Spatio-Temporal Approximate Reasoning over Hierarchical Information Maps

Andrzej Skowron[1], Piotr Synak[2], James Peters[3]

[1] Institute of Mathematics, Warsaw University
Banacha 2, 02-097 Warsaw, Poland
skowron@mimuw.edu.pl
[2] Polish-Japanese Institute of Information Technology
Koszykowa 86, 02-008 Warsaw, Poland
synak@pjwstk.edu.pl
[3] Department of Electrical and Computer Engineering
University of Manitoba Winnipeg, Manitoba R3T 5V6, Canada
jfpeters@ee.umanitoba.ca

Abstract. In the paper, we discuss the problems of spatio-temporal reasoning in the context of hierarchical information maps and approximate reasoning networks (AR networks). Hierarchical information maps are used for representations of domain knowledge about objects, their parts, and their dynamical changes. AR networks are patterns constructed over sensory measurements and they are discovered from hierarchical information maps and experimental data. They make it possible to approximate domain knowledge, i.e., complex spatio-temporal concepts and reasonings represented in hierarchical information maps.

Keywords: complex objects, concept approximation, spatio-temporal reasoning, information maps, AR networks

1 Introduction

One of the forms of data representation is an information system, where each investigated object is described by means of some attributes (features). Once some reflexive binary relation on a set of objects is given (e.g., relation of neighbourhood) one can consider new information systems with more complex objects that are clusters (clumps) of objects determined by this relation. In this case the attributes reflect some more general properties of objects, i.e., properties of sets of objects. This approach is typical for time series analysis, where attributes (features) are defined on the basis of relevant windows [2, 1, 16]. The chosen neighbourhoods and their properties should make it possible to induce the high quality approximations of a given concept. Observe that there are two problems in this approach: discovery of relevant neighbourhoods of objects and their properties. These are key problems of spatio-temporal data mining [2, 3].

In this paper, we extend this approach to the case of information maps and hierarchical information maps, where unstructured objects are substituted by more complex information granules corresponding to structural objects evolving in time. The paper is a continuation of [22, 23, 10].

The basic assumption we make is that the modelling process leading to hierarchical information maps is performed using soft concepts from a given domain knowledge language. The aim is to approximate these concepts and reasonings represented in hierarchical information maps using patterns, called AR networks. These networks represent properties of dynamic structures of complex objects and their parts over time. We assume that AR networks are induced from sensory (elementary, measurable) patterns using the hierarchical information maps and experimental data. Discovered AR networks create a knowledge base of relevant spatio-temporal patterns from which approximations of complex concepts from domain knowledge can be induced and approximate reasoning about them can be performed.

Any AR network is constructed out of production rules, i.e., dependencies with premises and conclusions consisting of definable patterns (from sensory patterns) labelled by some satisfiability degrees to which these patterns are included into the concepts. A given AR network is satisfiable in a given state if each pattern occurring in the network is satisfied at least to a degree specified by the labelling.

We also emphasise that information granulation, in passing from a lower level of hierarchy to a higher one, may be performed, e.g., by indiscernibility or similarity relation. Hierarchical information maps make it possible to model information granules relevant for the target tasks by taking into account functionality that the information granules should posses.

2 Preliminaries

In the paper, we use the notation of rough set theory [8, 4]. In particular, by $\mathbb{A} = (U, A)$ we denote an *information system* [7, 9] with the universe U of *objects* and the attribute set A. Each *attribute* $a \in A$ is a function $a : U \to V_a$, where V_a is the *value set* of a. For a given set of attributes $B \subseteq A$, we define the *indiscernibility relation* $IND(B)$ on the universe U that partitions U into classes of indiscernible objects. We say that objects x and y are *indiscernible* with respect to B if and only if $a(x) = a(y)$ for each $a \in B$.

Decision tables are denoted by $\mathbb{A} = (U, A, d)$, where $d \notin A$ is the *decision* attribute. The decision attribute d defines partition of the universe U into *decision classes*. An object x is *inconsistent* if there exists an object y such that $xIND(A)y$, but x and y belong to different decision classes, i.e., $d(x) \neq d(y)$. *Positive region* of a decision table $\mathbb{A}$ (denoted by $POS(\mathbb{A})$) is the set of all consistent objects.

Any pair $(\mathbb{A}, \mathbb{R})$, where $\mathbb{A} = (U, A, d)$ is a decision table and $\mathbb{R}$ is a set of binary and reflexive relations over $U \times U$, is called a *relational decision table*. For any $R \in \mathbb{R}$ by $R(x)$ we denote the *neighbourhood* of an object x, i.e., the set $\{y \in U : xRy\}$. One can consider a new decision table $\mathbb{A}_R = (U_R, A_R, d_R)$ obtained from $(\mathbb{A}, \mathbb{R})$, where $U_R = \{(x, R(x)) : x \in U\}$ is a family of object neighbourhoods, A_R is a set of attributes describing properties of objects and their neighbourhoods, and, e.g., $d_R((x, R(x))) = d(x)$. In this way one can consider attributes whose values depend on the context in which objects occur, i.e., on neighbourhoods of objects rather than on objects only. This approach is typical for time series analysis, where attributes (features) are defined on the basis of relevant windows [2, 1, 16]. It is also used in a multi-criteria decision making (see, e.g., [24]). The chosen neighbourhoods and their properties should make it possible to induce the high quality approximations of given target concepts. Observe that there are two problems in this approach: discovery of relevant neighbourhoods of objects and properties of such neighbourhoods defined by means of some new attributes. The former problem is related to the selection of $\mathbb{R}$ as well as $R \in \mathbb{R}$ for any object while the latter is based on discovery of a relevant language of formulas expressing properties of neighbourhoods and next on selection of relevant formulas from this language. Discovery of relevant neighbourhoods and their properties for proper object approximation is a key problem of spatio-temporal data mining [3]. From such a decision table there can be derived concept approximation classifiers by using strategies developed in rough sets or other areas like machine learning, information theory, or pattern recognition.

3 Information Maps

In this section we recall some basic concepts related to information maps. We also include some illustrative examples.

3.1 Basic definitions

Let us recall the definition of information maps [21, 23]. Such maps are usually generated from experimental data (e.g., information systems or decision tables) and are defined by some binary (transition) relations on the set of states. Any state consists of information label and the corresponding information extracted from a given data set. This kind of structures are basic models over which one can search for relevant patterns for many data mining problems [21, 23].

An *information map* $\mathcal{A}$ is a quadruple

$$\mathcal{A} = (E, \leq, I, f), \tag{1}$$

where E is a finite set of *information labels*, $\leq \subseteq E \times E$ is a binary *transition relation* on information labels, I is an *information set* and $f : E \to I$ is an *information function* associating any information label with the corresponding information.

In Figure 1a, we present an example of information map, where $E = \{e_1, e_2, e_3, e_4, e_5\}$, $I = \{f(e_1), f(e_2), f(e_3), f(e_4), f(e_5)\}$, and the transition relation $\leq$ is a partial order on E.

A *state* is any pair $(e, f(e))$, where $e \in E$. The set $\{(e, f(e)) : e \in E\}$ of all states of $\mathcal{A}$ is denoted by $S_{\mathcal{A}}$. The transition relation on information labels can

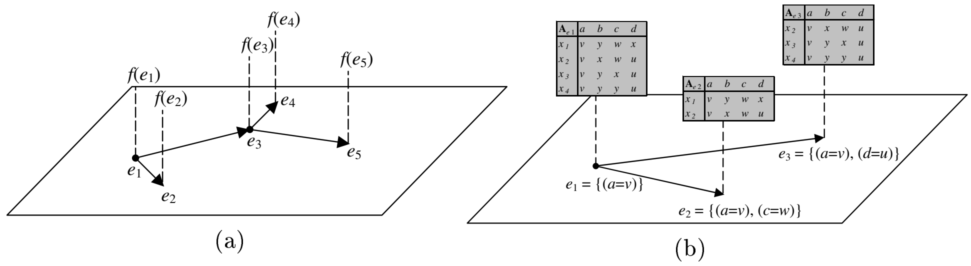

Fig. 1. (a) An information map; (b) An information map of an information system

be extended to the relation on states, e.g., in the following way: $(e_1, i_1) \leq (e_2, i_2)$ if and only if $e_1 \leq e_2$. A *path* in $\mathcal{A}$ is any sequence $s_0 s_1 s_2 \ldots$ of states such that for every $i \geq 0$:

1. $s_i \leq s_{i+1}$,
2. if $s_i \leq s \leq s_{i+1}$ then $s = s_i$ or $s = s_{i+1}$.

We say that a state s is *reachable* from a state s_0 if and only if there exists a path $s_1 s_2 \ldots s_n$ such that $s_1 = s_0$ and $s_n = s$. A *property* of $\mathcal{A}$ is any subset of $S_{\mathcal{A}}$. Let F be a set of temporal formulas. We say that the property φ is *expressible* in F if and only if $\varphi = \|\alpha\|$ for some $\alpha \in F$, where $\|\alpha\|$ is the semantics of α.

3.2 Information maps of data tables

Any information system $\mathbb{A} = (U, A)$ defines its information map as a graph consisting of nodes that are elementary patterns generated by $\mathbb{A}$, where an *elementary pattern* (or *information signature*) $Inf_B(x)$ is a set $\{(a, a(x)) : a \in B\}$ of attribute-value pairs over $B \subseteq A$ consistent with a given object $x \in U$. Thus, the set of labels E is equal to the set $INF(A) = \{Inf_B(x) : x \in U, B \subseteq A\}$ of all elementary patterns of $\mathbb{A}$. The relation $\leq$ is defined in a straightforward way, i.e., for $e_1, e_2 \in INF(A)$, $e_1 \leq e_2$ if and only if $e_1 \subseteq e_2$. Hence, relation $\leq$ is a partial order on E. Finally, the information set I is equal to $\{\mathbb{A}_e : e \in INF(A)\}$, where $\mathbb{A}_e$ is a sub-system of $\mathbb{A}$ with the universe U_e equal to the set $\{x \in U : \forall (a, t) \in e \; a(x) = t\}$. Attributes in $\mathbb{A}_e$ are attributes from $\mathbb{A}$ restricted to U_e. The information function f mapping $INF(A)$ into I is defined by $f(e) = \mathbb{A}_e$ for any $e \in INF(A)$ (see Figure 1b).

One can investigate several properties of such a system, e.g., related to the distribution of values of some attribute. Let α be a formula, such that $(e, \mathbb{A}_e) \models \alpha$ has the following intended meaning: "at least 75% of objects of system $\mathbb{A}_e$ has value u on attribute d." In our example $\|\alpha\|_{\mathbb{A}} = \{(e_1, \mathbb{A}_{e_1}), (e_3, \mathbb{A}_{e_3})\}$ [23].

One can consider other information functions for information maps over data tables. Such a function can be a kind of "view" of dependencies in the data table. Then, for example, $f(e)$ can be equal to the set of all dependencies in $\mathbb{A}_e$ that have sufficient support and confidence.

3.3 Decision tables over information maps

One of the typical schemes of object classification is based on the analysis of decision tables. From the given information about an object (object pattern), we try to classify it relative to a proper decision class. In many cases this scheme needs to be extended because together with the information given there has to be also considered its context. That means that instead of a single information signature, related to the investigated object x, we also have to examine some other objects that are in some relation to x. Properties of those objects can be important in order to extend information about x by information about the context in which x occurs. In a more complex case we can consider states of objects and relations between such states. Temporal relations between states, in the case of objects changing in time, provide another possible source of information about the context in which objects occur.

Thus, the scheme of object classification can be as follows. We are given a decision table. Next, it can be extended by some relations on objects (or values of attributes) to a relational decision table defining some neighbourhoods of objects (possibly overlapping each other). Thus, we construct a new decision table, where objects are pairs (*object*, *object_neighbourhood*), and attributes describe properties of the objects in the context of their neighbourhoods.

In the case of information maps, the above idea is generalised to more complex information granules that are pairs (*state*, *state_neighbourhood*), where *state* is a state of a given information map $\mathcal{A}$ and *state_neighbourhood* is the neighbourhood of this state in $\mathcal{A}$. A state can be identified by some information about an object and it determines some set of objects (a sub-table), e.g., set of objects indiscernible by means of some attributes. Thus, *state_neighbourhood* is a much more complex structure than *object_neighbourhood* in the previous case, because it is a set (defined by transition relation) of sub-tables satisfying some constraints. Also the attributes of the constructed decision table are more complex because they express properties of complex neighbourhoods. The decision attribute is complex too, because it classifies a state, that is a complex object (in our example – a sub-table). Thus, for a given state s, as the value of decision for s we can consider, e.g., the distribution of objects corresponding to s in decision classes.

4 The Structure of Static Complex Objects

Properties of the structured complex objects can be approximated by means of approximate reasoning schemes (AR schemes) [20, 13, 25]. Such schemes usually have a tree structure with the root labelled by the satisfiability degree of some feature by a complex object and leaves labelled by the satisfiability degrees of some other features by primitive objects (i.e., the most simple parts of a complex object). An AR scheme can have many levels. Then, from properties of basic parts we conclude about properties of more complex parts, and after some levels, about properties of the complex target object.

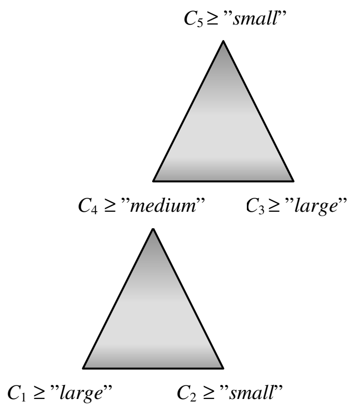

Fig. 2. A complex approximate reasoning rule built from concepts $C_1, \ldots, C_5$

Using carefully constructed derivations based on approximate reasoning rules one can estimate degrees of inclusion of objects into complex concepts on the basis of their elementary parts – input objects. Such derivations create AR schemes [12, 26, 20].

Example 1. Let $C_1, \ldots, C_5$ be some concepts. Let also an AR scheme be given like one in Figure 2 (a very important problem is related to developing efficient methods for inducing this kind of schemes from data). This scheme is a combination of two simpler production rules with the following interpretation: (1) if the inclusion degree in concept C_1 is at least *large* and in concept C_2 at least *small*, then the inclusion degree in concept C_4 is at least *medium*; (2) if the inclusion degree in concept C_4 is at least *medium* and in concept C_3 at least *large*, then the inclusion degree in concept C_5 is at least *small*. Observe that any AR scheme defines an input-output production. Its premises are composed from all premises of rules from AR scheme not used by other rules in the scheme, and the conclusion is at the root of the AR scheme.

The analysis of structural objects is very important because such objects can be found in many applications. Several problems for this kind of analysis can be formulated.

Thus, to reason about the structure of a complex object, we have to use composition schemes (one- or many-level) of its parts. This kind of structure should also make it possible to reason about inclusion degrees of the complex object in higher level concepts from inclusion degrees of its parts in lower level concepts, assuming that these parts satisfy some additional constraints (e.g., expressing closeness of parts in a considered space). These problems are considered in rough-mereological approach [14, 15].

However, in many cases of spatio-temporal reasoning it is necessary to perform reasoning about dynamically changing objects. Moreover, it can be necessary to use feedback from the root of an AR scheme to its input – the information about satisfiability degree of some feature by a complex object can be used as primitive feature in next iterations. Hence, we propose to use so called AR networks for this kind of reasoning.

5 The Structure of Dynamic Complex Objects

For our further discussion it is important to note that we describe states by some of their properties. Hence, we identify states with collections of objects defined by these properties rather than with single objects. The analogous assumption we make about parts of objects.

Another class of problems arises when a given object evaluates in time, what is measured by changes in time of some of its features [16]. Then, we consider different states of an object in different time points and transitions between states. In the case of a complex object, its structure is perceived in each state. This structure can evaluate from state to state what can be expressed in different ways. In the simplest case we observe changes of inclusion degrees in patterns (formulas) describing features of parts. In a more complex case, the structure itself can also change and some parts can be replaced by other ones. The language in which features are expressed can also evaluate, either in relation to the whole complex object, or to some of its parts. Problems related to this subject are widely studied in the literature (see, e.g., [17, 16]).

In our approach we assume that AR networks are discovered from data and domain knowledge. They make it possible to approximate (in a language definable in terms of attributes available from data, e.g., sensor measurements) reasonings about spatio-temporal properties of objects performed by an expert in his/her language (e.g., a simple fragment of natural language). The methods for discovery of AR schemes is discussed in [20, 5]. We are developing algorithmic methods for discovery of AR networks from data and a domain knowledge. This domain knowledge is guiding searching procedures aiming at discovery of AR networks. Observe, that AR networks, analogously to the case of AR schemes, are clusters of exact constructions of objects.

5.1 Modelling object changes

To model complex object changes in time, we use hierarchical information maps. AR networks, i.e., spatio-temporal AR schemes make it possible to approximate domain knowledge represented in information maps. They differ from the static case of AR schemes in that they are constructed also along the time dimension. Therefore, rules used in their construction link some spatio-temporal patterns describing spatio-temporal properties of objects or their parts. Moreover, premises can consist of some conditions expressing time constraints.

Observe that any AR network determines some pattern used for approximation of a given concept. To obtain good approximation several patterns (AR networks) should be discovered and next fused to create a classifier for the complex concept.

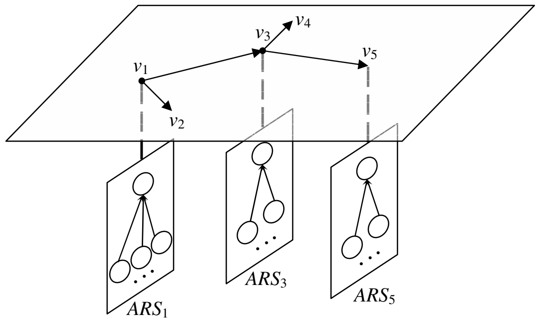

Fig. 3. An example of a simple AR network

We are going to present hierarchical information maps and AR networks.

In the presented case the information maps model different states of a complex object. Each state of the map is labelled by some formula specifying state of the object. The information corresponding to the given label is the minimal satisfaction degree of the formula. This degree can be interpreted as minimal one to identify (to a satisfactory degree) the situation observed with the state of an object. Let us note, that both formula and corresponding degree can be specified by an expert using soft concepts and, therefore, we have to use some approximation schemes.

Figure 3 illustrates a fragment of a simple AR network. A graph on a plane represents states of a complex object by means of some concepts labelled by minimal satisfaction degrees. With each node there is associated an AR scheme used to reasoning about the degree of matching of the complex concept by a complex object using information about satisfaction degrees of lower-level concepts by elementary (sensory) objects. Nodes are linked by edges labelled by temporal relations which can depend not only on time but also on the context in which objects appear. In the simplest case it can be a consequence relation.

One of the schemes of reasoning is the following one. A hierarchical information map represents the domain knowledge about different states of some complex object. Each state is identified by some property of the object, i.e., some concept. To measure degree of satisfaction of a concept the AR schemes are used. Suppose we have the measurements (e.g., obtained from some sensors) of the current (observed) situation. We test each AR scheme to determine a concept matched the best in order to identify the state of the complex object. Next, from the hierarchical information map, we conclude what are the possible next states, what gives the opportunity to undertake some action.

Let us note that AR network from Figure 3 obtained by a combination of AR schemes with a hierarchical information map can be treated also as an information map, where the information corresponding to each label (state of complex object) is the AR scheme.

In general, the information given about states can be broader. It might be based on some historical knowledge (training data) [16, 3]. Each state can be additionally described by some information system, where each attribute corresponds to some of the state's properties. In particular, an attribute can reflect satisfaction of some properties in the near past. Thus, given state can be described not only by spatial patterns but also by spatio-temporal ones.

In a more general case of hierarchical information maps each part of a complex object may have its own space of states together with corresponding transitions. Thus, we have a hierarchical structure with several graphs at each level - one graph corresponds to one part. The edges of these graphs are again labelled with some temporal relations, however, they are defined for particular parts. The lowest level corresponds to elementary (atomic) parts. The nodes of graphs from adjacent levels can be connected by some spatial relation defining scheme of constructing more complex object in a given state from its parts (that are also in some states). An example is presented in Figure 4. A complex object in state v_1 consists of two parts that are in states x_1 and y_1. The same object in state v_3 has three parts in states x_3, y_2, and z_2, respectively. With each non-atomic part x in some state x_i at any level, we can associate a decision table containing, e.g., information about historical observations of x in x_i. The rows (objects) of such a system correspond to different observations. The decision reflects the degree of satisfaction of a concept corresponding to x_i by x. The conditional attributes describe satisfaction of some concepts and some other relations by parts of x. From these information systems we can induce rules of reasoning about degrees of satisfaction of concepts from the upper level on the basis of such an information from a lower level. This is necessary since we usually cannot directly measure whether a complex object matches a complex concept. We can only measure degrees of satisfaction of lower level concepts by elementary objects.

This kind of construction corresponds to multi-level hierarchical information maps consisting of several information maps that are linked together by some relations on the sets of states. It is important to note that in modelling of such maps we express properties of states and relations between them using the language of domain knowledge (e.g., a simplified natural language). Next, using hierarchical information maps and experimental data we are searching for AR networks, representing relevant patterns for approximation of complex concepts that appear on different levels of maps. Such AR networks are constructed along the derivations performed in domain knowledge using the representation in hierarchical information maps.

In the most general case there can be also given some other relations defined between parts from the same level, e.g., spatial or temporal, reflecting some constraints that parts should satisfy in given states in order to reason about more complex object (see Figure 4).

Especially interesting in modelling of object changes are rules that describe how changes of some features (attributes) influence changes of some other ones. Let us consider an example related to information maps. Assume that with any label e it is associated an information $f(e)$ that is a pair $(T_1(e), T_2(e))$ of theories

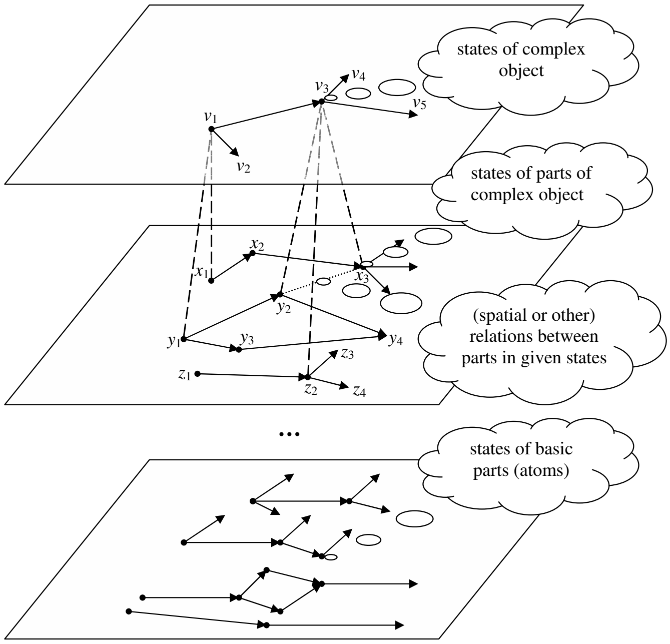

Fig. 4. An example of hierarchical information map

representing some view on knowledge represented in $\mathbb{A}_e$ consisting of the set of dependencies between conditional and decision attributes in the data table $\mathbb{A}_e$, respectively. Such a view can consists of association rules with sufficient support and confidence. Assume that e' is another label (e.g., an extension of e). Then one can consider rules making it possible to predict differences between $T_2(e)$ and $T_2(e')$ on the basis of differences between $T_1(e)$ and $T_1(e')$. Such rules are interesting on different levels of hierarchical modelling for spatio-temporal reasoning. Moreover, the laws for predicting changes in decisions quite often require to discover the relevant trends of conditional attribute changes (e.g., over some period of time) from data. We plan to develop algorithmic tools for discovery of such laws (dependencies) supported by hierarchical modelling. Observe that in searching for these laws one should, in particular, discover relevant "views" of sets of dependencies and measures of differences.

5.2 Constructing higher levels of hierarchical maps by information granulation

In this section we discuss an important role which the relational structure granulation [20, 11] plays in searching for relevant patterns in approximate reasoning, e.g., in searching for relevant approximation patterns (see Figure 5). For any object x there is defined a neighbourhood $I(x)$ specified by the value of the uncertainty function from an approximation space (see [18]). From these neighbourhoods some other, more relevant ones (e.g., for the considered concept approximation), should be found. Such neighbourhoods can be extracted by searching in a space of neighbourhoods generated from values of uncertainty function by applying to them some operations like generalisation operations, set theoretical operations (union, intersection), clustering and operations on neighbourhoods defined by functions and relations in the underlying relational structure.[1] Figure 5 illustrates an exemplary scheme of searching for neighbourhoods (patterns, clusters) relevant for concept approximation. In this example f denotes a function with two arguments from the underlying relational structure. Due to the uncertainty, we cannot perceive objects exactly but only by using available neighbourhoods defined by the uncertainty function from an approximation space. Hence, instead of the value $f(x, y)$ for a given pair of objects (x, y), one should consider a family of neighbourhoods $\mathcal{F} = \{I(f(x', y')) : (x', y') \in I(x) \times I(y)\}$. From this family $\mathcal{F}$, a subfamily $\mathcal{F}'$ of neighbourhoods can be chosen which consists of neighbourhoods with some properties relevant for approximation. Next, a subfamily $\mathcal{F}'$ can be, e.g., generalised to clusters that are relevant for the concept approximation, i.e., clusters sufficiently included into the approximated concept (see Figure 5). The inclusion degrees can be measured by granulation of the inclusion function from the relational structure. Using information granu-

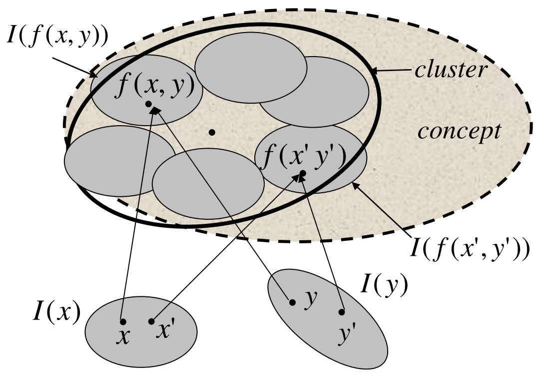

Fig. 5. Relational structure granulation

lation one can construct from a given information map a new one at the higher level which is simpler (more compact) but still sufficient for approximation of complex concepts with a satisfactory quality.

[1] Relations from such a structure may define relations between objects or their parts.

6 Conclusions

In the paper, we have discussed some problems related to hierarchical approximation of spatio-temporal knowledge by means of hierarchical information maps and AR networks. Hierarchical information maps can help to discover AR networks representing relevant spatio-temporal patterns from data and soft domain knowledge.

Several aspects of hierarchical information maps and AR networks were discussed, in particular, representation of spatio-temporal knowledge, its approximation by AR networks and the role of information granulation in constructing of AR networks from data and hierarchical information maps representing soft domain knowledge.

We have emphasised the need to develop laws describing trends of changing of decisions in terms of trends of changing of conditional attributes on different levels of hierarchical modelling.

In our further work we would like to develop a software system making it possible to model in the form of hierarchical information maps the domain knowledge and to discover from them and experimental data AR networks relevant for complex concept approximation. In particular, it will be necessary to extend methods for AR schemes synthesis (see, e.g., [19, 6]) to methods for AR networks synthesis.

Acknowledgements. The research has been supported by the grant 3 T11C 002 26 from Ministry of Scientific Research and Information Technology of the Republic of Poland.

References

1. Juan P. Caraa-Valente and Ignacio Lpez-Chavarras. Discovering similar patterns in time series. In Raghu Ramakrishnan, Sal Stolfo, Roberto Bayardo, and Ismail Parsa, editors, *Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD*, pages 497–505, Boston, MA, August 20-23 2000. ACM Press.
2. Gautam Das, King-Ip Lin, Heikki Mannila, Gopal Renganathan, and Padhraic Smyth. Rule discovery from time series. In Rakesh Agrawal, Paul E. Stolorz, and Gregory Piatetsky-Shapiro, editors, *Fourth International Conference on Knowledge Discovery and Data Mining KDD*, pages 16–22, New York, NY, August 27-31 1998. AAAI Press.
3. Will Kloesgen and Jan Żytkow, editors. *Handbook of Knowledge Discovery and Data Mining*. Oxford University Press, 2002.
4. Jan Komorowski, Lech Polkowski, and Andrzej Skowron. Rough sets: A tutorial. In Sankar K. Pal and Andrzej Skowron, editors, *Rough Fuzzy Hybridization: A New Trend in Decision-Making*, pages 3–98. Springer-Verlag, Singapore, 1999.
5. Sinh Hoa Nguyen, Jan Bazan, Hung Son Nguyen, and Andrzej Skowron. Layered learning for concept synthesis. *LNCS Transactions on Rough Sets*, 1 (LNCS 3100):193–214, 2004.
6. Sankar K. Pal, Lech Polkowski, and Andrzej Skowron, editors. *Rough-Neural Computing: Techniques for Computing with Words*. Cognitive Technologies. Springer Verlag, Heidelberg, Germany, 2004.

7. Zdzisław Pawlak. Information systems - theoretical foundations. *Information Systems*, 6:205–218, 1981.
8. Zdzisław Pawlak. Rough sets. *International Journal of Computer and Information Sciences*, 11:341–356, 1982.
9. Zdzisław Pawlak. *Rough Sets: Theoretical Aspects of Reasoning about Data*, volume 9 of *D: System Theory, Knowledge Engineering and Problem Solving*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 1991.
10. James F. Peters, Andrzej Skowron, Jarosław Stepaniuk, and Sheela Ramanna. Towards an ontology of approximate reason. *Fundamenta Informaticae*, 51(1-2):157–173, 2002.
11. James F. Peters, Andrzej Skowron, Piotr Synak, and Sheela Ramanna. Rough sets and information granulation. In T.B. Bilgic, D. Baets, and O. Kaynak, editors, *Tenth International Fuzzy Systems Association World Congress IFSA*, volume 2715 of *Lecture Notes in Artificial Intelligence*, pages 370–377, Istanbul, Turkey, June 30 - July 2 2003. Springer-Verlag.
12. Tomaso Poggio and S. Smale. The mathematics of learning: Dealing with data. *Notices of the AMS*, 50(5):537–544, 2003.
13. Lech Polkowski and Andrzej Skowron. Rough mereological approach to knowledge-based distributed AI. In J. K. Lee, J. Liebowitz, and J. M. Chae, editors, *Third World Congress on Expert Systems*, pages 774–781, Soeul, Korea, February 5-9 1996. Cognizant Communication Corporation.
14. Lech Polkowski and Andrzej Skowron. Rough mereology: A new paradigm for approximate reasoning. *International Journal of Approximate Reasoning*, 15(4):333–365, 1996.
15. Lech Polkowski and Andrzej Skowron. Rough mereology in information systems. A case study: Qualitative spatial reasoning. In Lech Polkowski, Tsau Young Lin, and Shusaku Tsumoto, editors, *Rough Set Methods and Applications: New Developments in Knowledge Discovery in Information Systems*, volume 56 of *Studies in Fuzziness and Soft Computing*, chapter 3, pages 89–135. Springer-Verlag/Physica-Verlag, Heidelberg, Germany, 2000.
16. John F. Roddick, Kathleen Hornsby, and Myra Spiliopoulou. YABTSSTDMR - yet another bibliography of temporal, spatial and spatio-temporal data mining research. In K. P. Unnikrishnan and R. Uthurusamy, editors, *SIGKDD Temporal Data Mining Workshop*, pages 167–175, San Francisco, CA, 2001. ACM Press.
17. Erik Sandewall, editor. *Features and Fluents: The Representation of Knowledge About Dynamical Systems*, volume 1. Oxford University Press, 1994.
18. Andrzej Skowron and Jarosław Stepaniuk. Tolerance approximation spaces. *Fundamenta Informaticae*, 27(2-3):245–253, 1996.
19. Andrzej Skowron and Jarosław Stepaniuk. Information granule decomposition. *Fundamenta Informaticae*, 47(3-4):337–350, 2001.
20. Andrzej Skowron and Jarosław Stepaniuk. Information granules and rough-neural computing. In Pal et al. [6], pages 43–84.
21. Andrzej Skowron and Piotr Synak. Patterns in information maps. In James J. Alpigini, James F. Peters, Andrzej Skowron, and Ning Zhong, editors, *Third International Conference on Rough Sets and Current Trends in Computing RSCTC*, volume 2475 of *Lecture Notes in Artificial Intelligence*, pages 453–460, Malvern, PA, October 14-16 2002. Springer-Verlag.
22. Andrzej Skowron and Piotr Synak. Complex patterns. *Fundamenta Informaticae*, 60(1-4):351–366, 2004.
23. Andrzej Skowron and Piotr Synak. Reasoning in information maps. *Fundamenta Informaticae*, 59(2-3):241–259, 2004.

24. Roman Słowiński, Salvatore Greco, and Benedetto Matarazzo. Rough set analysis of preference-ordered data. In James J. Alpigini, James F. Peters, Andrzej Skowron, and Ning Zhong, editors, *Third International Conference on Rough Sets and Current Trends in Computing RSCTC*, volume 2475 of *Lecture Notes in Artificial Intelligence*, pages 44–59, Malvern, PA, October 14-16 2002. Springer-Verlag.
25. Jarosław Stepaniuk. Approximation spaces, reducts and representatives. In Lech Polkowski and Andrzej Skowron, editors, *Rough Sets in Knowledge Discovery 2. Applications, Case Studies and Software Systems*, volume 19 of *Studies in Fuzziness and Soft Computing*, chapter 6, pages 109–126. Physica-Verlag, Heidelberg, Germany, 1998.
26. Peter Stone. *Layered Learning in Multi-Agent Systems: A Winning Approach to Robotic Soccer*. The MIT Press, Cambridge, MA, 2000.