, it may be expressed by a three-valued membership) function $f(X) \in \{0, 0.5, 1\}$ that is a direct generalization of set characteristic function.

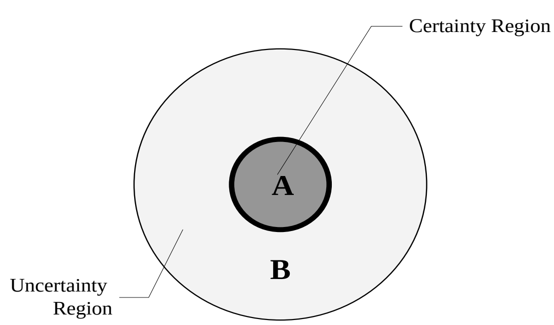

Figure 4. ***Non-Standard Set with Certainty and Uncertainty regions.***

Below we will specify two cases of data uncertainties by considering two different interpretations of X^0 region.

Over-Definite Set. In case of redundant and contradictory data about membership value for X^0 we propose the following interpretation for $f(X^0) = 0.5$: $x \in X^0$ and $x \notin X^0$ at the same time. For data, it corresponds to different valuation of the same attribute (like “resource is vacant” and “resource is occupied” at the same time).

*Sub-Definite Set***.** In case of lacking data about membership for X^0 the value $f(X^0) = 0.5$ is viewed as total uncertainty; nevertheless, information gathering may lead to the transition to well-defined set with either certainty or uncertainty regions. In particular, A.S.Narin’yani introduced lower l and upper u approximations for X^0 [40].

In data processing, sub-definite situation means the presence of unassigned attributes.

These two concepts of over-definiteness and sub-definiteness for sets correspond to the notions of gluts and gaps in Dunn-Belnap’s semantics for multi-valued logics.

The transition from two to n nested sets $A = (A_1,\ldots,A_n)$, $A_i \subset U$, i=1,.., n, $A_1 \subseteq \ldots \subseteq A_n$ leads to the concept of flou set.

Now let us discuss more sophisticated non-standard sets for building granular models, such as Rough Sets and Fuzzy Sets.

The concept of *Rough Set* [41] also has some information-grounded origins. The main assumption is that objects characterized by the same information are indiscernible. This indiscernibility can be defined relative to selected sets of functions representing objects attributes or features.

In other words, Rough Set emphasizes the idea of granulation as data approximation realized in terms of the indiscernibility relation. The roughness of data description is given by lower and upper approximations of a certain rough set.

More specifically, let U be the universal set, and denote by $R \subseteq U \times U$ an indiscernibility relation representing the lack of information about elements of U. It is usually assumed that R is reflexive, symmetrical and transitive, i.e. R is an equivalence relation. The pair $APR = (U, R)$ forms an approximation space.

Equivalence classes of the indiscernibility relation are called elementary sets in the approximation space, and every union of elementary sets in APR is referred as a composed set.

The lower approximation of a set X with respect to R denoted by $\underline{R}X=\{x||x|_R \subseteq X\}$ is the greatest composed set contained in X. It is viewed as the set of all data which can be certainly classified as X with respect to R (are certainly X with respect to R). The upper approximation of a set X with respect to R denoted by $\overline{R}X=\{x|\,|x|_R \cap X\}$ is the least composed set containing X. It is the set of all data which can be possibly classified as X in respect to R (are possibly X in view of R).

The boundary region of a set X in respect to R is the set of all data, which can be classified neither as X nor is not X with respect to R.

The rough set concept is illustrated in Figure 5.

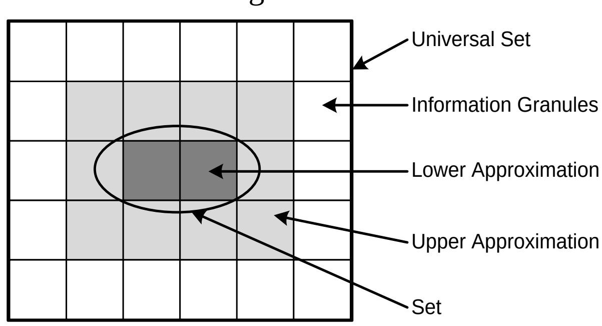

Figure 5. ***Rough Set Specification.***

Fuzzy Sets [42] provides an important conceptual and algorithmic generalization of sets by allowing transparent set boundaries with using partial membership of elements to a set. The description of a fuzzy set is given in terms of its membership function that takes values in the unit interval. Formally, a fuzzy set *A* is described by a membership function, mapping the elements of a universe ***U*** to the unit interval [0,1].

Fuzzy set techniques, allowing us to deal with uncertainties in raw data and perform data annotation, enable a valuable generation of both information and knowledge granules.

A natural way to constructing complex, higher level and multidimensional granules in the context of Bid Data usage is the formation of hybrid fuzzy sets. Here both internal and external hybridization are welcome. By internal hybridization we mean the construction of non-traditional fuzzy sets to deal with both data fuzziness and imprecision (interval-valued fuzzy sets), fuzziness and uncertainty (probabilistic sets, intuitionistic fuzzy sets), fuzziness and roughness (rough fuzzy sets and fuzzy rough sets), and so on. A typical example of external hybridization is a Soft Computing model, integrating fuzzy logic, neural network and evolutionary algorithm.

Following L.Zadeh [36], we shall consider procedures of fuzzification, randomization, usualization as granulation instances.

One can refer to a plethora of fuzzy sets applications in industrial automation in [30] and the proceedings of precedent WCIS conferences.

4. How to Form «Understanding Agents»?

The development of Industry 4.0 supposes a total enterprise agentification, where people, robots, industrial equipment (machines and materials), manufacturing software tools and even enterprise products form an Intelligent Organization as a System of Multi-Agent Systems (MAS). Most of these MAS include Cognitive Agents.

Let us mention some basic features of human cognition, which are of special interest for developers of artificial cognitive agents. Firstly, cognition is an open system based on both available knowledge and current data perception. Secondly, cognition does not make straight conclusions, but generates hypotheses, and these hypotheses should be confirmed or denied. Thirdly, agent cognition is intrinsically linked with the organization of action (as information process, physical movement or local environment change). And fourthly, cognition is iteratively connected with understanding. On the one hand, the cognitive capability itself and the result of action strongly depend on the reached understanding level (pre-understanding). On the other hand, human understanding is specified by cognitive capacities, available knowledge and language structure. Although, natural language understanding is driven not so much by purely linguistic factors as by extra-linguistic factors, including personal experience and presupposition.

Understanding is a necessary condition of communication between cognitive agents and their joint work. It is obvious that the human-robot cooperation, development of Social Cyberphysical Systems and Internet of Things require some mutual understanding.

Understanding is not a new problem for AI, but earlier it was mainly considered in the context of natural language processing and text analysis. Such understanding objects, as behavior, decisions, situations, remain almost unexplored.

Let us take the following basic definition: *Understanding* is a *universal cognitive process* (operation) that evaluates an analyzed object (text, behavior, situation, phenomenon) on the basis of some standard, norm, pattern.

This definition has an axiological nature. It is founded on value theory, because any evaluation implies some value (or logical inference from accepted values by using some general rules) [43]. Two basic operations to enable understanding are: a) the search for some norm and its formal representation; b) justification of the norm's applicability in a specific situation.

The level of agent task understanding can be specified by evaluating the results of his actions, which should not contradict the norms of agent behavior.

Norms are social bans and constraints imposed on an agent by an organization (community). They represent a special case of evaluations: these are socially tested and fixed assessments.

The formal model of norm viewed as a prescription to action is given by a quadruple

$$NORM = \langle AG, act, W, M \rangle,$$

where AG is a set of agents to whom a norm is addressed, $act \in ACT$ is an action being an object of normative regulation (the norm content), W is a set of worlds, where the norm is useful (application conditions or specific circumstances in which the action should be performed or not), M is a set of basic modal systems related to the action *act*, for example the system of deontic modalities $M_D = \{O,P,F\}$.

Here O stands for "obligatory", P means "permitted" and F is "forbidden".

An evaluation is transformed into a norm by threat of punishment, i.e. standardization of norms is made through sanctions. Here a typical sanction reasoning pattern is: $\Box q$ («obligatory *q*») and «if not *q*, then punishment or degradation».

So an information structure of cognitive agent combines both descriptive & normative models (Figure 6).

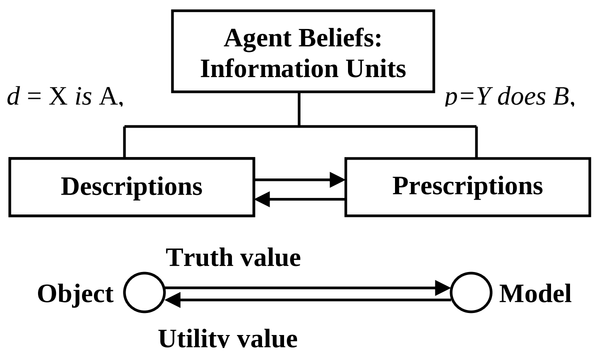

*Figure 6. **Two Sides of Agent Beliefs.***

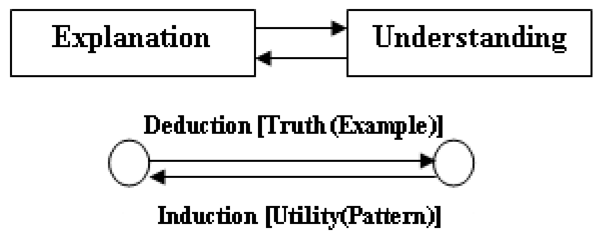

*Figure 7. **Illustration of Relationships Between Explanation and Understanding.***

Descriptions *d* contain the data on the states of environment perceived by the agent, and prescriptions *p* give the normative information about possible (permitted) actions or behavior patterns. Here a description is characterized by a truth value (possibly gradual) T(*d*) and a prescription has a worth (or utility) value W(*p*). Hence, a truth value is the correspondence between the object and its descriptive model (the object is primary), whereas an utility (worth) value gives the inverse mapping from normative model to this object (the norm is primary). A general agent understanding mechanism also has a dualistic «Description-Evaluation» nature. Every object having a standard prototype (pattern) is understandable, and the reason of misunderstanding consists in the lack of such pattern or its non-obviousness.

It is worth noticing that a pattern as a basis for understanding significantly differs from an example. The example refers to a real existing object, whereas the pattern shows what should be done ideally. The examples are taken to support descriptive models, but references to the patterns and standards serve as justifications of norms and prescriptions.

This dualism is used in «Explanation-Understanding» relationships (Figure 7).

Explanation, considered as the reduction of studied phenomenon to the scientific law, representative example or general truth, is based on descriptive model and helps to understand it, but understanding as searching for rule or standard has a normative basis.

Logical lattices and bilattices of strong and weak norms and anti-norms will be constructed to provide an artificial agent of Industry 4.0 enterprise with some basic understanding mechanisms.

Conclusion

Basic components and technologies of Industry 4.0 – a new paradigm of Industrial Automation, based on NBICS-convergence Basic components and technologies of Industry 4.0 – a new paradigm of Industrial Automation, based on NBICS-convergence concept and new generation integrated intelligent systems have been reviewed. The concept and new generation integrated intelligent systems have been reviewed. The emphasis has been made on the problem of coping with distributed industrial Big Data issued from various sources, including the objects of Industrial Internet. Some ways of hierarchical, multi-dimensional, multi-faced Big Data representation and analysis have been considered. In particular, advanced human-centered models and technologies of information granulation, founded on both classical mathematical approaches and non-standard sets, have been discussed. The problem of developing special cognitive tools to build artificial understanding agents for Integrated Intelligent Enterprises has been faced.

In our further work, we hope to develop some good examples of merging advanced information, intelligent and cognitive technologies in order to agentify Industry 4.0 enterprises. Specifically, the joint usage of many-valued and fuzzy reasoning models with granular clustering techniques seems to be a suitable way to develop an axiological approach in building an artificial agent capable to understand relations, patterns and situations in enterprise activities.

The work is supported by RFBR, project №17-07-01374.

References:

1. Горнев В.Ф., Тарасов В.Б. О проектах и мероприятиях по международной программе IMS// Автоматизация проектирования. – 1999. – №2. – С.42-46.
2. Brown R.H., Syntera H. Global Manufacturing in the 21st Centrury. Final Report. IMS, 1999.
3. Goldman S.L., Nagel R.N., Preiss K. Agile Competitors and Virtual Organizations: Strategies for Enriching the Customer. – New York: Van Nostrand Reinhold, 1995.
4. Тарасов В.Б. Агентно-ориентированный подход к формированию интеллектуальных и виртуальных предприятий// Реинжиниринг бизнес-процессов на основе современных информационных технологий. Системы управления знаниями. Сборник научных трудов международной научно-практической конференции (РБП-СУЗ-2005, Москва, 9-10 июня 2005 г.). – М.: МЭСИ, 2005. – С.37-44.
5. TOVE Ontology Project 2010. http://www.eil.utoronto.ca/enterprise-molelling.
6. Pedrycz W. Computational Intelligence: The Introduction. – Boca Raton: CRC Press, 1997.
7. Karpov V.E.,Tarassov V.B. Synergetic Artificial Intelligence and Social Robotics// Proceedings of the 2nd International Scientific Conference on Intelligent Information Technologies for Industry (IITI'17, Varna, September 14-16, 2017) /Ed. by A. Abraham, S.Kovalev, V.Tarassov et al. Advances in Intelligent Systems and Computing. Vol.679. – Cham, Switzerland: Springer International Publishing, 2018. – P.3-15.
8. Kagermann H., Helbig J., Hellinger A., Wahlster W. Recommendations for Implementing the Strategic Initiative Industrie 4.0: Securing the Future of German Manufacturing Industry. Final Report of the Industrie 4.0 Working Group 2013.
9. Schwab K. The Fourth Industrial Revolution. –Geneva: World Economic Forum, 2016.
10. 2016 Global Industry 4.0 Survey. Building the Digital Enterprise. – London: PwC, 2016.
11. Baur C., Wee D. Manufacturing's Next Act. McKinsey & Company, June 2015.
12. Suh S.C., Tanik U.J., Carbone J.N., Eroglu A.E. (Eds.) Applied Cyber-Physical Systems. Heidelberg: Springer-Verlag, 2014.
13. Gerbert Ph., Lorenz M., Rusmann M. et al. Industry 4.0: The Future of Productivity and Growth in Manufacturing Industries. BCG, April 2015.
14. Dietz J.L.G. et al. Enterprise Engineering Manifesto// Adavances in Enterprise Engineering I. Lecture Notes in Business Information Processing. Vol.10. – Heidelberg: Springer Verlag, 2011.
15. See also: www. ciaonetwork.org.
16. Тельнов Ю.Ф. Инжиниринг предприятий на основе интеллектуальных технологий// Информационно-измерительные и управля-ющие системы. – 2013. – №11, ч.2. – С.55-60.
17. Sundmaeker H., Guillemin P., Friess P., Woelffle S. (Eds.). Vision and Challenges for Realizing the Internet of Things. – Luxembourg: EU, CERP IoT, 2010.
18. Ashton K. That Internet of Things// RFID Journal, 22 June 2009.
19. Gartner, Internet of Things, http://www. gartner.com/it-glossary/internet-of-things/IDC, IDC Market in a Minute: Internet of Things http://www.ids/com/downloads/idc_market.

20. Mell P., Grance T. The NIST Definition of Cloud Computing. – Gaithersburg: National Institute of Standards and Technology, 2011.
21. Sehgal N.K., Bhatt P.Ch. Cloud Computing: Concepts and Practices. – Heidelberg: Springer-Verlag, 2018.
22. Rodic B. Industry 4.0 and the New Simulation Modeling Paradigm// Organizacija. – 2017. – Vol.50, №3. – P.193-207.
23. Евгенев Г.Б. Методы программирования комбинированной аддитивно-субтрактивной обработки// Известия высших учебных заведений. Машиностроение. – 2017. – №4. – С.47-56
24. Faneuff J. Designing for Collaborative Robotics. – Sebastopol Ca: O’Reilly, 2017.
25. Wang H., Xu Z., Pedrycz W. An Overview on the Role of Fuzzy Set Techniques in Big Data Processing: Trends, Challenges and Opportunities //Knowledge-Based Systems. – 2016. – P.1-16.
26. Miller H.J., Mork P. From Data to Decisions: a Value Chain for Big Data// IT Prof. – 2013. – Vol.15. – P.57-59.
27. Artiba A., Emelyanov V.V., Iassinovski S.I. Introduction to Intelligent Simulation: The RAO Language. – Dordrecht: Kluwer Academic Publishers, 1998.
28. Muller H. The Rise of Intelligent Cyberphysical Systems// Computer. – 2017. – Vol.50, №12. – P.7-9.
29. Aliev R.A., Huseynov O.H. Decision Theory with Imperfect Information. – Singapore: World Scientific Publishing Company, 2014.
30. Юсупбеков Н.Р., Алиев Р.А., Алиев Р.Р. и др. Интеллектуальные системы управления и принятия решений. – Ташкент: Узбекистон миллий энциклопедияси, 2014.
31. Pedrycz W. Granular Computing: Analysis and Design of Intelligent Systems. – Boca Raton: CRC Press, 2013.
32. Tarassov V.B., Svyatkina M.N. Monitoring Railway Infrastructure Objects with Using Cognitive Measurements// Proceedings of the VIII World Conference on Intelligent Systems for Industrial Automation – 2014 (WCIS-2014, Tashkent, Uzbekistan, November 25-27, 2014)/ Yusupbekov N.R., Aliev R.A., Kacprzyk J., Bonfig K. (Eds.) Kaufering: b-Quadrat Verlag, 2014. – P.194-204.
33. Федотова А.В.,Ануфриева Е.Ю.,Табаков В.В. Современные подходы к управлению большими данными в жизненном цикле изделия// Интеллектуальные системы и технологии: современное состояние и перспективы–2017. Сборник научных трудов IV-й международной летней школы-семинара по искусственному интеллекту для студентов, аспирантов, молодых учёных и специалистов (ISYT–2017, СПб, 30 июня-3июля 2017г.). – СПб.: Политехника-сервис, 2017. – С.178-188.
34. Ackoff R. From Data to Wisdom// Journal of Applied Systems Analysis. – 1989. – Vol.16, P.3-9.
35. Zeleny M. Human Systems Management: Integrating Knowledge, Management and Systems// World Scientific. – 2005. – P.15-16.
36. Zadeh L.A. Toward a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic// Fuzzy Sets and Systems. – 1997. – Vol.90. – P.111-127.
37. Yager R.R. Measures of Specificity// Computational Intelligence, Soft Computing and Fuzzy-Neuro Integration with Applications / Kaynak O., Zadeh L.A., Turksen B., Rudas I.J. (Eds.). – Heidelberg: Springer-Verlag, 1998. – P.94-113.
38. Guide to the Expression of Uncertainty in Measurement (GUM). – Geneva: International Organization of Standardization, 1995.
39. Yusupbekov N.R., Igamberdiev H.Z., Bazarov M.B. The Interval Approach to Designing of Control Systems of Technological Processes// Proceedings of the VIIth World Conference on Intelligent Systems for Industrial Automation/ (WCIS-2012, Tashkent, Uzbekistan, November 25-27, 2012). Yusupbekov N.R., Aliev R.A., Bonfig K. (Eds.). –Kaufering: b-Quadrat Verlag, 2012. – P.11-16.
40. Narin’yani A.S. Sub-Definiteness and Basic Means of Knowledge Representation// Computers and Artificial Intelligence (Bratislava). – 1983. – Vol.2, №5. – P.443-452.
41. Pawlak Z. Rough Sets// International Journal of Computer and Information Sciences. – 1982. –Vol.11. – P.341-356.
42. Zadeh L. Fuzzy Sets// Information and Control. – 1965. – Vol.№8. – P.338-353.
43. Ивин А.А. Логика оценок и норм. Философские, методологические и прикладные аспекты. – М.: Проспект, 2016.

Yusupbekov Nodirbek Rustamovich – Tashkent State Technical University, Tashkent city, Republic of Uzbekistan
E-mail: dodabek@mail.ru;
Tarassov Valery Borissovich – Bauman Moscow State Technical University, Moscow, Russia
E-mail: Vbulbov@yahoo.com.