Fundamenta Informaticae 47 (2001) 337–350
IOS Press

Information Granule Decomposition

Andrzej Skowron*
Institute of Mathematics
University of Warsaw
ul. Banacha 2, 02-097 Warszawa, Poland
skowron@mimuw.edu.pl

Jarosław Stepaniuk†
Institute of Computer Science
Białystok University of Technology
Wiejska 45A, 15-351 Białystok, Poland
jstepan@ii.pb.bialystok.pl

Abstract. Information sources provide us with granules of information that must be transformed, analyzed and built into structures that support problem solving. One of the main goals of information granule calculi is to develop algorithmic methods for construction of complex information granules from elementary ones by means of available operations and inclusion (closeness) measures. These constructed complex granules represent a form of information fusion. Such granules should satisfy some constraints like quality criteria or/and degrees of granule inclusion in (closeness to) a given information granule. Information granule decomposition methods are important components of those methods. We discuss some information granule decomposition methods.

Keywords: rough set theory, information granulation, granular computing, information fusion, decomposition, pattern, schemes of approximate reasoning

*This work was supported by the State Committee for Scientific Research of the Republic of Poland (KBN) research grant 8 T11C 025 19 and by the Wallenberg Foundation grant. Address for correspondence: Institute of Mathematics, University of Warsaw, ul. Banacha 2, 02-097 Warszawa, Poland

†This work was supported by by the State Committee for Scientific Research of the Republic of Poland (KBN) research grant 8 T11C 025 19. Address for correspondence: Institute of Computer Science, Białystok University of Technology, Wiejska 45A, 15-351 Białystok, Poland

1. Introduction

Information granulation belongs to intensively studied topics in soft computing (see, e.g., [24], [25], [26]). One of the recently emerging approaches to deal with information granulation is based on information granule calculi (see, e.g., [13], [17]). The development of such calculi is important for making progress in many areas like object identification by autonomous systems (see, e.g., [3], [22]), web mining (see, e.g., [5]) or spatial reasoning (see, e.g., [20], [4], [2]). In particular, reasoning methods using background knowledge as well as knowledge extracted from experimental data such as sensor measurements represented by concept approximations [3] are important for making progress in such areas.

Approximate Reasoning schemes (AR-schemes, for short) are obtained by means of relevant granules, called patterns for a given task decomposition of the identified or classified complex objects. The problem of deriving AR-schemes is closely related to perception [1], [26], [21], [15], [11].

In the paper, we assume the AR-schemes define parameterized operations on information granules. We discuss problems of tuning these parameters to derive from them relevant granules included in (or close to) target concepts to a satisfactory degree. Target concepts are assumed to be incomplete and/or vague.

One can distinguish two kinds of parts of AR-schemes represented, for example, by sub-formulas or sub-terms. Parts of the first type of scheme are represented by expressions from a language L_d (called the *domestic* language) that has known semantics (consider, for example, a given information system [10]). Parts of the second type of scheme are from a language L_f (called *foreign* language). Consider for instance, natural language, that has semantics definable only in an approximate way obtained by means of patterns extracted using rough, fuzzy, rough–fuzzy or other approaches). The parts of the second kind can be interpreted, for example, as soft properties of sensor measurements [3].

For a given expression e, representing a given AR-scheme that consists of sub-expressions from L_f, we propose to search L_d for relevant approximations of the foreign parts from L_f and next to derive global patterns from the whole expression after replacing the foreign parts by their approximations. This can be a multilevel process. That is we are facing propagation problems of discovered pattern through several domestic-foreign layers.

Let us consider some strategies for patterns construction from schemes.

The first strategy entails searching for relevant approximations of parts using a rough set approach. This means that each part from L_f can be replaced by its lower or upper approximation with respect to a set B of attributes. The approximation is constructed on the basis of relevant data table [3]. With the second strategy parts from L_f are partitioned into a number of sub-parts corresponding to cuts (or the set theoretical differences between cuts) of fuzzy sets representing vague concepts and each sub-part is approximated by means of rough set methods. The third strategy is based on searching for patterns sufficiently included in foreign parts. In all cases, the extracted approximations replace foreign parts in the scheme and candidates for global patterns are derived from the scheme obtained after the replacement. Searching for relevant global patterns is a complex task because many parameters should be tuned, e.g., the set of relevant features used in approximation, relevant approximation operators, the number and distribution of objects from the universe of objects among different cuts and so on. We plan to use evolutionary techniques for relevant pattern searching to obtain optimal parameters with respect to the quality of synthesized patterns.

We propose an approach for extracting from data patterns relevant to a target concept α. The approach is based on information granule decomposition strategies. It is shown that these strategies can be based on available rough set methods for decision rules generation and Boolean reasoning [6]. We discuss in particular methods for decomposition which can be based on background knowledge. In [9], [17], the reader can find another approach to decomposition.

The paper is structured as follows. In Sections 2 and 3 we discuss parameterized rough and fuzzy information granules. In Section 4 methods for decomposition of information granules are outlined.

2. Approximation Granules

We use standard notation (see [10]). In particular let $IS = (U, A)$ denote an information system; (a, v), a descriptor defined by the attribute a and its value v. In addition, let α denote a Boolean combination of descriptors, $[\alpha]_{IS}$ (or $[\alpha]_A$), the meaning of α in IS, i.e., the set of all objects from U satisfying α. The lower and upper approximations of $X \subseteq U$ are denoted by $\underline{A}X$ and $\overline{A}X$, respectively.

An elementary granule (in IS) is any pair $(\alpha, [\alpha]_{IS})$ where for simplicity we assume $[\alpha]_{IS} \neq \emptyset$.

One can extend the set of elementary granules by assuming if α is any Boolean combination of descriptors over A, then $(\overline{B}\alpha, \overline{B}[\alpha]_{IS})$ and $(\underline{B}\alpha, \underline{B}[\alpha]_{IS})$ are elementary granules too, for any $B \subseteq A$.

The inclusion and closeness are the basic concepts related to information granules [13], [17]. Using them one can measure the closeness of the constructed granule to the target granule and robustness of the construction scheme with respect to deviations of information granules that are components of the construction. For details and examples of closeness relations, the reader is referred to [13], [17]. Here, we present only necessary introductory remarks. Let us consider an example. For $X, X' \subseteq U$, let us assume, e.g., $\nu(X, X') = card(X \cap X')/card(X)$ if $X \neq \emptyset$ and 1, otherwise. By $\nu_p(X, X')$, we denote $\nu(X, X') \geq p$. We assume $cl_p(X, X')$ if and only if $\nu_p(X, X')$ and $\nu_p(X', X)$. The inclusion relation can be extended to inclusion on information granules and to the closeness relation: $\nu_p(g, g')$ means that the information granule g is included in g' to a degree at least p and $cl_p(g, g')$ means that the information granules g, g' are close in degree at least p [17].

It is worth mentioning that instead of Boolean propositional connectives one can consider other operations to create granule descriptions. Operations on information granules have been discussed, e.g., in [13], [14], [17]. Among these operation are set theoretical operation of union and intersection, operations defined by data tables, operations generating classifiers, etc. They are used to define complex information granules from elementary ones. These complex information granules constitute a form of information fusion. Other operations can be used to define granules corresponding to classifiers. Let us observe that inclusion (closeness) measures can be used to define new granules that are approximations or generalizations of existing ones. Assume g, h are given information granules, and ν_p is an inclusion measure (where $p \in [0, 1]$). A (h, p)- approximation of g is an information granule $g_{\nu,h,p}$ represented by a set $\{h' : \nu_1(h', h) \wedge \nu_p(h', g)\}$. Now one can easily define the lower and upper approximations of information granules [16].

Let us now denote by L_e^B the language of elementary granules consisting of Boolean combinations of expressions of the following two kinds:

- descriptors over attributes from $B \subseteq A$;
- formulas of the form $\overline{B}\alpha$, $\underline{B}\alpha$ where α is a Boolean combination of descriptors over attributes from A.

The formulas from L_e^B describe the syntax of B-elementary granules in IS. By Sem_B we denote the semantics of formulas from L_e^B, i.e., the function assigning to any formula $\alpha \in L_e^B$ the set $[\alpha]_{IS}$ of objects satisfying α in IS. Any pair $(\alpha, [\alpha]_{IS})$, where $\alpha \in L_e^B$ is called a B-elementary granule in IS.

Any elementary granule $(\alpha, [\alpha]_{IS})$ in IS defines the granule

$$g = (Approx_B(\alpha), \{[\beta]_{IS} : \beta \in Approx_B(\alpha)\})$$

where $Approx_B(\alpha) \subseteq L_e^B$ is the set of all formulas which can be obtained from α, in a finite number of steps, by application of the following rule:

if $\beta \notin L_e^B$ is a subformula of α (not prefixed by $\overline{B}$ or $\underline{B}$), then β is replaced by $\overline{B}\beta$ or $\underline{B}\beta$.

Hence, g consists of all granules obtained from α by replacing foreign parts by approximations using attributes from B. An important task is to define among components of g those which are relevant for matching the target information granule, i.e., those patterns returned by the scheme which are sufficiently close to (or included in) the target granule.

Now, let us consider a target D-elementary granule $g_t = (\delta, [\delta])_{IS}$ where $D \subseteq A$ and $D \cap B = \emptyset$. For a given degree p, one can consider a new granule that is in a sense a part of

$$g = (Approx_B(\alpha), \{[\beta]_{IS} : \beta \in Approx_B(\alpha)\})$$

and defined by

$$g' = (Approx'_B(\alpha), \{[\beta]_{IS} : \beta \in Approx'_B(\alpha)\})$$

where $Approx'_B(\alpha)$ is the set of all formulas γ from $Approx_B(\alpha)$ close to g_t (or included in g_t) in degree at least p, i.e., $cl_p([\gamma]_{IS}, [\delta]_{IS})$ (or $\nu_p([\gamma]_{IS}, [\delta]_{IS})$).

The target information granule g_t is approximated by means of granules sufficiently close to it. The granules are extracted from a given data sample represented by information system IS. Approximating granules are constructed not only from descriptors over B but also by means of the B-lower and B-upper approximations of some formulas consisting of attributes not necessarily from B.

Certainly, one should develop methods for inducing target concept approximation considered on extensions of IS. One possible way to induce target concept approximation is to use evolutionary methods [7] to solve the problem by optimization of parameters involved in the approximate description of the target concept.

Some parameters to be optimized express the approximation quality of concepts represented by subformulas with approximation operators. These coefficients can be also used to measure the robustness of the target concept approximation.

Let us consider some examples. Any pair

$$((\alpha, [\alpha]_{IS}), (\beta, [\beta]_{IS}))$$

of elementary granules in IS such that $[\alpha]_{IS} \subseteq [\beta]_{IS}$, $[\alpha]_{IS} = \underline{A}X$ and $[\beta]_{IS} = \overline{A}X$, for some $X \subseteq U$, can be considered as a granule. Let us call it an approximation granule. If α is any Boolean combination of descriptors over A (not necessarily over B), then

$$g = ((\underline{B}\alpha, \underline{B}[\alpha]_{IS}), (\overline{B}\alpha, \overline{B}[\alpha]_{IS}))$$

is called the B-approximation granule. By $Sem(g)$ we denote a pair $(\underline{B}[\alpha]_{IS}, \overline{B}[\alpha]_{IS})$.

For simplicity of exposition and to avoid unnecessarily tedious notation, we sometimes use only the semantic part $Sem(g)$ of g to denote the granule g, if this does not lead to a misunderstanding.

Let $g = ((\alpha, [\alpha]_{IS}), (\beta, [\beta]_{IS}))$ and $g' = ((\alpha', [\alpha']_{IS}), (\beta', [\beta']_{IS}))$ be two approximation granules. They are close in degree at least p (in IS), in symbols $cl_p(g, g')$ if and only if

- $cl_p([\alpha]_{IS}, [\alpha']_{IS})$;
- $cl_p([\beta]_{IS} - [\alpha]_{IS}, [\beta']_{IS} - [\alpha']_{IS})$.

This means that two approximation granules g, g' are close in degree at least p (in IS) if and only if the lower approximations, upper approximations and boundary regions defined by them are close in degree at least p, respectively. The degree of closeness describes the quality of approximation of sets X satisfying the above conditions. Another coefficient which can be used for expressing closeness of approximation granules to sets they approximate is

$$Q(g) = \frac{card([\alpha]_{IS})}{card([\beta]_{IS})}$$

where $((\alpha, [\alpha]_{IS}), (\beta, [\beta]_{IS}))$ is an approximation granule in IS.

The quality of a sequence $g = (g_1, ..., g_k)$ of approximation granules $g_i = ((\alpha_i, [\alpha_i]_{IS}), (\beta_i, [\beta_i]_{IS}))$ where $i = 1, \ldots, k$ can be defined by

$$Q(g) = \frac{card(\bigcup_{i=1}^{k}[\alpha_i]_{IS})}{card(\bigcup_{i=1}^{k}[\beta_i]_{IS})}.$$

3. Rough–Fuzzy Granules

In this section, we will briefly discuss approximation schemes of granules and methods for extracting from them relevant patterns in case where these schemes include fuzzy concepts [23] as foreign parts. We propose to use the rough set approach to define approximations of fuzzy concepts in a constructive way. The rough set approximations of the fuzzy cuts are used in searching for constructive definition of fuzzy sets. We use the cut approximations to derive from schemes patterns relevant for the target concept. In the process of searching for high quality patterns, evolutionary techniques can be used.

Let $DT = (U, A, d)$ be a decision table where the decision is restricted to the objects from U of the fuzzy membership function $\mu : U \to [0, 1]$. Consider reals $0 < c_1 < \ldots < c_k$ where $c_i \in (0, 1]$ for $i = 1, \ldots, k$. Any c_i defines c_i-cut by $X_i = \{x \in U : \mu(x) \geq c_i\}$. Assume $X_0 = U, X_{k+1} = X_{k+2} = \emptyset$.

A *rough–fuzzy granule* (*rf–granule*, for short) corresponding to $(DT, c_1, \ldots, c_k)$ is any granule $g = (g_0, \ldots, g_k)$ such that for some $B \subseteq A$

1. $Sem_B(g_i) = (\underline{B}(X_i - X_{i+1}), \overline{B}(X_i - X_{i+1}))$
   for $i = 0, \ldots, k$;
2. $\overline{B}(X_i - X_{i+1}) \subseteq (X_{i-1} - X_{i+2})$
   for $i = 1, \ldots, k$.

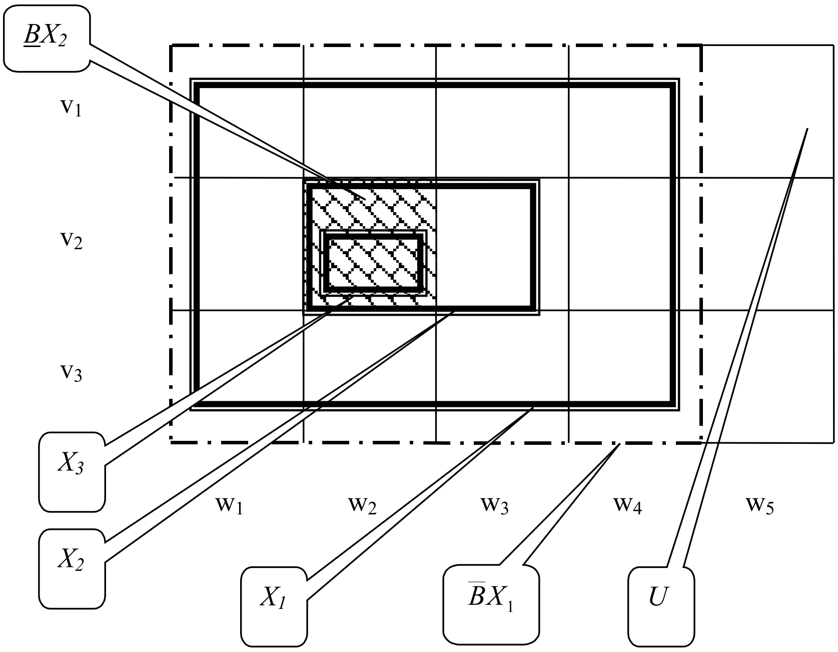

Figure 1. Example of Rough–Fuzzy Granule

An example of rough–fuzzy granule is presented in Figure 1. We assume that there are three cuts, i.e., $k = 3$ and we obtain the following sequence of sets $\emptyset = X_5 = X_4 \subset X_3 \subset X_2 \subset X_1 \subset X_0 = U$. We also assume that a set of attributes B is equal to $\{a_1, a_2\}$. There are three possible values v_1, v_2, v_3 of attribute a_1 and five possible values w_1, w_2, w_3, w_4, w_5 of attribute a_2. Thus we obtain fifteen indiscernibility classes described by descriptor conjunctions of the form $a_1 = v_i \wedge a_2 = w_j$, where $i = 1, 2, 3$ and $j = 1, 2, 3, 4, 5$. The lower approximation of X_3 is equal to the empty set and the upper approximation of X_3 is equal to the lower approximation of X_2. The upper approximation of X_2 is equal to the lower approximation of X_1.

Any function $\mu^* : U \to [0, 1]$ satisfying the following conditions:

1. $\mu^*(x) = 0$ for $x \in U - \overline{B}X_1$;
2. $\mu^*(x) = 1$ for $x \in \underline{B}X_k$;
3. $\mu^*(x) = c_{i-1}$ for $x \in \underline{B}(X_{i-1} - X_i)$
   and $i = 2, \ldots, k - 1$;
4. $c_{i-1} < \mu^*(x) < c_i$ for $x \in \overline{B}X_i - \underline{B}X_i$,
   $i = 1, \ldots, k$, and $c_0 = 0$;

is called a B-approximation of μ.

Now one can choose either the lower or upper approximations of parts, i.e., the set theoretical differences between successive cuts, and propagate them along the scheme in searching for relevant patterns. Another strategy is to propagate the global approximation of foreign fuzzy concepts through the scheme.

The approach presented here can be generalized. One can consider approximations of relevant patterns instead of cut approximations and use them for deriving AR-rules. One of such methods for extracting patterns from data is presented in [9]. The method is based on decomposition of data tables with respect to attributes aimed at finding the largest Cartesian products of local patterns sufficiently included in the target concept. In the following section we discuss a decomposition problem based on background knowledge. This problem is of a great importance in classification of situations by autonomous systems on the basis of sensor measurements [22].

4. Decomposition

In this section, we show that in some cases decomposition can be performed using methods for specific rule generation based on Boolean reasoning [6]. Moreover, we present how the decomposition stable with respect to information granule deviations can be obtained.

4.1. Granule decomposition problems

First, the representation problem for operations on information granules will be discussed. We assume, any (partial) operation $f : G_1 \times \ldots \times G_k \rightarrow H$ with arguments from the sets $G_1, \ldots, G_k$ of information granules and values in the set H of information granules is partially specified by a data table (information system [10]). Any raw datum in the table corresponds to an object that is a tuple $(g_1, \ldots, g_k, f(g_1, \ldots, g_k))$, where $(g_1, \ldots, g_k)$ belongs to the domain of f. The attribute values for a given object consist of

1. values of attributes from sets $A_{G_1}, \ldots, A_{G_k}$ on information granules $g_1, \ldots, g_k$ (attributes are extracted from some preassumed feature languages $L_1, , \ldots, L_k$);
2. values of attributes characterizing relations between information granules $g_1, \ldots, g_k$ specifying conditions under which the tuple $(g_1, \ldots, g_k)$ belongs to (a relevant part of) the domain of f;
3. values of attributes selected for the information granule $f(g_1, \ldots, g_k)$ description.

In this way a partial information about the function f is given. In our considerations, we assume, objects indiscernible by condition attributes are indiscernible by decision attribute, i.e., the considered decision table $DT = (U, A, d)$ is consistent [10]. We assume also the representation is consistent with a given function on information granules, i.e., any f-image of Cartesian product of indiscernibility classes defined by condition attributes is included in a decision indiscernibility class.

Now, we explain in what sense the decision table $DT = (U, A, d)$ can be treated as a partial information about the function $f : G_1 \times \ldots \times G_k \rightarrow H$. Let for $i = 1, \ldots, k$

$$G_i^{DT} = \{g_i \in G_i : \text{ there exists in } DT \text{ an object } (g_1, \ldots, g_i, \ldots, g_k, h)\}.$$

One can define H^{DT} in an analogous way. The decision table DT defines a function

$$f_{DT} : G_1/IND(A_{G_1}) \times \ldots \times G_k/IND(A_{G_k}) \rightarrow H^{DT}/IND(d)$$

by

$$f_{DT}([g_1]_{IND(A_{G_1})}, \ldots, [g_k]_{IND(A_{G_k})}) = [h]_{IND(d)}$$

if and only if

$(g_1, \ldots, g_k, h)$ is an object of DT.

We assume a consistency modeling condition for f is satisfied, namely

$$f([g_1]_{IND(A_{G_1})} \times \ldots \times [g_k]_{IND(A_{G_k})}) = f_{DT}([g_1]_{IND(A_{G_1})}, \ldots, [g_k]_{IND(A_{G_k})})$$

for any $(g_1, \ldots, g_k) \in G_1^{DT} \times \ldots \times G_k^{DT}$.

The function description can be induced from such a data table by interpreting it as a decision table with the decision corresponding to the attributes specifying the values of the function f.

We assume a family of inclusion relations $\nu_p^i \subseteq G_i \times G_i$, $\nu_p^H \subseteq H \times H$ and a family of closeness relations $cl_p^i \subseteq G_i \times G_i$ (for every $p \in [0,1]$ and $i = 1, \ldots, k$), $cl_p^H \subseteq G \times G$, are given [13]. Let us also assume two thresholds t, p are given. We define a relation $Q_{t,p}^{DT}(Pattern_1, \ldots, Pattern_k, \overline{v})$ between granules $Pattern_1, \ldots, Pattern_k$, called patterns, and the target pattern $\overline{v}$ representing the decision value vector in the following way:

$$Q_{t,p}^{DT}(Pattern_1, \ldots, Pattern_k, \overline{v})$$

if and only if the following two conditions are satisfied

$$\nu_p^H(f(Sem_{DT}(Pattern_1) \times \ldots \times Sem_{DT}(Pattern_k)), [\overline{v}]_{IND(d)})$$
$$card(Sem_{DT}(Pattern_1) \times \ldots \times Sem_{DT}(Pattern_k)) \geq t.$$

Let us consider the following decomposition problem:

Granule decomposition problem

Input:

- two thresholds t, p;
- a decision table $DT = (U, A, d)$ representing an operation $f : G_1 \times \ldots \times G_k \rightarrow H$ where $G_1, \ldots, G_k$ and H are given sets of information granules;
- a fixed decision value vector $\overline{v}$ represented by a value vector of decision attributes.

Output:

- a tuple $(Pattern_1, \ldots, Pattern_k)$ of patterns such that $Q_{t,p}^{DT}(Pattern_1, \ldots, Pattern_k, \overline{v})$.

We consider a description given by means of decision rules extracted from the data table specifying the function f. Any left hand side of a decision rule can be divided into parts corresponding to different arguments of the function f. The i-th part (denoted by $Pattern_i$), specifies a condition which should be satisfied by the i-th argument of f to obtain the function value specified by the decision attributes. In this way the left hand sides of decision rules describe patterns $Pattern_i$ where $i = 1, \ldots, k$. The semantics of extracted patterns relevant for the target can be defined as the image with respect to f of the Cartesian product of sets $Sem_{DT}(Pattern_i)$, i.e., by $f(Sem_{DT}(Pattern_1) \times \ldots \times Sem_{DT}(Pattern_k))$.

One can use one of the methods for decision rule generation, e.g., for generation of minimal rules or their approximations (e.g., in the form of association rules) [6] to obtain such decision rules.

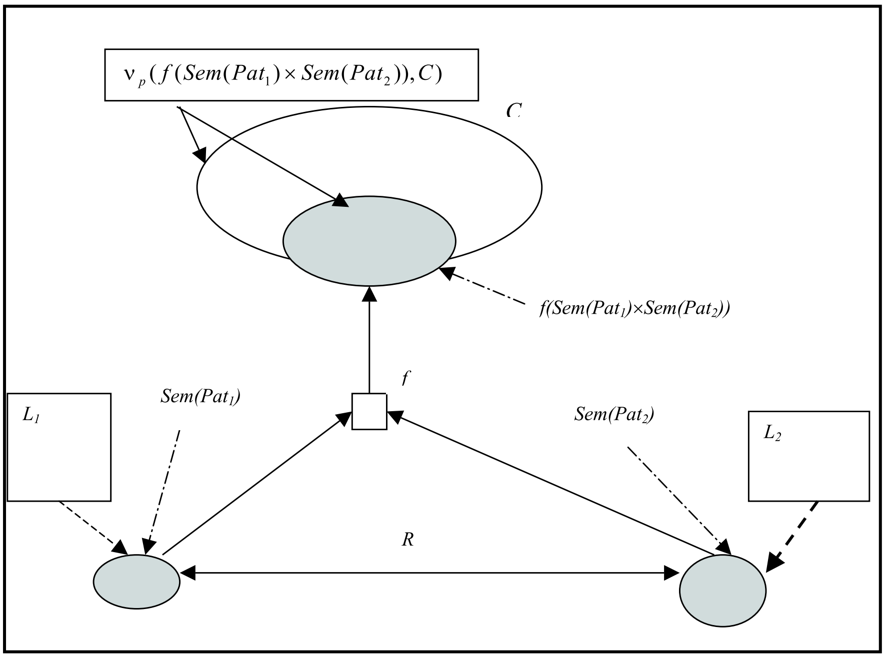

Figure 2. Decomposition of Information Granule

In the former case we obtain the most general patterns for function arguments consistent with a given decision table, i.e., the information granules constructed by means of the function f from patterns extracted for arguments are included exactly in the information granule represented by a given decision value vector in the data table.

In the latter case we obtain more general patterns for function arguments having the following property: information granules constructed by means of f from such patterns will be included in a satisfactory degree in the information granule represented by a given decision value vector in the data table.

One of the very important properties of the above discussed operations on information granules is their robustness with respect to the deviations of arguments (see, e.g., [14]). This property can be formulated as follows: if information granule constructed by means of f from the extracted patterns $Pattern_1, \ldots, Pattern_k$ is satisfying the target condition then the information granule constructed from patterns $Pattern'_1, \ldots, Pattern'_k$ sufficiently close to $Pattern_1, \ldots, Pattern_k$, respectively, satisfies the target condition too. In this way we obtain the following problem:

Robust decomposition problem (RD-problem)

Input:

- thresholds t, p;
- a decision table $DT = (U, A, d)$ representing an operation $f : G_1 \times \ldots \times G_k \rightarrow H$ where $G_1, \ldots, G_k$ and H are given sets of information granules;
- a fixed decision value vector $\overline{v}$ represented by a value vector of decision attributes.

Output:

- a tuple $(p_1, \ldots, p_k)$ of parameters;
- a tuple $(Pattern_1, \ldots, Pattern_k)$ of patterns such that
  if $cl^i_{p_i}(Sem_{DT}(Pattern_i), Sem_{DT}(Pattern'_i))$ for $i = 1, \ldots, k$
  then $Q^{DT}_{t,p}(Pattern'_1, \ldots, Pattern'_k, \overline{v})$.

It is possible to search for the solution of the RD-problem by modifying the previous approach of decision rule generation. In the process of rule generation one can impose a stronger discernibility condition by assuming objects to be discernible if their tolerance classes are disjoint. Certainly, one can tune parameters of tolerance relations to obtain rules of satisfactory quality. We would like to stress that efficient heuristics for solving these problems can be based on Boolean reasoning [6].

4.2. Granule decomposition based on background knowledge

In this section, we discuss briefly a granule decomposition problem supported by background knowledge. This is one of the basic problems in synthesis of approximate schemes of reasoning from experimental data.

Assume knowledge base consists of a fact expressing that if two objects belong to concepts C_1 and C_2, then the object constructed out of them by means of a given operation f belongs to the concept C (see Figure 3) provided that the two objects satisfy some constraints. However, we can only approximate these concepts on the basis of available data. Using a (generalized) rough set approach [16] one can assume that it an inclusion measure v_p for $p \in [0, 1]$ is given that makes possible to estimate the degree of inclusion of data patterns Pat, Pat_1, and Pat_2 from languages L, L_1, and L_2 in the concepts C, C_1, and C_2, respectively. Patterns included in a satisfactory degree p in a concept are classified as belonging to its lower approximation while those included to a degree less than a preset threshold are classified as belonging to its complement. The decomposition problem is a searching problem for patterns Pat of *high* quality (e.g., supported by a large number of objects) and included in a satisfactory degree in the target concept C. These patterns are obtained by performing a given operation f on some input patterns Pat_1 and Pat_2 (from languages L_1 and L_2, respectively) sufficiently included in C_1 and C_2, respectively.

One can develop a searching method for such patterns Pat based on tuning of inclusion degrees p_1, p_2 of input patterns Pat_1, Pat_2 in C_1, C_2, respectively, to obtain patterns Pat (constructed from Pat_1, Pat_1 by means of a given operation f) included in C in a satisfactory degree p and of acceptable quality (e.g., supported by the number of objects larger than a given threshold).

Assume degrees p_1, p_2 are given. There are two basic steps of searching procedures for relevant pairs of patters (Pat_1, Pat_2) :

1. Searching in languages L_1 and L_2 for sets of patterns included in degree at least p_1 and p_2 in concepts C_1 and C_2, respectively.
2. Selecting from sets of patterns generated in Step 1 of satisfactory pattern pairs.

We would like to add some general remarks on the above steps.

One can see that our method is based on a decomposition of degree p into degrees p_1 and p_2 under some constraints. In Step 2 we search for a relevant constraint relation R between patterns. The goal is to extract the following approximate rule of reasoning:

if

$R(Sem(Pat_1), Sem(Pat_2))\wedge$
$\nu_{p_1}(Sem(Pat_1), C_1)\wedge$
$\nu_{p_2}(Sem(Pat_2), C_2)$

then

$\nu_p(f(Sem(Pat_1) \times Sem(Pat_2)), C)\wedge$
$Quality_t(f(Sem(Pat_1) \times Sem(Pat_2)))$

where p is a given inclusion degree; t, - a threshold of pattern quality measure $Quality_t$; f, - an operation on objects (patterns); Pat, - a target pattern, C, C_1, C_2, - given concepts; and R, p_1, p_2 are expected to be extracted from data and (Pat_1, Pat_2) is satisfying R (in our case R is represented by a finite set of pattern pairs).

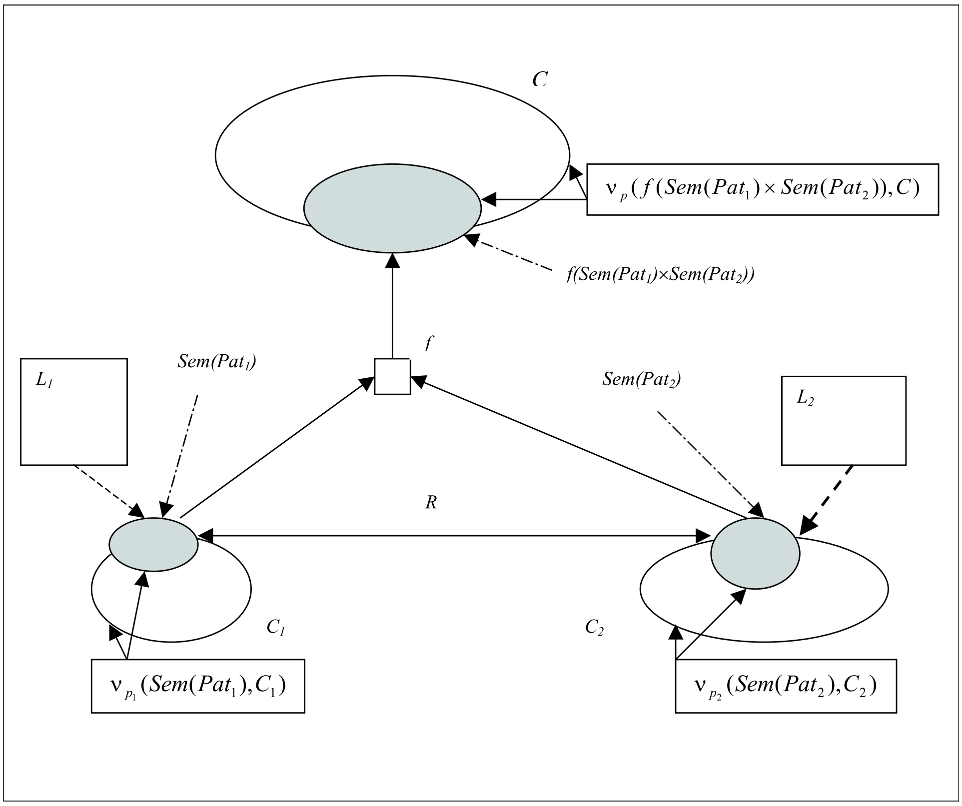

Figure 3. Granule Decomposition Based on Background Knowledge

One can consider soft constraint relations R_r where $r \in [0, 1]$ is a degree of truth to which the constraint relation holds.

Two sets P_1, P_2 are returned as the result of the first step. They consist of pairs (*pattern*, *degree*) where *pattern* is included in C_1, C_2, respectively in degree at least *degree*.

These two sets are used to learn the relevant relation R. We outline two methods.

The first one is based on an experimental decision table (U, A, d) where U is a set of pairs of discovered patterns in the first step; $A = \{deg_1, deg_2\}$ consists of two attributes such that $deg_i((Pat_1, Pat_2))$ is equal to the degree to which Pat_i is at least included in C_i for $i = 1, 2$; the decision d has value p to which the granule composed by means of operation f from (Pat_1, Pat_2) is at least included in C. From this decision table the decision rules of a special form are induced:

$$\textbf{if } deg_1 \geq p_1 \wedge deg_2 \geq p_2 \textbf{ then } d \geq p$$

where (p_1, p_2) is the minimal degree pair such that if $p'_1 \geq p_1$ and $p'_2 \geq p_2$, then the decision rule obtained from the above rule by replacing p'_1, p'_2 instead of p_1, p_2, respectively, is also true in the considered decision table.

A version of such a method has been proposed in [12]. The relation R consists of the set of all pairs (Pat_1, Pat_2) of patterns with components included in C_1, C_2, respectively in degrees $p'_1 \geq p_1, p'_2 \geq p_2$ where p_1, p_2 appear on the left hand side of some of the generated decision rules.

The second method is based on another experimental decision table (U, A, d) where objects are triplets $(x, y, f(x, y))$ composed out of objects x, y and the result of f on arguments x, y; attributes from A describe features of arguments of objects and the decision d is equal to the degree to which the elementary granule corresponding to the description of $f(x, y)$ by means of attributes is at least included in C. This table is extended by adding new features that are characteristic functions a_{Pat_i} of patterns Pat_i discovered in the first step. Next the attributes from A are deleted and from the resulting decision table the decision rules of a special form are induced:

$$\textbf{if } a_{Pat_1} = 1 \wedge a_{Pat_2} = 1 \textbf{ then } d \geq p$$

where if Pat_1, Pat_2 are included in C_1, C_2, in degree at least p_1, p_2, respectively and Pat'_1, Pat'_2 are included in C_1, C_2 in degree $p'_1 \geq p_1$ and $p'_2 \geq p_2$, respectively then a decision rule obtained from the above rule by replacing Pat'_1, Pat'_2 instead of Pat_1, Pat_2 is also true in the considered decision system.

The decision rules again describe constraints specifying the constraint relation R.

Certainly, in searching procedures one should also consider constraints for the pattern quality.

The searching methods discussed in this section return local granule decomposition schemes. These local schemes can be composed using techniques discussed in [13]. The resulting schemes of granule construction (these schemes can be also treated as approximate reasoning schemes) have also the following property: if the input granules are sufficiently close to input concepts, then the output granule is sufficiently included in the target concept provided this property is preserved locally [13].

Searching for relevant patterns for information granule decomposition can be based on methods for tuning parameters of rough set approximations of fuzzy cuts or concepts defined by differences between cuts (see Section 3). In this case, pattern languages consist of parameterized expressions describing the rough set approximations of *parts* of fuzzy concepts that are fuzzy cuts or differences between cuts. Hence, an interesting research direction related to the development of new hybrid rough-fuzzy methods arises aimed at developing algorithmic methods for rough set approximations of such parts of fuzzy sets relevant for information granule decomposition.

Conclusions

We have discussed an approach for extracting relevant patterns from parameterized schemes of information granule construction (AR-schemes) consisting of parts from different information sources. The schemes can be also treated as schemes of approximate reasoning built on the basis of perception by means of information granule calculi. Relevant output patterns (information granules) can be obtained by tuning of the scheme parameters. We have emphasized the necessity of approximation (in an accessible language) of information granules that are parts of schemes and expressed in another language called foreign language. In our further study we plan to develop evolutionary searching techniques for optimal parameters of information granule construction schemes extracted from data and background knowledge.

Acknowledgements

We would like to thank to Professor James F. Peters for insightful comments and a number of clarifying discussions on the presented paper.

References

[1] Barsalou, L. W.: Perceptual symbol systems, *Behavioral and Brain Sciences* **22**, 1999, 577–660.

[2] Escrig, M. T., Toledo, F.: *Qualitative Spatial Reasoning: Theory and Practice*, IOS Press, Amsterdam, 1998.

[3] Doherty, P., Łukaszewicz, W., Skowron, A., Szałas, A.: Knowledge representation + data mining = knowledge mining (submitted).

[4] Düntsch, I. (Ed.): Spatial Reasoning , *Fundamenta Informaticae* (special issue, in press).

[5] Kargupta, H., Chan, Ph.: *Advances in Distributed and Parallel Knowledge Discovery*, AAAI Press/MIT Press, Cambridge 2001.

[6] Komorowski, J., Pawlak, P., Polkowski, L., and Skowron, A.: Rough sets: A tutorial, *Rough fuzzy hybridization: A new trend in decision-making* (S. K. Pal and A. Skowron, Eds.), Springer-Verlag, Singapore, 1999, 3–98.

[7] Koza, J. R.: *Genetic Programming II: Automatic Discovery of Reusable Programs*, MIT Press, Cambridge, MA, 1994.

[8] Nguyen, H. S., Nguyen, S. H., Skowron, A.: Decomposition of task specification, *Proc. of the 11th International Symposium on Foundations of Intelligent Systems*, June 1999, Warsaw, Poland, LNCS 1609, Springer-Verlag, Berlin 1999, 310–318.

[9] Nguyen, H. S., Skowron, A., Stepaniuk, J.: Granular computing: A rough set approach, *Computational Intelligence* (to appear).

[10] Pawlak, Z.: *Rough Sets. Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Dordrecht, 1991.

[11] Pirri, F., Finzi, A.: An approach to perception in theory of actions, *Linköping Electronic Articles in Computer and Information Science*, **4**(41), 1999, http://www.ep.liu.se/ea/cis/1999/041

[12] Polkowski, L., Skowron, A.: Rough mereological approach to knowledge-based distributed AI, *Critical Technology, Proc. of the Third World Congress on Expert Systems* (J.K. Lee, J. Liebowitz, and J.M. Chae, Eds.), February 5-9, Seoul, Korea, Cognizant Communication Corporation, New York, 1996, 774–781.

[13] Polkowski, L., Skowron, A.: Towards adaptive calculus of granules, In: [25], vol. I, 201–227.

[14] Polkowski, L., Skowron, A.: Rough mereological calculi of granules: A rough set approach to computation, *Computational Intelligence* (to appear).

[15] Rock, I.: *Indirect Perception*, The AAAI Press/ MIT Press, Cambridge 1996.

[16] Skowron, A., Stepaniuk, J.: Tolerance approximation spaces *Fundamenta Informaticae*, **27**(2-3), 245–253.

[17] Skowron, A., Stepaniuk, J.: Information granules: Towards foundations of granular computing, *International Journal of Intelligent Systems*, **16**(1), 2001, 57–86.

[18] Stepaniuk, J.: Knowledge discovery by application of rough set models, *Rough Set Methods and Applications. New Developments in Knowledge Discovery in Information Systems* (L. Polkowski, S. Tsumoto, T.Y. Lin, Eds.), Physica–Verlag, Heidelberg, 2000, 137–233.

[19] Stone, P.: *Layered Learning in Multiagent Systems: A Winning Approach to Robotic Soccer*, MIT Press, Cambridge 2000.

[20] WWW SPACENET page: http://agora.scs.leeds.ac.uk/spacenet/.

[21] Waltz D.: The importance of importance, *AI Magazine*, **20**(3), 1999, 18–26.

[22] WITAS project web page: http://www.ida.liu.se/ext/witas/eng.html.

[23] Zadeh, L. A.: Fuzzy sets, *Information and Control*, **8**, 1965, 333–353.

[24] Zadeh, L. A.: Fuzzy logic = Computing with words, *IEEE Trans. on Fuzzy Systems*, **4**, 1996, 103–111.

[25] Zadeh, L. A., Kacprzyk, J. (Eds.): *Computing with Words in Information/Intelligent Systems*, vol. I–II, Physica–Verlag, Heidelberg, 1999.

[26] Zadeh, L. A.: A new direction in AI: Toward a computational theory of perceptions, *AI Magazine*, **22**(1), 2001, 73–84.