Relation Recognition Problems and Algebraic Approach to Their Solution

Juliusz L. Kulikowski(✉)

Nalecz Institute of Biocybernetics and Biomedical Engineering PAS.,
4 Ks. Trojdena Street, 02-109 Warsaw, Poland
juliusz.kulikowski@ibib.waw.pl

Abstract. Relation recognition as an extension of the well known pattern recognition problem is presented in the paper. Four types of such problems: simple, extended, matching and constructive relation recognition problems are considered. It is shown that such problems may arise in various application areas. There are presented possible approaches to the solution of the problems under consideration. It is shown that the extended algebra of relations is suitable as an universal tool to the relation recognition problems exact formulation and to description of the methods of their solution. Suitability of a concept of general (multi-aspect) similarity measure to the solution of constructive relation recognition problems based on a concept of covering by similarity spheres is also shown. Suggestions concerning desired future works in the domain of relation recognition theory and applications are given.

Keywords: Relation recognition · Pattern recognition · Extended algebra of relations · Learning systems · Artificial intelligence

1 Introduction

The notion of *relation* belongs to basic concepts of modern mathematics. It formally is defined as any subset of Cartesian product of a linearly ordered family of sets. It also plays a significant role in computer science as a formal model of data structures (Codd 1970), description of composite real objects or situations (Bagui 2005), description of program structures (Wirth 2002) etc. The well known *pattern recognition* (*PR*) problems are based on the concept of *similarity* being in fact a bi-variable reflexive, symmetrical and transitive relation described on a set of some (abstract or real) objects. An important property of similarity relation consists in partition of the given set of objects into mutually disjoint subsets called *similarity classes* (Kulikowski 2003). It was remarked in Kulikowski (1987) that *PR* can be interpreted as *checking* the fact that a given object $\boldsymbol{x}$ satisfies the relation of similarity to other objects belonging to a similarity class of objects. The pairs (in general – n-tuples) of elements satisfying a given relation Ξ are called *syndromes* of Ξ. The *PR* problem can thus be also interpreted as proving whether a given object $\boldsymbol{x}$ forms syndromes of similarity with the elements of some similarity classes established by the relation.

This brings to mind a more general concept of *relation recognition* (*RR*) consisting in identification of syndromes of a relation (Kulikowski 2002). In this case no

J. Świątek and J.M. Tomczak (eds.), *Advances in Systems Science*, Advances in Intelligent Systems and Computing 539, DOI 10.1007/978-3-319-48944-5_3

homogeneity of the sets constituting the space U of n-tuples is assumed. This means that in general objects of different formal nature: arithmetical, algebraic, geometrical, Boolean, topological or symbolic denotations of qualitative features of real objects etc. may be linked together by the relations. Under such assumption several variants of *RR* problem can be formulated:

1st, it is given a relation Ξ described on a Cartesian product U of n non-empty sets and a n-tuple of objects $\boldsymbol{x}$ belonging to U; check whether Ξ is satisfied by $\boldsymbol{x}$. This can be called a *simple RR* (*sRR*) problem.
2nd, it is given a relation Ξ and a finite subset $X \subset U$ of n-tuples; find in X all n-tuples being syndromes of Ξ. This can be called an *extended RR* (*eRR*) problem.
3rd, it is given a relation Ξ described on a Cartesian product U of n non-empty sets and its incomplete syndromes $\boldsymbol{x}$, whose some components are unknown; restore their possible complete forms matching the relation. This can be called a *matching RR* (*mRR*) problem.
4th, it is given a Cartesian product U of n non-empty sets and a finite subset $X \subset U$ of n-tuples; find formal rules describing a relation Ξ in U containing X as a subset of its syndromes. This can be called a *constructive RR* (*cRR*) problem.

Some *RR* problems are closely connected with machine learning problems. This, for example, takes place in the *eRR* problem consisting in detection of tracks of particles in a series of snapshots (Sect. 3, Example 2) or in *mRR* problem consisting in meteorological prognosis based on a series of past observations. However, the paper does not present any extended example of *RR* application; it is rather aimed at showing the *RR* as a large and interesting area of investigations.

This paper is aimed at presentation of an approach to the solution of the above-mentioned classes of *RR* problems. However, the admitted heterogeneity of the space V of n-tuples limits the ability of a geometrical- or vector-space models to be used as a basis of the solution. That is why in this work an approach based on the extended algebra of relations (Kulikowski 1992) and on a general concept of similarity measure (Kulikowski 2001) is proposed.

The paper is organized as follows. Basic notions, concerning mainly the used in this work extended algebra of relations, are presented in Sect. 2. Section 3 presents solution of the *sRR* and *eRR* problems. An approach to the solution of *mRR* problems is presented in Sect. 4, while Sect. 5 presents a proposed solution of the *cRR* problem based on a concept of covering by similarity spheres. Concluding remarks and suggestions concerning future works are presented in Sect. 6.

2 Basic Notions

It will be taken into consideration a finite, linearly ordered family F of non-empty sets $[\Omega^{(i)}]$, $i = 1,2,\ldots,n$. As mentioned above, no restrictions on formal nature of the elements of the sets $\Omega^{(i)}$ will be imposed. The Cartesian product of the sets $\Omega^{(i)}$ of F will be denoted by U and will be called an *universe*. In order not to suggest any connection with particular mathematical objects, the elements of the sets $\Omega^{(i)}$ are called *symptoms*, while the elements of U (strings of symptoms, n-tuples) are called *syndromes*.

Any subset $\Xi \subseteq U$ is called a *relation* described on U. The relation is called *trivial* if $\Xi \equiv U$ and *empty* if $\Xi \equiv \varnothing$ (empty set). For a given universe U the family of all possible relations described on U is denoted by Φ. In this family the well-known algebraic operations of sets sum ($\cup$), intersection ($\cap$) and asymmetrical (\) and symmetrical ($\div$) difference can directly be applied to the relations. In particular, a *negation* of relation $\neg\Xi$ is defined as asymmetrical difference $U\backslash\Xi$. It directly follows that for any Ξ it is $\Xi \cup (\neg\Xi) = U$ and $\Xi \cap (\neg\Xi) = \varnothing$.

If Ξ is a relation described on U then any subset $S \subset \Xi$ is called a *sub-relation* of Ξ. On the other hand, if F' is a subfamily of sets, $F' \subset F$, (preserving their order in F), U' is a Cartesian product of the sets of F' (a *sub-universe* of U) and Ξ is a relation described on U then a set Ξ' consisting of all intersections of the syndromes of Ξ and of U' will be called a *projection* of Ξ onto U' and will be also denoted by $\Xi|_{U'}$.

For a given family F of consisting of n sets it is possible to extract from it 2^n–1 non-empty sub-families of sets (including F itself). On each such sub-family of selected sets a family of all described on them relations, called *partial relations* described in U, can be considered. Consequently, it is possible to take into consideration a family Φ of all partial relations described in U (including also the relations described on U). An extension of the algebra of relations described on a fixed family of sets F on partial relations described on different subfamilies of F was given in Kulikowski (1992). If F' and F'' are any two subfamilies of F and Ξ', Ξ'' are some relations described, respectively, on the Cartesian products U' of the sets of F' and U'' described on F'' then there can be established the following:

Extended algebraic operations:

(a) The sum $\Xi' \underline{\cup} \Xi''$ is a relation described in $U' \cup U''$, consisting of all syndromes whose projection on U' satisfies Ξ' ***or*** the projection on U'' satisfies Ξ''.
(b) The intersection $\Xi' \underline{\cap} \Xi''$ is a relation described in $U' \cup U''$, consisting of all syndromes whose projection on U' satisfies Ξ' ***and*** the projection on U'' satisfies Ξ''.
(c) The asymmetrical difference $\Xi' \underline{\backslash} \Xi''$ is a relation described in $U' \cup U''$, consisting of all syndromes whose projection on U' satisfies Ξ' ***and*** the projection on U'' ***does not*** satisfy Ξ''.

Let us remark that $U' \cup U''$ is an universe constructed as a Cartesian product of the family of sets $F' \cup F''$ (not as $U' \times U''$). Then, on the basis of the above-given operations the following notions can also be defined:

(d) The difference $2^{(U' \cup U'')} \underline{\backslash} \Xi$ is called an *extended negation* of Ξ and will be denoted by $\underline{\neg}\Xi$.
(e) The asymmetrical difference $\Xi' \underline{\div} \Xi''$ is a relation described in $U' \cup U''$, consisting of all syndromes satisfying $\Xi' \underline{\cup} \Xi''$ ***and not*** satisfying $\Xi' \underline{\cap} \Xi''$.
(f) For a relation Ξ' described in a sub-universe U' and another relation Ξ'' described in a sub-universe U'' the projection $(\Xi' \underline{\cap} \Xi'')|_{U'}$ will be called a *relative relation* Ξ' *assuming that* Ξ'' *holds*. It will be shortly denoted by Ξ'/Ξ''.

It can be shown that the quintuple $[U, \underline{\cup}, \underline{\cap}, \underline{\backslash}, \varnothing]$ constitutes a Boolean algebra (Rudeanu 2012) with U as its “unity” and $\varnothing$ as its “null relation” (no syndromes of any length); this will be called an *extended algebra of relations*. It thus introduces algebraic operations not only on the relations defined on the universe U but also on any their sub-relations and/or projections defined in U.

On the basis of the property (f) it can be established the following

Extension property:

For any relation Ξ' and Ξ'' described, respectively, in the universes U' and U'', the following property holds:

$$\Xi' \underline{\cap} \Xi'' \equiv \Xi'/\Xi'' \underline{\cap} \Xi''. \tag{1}$$

This property, seeming to be a trivial one, will be shown to be useful in solving some *RR* problems.

3 Solution of *Simple* and *Extended RR* Problems

Four basic methods can be used to description of relations:

(1) Analytical (functional) description,
(2) Logical description,
(3) Presentation by algebraic composition of other known (simpler) relations,
(4) Characterization by list of syndromes.

Analytical description takes place when syndromes of a relation should by definition satisfy some algebraic or analytical (functional, differential, integral etc.) equations. Solution of the equation provides all syndromes of the relation. On the other hand, an assumed solution can be proven as syndrome of the relation by substitution to the equation and checking whether it is satisfied.

Logical description of a relation may be given in the form of a formula:

$$\boldsymbol{if}\ T(\boldsymbol{x})\boldsymbol{then}\ \boldsymbol{x} \in \Xi \tag{2}$$

where $T(\boldsymbol{x})$, is a logical predicate described in the universe U. The predicate can thus be used as a basis of construction of logical tests of the syndromes' validity. However, it does not deny existence of some other syndromes of Ξ, not satisfying $T(\boldsymbol{x})$.

Algebraic compositions of relations can be constructed on the basis of the above (in Sect. 2) described algebraic operations and concepts.

Characterization of a relation by a complete list of its syndromes is possible only in particular cases. In a more general case the list may be incomplete and it can be used as a rough representation of the relation.

Taking this into account we can more exactly formulate:

The *sRR* problem:

It is given an universe U and a relation Ξ described in it by some of the above-mentioned methods (1)–(2). It is also given an element $\boldsymbol{x} \in U$. Check whether $\boldsymbol{x} \in \Xi$.

Computer-aided solution of this problem depends on the method the relation has been described. In the case of analytical description it needs substitution of x into the corresponding equation and proving whether the equation is satisfied. This seems to be a rather trivial task. A less trivial task arises if the relation is described in a logical form (2). The predicate $T(\boldsymbol{x})$ may not directly characterize the relation, as it illustrates the following example.

Example 1. It is given an Euclidean plane E^2 with a system of Cartesian coordinates (u, v). It is taken into consideration an universe $U = E^2 \times E^2 \times E^2$ of triplets of points laying on E^2. It is also given a relation $\Xi \subset U$ described by the predicate:

"The triplet of points $[P_i, P_j, P_k]$ constitute vertices of an equilateral triangle",

Let it be given a triplet of points $x = [(u_i, v_i), (u_j, v_j), (u_k, v_k)]$. Check whether $\boldsymbol{x}$ is a syndrome of Ξ.

In this case a direct assessment of logical value of the statement:

"$[(u_i, v_i), (u_j, v_j), (u_k, v_k)]$ constitute vertices of an equilateral triangle"

is impossible. In this case it is necessary to use the definition of an equilateral triangle and to replace the statement by the following, semantically equivalent:

$$[P_i, P_j, P_k] \text{ constitute vertices of an equilateral triangle}$$
$$\textbf{\textit{if and only if}}\left[d(P_i, P_j) = d(P_j, P_k)\right] \wedge \left[d(P_j, P_k) = d(P_k, P_i)\right]$$

where $d(P_i, P_j)$ (similarly, $d(P_j, P_k)$ and $d(P_k, P_i)$) denotes an Euclidean distance between the points P_i, P_j. Now, the relation description leads to the following logical test of a triplet of points satisfying the relation:

$$\textbf{\textit{if}}\left[d(P_i, P_j) = d(P_j, P_k)\right] \wedge \left[d(P_j, P_k) = d(P_k, P_i)\right]$$
$$\textbf{\textit{then}}\left[P_i, P_j, P_k\right] \text{ satisfies the relation } \Xi\text{''},$$
$$\textbf{\textit{otherwise}} \text{ it does not satisfy it.}$$

In fact, if the relations Ξ are interpreted as formal characteristics of some classes of objects then the corresponding *sRR* problems become widely interpreted *PR* problems. However, the Example 1 shows that solution of some, apparently simple, *sRR* problems may need some wider knowledge about the application area in order to reformulate the primary logical description of the relation.

In some application problems (say, in management, medical treatment etc.) the *sRR* may take a more general form:

1st 'It is given a relation Ξ described by an algebraic combination F of m partial relations $\Xi^{(1)}, \ldots, \Xi^{(m)}$ described on a Cartesian product U of n non-empty sets and a n-tuple of objects $\boldsymbol{x}$ belonging to U. Check whether Ξ is satisfied by $\boldsymbol{x}$.

Let us remind (see Sect. 2) that the partial relations can be defined on different subsets of variables. Their algebraic combination may be given in a canonical form, as an (extended) sum of products of selected partial relations or of their negations (denoted both by the brackets $\langle\,\rangle$):

$$\Xi = F(\Xi^{(1)}, \ldots, \Xi^{(m)}) = \underline{\cup\cap}\langle \Xi^{(v)} \rangle \tag{3}$$

This form of the relation Ξ directly leads to a decomposition of the *sRR* problem into a finite set of simpler *sRR* sub-problems and then proving the validity of the corresponding composite logical predicate.

The calculation complexity of solution of a *sRR* problem depends linearly on the number of additive terms in expression (3). Moreover, each product-term $\underline{\cap}\langle \Xi^{(v)} \rangle$ in this expression needs proving the validity of a subset of partial relations described on various subsets of x. This leads to an exponential complexity of calculation of the product-terms.

An *eRR* problem can be considered as simple extension of a corresponding *sRR* problem on several elements of the universe U; in such case its solution may consist of multiple solution of the *sRR* problem for assumed syndromes x_1, x_2,… etc. However, it also can exactly be formulated as follows:

The *eRR* problem:

It is given an universe U and two described on it relations: Ξ described by the above-mentioned methods (1)–(2) and a finite relation χ described by method (4). Find $\Psi = \Xi \underline{\cap} \chi$

Really, Ψ is a finite set of n-tuples in U, constituting by definition a relation. We are looking in χ for all syndromes satisfying also the given relation Ξ. The following example should illustrate this idea.

Example 2. It is given a linearly ordered set of N photo snapshots. In any snapshot a finite set S_v, $v = 1, 2, \ldots, N$, of detected in it tracks $\omega_{v,p}$ of moving objects is given. Each track $\omega_{v,p}$ is presented by a sequence of m numerical and/or qualitative features describing its spatial coordinates and other (if any) its detected individual characteristics. Find the sequences $\boldsymbol{x} = [\omega_{1,p}, \omega_{2,q}, \ldots, \omega_{N,s}]$, where $p, q, \ldots, s = 1, 2, \ldots, K$, denoting numbers assigned to the detected tracks of objects in the snapshots (K being a total number of objects under observation). The tracks illustrate position of selected real objects in consecutive phases of their motion. The example presents a simplified case of a more general problem of objects' action recognition and trajectories description (Wang et al. 2013). This may correspond in practice to monitoring (or back-monitoring) of selected individual vehicles in road traffic.

The universe has thus the form of a Cartesian product

$$U = S_1 \times S_2 \times \ldots \times S_N \tag{4}$$

consisting of K^N N-tuples and the relation Ξ being to be recognized is described by the statement:

"$\boldsymbol{x} \in \Xi$ ***iff*** $\boldsymbol{x}$ consists of elements representing the same real object"

Solution of the problem consists thus in finding in U all N-tuples $\boldsymbol{\omega}$ satisfying a relation Ξ of "representing the same real object". This needs, first of all, establishment the criteria of assigning tracks to fixed, individual objects Two such criteria are possible: 1st, the similarity of tracks if they present the appearance of real objects, 2nd, proximity of tracks in consecutive snapshots.

Two possible approaches to solution of the given *eRR* problem then can be taken into consideration:

(a) Solving the *eRR* problem by its consideration as multiple *sRR* problems. This means that the elements of U are separately taken into consideration and the property "representing the same real object" of consecutive pairs of tracks is proven. Any sequence $\boldsymbol{x} = [\omega_{1,p}, \omega_{2,q}, \ldots, \omega_{N,s}]$ is thus subjected to a series of tests consisting in proving logical value of the statements:

"The pair $[\omega_{i,p}, \omega_{i+1,q}]$ represents the same real object"

for $i = 1, 2, \ldots, N–1$. First "*false*" statement in the series causes rejection of the given N-tuple $\boldsymbol{x}$ as not satisfying the relation Ξ. This, apparently simple approach does not guarantee uniqueness of the *eRR* problem's solution. The partial decisions being based on separately taken pairs of tracks neglecting wider context may lead to ambiguous assigning tracks to objects if the tracks are similar or laying close each to the other.

(b) Alternative approach to the *eRR* problem is based on the concept of conditional relation. For this purpose the following relations will be defined:

Ξ'_i: "The tracks $\omega_{1,p}, \ldots, \omega_{i,p}$ represent the same real object"

Ξ''_i: "The tracks $\omega_{i,p}, \ldots, \omega_{i+1,p}$ represent the same real object assuming that $\omega_{1,p}, \ldots, \omega_{i,p}$ represent the same real object"

for $2 \leq i \leq N–1$ and $1 \leq p \leq K$. Finally, Ξ'_N contains K syndromes $\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_K$, each containing N tracks assigned to a given object.

The problem is then solved according to the following pseudo- procedure:

Input: N lists of K tracks $\omega_{v,p}$, $1 \leq v \leq N$, $1 \leq p \leq K$;
Output: $\Xi'_N = (\boldsymbol{\omega}_1, \ldots, \boldsymbol{\omega}_K)$;
for $i = 2$ **to** $N–1$ **do**
1. Find all syndromes of Ξ'_i;
2. Find all syndromes of Ξ''_{i+1};
3. Create the syndromes of $\Xi'_{i+1} = \Xi'_i \cap \Xi''_{i+1}$;

end for

The procedure starts by finding K syndromes of Ξ'_2 consisting of ordered pairs $[\omega_{1p}, \omega_{2,q}]$ and K syndromes of Ξ'_3 consisting of ordered pairs $[\omega_{2q}, \omega_{3,r}]$ assigned to the same object. Next, the syndromes of Ξ' are extended by adding tracks assigned to formerly fixed objects.

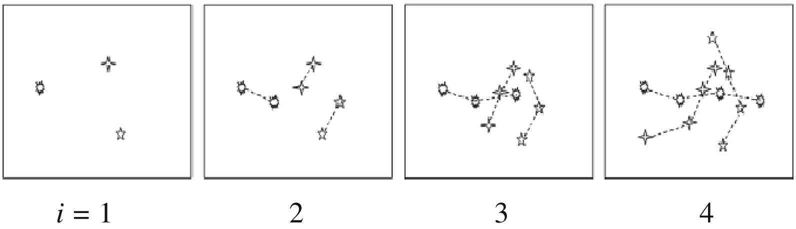

Fig. 1. Construction of moving particles' trajectories as a solution of the *eRR* problem for the relation Ξ = "assigned to the same real object".

At each step the similarity of tracks in the pairs $[\omega_{i,p}, \omega_{i+1,q}]$ is checked, however, it is established by selection of K most similar from $½K^2$ possible pairs of tracks in two consecutive snapshots. For this reason, this approach guarantees not worse solution than this provided by the approach (a). The way of step-wise construction of the trajectories of some moving particles is illustrated in Fig. 1.

4 Solution of *Matching RR* Problems

A *mRR* problem can be simply and exactly formulated in the extended relation algebra terms. Let us assume that it is given a family F of sets composed of two disjoint sub-families:

$$F = F' \cap F'', F' \cap F'' = \emptyset \tag{5}$$

and the described on them universe U and sub-universes U' and U''.

The *mRR* problem:

Let it be given a relation Ξ described on the universe U and a sub- relation $\chi \subset \Xi|_{U'}$ described by a list of syndromes in a sub-universe U'. Find the sub-relation $\Psi = \Xi \cap \chi$.

The sub-relation Ψ contains all syndromes satisfying both, Ξ and χ. Its projection $\Psi|_{U''}$ consists of the lacking elements in the syndromes of χ that transform them into syndromes of Ξ. There are lot of real situations leading to *mRR* problems.

Example 3. In a meteorological database data concerning observations of minimal and maximal air temperature, speed and direction of wind, type and intensity of falls, type and intensity of clouds, etc. in a certain geographical region are collected. The data have been daily recorded for several years. From this long time-interval shorter (say, $N = 10$ days long) sub-intervals have been cut out and the sequences of meteorological data in the sub-intervals as syndromes of a relation H describing admissible sequences of weather conditions have been extracted. Moreover, the relation H is presented as a sum of sub-relations:

$$H = H_w \cup H_{sp} \cup H_{su} \cup H_{au} \tag{6}$$

admitting various types of data to be recorded, correspondingly, in the winter, spring, summer and autumn seasons.

It is also given a sequence $\boldsymbol{x}$ of meteorological data recorded for the last (say, n') several days in a current season of the year. It is desired to forecast possible sequence of weather conditions in the next n'' days.

The task can be solved in three basic steps:

1. Construction, on the basis of the relation H, of a relation Ξ describing admissible sequences of weather conditions for $n' + n''$ days in the given season of the year;
2. Finding a sub-relation $\chi \subset \Xi\,|_{U'}$ where U' denotes a sub-universe of formally possible sequences of weather conditions for n' days in the given season of the year; χ being reduced to the given single sequence $\boldsymbol{x}$.
3. Finding the sub-relation $\Psi = \Xi \underline{\cap} \chi$.

The way of construction of the relation Ξ (step 1) depends on the length $n' + n''$ with respect to N:

a. If $n' + n'' < N$ then Ξ is given as a projection $H|_{U' \cup U''}$;
b. If $n' + n'' = N$ then $\Xi = H$;
c. If $n' + n'' > N$ then Ξ should be given by an intersection of s, $s > N - (n' + n'')$, seasonal sub-relations of Ξ:

$$\Xi = \Xi^{(1)} \underline{\cap} \Xi^{(2)} \underline{\cap} \ldots \underline{\cap} \Xi^{(s)} \tag{7}$$

corresponding to shifted overlapping time-intervals covering $n' + n''$ days. The solution of the problem is not unique; it provides a set of possible extensions of the given sequence $\boldsymbol{x}$ of weather conditions satisfying the relation Ξ. Let us also remark that the above-presented prognostic model is strongly deterministic. In fact, description of meteorological processes needs more sophisticated, randomized models which are not a subject of this paper.

5 Solution of *Constructive RR* Problems

Construction of a relation satisfied by a given set of syndromes belongs to a large class of knowledge discovery problems (Maimon and Rokach 2005). However, the problem needs to be more exactly formulated. If U is a Cartesian product of n non-empty sets and X, $X \subset U$, denotes a finite subset of n-tuples $\boldsymbol{x}$ in U then X, by definition, constitutes a "minimal" solution of the problem while U constitutes its "maximal" solution. From a practical point of view, both solutions are useless. On the other hand, the problem of reconstruction of an assumed "hidden" relation governing some social, economical, political etc. events on the basis of recorded symptoms preceding some other, similar-type events in the past is of high importance. It is thus necessary to put some constraints on the *cRR* problem in order to make it non-trivial and more useful.

One of possible ways to do it consists in introducing into the universe U a *similarity measure*. Lot of various similarity measures have been proposed in the literature (see, e.g. Sobecki (2009)). However, not all of them satisfy the requirement of ability to be used to multi-aspect assessment of similarity of composed objects. For this purpose it is here proposed using a similarity measure defined generally as a function:

$$\sigma : U \times U \to [0, 1] \tag{8}$$

satisfying the conditions: for any $\boldsymbol{x}', \boldsymbol{x}'', \boldsymbol{x}''' \in U$ the following properties hold:

I. $\sigma(\boldsymbol{x}', \boldsymbol{x}') = 1$,
II. $\sigma(\boldsymbol{x}', \boldsymbol{x}'') = \sigma(\boldsymbol{x}'', \boldsymbol{x}')$,
III. $\sigma(\boldsymbol{x}', \boldsymbol{x}'') \cdot \sigma(\boldsymbol{x}'', \boldsymbol{x}''') \leq \sigma(\boldsymbol{x}', \boldsymbol{x}''')$.

A very important property of so-defined similarity measure consists in its extensibility: if σ_1, σ_2 are any two similarity measures satisfying the conditions I – III, then their product $\sigma = \sigma_1 \cdot \sigma_2$ also satisfies the conditions and as such it constitutes a similarity measure (Kulikowski 2001). Consequently, any power σ^μ for $\mu \geq 1$ satisfies the conditions I – III of similarity measure, and this makes us able to modify some similarity measures so as to make them less or more sensible to some parameters of the objects. Moreover, taking into account that the above-given properties may be satisfied by a large class of mathematical objects: natural or real numbers, vectors, Boolean variables, etc., a similarity measure can be constructed for syndromes composed of various features of objects.

For a given universe U, described in it similarity measure σ, a fixed number d, $0 < d \leq 1$, and any element $\boldsymbol{u} \in U$ the following relation can be defined:

"$\boldsymbol{x} \in S(\boldsymbol{u},d)$ ***if and only if*** $\sigma(\boldsymbol{u}, \boldsymbol{x}) \geq d$"

The relation $S(\boldsymbol{u},d)$ can be called a *similarity sphere of center **u** and range d.* This is an analogue of a sphere in a metric space (existence of any metric in the universe U has not been assumed). Remark that the lower is d, the larger is the similarity sphere. The concept of similarity sphere makes us able to propose a solution of the following problem:

6 The *CRR* Problem

Let it be given an universe U, a described in it similarity measure σ, a finite set G, $G \subset U$, of syndromes and a constant δ such that $\delta_{min} \leq \delta \leq 1$ where δ_{min} denotes minimal similarity between the syndromes in G. Find a minimal relation $\Xi\ (G, d)$ such that for any $\boldsymbol{x} \in G$ and any other $\boldsymbol{y} \in \Xi\ (G, d)$ the inequality $\sigma\ (\boldsymbol{x}, \boldsymbol{y}) \leq \delta$ is satisfied.

We call G a *germ* of the relation $\Xi\ (G, d)$. For solution of the *cRR* problem there will be taken into consideration all unordered pairs $(\boldsymbol{x}_i, \boldsymbol{x}_j)$ of different syndromes in G and their similarity measures $\delta_{i,j} = \sigma\ (\boldsymbol{x}_i, \boldsymbol{x}_j)$ will be calculated. Then for any $\boldsymbol{x}_i$ its minimal similarity to other syndromes in X:

$$d_i = min_{(j)}(\delta_{i,j}) \tag{9}$$

will be found. Let us take into consideration a relation:

$$\Xi = \underline{\cap}_i S(x, di) \tag{10}$$

It can be shown that so-defined relation Ξ is solution of the *cRR* problem for $\delta = \delta_{min} = min_{(i)}(d_i)$. Really, any $\boldsymbol{x} \in \Xi$ is a syndrome belonging to all similarity spheres of syndromes in *X*. The largest similarity spheres are $S(\boldsymbol{x}_*, \delta_{min})$ where $\boldsymbol{x}_*$ is one of (at least two) the less similar to any other syndrome in *X*. Therefore, if $\boldsymbol{x}$ belongs to $S(\boldsymbol{x}^*, \delta_{min})$ then the more its similarity to any other syndrome in *X* is not less than δ_{min}. Geometrically, this situation is illustrated in Fig. 2.

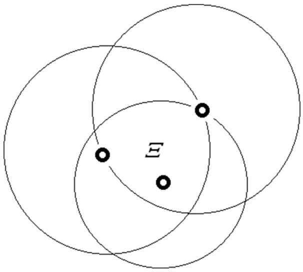

Fig. 2. Geometrical illustration of the solution of a *cRR* problem based on covering by similarity spheres.

The set *X* consists here of 3 syndromes denoted by **o**. The area of problem solution Ξ lies in the intersection of three similarity circles. Its maximal range is thus limited by the intersection of the two largest similarity circles.

The spherical covering method of *cRR* problem solution is a cautious one. In case of a large number of syndromes in *G* and small δ it may happen that the intersection of similarity spheres is empty and the solution of the problem does not exist. This in particular may happen in learning *cRR* problems when the elements of the germ are step-by-step acquired. Then, instead of a single germ we have an increasing sequence of germs:

$$G^{(1)} \subset G^{(2)} \subset \ldots G^{(t)} \subset \ldots \tag{11}$$

Let $G^{(1)}$ be a maximal germ that provides a non-empty solution $\Xi^{(1)}$ of the *cRR* problem. Let us assume that a next germ, $G^{(2)}$ is acquired. It can be presented as a sum:

$$G^{(2)} = H^{(2,1)} \cup \Delta^{(2)} \tag{12}$$

where $H^{(2,1)}$ denotes a subset of syndromes such that a modified germ:

$$G^{(1,2)} = G^{(1)} \cup H^{(2,1)} \tag{13}$$

provides a still non-empty solution $\Xi^{(1,2)}$ of the *cRR* problem, while $\Delta^{(2)}$ contains the remaining syndromes of $G^{(2)}$. Then $\Delta^{(2)}$ can be used as a germ to construction of the next non–empty sub-relation $\Xi^{(2)}$ (or a set of non-empty sub-relations) which can be added to $\Xi^{(1,2)}$ giving the second approximation of the solution:

$$\Psi^{(2)} = \Xi^{(1,2)} \underline{\cup} \Xi^{(2)} \tag{14}$$

Similar procedure can be continued in order to get next approximations of the solution of the learning *cRR* problem. The calculation complexity of the above-proposed method of the *cRR* problem solution is polynomial because each next syndrome acquired as a germ should be matched to the formerly acquired germs.

7 Conclusions

Relation recognition (*RR*) is a formal model of a large class of decision making problems. Its widely known example is any pattern recognition problem. However, the class of *RR* problems is much larger and it contains, in particular, the simple, extended, matching and constructive *RR* problems. Each type of the *RR* problems needs some specific methods of their solution. Both, the types of *RR* problems and the methods of their solution can be strongly described using the formalism of the extended algebra of relations. Also, the general concept of similarity measure are useful in some types of *RR* problems solution. Examples of such solution methods have been presented in the paper. However, they do not exhaust the *RR* area of investigation. In particular, the fuzzy *RR* problems (Rutkowski 2006), the problems following from using based on relations ontologies to describe real systems and processes (Abdoullayev 2008) and the problems of extension of the *RR* problems on the hyper-relation recognition problems (as it follows from the comparison of the Example 2 in this paper with a similar problem described in Kulikowski (2006) are worthy to be undertaken and deeply investigated.

Acknowledgement. This work was supported by the Nalecz Institute of Biocybernetics and Biomedical Engineering, Polish Academy of Sciences, in Warsaw.

References

Codd, C.J.: A relational model of data for large shared banks. Commun. ACM **13**, 377–387 (1970)

Bagui, S.: Extended entity relationship modeling. In: Rivero, L.C., et al. (eds.) Encyclopedia of Database Technologies and Applications, Idea Group Reference, Hershey, pp. 233–239 (2005)

Wirth, N.: Algorithms + Data Structures = Programs. Prentice Hall Inc., Englewood Cliffs (1976). (Polish translation WNT, Warsaw, 2002)

Kulikowski, J.L.: Rozpoznawanie obrazów. In: Nałęcz, M. (ed.) Biocybernetyka i Inżynieria Biomedyczna 2000, vol. 8, Obrazowanie Biomedyczne, AOW EXIT, Warsaw, Chap. 7, pp. 193–237 (2003). (in Polish)

Kulikowski, J.L.: Structural analysis of biomedical data. Biocybernetics Biomed. Eng. **7**(1–4), 143–154 (1987)

Kulikowski, J.L.: From pattern recognition to image interpretation. Biocybernetics Biomed. Eng. **22**(2–3), 177–197 (2002)

Kulikowski, J.L.: Relational approach to structural analysis of images. Mach. Graph. Vis. **1**(1/2), 299–309 (1992)

Rudeanu, S.: Sets and Ordered Structures, Sect. 4.2, pp. 121–138. Bentham Science Publishers, Oak Park (2012)

Wang, H., Kläser, A., Schmid, C., Liu, C.-L.: Dense trajectories and monitoring boundary descriptors for action recognition. Int. J. Comput. Vis. **103**, 594 (2013). SpringerLink

Maimon, O., Rokach, L.: The Data Mining and Knowledge Discovery Handbook. Springer, New York (2005)

Sobecki, J.: Rekomendacja Interfejsu Użytkownika w Adaptacyjnych Webowych Systemach Informacyjnych. Of. Wyd. Politechniki Wrocławskiej, Wrocław (2009). (in Polish)

Kulikowski, J.L.: Pattern recognition based on ambiguous indications of experts. In: Kurzyński, M. (ed.) KomPuterowe Systemy Rozpoznawania KOSYR 2001, pp. 15–22. Wyd. Politechniki Wrocławskiej, Wroclaw (2001)

Rutkowski, L.: Metody i Techniki Sztucznej Inteligencji, Sect. 4.7, pp. 90–94. Wyd. Naukowe PWN, Warszawa (2006). (in Polish)

Abdoullayev, A.: Reality, Universal Ontology, and Knowledge Systems, Toward the Intelligent World. IGI Publishing, Hershey (2008)

Kulikowski, J.L.: Description of irregular composite objects by hyper-relations. In: Proceedings of the International Conference on ICCVG 2004, Computer Vision and Graphics, Warsaw, Poland, September 2004, Springer CIaV vol. 32, Dordrecht, pp. 141–146 (2006)