Rough Sets Turn 40:
From Information Systems to Intelligent Systems

Andrzej Skowron
Systems Research Institute
Polish Academy of Sciences, Warsaw, Poland
Cardinal Stefan Wyszyński University, Warsaw, Poland
Email: andrzej.skowron@gmail.com

Dominik Śleza
Institute of Informatics
University of Warsaw, Warsaw, Poland
QED Software Sp. z o.o., Warsaw, Poland
Email: slezak@mimuw.edu.pl

Abstract—The theory of rough sets was founded by Zdzisław Pawlak as a framework for data and knowledge exploration. His seminal paper titled “Rough Sets” was published in 1982, in International Journal of Computer and Information Sciences. One of the key aspects that lets us use rough sets in practical scenarios is the notion of information system, which comes from earlier Professor Pawlak’s works. Information systems are the means for data and knowledge representation. They constitute the input to rough set mechanisms aimed at computing approximations of concepts and deriving compacted, interpretable decision models. In particular, the fundamental notion of the indiscernibility relation is defined on the basis of a given information system. Accordingly, we discuss to what extent information systems can serve as the basis for intelligent systems. We claim that in many cases it is not enough to treat a data set – represented as an information system – as a purely abstract object with no linkage to the data origins. Oppositely, we should give ourselves a technical possibility to construct information systems dynamically, taking into account interaction with physical environments where the data comes from. With this respect, we refer to the notions of interactive granular computing and we generally consider together the paradigms of rough sets, information systems, and information granulation.

Index Terms—Rough Sets; Information Systems; Data Mining; Big Data; Interactive Granular Computing; Intelligent Systems

I. INTRODUCTION AND BASIC CONCEPTS

Zdzisław Pawlak (1926-2006) founded the theory of rough sets with the aim of analyzing incomplete information by means of approximations [1]. His book published in 1991 [2] established worldwide recognition of rough sets as an approach which can model complex problems using simple constructs. The idea of rough sets originated from earlier Professor Pawlak’s research on knowledge representation and information retrieval [3], [4]. The key observation was that we often operate with objects (cases, instances) which are indistinguishable from each other, so approximating the sets of objects by using indistinguishable “blocks” is the most reasonable thing we can do. A need for reaching such simple solutions was important for Professor Pawlak during his entire scientific career which included far more achievements than just rough sets [5]. It is also interesting to note that the basic models he used working in different fields, such as conflict analysis [6] or concurrency [7], were based on information systems. In the same time, his personal interests were going far beyond computer science and science in general [8].

The updated (after 25 years) viewpoint on rough sets was presented in [9], [10], [11]. As it happens with every theory, it took time to understand the actual contribution of rough sets with respect to other approaches. Firstly, rough sets were compared with the theory of fuzzy sets [12]. Over the years, it turned out that these two theories can be successfully combined because they offer complementary granules understood as computational building blocks for approximations [13], [14]. In the next sections, one can see a number of examples of such combinations. Another thread of comparisons was devoted to the principles of information granulation and granular computing (GrC) [15], [16], whereby it is generally assumed to operate on groups of objects / instances / items gathered together into granules. In this case, the relationship turned out to be even more natural because such granules (or generally, various frameworks and layers of granular information) are a natural input for calculations related to rough sets.

Nowadays, rough sets are particularly popular in the area of learning decision models from the data. There are many rough set methods aiming at feature selection and interpretable decision model construction [17], [18], [19]. However, the principles of rough sets have much broader influence. They have an impact on various methodologies of decision making, e.g. multi-criteria decision making [20] or three-way decision making which strongly relies on rough set positive, negative and boundary regions [21]. Moreover, rough sets are employed to enhance expressive and algorithmic capabilities of many approaches to data mining and knowledge discovery. A good example here is an extension of standard data clustering toward rough clustering (whereby we search and operate with lower and upper approximations of rough data clusters) and its fuzzy hybridizations [22], [23], [24]. Another example corresponds to formal concept analysis [25] which has this year its 40th anniversary exactly like rough sets. With that respect, there are many approaches to constituting rough set approximations of formal concepts or in other words, formal concepts which comprise of rough set approximations [26]. Rough sets turned out to be useful also in other fields of science and industry. For instance, they were adopted in design of database engines and solutions, with respect to query language extensions (which can be referred as rough querying) and acceleration of commercial database software performance [27], [28].
In all above scenarios, rough set methods (or mechanisms that adopt rough set principles) require an input to derive approximations of concepts, relations, etc. In [1], Professor Pawlak discussed several examples of domains, whereby the inputs to rough set methods could take different forms. One of those forms corresponded to the notion of an information system [3]. In the framework proposed by Professor Pawlak, information systems are aimed at representing the underlying data, information or knowledge that we want to use to describe (and approximate) the concepts of practical interest. An information system comprises of objects, attributes and the values of attributes over those objects. Information systems resemble data tables in the theory of relational databases. It should be noted that data tables, in particular decision tables, have been studied and used in many applications since 60-ties of the XX-century [29], [30]. Still, we need to remember that Professor Pawlak’s goal to operate with information systems was to represent information (including information derived from the data) rather than focus on the data itself [2]. This topic is actually wider and one can find analogies between information systems and data / information / knowledge representation frameworks in some other theories [25], [31], [32].

The complete idea of information systems has been created thanks to cooperation of Professor Pawlak with other scientists [4], [33]. At that time, there was a great demand for constituting the foundations for information representation and retrieval [34]. All those works were reaching beyond a standard understanding of data storage and processing, particularly with respect to incompleteness, imprecision and indeterminism of information that one needs to handle in practice. Over the years, the concept of non-deterministic information systems evolved in many interesting directions [35], [36]. Going back to the principles of information granulation [16], one may say that such information systems comprise of descriptions – also called signatures – of granules (composed in different ways, specific to different applications) in terms of available attributes, where those signatures are not always precise. Operating with granules introduces a useful abstraction, a kind of border between granules’ signatures (being the inputs to further computations) and granules’ internals, whereby one can locate guidelines about how and / or what for those signatures are derived from physical data sources. This is why in this paper, we are interested equally in: (i) rough sets as the methodology for deriving concept approximations and compacted decision models, (ii) information systems which provide the input to those derivations, and (iii) the paradigms of GrC and information granulation, as various forms of granules can be “hidden right under” the abstract descriptions that information systems consist of. We focus particularly on the methodology of interactive granular computing (IGrC) [37] and interactive information systems [38], where the above-mentioned abstraction was explicitly introduced in terms of complex granules (c-granules) with the embedded control mechanisms that decide how their signatures are derived.

IGrC raised from the observation that traditional ways of computing do not take into account how the process of perceiving attribute values is realized, where and when to access the concerned objects in a physical space, and why particular attributes are selected. This kind of awareness / attention / agency [39] is important for designing intelligent systems which are supposed to deal with complex phenomena in the real physical world [40]. This becomes even more important when one realizes that unexperienced data scientists often treat the available data sets as an ultimate baseline without investigating how those sets were collected.

In the above considerations, we evolved from information systems toward intelligent systems. We referred to the concepts of awareness and perception (perceiving attribute values), we can also refer to the concept of cognition. With this respect, let us cite the following statement of Leslie Valiant1, which is particularly relevant when extending GrC toward IGrC:

A fundamental question for artificial intelligence is to characterize the computational building blocks that are necessary for cognition.

The above computational building blocks can be treated as a generalization of granules known from GrC [15]. (In particular, indistinguishable blocks / indiscernibility classes known from rough sets can be treated as atomic granules.) Naturally, such granules / blocks / classes have been already studied within GrC in the context of cognition [14]. However, when moving from blocks to computational building blocks, it is indeed useful to rely on IGrC because therein, the aforementioned c-granules are aimed at more tasks than just storing their signatures. Such c-granules build the relevant configurations of physical objects, initiate and modify interactions between them, so they are generally responsible for perceiving the physical world. They link physical objects with abstract objects used to represent the instances of decision making from the viewpoint of models working on information systems.

In the next sections, we consider some examples of challenges with respect to which IGrC can be worth adapting. For now, let us mention just one of them, namely hierarchical learning [41], [42]. From a logical viewpoint, one can think about it as learning satisfiability relations at different levels of hierarchy [43], [44]. This includes learning logical structures (e.g. relational structures or models), as well as logical formulas and their semantics expressed using those structures. The current methods of hierarchical learning are often based on GrC, with a special emphasis on designing hierarchies of information systems by basing on domain knowledge [45], [46]. However, granules on which the corresponding reasoning pipelines are performed, cannot neglect the underlying hierarchies of physical objects that are crucial for perception processes. Also, different layers of hierarchy can be connected to different types of sensors and actuators. Thus, the IGrC framework can be helpful indeed to embrace both, the relationships between information systems at different levels of hierarchy and the relationships between particular systems and the associated physical-object-related information sources.

1people.seas.harvard.edu/~valiant/researchinterests.htm
In the rest of the paper, in Section II, we refer to some selected literature on rough sets. This section is quite extensive given the anniversary flavor of the paper. In Section III, we go back to the discussion about the importance of information systems. We emphasize a need of operating with information systems (and the results of rough set computations over information systems) considered in a wider context of interactions between abstract and physical objects of different sorts. We go through several aspects of applications, whereby this kind of interaction is needed. We show to what extent the paradigms of IGrC can be helpful. We also refer to some concepts known from the domain of big data in order to put our discussion into other contexts. In Section IV, we conclude the paper.

Let us reemphasize the retrospective context of this paper. Besides the 40th anniversary of rough sets (which is our major focus) and the 40th anniversary of formal concept analysis (which was mentioned above), there are two more celebrations worth mentioning. The first of them refers to the rough set workshop series which “visited” the FedCSIS conferences for the first time 10 years ago and which is now back to the program of technical sessions. Secondly, this year’s FedCSIS hosts the 30th International Symposium on Concurrency, Specification and Programming (CS&P 2022). The topics related to rough sets and information systems have been always visible at the CS&P events. In particular, the above-cited papers [36], [40], [44] come from CS&P.

II. SELECTED RELATED WORK ON ROUGH SETS

In order to provide a better viewpoint of the theory and applications of rough sets, we refer to two events from the past. These references will also constitute a better background for our major goal in this paper, which is the review of new advances on rough-set-related information systems.

A. Rough Sets at FedCSIS 2012

The first considered event is FedCSIS 2012 held 10 years ago in Wrocław, Poland. As already mentioned, that was the first time when rough sets occurred so intensively at a FedCSIS conference. Let us start outlining the FedCSIS 2012 rough-set-related publications from [47], [48]. The first paper equips the Variable Precision Rough Set (VPRS) approach with a Bayesian background. The second paper combines VPRS with fuzzy rough set methods in order to produce flexible decision rules. In summary, both papers deal with information imprecision—not modeled by probabilities (which is the domain of VPRS) and fuzziness (which can be used to work e.g. with partial matching of rules’ antecedents) – and attempt to extract interpretable decision models from the data [11].

The topic of rough-set-driven decision rules is considered also in one more FedCSIS 2012 publication [51]. In general, one will see throughout our whole paper that rough set principles fit the field of rule induction very well [18], [31].

This relationship is evident not only at a technical algorithmic level but also with respect to the common assumption of looking at the data through the glasses of information granules [16]. For more examples of connections between the worlds of rules and rough sets, let us refer e.g. to [17], [52].

Going further, papers [53], [54] introduce new heuristic measures that can be used during attribute reduction. It is worth noting that attribute reduction – or in other words algorithmic elimination of redundant attributes from the constructed set of attributes – is an important contribution of rough set research to knowledge discovery and in particular to its phase of feature selection [55]. As a complement to typical feature selection algorithms which attempt to add the most useful attributes, rough set methods take as input the sets of attributes produced by those typical algorithms and attempt to additionally compact them by eliminating unnecessary or approximately unnecessary elements. The additional aspect of attribute elimination occurs in just a few machine learning methodologies worldwide [56], so we can indeed say that this is an important rough sets’ contribution to this area.

Papers [57], [58] continue the topic of attribute reduction. The first of them proposes greedy algorithms for deriving so-called superreducts from data sets with multivalued decision attributes (target variables). It is one more example of dealing with information imprecision in rough set frameworks. Superreducts are the subsets of attributes which are sufficient to induce values (or as in this case, the sets of possible values) of decision attributes. It is also important to note that the notion of superreduct is equivalent to the notion of test in the test theory [31]. The second paper compares the notions of decision bireduct (aimed at deriving both the sets of attributes and the sets of data objects for which those attributes are sufficient to construct rule-based decision models) and approximate reduct (aimed at eliminating as many attributes as possible, even if the ability to induce decisions is not fully preserved). This comparison was later extended in [17].

The next two papers extend the topic of feature selection toward some of modern data challenges, namely high-dimensionality and large data volumes. Paper [59] combines attribute reduction with attribute clustering. Attributes are first grouped using some rough-set-inspired measures and then the methods of attribute reduction work iteratively on cluster representatives. This allows for decreasing the complexity of attribute reduction for high numbers of attributes and it also improves interpretability of results. These methods were later extended to let them work with attribute groups which can be set up for many reasons, including heterogeneity of data sources that are required to derive attribute values [60].

Paper [61] copes with big data volumes by putting attribute reduction and decision tree induction into a relational database framework, whereby the corresponding algorithms are implemented in SQL. The authors extend some previous ideas in this field [19] and, in particular, employ an open source database engine called Infobright Community Edition to run experiments. Infobright Community Edition is an example of using rough sets to optimize other types of data computations,
in this case – query execution in relational databases. This emphasizes that rough sets can be successfully used not only for machine learning and data mining but also for other tasks of big data processing. We refer to [62] for current developments related to Infobright Community Edition. We also refer to [63], where the Infobright’s technology performance is explained in terms of rough set operations on specifically aggregated (granulated) multivalued information systems.

The last two rough-set-related publications are interesting from the information systems’ viewpoint as well. In [64], the source of building an information system is a transformed ontological graph which encodes our knowledge about a given area [65]. The rules derived using the Dominance Rough Set Approach (DRSA) [20] express useful regularities within the original graph. This is actually an illustration of the fundamental idea behind information systems, namely, that such systems may contain not only the empirical data but they may also integrate it with domain knowledge [11], [43].

Finally, paper [66] presents the real-world application of rough sets to explore medical data. Herein, the information system – the input for rough-set-based model learning methods – does not correspond directly to the original data measurements. It is rather a result of a sequence of time-window-driven data aggregations which are typical for building hierarchical information systems describing complex objects [45], [67]. This work applied in particular a rough-set-based software system for machine learning and data mining – called RSES – which is now available in a library format [68] (see also the RSES extension targeted at spatio-temporal concepts). It is also one more practical use case of deploying the Infobright Community Edition database engine to run the underlying operations over granulated and compressed data sets.

**B. Rough Set Contest at PP-RAI 2022**

The second considered event is the PP-RAI 2022 conference held this year in Gdynia, Poland. The chairs of PP-RAI 2022 decided to celebrate the 40th anniversary of rough sets by organizing the contest for the most influential article on rough sets co-authored by Polish researchers in 2020 or later. Let us discuss below the articles submitted to this contest.

Papers [69], [70] operate at the edge of rough sets and formal concept analysis [25]. The first paper adopts the principles of attribute reduction (or more generally, model compactness) to simplify so-called fuzzy concept lattices, introduced as the means for representing patterns and regularities hidden in numerical data [71]. The second paper is actually an extension of the previously-cited work [36]. The authors attempt to put classical rough sets, formal concept analysis and the DRSA-style extensions of rough sets [20] into a unified conceptual pipeline aimed at transforming the data – through various forms of (possibly multivalued) information systems [33] – to knowledge. Within such a universal framework, the authors reconsider special cases of rough set operators known from different approaches. Therefore, one may say that this paper is a direct continuation of the ideas introduced in [1].

Papers [72], [73] link rough sets with logical foundations. The first paper shows how to express reasoning based on the VPRS-style extensions of rough sets [49] within the framework provided by a probabilistic extension of PROLOG [74]. The second paper shows how to reason about the properties of various types of rough set approximations within the framework provided by Mizar – a powerful system for automated proving [75]. Needless to say, such foundations are crucial for every theory, including reasoning within the theory and reasoning about the theory. We refer to [76] for more information about logical background of rough sets.

Papers [77], [78] present further advances in the previously-discussed popular rough set approaches such as the above-mentioned DRSA and fuzzy rough sets, respectively. The first paper uses the statistical learning machinery [79] to give new insights into parameters of probabilistic extensions of DRSA. The second paper, somewhat analogously, attempts to provide new interpretation of fuzzy rough set parameters. This is done by considering a new form of fuzzy granules [80], which consequently leads toward more intuitive derivation of fuzzy rough decision rules. One can say that these two articles fall into the same thematic categories as the previously-considered FedCSIS 2012 publications [47] and [48], respectively.

Papers [81], [82] continue the topic of feature selection. The first paper refers to heuristic attribute evaluation measures and data discretization techniques analogous to those reported in [10], [19]. The second paper seems to be particularly interesting as it extends the already-discussed topic of rough set software packages and libraries [24], [68] toward hardware optimizations that are specific for high performance computing. Such optimizations should be further compared and integrated with other acceleration opportunities, e.g., adaptation of MapReduce [60] and analytical database engines [61].

Papers [83], [84] refer to rough set software too. The first paper reports one more package delivering rough set methods for data mining and knowledge discovery. The second paper is about the application of that package to biomedical data mining. This second paper – besides its important experimental results – touches the aspects of visual data analytics [85], [86] and a need of understanding both, the analytical processes and their outcomes by subject matter experts [87], [88].

Papers [89], [90], [91] illustrate more real-world applications of rough set methods in the area of biomedicine. The first paper uses rough set approximations built over neighborhood-based information granules [92]. The remaining two papers confirm the expressive power of the DRSA-based decision rules. They also compare the accuracy of rule-based models with other approaches (such as random forests and logistic regression [93]) and show how to derive the attribute importance (see e.g., [94]) from the considered rules.

Papers [95], [96] continue the topic of rule induction. The first paper can be compared to [35], as both of them deal with deriving probabilistic rules from incomplete information sys-

---

6 en.wikipedia.org/wiki/Infobright
7 mimuw.edu.pl/~bazan/roughice/?sLang=en
8 pp-rai2022.umg.edu.pl/
9 roughsets.org/newspage/events/
tems, assuming several types of incompleteness. The second paper applies both rough-set-based [18] and fuzzy-set-based [12] rules in the task of posture detection. This is an example of real-world application, whereby the multi-stage solution needs to integrate sensor calibration, sensor data acquisition, inducing rules from the acquired data, as well as rule-based inference. With respect to making all such layers working together, this work can be compared to [43], [66], [67].

Papers [97], [98] deal with ensembles of decision models. The first paper employs so-called Dominance-based Rough Set Balanced Rule Ensemble for fraud detection. Herein, it is worth adding that rough set methods and applications include also examples of operating with ensembles of the aforementioned approximate reducts [99] and bireducts [17], [58] which correspond to bigger collections of rules. The second paper shows how to negotiate between classifiers and actually refers to the aforementioned conflict analysis model proposed by Professor Pawlak [6], [100]. On the other hand, the mechanism of voting in the third paper relies on the aforementioned three-way decision making [21]. It is worth emphasizing that solutions described in both papers attempt to provide a deeper insight into the ensemble decisions.

Paper [101] remains in the area of ensembles of decision models but it also touches an important aspect of incremental learning in dynamic data environments [102]. Herein, it is worth recalling a gentle difference between reasoning about objects or states in a repetitive fashion (whereby the values of attributes in information systems need to be cyclically updated) and reasoning about temporal objects or phenomena (which require different construction of information systems with attributes reflecting changes and trends) [103], [104].

Finally, papers [105], [106] combine the principles of rough sets and GrC with popular machine learning methods, referring to decision model ensembles as well. The idea is to prepare compacted data inputs – called granular reflections [15] – for the algorithms responsible for learning decision models such as e.g. neural networks or random forests (see [93] again). From a conceptual perspective, it corresponds to the aforementioned studies on aggregated / granulated / summarized information systems [63], [66]. This topic has also interesting relationships with some branches of approximate computing [107] and compressed image recognition [108].

III. INFORMATION SYSTEMS AND IGrC

As we have already emphasized, rough sets are based on data / information granulation. Both the original rough set approach and its extensions, need granules (and their descriptions / signatures) as inputs to compute approximations. The same applies to rough set methods of constructing decision models, e.g. rule-based models [52], [95]. On the other hand, granules can take different forms and have different origins. They can be partition blocks (induced by combinations of attribute values or ranges), dominance classes or neighborhoods [20], [89], relationships based on fuzzy (dis)similarity and (in)discernibility [78], [80] and so on. In information systems, granules can take different information signatures such as precise values, value sets, ranges and distributions [35], [57]. Those signatures can be computed using different aggregation mechanisms, often assuming non-trivial interdependencies with processes and devices that produce the data [63], [66]. The reliability and accessibility of information – therefore also reliability and accessibility of the outcomes of calculations over information systems – requires a careful analysis of all phases of forming the contents of such systems.

In Section I, we highlighted that the IGrC framework [37], [40] could be helpful to keep information systems aligned with respect to practical needs of operations on them in different contexts. In the next subsections, we will elaborate on several aspects of such alignment. As already discussed, IGrC uses so-called c-granules in order to create configurations of physical objects and control interactions between them so as to achieve computational objectives. Now let us add that the control mechanisms embedded within c-granules rely on one more type of granules – informational c-granules (ic-granules) which include both abstract (informational) and physical layers. They contain specifications how to link the abstract and physical worlds, whereby the abstract world corresponds in particular to (the networks of) information systems. The perceived properties of physical objects can be used to transform the current configurations of ic-granules, i.e. to modify interactions between objects. Such mechanisms require a design of new methods of reasoning about where, when, what, and how to perceive using different sensors or actuators. New methods for judging membership (alignment, matching) of the perceived situations in (with) rough set approximations of complex concepts are needed too.

A. Reliability of Information

This kind of reliability is studied in many fields. In the domain of big data, it is referred as one of the “V’s” – Veracity [109]. Actually, we have already dealt with some of other “V’s” in the previous sections, e.g. Volume [61], Velocity [101] and Variety [60]. However, without addressing Veracity, i.e. assuring data quality that is transformed into information reliability, any solutions focused on those other “V’s” cannot guarantee anything useful. Another popular term related to this problem is “garbage data”. It refers to the fact that if a machine learning method is executed on improper data, then the resulting models cannot be expected to work successfully. The causes of data being garbage data may be connected to problems with e.g. sensor measurements, data parsing, or even data labels acquired from human experts [67], [110]. A technical solution to cope with garbage data is often to filter them out by using validation procedures (e.g. checking sensor scales) [99]. However, in many applications – such as [66] (Subsection II-A), [96] (Subsection II-B) or just-mentioned [67] – it would lead toward disqualifying too broad data fragments, if any formalized validation is possible at all. Another approach is to live with the unreliable data and moreover, to take such unreliability into account while conducting any computations. With this respect, non-deterministic information systems have some tools to express uncertainty.
by replacing precise values with sets, intervals, etc. [35], [63]. However, (a degree of) reliability remains something different, as it refers to the way the data was acquired from the physical world rather than the specification of attribute values.

The IGrC framework is quite natural when it comes to reasoning about such an additional layer of information. Interactive granular computations can be actually extended toward adaptive searching strategies for the most relevant and reliable data, spatio-temporal windows pointing out to fragments of the physical world where the most reliable measurements and/or actions should be performed, and so on. This kind of reasoning may be also associated with the domain of data governance which extends towards data and information security, accessibility, as well as the protocols of interactions between intelligent systems and humans [111], [112]. On the other hand, the discussed physical-world-related aspects can be an additional contribution of IGrC to data governance. Moreover, IGrC can be helpful to operate with often softly expressed regulations about data integrity and timeliness.

It is also worth referring the above discussion to the meanings of aleatoric and epistemic uncertainties in machine learning [113]. From this perspective, a limited reliability of the contents of information systems can be treated as one of ingredients of the epistemic uncertainty, as it puts together both, the model and the data deficiencies. However, we believe that these two sources of deficiencies should be kept separately, with the third type of experimental/physical uncertainty explicitly considered. The analysis of this third type of uncertainty should be taken into account when assessing the efficiency and stability of machine learning models, especially given the fact that in some practical scenarios the inputs to the learning algorithms can be unreliably extracted for the purpose of e.g. accelerating computations [114].

B. Acquisition of Information

In order to talk about information reliability, we first need to assure that information can be gathered at all. In practice, there is often a great variety of data available but it does not mean yet that the corresponding information is sufficiently complete to perform any kind of analysis. (This relates to one more “V”: Value.) Some promising approaches to data enrichment refer to the paradigm of active learning [115], which can be further extended toward establishing an interactive loop within which subject matter experts label data objects that are of the highest interest to the machine learning algorithms. One just needs to think about controlling the quality of such labels [110].

Similarly, the data enrichment processes can rely on connecting information systems with physical systems [116]. Actually, it is worth pointing out that humans can be considered as a special kind of physical objects that interact with decision support systems and/or intelligent systems. This refers to a broader topic of the information and communication technologies (ICT) systems [39] which put together the aspects of hardware (e.g. sensors), software (e.g. machine learning methods), the data (including domain knowledge), and the system users (in particular subject matter experts).

The above ideas require a firm layer that connects information systems with the physical world where the data comes from. In IGrC, every granule should have an access to instructions how to compose the values of particular attributes for particular objects [38], [40]. Moreover, it is important for this layer to log a history of attempts to calculate particular fragments of an information system. Such history may let us avoid mistakes and misinterpretations related to the data acquisition processes. That history may be also useful while assessing reliability of the current contents of an information system. Such mechanisms can be adopted also from the architectures of granular database engines [62], [107], whereby the aspects of information completeness and reliability are equally important as in the field of machine learning.

Going back to the framework of active learning, let us claim that subject matter experts can assist us not only in enriching the data with labels but also enriching the data mining algorithms with domain insights. As an example, let us think one more time about the task of feature selection. There are various techniques of measuring and visualizing attribute importance [88], [90], [94] but they are usually applied to report to humans the final results instead of “inviting” them into a more interactive dialogue on feature selection process. In this regard, we refer to [85] where incrementally constructed information systems are employed to guide subject matter experts through such an interactive process, letting them share their recommendations about the most relevant attributes.

Last but not least, when it comes to decision problems related to complex phenomena, it is worth attempting – using the elements of active learning and human-computer interaction – to acquire from subject matter experts even more advanced knowledge, expressed in terms of hierarchical structures and dependencies. This fits the paradigm of computing with words [117] (which also corresponds to the foundations of information granulation with respect to decomposing complex problems onto their smaller components) and, in particular, the following challenge formulated by Judea Pearl [118]:

Traditional statistics is strong in devising ways of describing data and inferring distributional parameters from sample. Causal inference requires two additional ingredients: a science-friendly language for articulating causal knowledge, and a mathematical machinery for processing that knowledge, combining it with data and drawing new causal conclusions about a phenomenon.

C. Accessibility and Cost of Information

In practical deployments, there is always a risk that some of data sources – which are needed to calculate some of attributes being inputs to a decision model – will be temporarily unavailable (because of e.g. physical connection problems or dissatisfaction of some data governance rules) or unreliable (as discussed in Subsection III-A). Feature selection [56], including contribution of rough set methods to elimination/reduction of redundant attributes [10], can be a remedium to this problem – less attributes require less aspects of data to be
calculated. Moreover, it is possible to diversify data sources needed to derive attributes that are used by particular models in an ensemble [17], [55]. This increases a chance that at least some of models would be usable in a given situation.

However, it is not only about the accessibility, sometimes it is also about the cost. For instance, in the recent contests organized at one of online data mining competition platforms\textsuperscript{10} [119], [120], the participants purposefully did not take into account some of the available data sources (modalities) because – according to them – derivation of attributes from those sources would be too expensive computationally. Going further along this path, one may say that even if such “expensive” attributes are included into a model, its deployed version should have a choice to decide dynamically which of them (and at what level of precision [114]) are necessary to be calculated. This is important the more so as in intelligent systems dealing with complex phenomena, the most adequate selections and meanings of attributes can be changed over time [40].

Going even further, if appropriate metadata is maintained on the side of an information system, the same attribute values can be derived from different information sources or using different data modalities, perhaps subject to different cost and precision [43], [67]. This is well-aligned with the IGrC assumptions discussed in the previous subsections. The point is to pass to information granules the decision power in regard to how they produce information that they are responsible for, and make decision models responsible for timely asking those granules for particular information pieces [38], [116].

The above discussion can be extended toward a broader topic of whether the data / information updates should be rather “pushed” or “pulled” in the computational pipelines which involve learning and applying the learnt models. A common assumption is that any change in the underlying information should be more or less quickly transmitted to the inputs of a model, causing its recalculations or at least modification of its behaviors [101], [102]. However, in big data scenarios it is not so obvious – it may be safer to leave such decisions in hands of information granules equipped with well-designed triggers and internal cost models [37], [107].

D. Networks of Information Systems

Continuing with the topic of intelligent systems aimed at reasoning about complex phenomena, we already know that data sources required to learn the underlying decision models cannot be acquired in a single-step process. It is necessary to provide such systems with permanent links to relevant fragments of the physical world and keep adapting (actively but also reasonably, from the computational perspective) the induced models following changes in the perceived situation. Recalling our comments about logical reasoning related to hierarchical systems [43] (see the end of Section I), one should also be aware that these hierarchical structures are dynamically changing in time and the relevant reasoning methods should allow to the system to perform the necessary reasoning about

\textsuperscript{10}knowledgepit.ai

such dynamical structures. This seems to be aligned with the following opinion expressed by Frederick Brooks [121]:

\textit{Mathematics and the physical sciences made great strides for three centuries by constructing simplified models of complex phenomena, deriving, properties from the models, and verifying those properties experimentally. This worked because the complexities ignored in the models were not the essential properties of the phenomena. It does not work when the complexities are the essence.}
selection and construction of sensors. At a more general level, the whole idea requires a distributed control of c-granules, whereby specific reasoning methods (related to cooperation / competition between granules) need to express the expected behavioral patterns of the whole “society” of granules. Herein, one can seek for inspirations in the previously discussed conflict analysis [6]. The requirements of the aforementioned ICT systems [39], web intelligence [65] or e.g. IoT analytics [124] – whereby there are a number of federated learning scenarios involving distributed agents (and their underlying information systems) – can be a useful analogy as well.

IV. Conclusions and Future Directions

The first goal of this paper was to expose the current progress of the theory and applications of rough sets – the methodology founded by Zdzisław Pawlak with the aim of deriving and expressing important patterns and dependencies subject to limited (incomplete, imprecise) information about the concepts of practical interest [1], [9]. We examined connections of rough sets with decision making [20], [21], logics [73], [76], probability [47], [72], statistical / machine learning [77], [105], data mining [19], [23], fuzzy sets [13], [48], formal concept analysis [26], [69], and other data / information / knowledge representation methodologies [31], [32]. We discussed some of rough set techniques aimed at attribute selection / reduction treated as a component of knowledge discovery processes [10], [17], with particular emphasis on computational scalability challenges [60], [82]. We paid special attention to rough set approaches to construction of interpretable (explainable by design) rule-based decision models [18], [48], [52], [97]. We referred to rough set software packages for data mining and machine learning [24], [68], [83], as well as other technologies which utilize rough set approximation principles for their internal purposes [62]. We also recalled several (out of many) applications of rough set methods in real-world data analysis, including biomedical and healthcare applications whereby interpretability of decision models is of special importance [66], [67], [84], [91].

Our second goal was to address the progress in the area of information systems [3], [4]. We referred to their extensions [33], [70] and we outlined a number of applications which use specifically formed information systems as the means for representing (granulated / aggregated) data, (uncertain / imprecise) information, and (appropriately transformed) knowledge structures [35], [43], [63], [89]. We pointed out that information systems – especially their hierarchies and networks – constitute the means for reasoning about complex spatio-temporal phenomena [45], [104]. We also claimed that information systems can be a medium to conduct interactive data analytics involving subject matter experts [85] and support interactions between multiple data exploration processes [123]. That led us toward discussing the current challenges (often referred as the big data “V’s”) in front of information systems understood as the means for representing and delivering data required for the learning processes [102], [109], [110], [119]. Accordingly, we examined whether the principles of so-called interactive granular computing (IGrC) [37], [116] can help us to face those challenges and to what extent they are aligned with some of emerging trends in machine learning [115], [124].

It was important for us to discuss the principles of granular computing – including IGrC – together with rough sets and information systems, as these three domains interfere with each other in many interesting ways [11], [14], [15], [40]. In particular, IGrC may have future implications for the design of intelligent systems, e.g. when it comes to so-called perceptual rough sets [11]. If one wants to build rough set approximations of complex concepts in real-world environments, then it is required to design a dynamic space of granules which are able to reason about complex approximation constructions. The corresponding reasoning methods will need to be far richer than the ones considered so far in rough set applications.

Some other future directions for rough sets and information systems refer to continuation of development of real-world applications, focused on e.g. images and video recordings [22], [46], [120], as well as signals and sensor measurements [99], [96], [103]. This kind of development should emphasize strong assets of rough sets, such as straightforward interpretability of the derived decision models, even when it comes to modeling very complex and dynamic situations [11], [101]. Needless to say, interpretability is now the key objective for a great majority of machine learning applications [87], [88].

When thinking about the future it is also worth referring to the history. That reflected one more objective: acknowledging the 40th anniversary of rough sets [1], their founder [5], as well as some of relevant past and present events such as FedCSIS 2012 (rough set papers published exactly 10 years ago) [47], [48], [51], [53], [54], [57], [58], [59], [61], [64], [66], the PP-RAI 2022 rough set contest [69], [70], [72], [73], [77], [78], [81], [82], [83], [84], [89], [90], [91], [95], [96], [97], [98], [101], [105], [106] and celebration of the 30th CS&P – the event series whereby this paper’s topics have been regularly addressed [36], [40], [44], [103], [116], [123].

In the end, let us recall that this is not the first anniversary corresponding to rough sets in the history of the FedCSIS conferences. Indeed, FedCSIS 2016 (Gdańsk, Poland) hosted the international panel discussion in memoriam of the 90th anniversary of the birth and the 10th anniversary of the death of Professor Pawlak [12]. The previously-cited publications [4], [8], [100] were prepared specially for that panel.

REFERENCES

[1] Z. Pawlak, “Rough Sets,” International Journal of Computer and Information Sciences, vol. 11, no. 5, pp. 341–356, 1982. [Online]. Available: doi.org/10.1007/BF01001956
[2] ——, Rough Sets – Theoretical Aspects of Reasoning about Data, set. Theory and Decision Library D. Springer, 1991. [Online]. Available: doi.org/10.1007/978-94-011-3534-4

11See e.g. A. Skowron: Perceptual Rough Set Approach in Interactive Granular Computing (keynote). Information Processing and Management of Uncertainty in Knowledge-Based Systems (IPMU 2022), July 11-15, 2022, Milan, Italy. roughsets.org/bnr/1094938771802d02dec99c66823875c.PDF

12fedciss.org/2016plenary_panel
A. Skowron and J. Stepaniuk, “Information Granules: Towards Foundations and Applications in Data Mining,” Transactions on Rough Sets, vol. 5, pp. 334–356, 2006. [Online]. Available: doi.org/10.1007/11847465_16

S. Greco, B. Matarazzo, and R. Slowiński, “Rough Sets Theory for Multicriteria Decision Analysis,” European Journal of Operational Research, vol. 129, no. 1, pp. 1–47, 2001. [Online]. Available: doi.org/10.1016/S0377-2217(00)00167-3

Y. Yao, “Three-Valued Decision and Imporiori-based Rule Generation from Table Data Sets: A Survey and a Combination,” CAII Transactions on Intelligence Technology, vol. 4, no. 4, pp. 203–213, 2019. [Online]. Available: doi.org/10.1004/tr.2019.0001

M. Wolski and A. Gomolińska, “Semantic Rendering of Data Tables – Multivalued Information Systems Revisited,” in Proceedings of the 25th International Workshop on Concurrency, Specification and Programming, Rostock, Germany, September 28–30, 2016, ser. CEUR Workshop Proceedings, B. Schlingloff, Ed., vol. 1698. CEUR-WS.org, 2016, pp. 113–124. [Online]. Available: ceur-ws.org/Vol-1698/CS&P2016_11_Wolski&Gomolinska_Semantic-Rendering-of-Data-Tables-Multivalued-Information-Systems-Revisited.pdf

A. Skowron and J. Jankowski, “Rough Sets and Interactive Granular
Computing,” Fundamenta Informaticae, vol. 147, no. 2-3, pp. 371–385, 2016. [Online]. Available: doi.org/10.3233/FI-2016-1413

[38] A. Skowron and P. Wasilewski, “Interactive Information Systems: Toward Perception Based Computing,” Theoretical Computer Science, vol. 454, pp. 240–260, 2012. [Online]. Available: doi.org/10.1016/j.tcs.2012.04.019

[39] M. Lippi, S. Mariani, M. Martinelli, and F. Zambonelli, “Individual- and Collective Self-Development: Concepts and Challenges,” in Proceedings of the 2022 Federated Conference on Computer Science and Intelligence Systems, Sofia, Bulgaria, September 4-7, 2022, ser. Annals of Computer Science and Information Systems, M. Ganzha, M. Paprzycki, and D. Ślęzak, Eds., vol. 30, 2022.

[40] A. Jankowski, A. Skowron, and R. W. Świarniak, “Interactive Complex Granules,” in Proceedings of the 22nd International Workshop on Concurrency, Specification and Programming, Warsaw, Poland, ser. CEUR Workshop Proceedings, M. S. Szczuka, L. Czaja, and M. Kaczprzak, Eds., vol. 1032. CEUR-WS.org, 2013, pp. 206–218. [Online]. Available: ceur-ws.org/Vol-1032/paper-18.pdf

[41] T. A. Poggio and S. Smale, “The Mathematics of Learning: Dealing with Data,” Notices of the American Mathematical Society, vol. 50, no. 5, pp. 537–544, 2003.

[42] S. H. Nguyen, T. T. Nguyen, M. S. Szczuka, and H. S. Nguyen, “On Elimination of Redundant Attributes from Decision Table,” in Federated Conference on Computer Science and Information Systems – FedCSIS 2012, Wroclaw, Poland, 9-12 September 2012, Proceedings, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., 2012, pp. 317–322. [Online]. Available: fedcsis.org/proceedings/2012/pliks/324.pdf

[43] M. L. Nguyen, “Sequential Optimization of γ–Decision Rules,” in Federated Conference on Computer Science and Information Systems – FedCSIS 2012, Wroclaw, Poland, 9-12 September 2012, Proceedings, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., 2012, pp. 339–346. [Online]. Available: fedcsis.org/proceedings/2012/pliks/87.pdf

[44] M. Kryszkiewicz, “ACBC–Adequate Association and Decision Rules Versus Key Generators and Rough Sets Approximations,” in Fundamenta Informaticae, vol. 148, no. 1-2, pp. 65–85, 2016. [Online]. Available: doi.org/10.3233/FI-2016-1423

[45] L. G. Nguyen, “Metric Based Attribute Reduction in Decision Tables,” in Federated Conference on Computer Science and Information Systems – FedCSIS 2012, Wroclaw, Poland, 9-12 September 2012, Proceedings, M. Ganzha, L. A. Maciaszek, and M. Paprzycki, Eds., 2012, pp. 311–316. [Online]. Available: fedcsis.org/proceedings/2012/pliks/311.pdf

[46] M. Ganzha, L. A. Maciaszek, and M. Paprzycki, and D. Ślęzak, Eds., vol. 30, 2022.

[47] J. G. Bazan, S. Bazan-Socha, S. Buregwa-Czuma, P. W. Pardel, and B. Sokolowski, “High Performance Predicting the Presence of Coronary Artery Disease Based on 24 hour Holter ECG Monitoring,” in Federated Conference on Computer Science and Information Systems – FedCSIS 2012, Wroclaw, Poland, 9-12 September 2012, Proceedings,
A. Wojna and R. Łatkowski, “Toward Automatic Assessment of a Risk of Women’s Health Disorders Based on Ontology Decision Models and Mensural Cycle Analysis,” in 2021 IEEE International Conference on Big Data (Big Data), Orlando, FL, USA, December 15-18, 2021, Y. Chen, H. Ludvig, Y. Tu, U. M. Fayyad, X. Zhu, X. Hu, S. Byna, X. Liu, J. Zhang, S. Pan, V. Papalexakis, J. Wang, A. Cuzzocrea, and C. Ordenez, Eds. IEEE, 2021, pp. 5544–5552. [Online]. Available: doi.org/10.1109/BigData52589.2021.9671481

A. Wojna and R. Łatkowski, “Reselib 3: Library of Rough Set and Machine Learning Methods with Extensible Architecture,” Transactions on Rough Sets, vol. 21, pp. 301–332, 2019. [Online]. Available: doi.org/10.1007/978-3-030-52589-2_38

M. I. Benítez-Caballero, J. Medina, E. Ramírez-Poussa, and D. Ślezak, “Rough-set-driven Approach for Attribute Reduction in Fuzzy Formal Concept Analysis,” Fuzzy Sets and Systems, vol. 391, pp. 117–138, 2020. [Online]. Available: doi.org/10.1016/j.fss.2019.11.009

M. Wolski and A. Gomolinska, “Data Meaning and Knowledge Discovery: Semantical Aspects of Information Systems,” International Journal of Approximate Reasoning, vol. 119, pp. 40–57, 2020. [Online]. Available: doi.org/10.1016/j.ijar.2020.01.002

R. Belohlávek and V. Vychodil, “What is a Fuzzy Concept Lattice?” in Proceedings of the 2005 International Workshop on Concept Lattices and their Applications Olomouc, Czech Republic, September 7-9, 2005, ser. CEUR Workshop Proceedings, R. Belohlávek and V. Snášel, Eds., vol. 162. CEUR-WS.org, 2005. [Online]. Available: ceur-ws.org/Vol-162/paper4.pdf

P. Doherty and A. Szalas, “A Landscape and Implementation Framework for Probabilistic Rough Sets Using ProbLog,” Information Sciences, vol. 593, pp. 546–576, 2022. [Online]. Available: doi.org/10.1016/j.ins.2021.12.062

A. Grabowski, “Automated Comparative Study of Some Generalized Rough Approximations,” Fundamenta Informaticae, vol. 179, no. 2, pp. 165–182, 2021. [Online]. Available: doi.org/10.3233/FI-2021-2019

L. De Raedt, A. Kimmig, and H. Toivonen, “ProbLog: A Probabilistic Logic and Its Application in Link Discovery,” in IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007, M. M. Veloso, Ed., 2007, pp. 2462–2467. [Online]. Available: ijcai.org/Proceedings/07/Papers/P96.pdf

A. Grabowski, A. Korniłowicz, and A. Naumowicz, “Four Decades of Mizar – Foreword,” Journal of Automated Reasoning, vol. 55, no. 3, pp. 191–198, 2015. [Online]. Available: doi.org/10.1007/s10817-015-9345-1

P. Pagliani and M. K. Chakraborty, A Geometry of Approximation – Rough Set Theory: Logic, Algebra and Topology of Conceptual Patterns, ser. Trends in Logic. Springer, 2008. [Online]. Available: doi.org/10.1007/978-1-4020-8622-9

Y. Kusunoki, J. Blaszczynski, M. Inuguchi, and R. Slowiński, “Empirical Risk Minimization for Dominance-based Rough Set Approaches,” Information Sciences, vol. 567, pp. 395–417, 2021. [Online]. Available: doi.org/10.1016/j.ins.2021.02.043

M. Palangetić, C. Cornelis, S. Greco, and R. Slowiński, “Granular Representation of OWA-based Fuzzy Rough Sets,” Fuzzy Sets and Systems, vol. 440, pp. 112–130, 2022. [Online]. Available: doi.org/10.1016/j.fss.2021.04.018

V. Vapnik, “Principles of Risk Minimization for Learning Theory,” in Advances in Neural Information Processing Systems 4, [NIPS Conference, Denver, Colorado, USA, December 2-5, 1991], J. E. Moody, S. J. Hanson, and R. Lippmann, Eds. Morgan Kaufmann, 1991, pp. 831–838. [Online]. Available: papers.nips.cc/paper/506-principles-of-risk-minimization-for-learning-theory

L. A. Zadeh, “Toward a Theory of Fuzzy Information Granulation and its Centrality in Human Reasoning and Fuzzy Logic,” Fuzzy Sets and Systems, vol. 90, no. 2, pp. 111–127, 1997. [Online]. Available: doi.org/10.1016/S0165-0114(97)00077-8

U. Stańczyk and B. Zielosko, “Heuristic-based Feature Selection for Rough Set Approach,” International Journal of Approximate Reasoning, vol. 55, pp. 187–202, 2015. [Online]. Available: doi.org/10.1016/j.ijar.2020.07.005

M. Kopczyniś and T. Grzeźi, “Hardware Rough Set Processor Parallel Architecture in FPGA for Finding Core in Big Datasets,” Journal of Artificial Intelligence and Soft Computing Research, vol. 11, no. 2, pp. 99–110, 2021. [Online]. Available: doi.org/10.2478/jascr-2021-0007

S. M. Lundberg and S. Lee, “A Unified Approach to Interpreting Model Predictions,” in Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA, I. Guyon, U. von Luxburg, S. Bengio, H. M. Wallach, R. Fergus, S. V. N. Vishwanathan, and R. Garnett, Eds., 2017, pp. 4765–4774. [Online]. Available: proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html

K. Pancerz, “Rough Set Based Description of Plasmodium Propagation,” International Journal of Unconventional Computing, vol. 15, no. 4, pp. 287–297, 2020. [Online]. Available: oldcitupypublishing.com/journals/jicuc-home/jicuc-issue-contents/jicuc-volume-15-number-4-2020/jicuc-15-4-p-287-297

L. Pańkowski, M. Karolak, J. Blaszczynski, J. Krysiński, and R. Slowiński, “Structure-Activity Relationships of the Imidazolium Compounds as Antibacterials of Staphylococcus aureus and Pseudomonas Aeruginosa,” International Journal of Molecular Sciences, vol. 22, no. 15, 2021. [Online]. Available: mdpi.com/1422-0066/22/15/7967

M. Karolak, L. Pańkowski, B. Kubiak, J. Blaszczynski, R. Luniu, W. Sawicki, R. Slowiński, and J. Krysiński, “Application of Dominance-based Rough Set Approach for Optimization of Pellets Tableting Process,” Pharmaceutics, vol. 12, no. 11, p. 1024, 2020. [Online]. Available: doi.org/10.3390/pharmaceutics12111024

T. Y. Lin, “Neighborhood Systems and Mathematical Models of Information Granulations,” in Proceedings of the IEEE International Conference on Systems, Man & Cybernetics: Washington, D.C., USA, 5-8 October 2008, IEEE, 2003, pp. 3188–3193. [Online]. Available: doi.org/10.1109/CSCC.2003.1244381

E. Frank, M. A. Hall, and I. H. Witten, “The WEKA Workbench,” Online Appendix for “Data Mining: Practical Machine Learning Tools and Techniques”, Morgan Kaufmann, Fourth Edition, 2016.

R. Rudnicki, M. Kierczak, J. Koronacki, and H. J. Komorowski, “A Statistical Method for Determining Importance of Variables in an Information System,” in Rough Sets and Current Trends in Computing, 5th International Conference, RSCTC 2006, Kobe, Japan, November 6-8, 2006, Proceedings, ser. Lecture Notes in Computer Science, S. Greco, Y. Hata, S. Hirano, M. Inuguchi, S. Miyamoto, H. S. Nguyen, and R. Slowinski, Eds., vol. 4259. Springer, 2006, pp. 557–566. [Online]. Available: doi.org/10.1007/11890829_58

P. G. Clark, C. Gao, J. W. Grzymała-Busse, T. Mroczek, and R. Niemiec, “Complexity of Rule Sets. Incomplete Data Using Characteristic Sets and Generalized Maximal Consistent Blocks,” Logic Journal of the IGPL, vol. 29, no. 2, pp. 124–137, 2021. [Online]. Available: doi.org/10.1093/jigpal/jzaa041
