Fundamenta Informaticae 60 (2004) 351–366
IOS Press

Complex Patterns

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

Piotr Synak
Polish-Japanese Institute of Information Technology
Koszykowa 86, 02-008 Warsaw, Poland
synak@pjwstk.edu.pl

Abstract. We outline some results of our current research on developing a methodology for solving problems of spatio-temporal reasoning. We consider classifiers for complex concepts in spatio-temporal reasoning that are constructed hierarchically. We emphasise the fact that the construction of such hierarchical classifiers should be supported by domain knowledge. Approximate reasoning networks (AR networks) are proposed for approximation of reasoning schemes expressed in natural language. Such reasoning schemes are extracted from knowledge bases representing domain knowledge. This approach makes it possible to induce classifiers for complex concepts by constructing them along schemes of reasoning extracted from domain knowledge.

Keywords: complex object, concept approximation, pattern, rough sets, spatio-temporal reasoning, classifier, network of classifiers

1. Introduction

The existing approaches are not sufficient for solving many real-life problems of spatio-temporal reasoning (e.g., related to autonomous systems). It is not possible to induce concept approximations directly

[*] Address for correspondence: Institute of Mathematics, Warsaw University, Banacha 2, 02-097 Warsaw, Poland

from concepts close to measurements (see, e.g., [40]). Our main goal is to show that to solve such problems some non-conventional approaches to discovery of spatio-temporal patterns relevant for complex concept approximation should be developed.

We outline some results of our current research on developing a methodology for solving such problems. We extend the existing approach (see, e.g., [26, 36, 30]) based on the assumption that together with experimental data, soft domain knowledge is also given. Such domain knowledge consists of domain ontologies including concepts expressed in natural language and dependencies between them. One can construct the domain reasoning schemes derived from such dependencies linking concepts close to measurements with complex target concepts. The domain reasoning schemes are used as hints in searching for approximate reasoning networks (AR networks) representing construction of complex patterns for complex target concepts from some simpler patterns and dependencies between them. AR networks make it possible to carry out reasoning along the domain reasoning schemes leading to conclusions sufficiently close to the conclusions of such schemes. Several results have been reported on applications of AR schemes ([18, 20, 22, 26, 30]), that are special cases of AR networks, for reasoning about complex objects. AR networks used for spatio-temporal reasoning represent patterns (or clumps of such patterns) included to a satisfactory degree in concepts from domain knowledge describing properties of spatio-temporal processes. An example of such concept can be: *On the road car B is overtaking car A.*

In the paper we discuss the foundations for complex pattern extraction from temporal data and soft domain knowledge. Such patterns are next used in spatio-temporal reasoning. Perception of parts and their properties in the case of spatio-temporal reasoning is a much more complex process than in case of static spatial objects. This follows from, for example, the necessity to learn rules expressing relevant changes of object structures in time. In our approach we extend relations defining spatial object structures to spatio-temporal relations. Approximate reasoning schemes (AR schemes) (see, e.g., [20, 21, 26, 29, 30]) are patterns expressing properties of static objects by means of properties of their parts. Such patterns are assumed to be discovered from data and domain knowledge. Analogously, we introduce AR schemes at time dimension, called spatio-temporal AR networks (AR networks, for short). This makes it necessary to introduce relevant relations for expressing time relations. For example, AR networks can represent patterns: (i) defining (to a degree) soft properties of parts of objects or the whole objects in a given moment (or period) of time by means of degrees to which soft properties of some other parts of objects observed in the past are satisfied; (ii) used for prediction of properties of the whole object in the near future from properties of the behaviour of the object and its parts in the past; (iii) defining dynamic properties of objects or their parts in time. For modelling of temporal patterns the recurrent networks of patterns can be used rather than acyclic AR schemes.

In the following sections we discuss some problems related to induction of AR networks from experimental data and domain knowledge representing spatio-temporal patterns (or their clusters) that are relevant for approximation of complex spatio-temporal concepts [5, 23, 34].

2. Outline of the approach

Let us describe the main idea in more detail. We consider hierarchical classifiers for complex concepts that are constructed gradually from simpler classifiers related to parts of objects and spatio-temporal relations between them. We emphasise the fact that the construction of such hierarchical classifiers should be performed along schemes of reasoning extracted from domain knowledge.

One of the fundamental features of our approach is that on each level of hierarchical construction we use information systems (decision tables) [15]. Information systems are defined by a universe of objects described in terms of attributes. The objects usually correspond to some real entities, e.g., bank customers, situations on stock market, or patients. In the paper we consider structural objects (i.e., objects having complex structure) and information systems built over such objects. In this case the values of attributes depend on the object structure. Discovery of relevant structure for particular tasks is a complex problem strongly related to perception [8, 6, 40]. The object structure can be represented, e.g., by some relations on object parts (like rough inclusions [21] or partial orders [32]). Properties of objects embedded in such structures can be dependent on the context [11, 10], i.e., on neighbourhoods in which these objects or their parts appear. These properties can be used for prediction of behaviour of objects in time. Properties of neighbourhoods are expressed by formulas of some languages like those used in modal or temporal logics (see, e.g., [3]). By discovering structures of complex objects, one can build and investigate new information systems over these objects. Observe that several problems related to construction of such information systems should be solved, including the problems of specification of neighbourhoods and selection of the relevant language for expressing their properties. The choice of neighbourhoods raises the problem of their structure perception. Partial information about the objects themselves, and hence their neighbourhoods, is another source of difficult problems.

Complex concepts (e.g., the result of perception of situation on a road) can be vague, linked with other simpler concepts represented in domain knowledge and expressed in natural language. Moreover, concepts close to measurements may not be directly linkable with complex target concepts. Hence, it is necessary to develop methods for hierarchical learning (layered learning) [17, 35] of such concepts, since inducing high quality classifiers for such concepts directly from measurements is infeasible due to, in a sense, the large distance of target concepts from directly available patterns. This means that the size of search space for such relevant patterns can be huge. Hence, we propose to use domain knowledge to create constraints for such a search. We assume that domain knowledge consists of concepts as well as dependencies between them, expressed in a natural language. Reasoning schemes based on such dependencies are used as hints in searching for patterns relevant to complex target concepts that are constructed from patterns close to measurements. The concepts on the level below a given level correspond to properties of parts of objects considered on this level.

Local dependencies from domain knowledge, often expressed in a natural language, are approximated by a set of patterns that are called productions. Any production corresponding to the dependency

$$concept_1 \wedge \ldots \wedge concept_k \rightarrow concept$$

is of the following form:

$$(concept_1, pattern_1, degree_1) \wedge \ldots \wedge (concept_k, pattern_k, degree_k) \rightarrow (concept, pattern, degree).$$

The production is valid in a given data model if and only if the following condition is satisfied: if degree of inclusion of $pattern_i$ into $concept_i$ is at least $degree_i$ for any $i = 1, \ldots, k$ then $pattern$ is included into $concept$ to a degree at least $degree$.

Certainly, it is necessary to define inclusion measures for measuring degrees of inclusion of patterns into concepts [30]. Hence, any production describes a target pattern in terms of patterns from its premise. It guarantees that if these patterns are included to satisfactory degrees into corresponding concepts (from the left hand sides of dependencies) then the target pattern (on the right hand side of the production) is included to a satisfactory degree into the target concept (of dependency). In other words, each production specifies a target pattern included to a satisfactory degree into the approximated concept (from the right hand side of dependency) provided that all patterns from which the target pattern was constructed (and which are on the left hand side of the production) are included to satisfactory degrees into corresponding concepts. Patterns in productions can have different form. In the most general case they are represented by classifiers. In such an approach the induction is expected to be feasible, i.e., patterns relevant to the target concepts can be induced efficiently by learning strategies from patterns sufficiently close to concepts from the left hand sides of these dependencies. Observe that typical information about patterns in productions are degrees to which they are satisfied in a given global state. Using the rough mereological approach (see, e.g., [18, 20, 22, 26, 30]) we assume that rough mereological connectives, describing propagation of uncertainty, are extracted from data. Next, such connectives can be fuzzified. Several methods for inducing productions from data and domain knowledge have been developed [13, 28].

More general productions are production granules that are granules (clumps) of patterns represented by productions. Each production in such a granule corresponds to a different inclusion degree of its target pattern into the target concept. Such production granules make it possible to analyse the deviation of inclusion degrees of target patterns into the target concepts using degrees of inclusion of patterns from the left hand sides of productions to corresponding (simpler) concepts from dependencies. Within production granules, we can search for robust granules of productions. Moreover, they can be used for analysis of the dynamics of networks of productions [13].

Reasoning schemes based on domain knowledge are approximated by AR schemes derived from relevant productions. Sufficient conditions for derivation of relevant AR schemes from productions have been developed [18, 20, 22, 26, 30].

We can summarise the above discussion by observing that natural language reasoning schemes that represent domain knowledge can be, in a sense, simulated by AR networks and that it leads to two-tiered schemes of reasoning: natural language reasoning schemes and their approximations by AR networks.

In the case of spatio-temporal reasoning the productions have more complex form than in spatio reasoning because they represent dependencies between more complex spatio-temporal concepts. Their construction can be outlined as follows. Objects over which approximations of concepts are constructed have some structure expressed by properties of their parts and spatio-temporal relations between them. Objects are represented by structures, called neighbourhoods. Next, attributes (properties) on such neighbourhoods are defined (in some language) using already defined or approximated concepts and relations between them. One should also define relevant relations between neighbourhoods using already defined relations on objects. Certainly, it is useful to consider information systems (decision tables) over neighbourhoods of some fixed type, e.g., specified by the object structure and some of its properties.

In this way the patterns constructed on higher level of hierarchical construction can be used to express different kinds of neighbourhoods, e.g., properties of sequences of concepts linked by spatio-temporal relations, sets of such sequences, or properties of concurrent processes defined by Petri nets. Properties of such neighbourhoods are defined by means of constructed concepts (patterns), relations between them and degrees of inclusion of one concept into another. Moreover, some structures like information maps [33] or Petri nets can be used on one level, and their properties can be included on the next level of the hierarchical construction.

In the following sections we discuss the approach to hierarchical modelling in more detail.

3. Information systems for structural objects

For structural objects we usually have more complex relational structures than those represented so far by information systems [15]. Starting from the basic level of hierarchical modelling we often have to deal with relations (on objects) of arity higher than one, together with unary predicates corresponding to descriptors widely used in information systems. For example, often are used *to be a part to a degree* relation [18, 20, 22, 26, 30] or some time related relations. Approximations of concepts on this level are derived by means of neighbourhoods of objects defined by the uncertainty function and the rough inclusion [27, 30].

Hence, we propose the following definition of decision systems for structural objects.

Definition 3.1. Let $\mathcal{R}$ be a relational structure over a (finite) universe U and let $\mathcal{N}_{\mathcal{R}}$ be a family of neighbourhoods, i.e., restrictions of $\mathcal{R}$ to subsets of U. A decision system over the relational structure R and the family of neighbourhoods $\mathcal{N}_{\mathcal{R}}$ is any decision system $DT = (\mathcal{N}_{\mathcal{R}}, A, d)$ [15].

Let us consider an example of a neighbourhood family $\mathcal{N}_{\mathcal{R}}$. Assume $N(x) \subseteq U$ is selected for any object $x \in U_0 \subseteq U$ where U_0 is a finite sample of U. Then $\mathcal{N}_{\mathcal{R}}$ is equal to the set of all restrictions of $\mathcal{R}$ to $N(x)$ for $x \in U_0$. For real-life applications it is necessary to discover (from given data and domain knowledge) relevant relational structure $\mathcal{R}$, family of neighbourhoods $\mathcal{N}_{\mathcal{R}}$ as well as the set of conditional attributes A over such neighbourhoods.

Higher arity relations on objects can be often approximated from data. However, in some cases such relations are explicitly defined on basic objects (e.g., using a distance between objects) that can be indiscernible [15]. Then a granulation of these relations should be performed what leads to relations defined on neighbourhoods of objects rather than on objects [16].

For decision problems with complex structural objects one should consider hierarchical structures of information systems over different neighbourhood families representing parts of different relational structures. Any higher level of such a hierarchy is defined over the relational structures of the lower levels. The above definition of decision systems can be also used on higher levels of hierarchical modelling.

The relational structures constructed on the lower level of hierarchical construction are used to define new information systems on the next level of construction. Such information systems for more complex objects are defined by a composition of information systems from lower level of hierarchy representing parts of these more complex objects [31]. Each object on a higher level of hierarchical construction represents partial information about neighbourhoods (relational structures) of composed information systems, i.e., it consists of a subset of the universe together with object relationships defined by relations from the underlying relational structures. In specification of objects on higher levels some constraints between composed neighbourhoods from lower levels are also used. In this way neighbourhoods are generalisations of windows, widely used in temporal reasoning (e.g., time windows in time series analysis [17, 7]).

The neighbourhoods of objects from the universe on a higher level of hierarchy are constructed using the following information:

1. parts of the object structure represented by neighbourhoods on lower level of hierarchy (by applying some operations to them);

2. attributes (formulas) defined over the neighbourhoods constructed on the lower level of hierarchy;

3. formulas describing constraints between composed neighbourhoods from the lower level of hierarchy (that are also based on new conditional attributes for the higher level);

4. degrees to which (at least) the considered above formulas are satisfied.

In spatio-temporal reasoning we often have to deal with information systems called decision tables, i.e., information systems with a distinguished decision attribute [15]. The approximation of decision classes is expressed by conditional attributes of the decision system. The conditional attributes over neighbourhoods representing objects in such decision tables should be relevant for approximation of decision classes defined by the decision. For structural objects these conditional attributes are dependent on the neighbourhood structure. Important problems for spatio-temporal reasoning include the discovery of neighbourhoods and their properties relevant for decision classes approximation. For other applications, such as multi-criteria decision making, the relevant neighbourhoods are given and only their relevant properties should be discovered. The conditional attributes of decision systems on a higher level are defined over neighbourhoods available on that level. Such conditional attributes can also be defined by classifiers, in particular, rough-fuzzy classifiers [30].

Important examples of operations on neighbourhoods are related to operations of composition of information systems [31] and to the neighbourhood granulations [16]. The latter case is related to the necessity of defining the relations between objects-neighbourhoods on a higher level using relationships between objects included in their parts represented by neighbourhoods on a lower level. That means that we have to granulate the relational structure each time we create a new level. Some examples of relational structures granulation are presented in [16].

The approach presented makes it possible to realise the following optimisation principle of approximate reasoning. Reasoning is simplified due to information granulation to such a degree that it is still possible to obtain a solution of satisfactory quality.

Some relationships of this approach with information networks, investigated in [2], are reported in [31].

The structure of objects is defined by their decomposition into parts. This is strongly related to perception. In the rough mereological approach parts are defined by rough inclusion relations that are relations *to be a part to a degree*. Strategies searching for rough inclusions that make it possible to decompose objects into relevant structures are crucial for understanding perception. Domain knowledge can substantially simplify searching by such strategies for relevant rough inclusions.

4. Concept approximation

Discovery of relevant concepts for describing the investigated phenomena, and rules expressing dependencies between them from available data and domain knowledge, is one of the main tasks of machine learning, pattern recognition, as well as data mining and knowledge discovery [9, 12]. This task is very hard not only from a computational complexity point of view. Usually it is not possible to find an exact concept description, because of, e.g., uncertainty of concept specification and concept vagueness. Moreover, to approximate complex concepts one should develop hierarchical classifiers (see, e.g., [35, 17]). AR networks used in this paper, make it possible to link lower level concepts with complex concepts to be approximated. Each AR network can be interpreted as a cluster of complex objects close to some

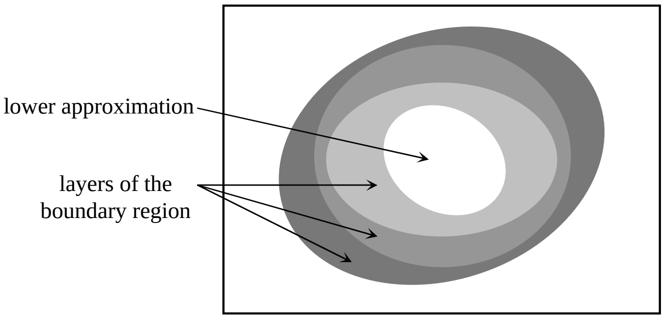

Figure 1. Concept approximation: lower approximation and layered boundary region

prototype (standard) to a satisfactory degree. It is assumed that such AR networks are induced using experimental data and soft domain knowledge.

Observe that in many tasks the exact concept descriptions are not required [40], and due to this, satisfactory and robust solutions with much lower cost can be provided [40]. One can consider the task of parking a car as an illustrative example.

To describe laws for a given phenomenon one should induce approximations of the relevant concepts and create tools for approximate reasoning about them. We propose to use approximate reasoning rules of inference called also productions (see, e.g., [26, 20]). Such rules are implications with premises and conclusion built from patterns, labelled by inclusion degrees of these patterns into the concept approximations from premises and conclusion, respectively [26]. The pattern from the conclusion (target pattern) of a given rule is constructed out of patterns (input patterns) from the rule premise. Estimation of the inclusion degrees from data, labelling target patterns on the basis of inclusion degrees labelling input patterns, is one of the important tasks of data mining (see, e.g., [28, 20]).

Approximation of a given concept can be expressed using the rough set, fuzzy set, or rough-fuzzy approaches [15, 39, 14]. In the rough set approach the approximation is expressed by the lower approximation, i.e., the set of all objects belonging to the concept with certainty (to degree 1); the upper approximation, i.e., the set of all elements belonging to the concept to a degree $0 < p < 1$; and the complement – the set of all elements not belonging to the concept with certainty (i.e., belonging to the concept to degree 0). In many cases the lower approximation can be relatively small, contrary to the boundary region. Therefore, it can be necessary to take into account the information about different fragments of the boundary region to discover relevant patterns. To provide relevant rules for approximate reasoning, the inclusion degrees of patterns in concepts can be granulated (e.g., they can be discretised). This leads to some patterns in the form of layers in the boundary region (see Figure 1). In some cases, one can label them by linguistic terms, e.g., small, medium, large, by analogy to fuzzy sets. The layers can be also specified by the lower and upper approximations. This means that some elements may belong with certainty to some layer, while some others to the boundary of successive layers. The choice of such relevant patterns depends on a particular application and, in particular, on the perception of structures of objects. Let us observe that for reasoning about changes, it is useful to consider some order on such patterns, e.g., a linear order [30].

There are several problems related to reasoning with approximate reasoning rules. In particular, the problem of estimation of inclusion degrees of input objects to the concepts. A typical scheme of solving

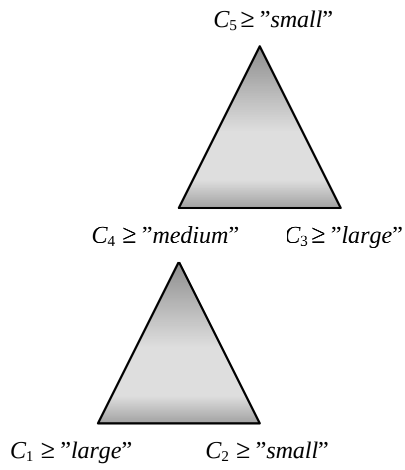

Figure 2. A complex approximate reasoning rule built from concepts $C_1, \ldots, C_5$

such problem is the following. There are given some patterns together with their estimated inclusion degrees in the target concept (e.g., inclusion of the left hand sides of decision rules in the concepts defined by decision classes). The available information about investigated objects is used to determine their inclusion degrees in the patterns from premises of the rule. The degrees of objects inclusion in patterns together with inclusion degrees of patterns in a given concept and its complement, are used for the estimation of the object inclusion degree in the concept. This is a standard scheme in the classifier construction [12].

The problem of discovering approximate reasoning rules from data becomes a challenge nowadays. Such rules are basic components, e.g., in searching for more complex structures, like hierarchical classifiers, classifier networks, AR schemes [30], and AR networks.

5. Complex concepts and structural objects

Using carefully constructed derivations based on approximate reasoning rules one can estimate degrees of inclusion of input objects in complex concepts from the degrees of inclusion in the primitive (e.g., sensory) concepts. Such derivations create AR schemes [17, 35, 30].

Example 5.1. Let $C_1, \ldots, C_5$ be some concepts approximated by the lower approximations and three linearly ordered layers *small*, *medium*, and *large*. Let an AR scheme like the one in Figure 2 also be given (a very important problem is related to developing efficient methods for inducing such schemes from data). This scheme is a combination of two simpler rules with the following interpretation: (1) if 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.

In particular, by using the above scheme, we can reason about structural objects on the basis of properties of their immediate parts that can also be built from simpler parts, and so on. For example, a car is a structural object built from parts being also complex objects.

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.

5.1. Feature selection and extraction

The key point is to develop methods searching for relevant features for concept approximation. Such features can be selected from a given set of available features. In a more general case, one can first try to discover a set of features from which the relevant features are next selected. This is due to the fact that the space of all features can be very large and from practical point of view searching for relevant features in the whole space is not feasible. These are the two basic steps in classifiers construction in machine learning, pattern recognition and data mining known as the feature selection and extraction [12, 7, 9]. Features used for object description are expressed by formulas from some language – its choice depends on the investigated problems and, in particular, on a context in which we should consider objects. The problem of relevant language discovery is a challenging problem in data mining.

Note that we propose to use soft domain knowledge for relevant feature extraction for complex concept approximation. Domain reasoning schemes are treated as hints in searching for such relevant features. On the basis of such schemes we propose to induce AR networks representing patterns relevant for complex concepts. Characteristic functions of such patterns can be used as relevant features. By collecting AR networks, relevant for a given concept, we can gradually approximate the concept. Searching for relevant features for complex concept approximation without domain knowledge is infeasible because of the very huge search space from which one should extract such features. This aspect is also related to a general discussion in statistical community about the necessity to develop a new statistical approach to deal with complex real-life problems (see, e.g., [37, 4]).

Observe that in many cases, we can measure only selected number of features (satisfiability degrees of some concepts) what makes the given information about objects incomplete. Moreover, satisfiability degree of some features may be estimated only partially (i.e., to some degree), and therefore for expressing the satisfiability of corresponding formulas, multi-valued logic should be used.

5.2. Structure discovery

Another very important problem is related to the relevant structure discovery of complex objects, taking into account the context in which they appear. Such a structure description, usually unknown and very complex, depends on the chosen description language. The choice of another description language can change the object perception. To discover the relevant perception of the object structure is very hard, because of very huge space of possible structures from which it is necessary to search for the relevant ones. Moreover, in many cases we can observe only some features of some parts of objects and relations (or their approximations) they satisfy; the features of different parts can be expressed in different languages. All these aspects are closely related to perception problems (see, e.g., [1, 8, 40, 6]). Note that nowadays it is a growing expectation that existing knowledge about perception in psychology and neuroscience can lead to feasible searching strategies for relevant perception of complex objects. In our approach we propose to support the object perception using soft domain knowledge. The developed AR schemes on the basis of domain knowledge can be considered as perception schemes [40, 8, 6, 30].

To reason about the structure of a complex object we have to use composition schemes (one- or many-level) of its parts. Such 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 [19, 21].

A structure of complex objects can be expressed by means of AR schemes. Such schemes usually have a tree structure with the root, corresponding to satisfiability degree of some feature by complex object, and leaves – to the satisfiability degrees of some other features by primitive objects (i.e., the most simple parts of complex object). An AR scheme can have many levels. Then from features of basic parts we conclude about features of more complex parts, and after some levels, about features of the complex target object. However, in many cases of spatio-temporal reasoning it can be necessary to use feedback from the root of an AR scheme – the information about satisfiability degree of some feature by a complex object can be used by primitive objects in next iterations. Hence, we propose to use AR networks for such reasoning.

6. Structural objects changing in time

Another class of problems arises when a given object evaluates in time, what is measured by changes in time of some of its features [24]. 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., [25, 24]).

6.1. Modelling object changes

Reasoning about objects evaluated in time is usually based on some historical knowledge (training data) [24, 9]. In the learning process the state space is constructed, where states correspond to the observed objects and transitions to their behaviour in time. The states can be ordered by transition relation – sometimes it is a linear order, but a more general case is possible, too, e.g., a partial order specifying possibilities of moving to the next state. By adding a time dimension to patterns we obtain temporal patterns, describing properties of objects or their parts related to time. Then, in particular, we are interested in discovery of spatio-temporal rules for prediction of inclusion degrees of objects into some patterns related to the future, from information about degrees of inclusion of objects in some other spatio-temporal patterns in the past. Patterns are expressed by formulas of some languages and some temporal operators can be used for expressing their relationships in time.

The AR networks for complex objects evaluated in time differ from the static case of AR schemes. They are constructed also along the time dimension, i.e., rules used in their construction link some spatio-temporal patterns, describing spatio-temporal properties of objects or their parts. Moreover, premises can

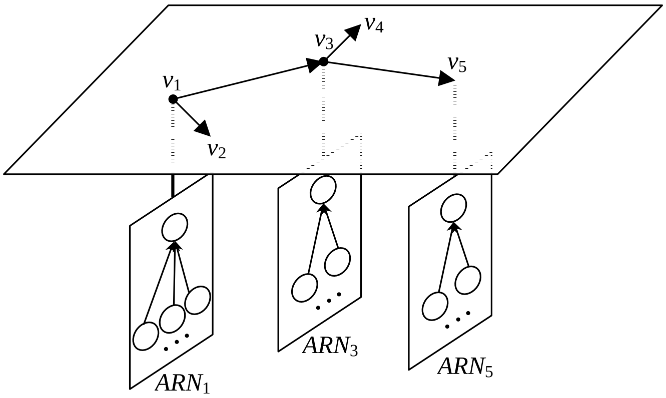

Figure 3. An example of AR network

consist of some conditions expressing time constraints. Figure 3 illustrates a fragment of spatio-temporal AR scheme construction. A graph on a plane represents, for example, a partial order with nodes, labelled by AR schemes matching (to a degree) some complex concepts. Nodes are linked by edges labelled by spatio-temporal relations. In the case of temporal relations they can depend not only on time but also on the context in which objects appear. On the top of such a structure the next layer of AR network is constructed with the conclusion dependent not only on concepts delivered by AR schemes but also on time relations between them. In particular, such spatio-temporal patterns can express information about the history of changes of complex objects or their parts. In a more general case graphs, as the one mentioned above, are composed from some simpler patterns (or concepts) and spatio-temporal relations between them. AR networks constructed over spatio-temporal patterns are stored in a spatio-temporal knowledge base that can support approximate reasoning.

The AR networks represent properties of processes. Hence, they can have a more general structure than the AR schemes considered so far, represented by trees or acyclic graphs of dependencies. We return to this issue in the next part of the paper.

One can also use another structure of AR networks in which nodes are labelled by decision systems. Attributes of such a decision system correspond to spatio-temporal patterns defined by AR networks and the attribute values represent inclusion degrees of objects into such patterns. Decision rules generated from such decision system are spatio-temporal rules.

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

Example 6.1. Let us consider a problem of an automatic situation tracking on the road by the unmanned aerial vehicle (UAV) [38]. Suppose an UAV monitors two vehicles going on the road close to each other, and using the observation is supposed to evaluate their situation in the context of, e.g., safety. Different approximation schemes describe their mutual location, for example, “first behind second”, “first on the left of second”. Thus, we have a collection of AR networks $ARN_1, \ldots, ARN_n$, describing typical situations in which two vehicles may occur. The UAV, while monitoring the vehicles, tries to find an AR network matching the observed situation to the highest degree. The result of monitoring the vehicles in time is then a sequence

$$\left\langle (ARN_{i_1}, d_1, t_1, Time(t_1, t_2)), \ldots, (ARN_{i_{k-1}}, d_{k-1}, t_{k-1}, Time(t_{k-1}, t_k)) \right\rangle,$$

where for $j = 1, \ldots, k-1$, the network ARN_{i_j} is matched by the observed situation to the highest degree d_j (among all considered AR networks) at time t_j, and $Time(t_j, t_{j+1})$ are time constraints satisfied for $j = 1, \ldots, k-1$. These time constraints can have a more complex form. In particular, they can depend on properties of AR networks linked with them.

6.2. Invariants and protocols

The investigation of a structural object evolving in time may be often based on the analysis of some of its parts. Patterns related to parts are combined into more complex patterns by means of approximate reasoning rules. Degrees to which patterns are satisfied by a given object (situation) can be combined into degrees to which these more complex patterns are satisfied. Finally, one can estimate inclusion degrees of objects in patterns related to complex concepts.

In spatio-temporal reasoning, we often would like to preserve some soft invariants to at least a satisfactory degree. These invariants can be expressed by complex vague concepts, e.g., *safe situation on the road*. In many cases parts of a complex object (situation) can be treated as independent local processes having its own space of states together with transition relations describing possible next states of parts reachable by some actions. Then, for many such parts the choice of their next states must be taken by considering the constraints that the whole complex object should satisfy. That makes it possible to achieve the considered goal (e.g., to preserve to a satisfactory degree global invariants). Thus, some protocols have to be discovered, controlling the transitions among parts. For example, such protocols can have a form of a negotiating scheme returning transitions relevant for each of parts. One of the challenging problems is to develop learning strategies for such protocols (from experimental data and soft domain knowledge), making it possible to preserve global invariants.

Example 6.2. Let us consider the case of approximate reasoning schemes for a complex object changing in time. An example of such an object is a situation on the road with different vehicles, or their groups, moving along. Such a complex object evaluates in time – at time t_1 the situation on the road can be safe, at t_2 unsafe, while at t_3 – dangerous. For each of the parts of a given situation on the road, it can be given an information (based on experts' knowledge or learnt from the observations in the past), describing their possible next states. The protocol should force all of the situation parts to choose such transitions that the next state of the whole complex object will be *safe*. For example, assume two vehicles are in a close distance and they are going in opposite directions on a road with two lanes. The protocol should guarantee that if one car starts the manoeuver of overtaking of a vehicle then the second does not.

6.3. Complex patterns describing behaviour of objects in time

An important problem related to the analysis of approximate reasoning about complex objects changing in time is related to the estimation of degrees to which spatio-temporal patterns of behaviour are matched by the behaviour of a given object. For example, let some patterns of typical situations on the road be given, e.g., overtaking, or too close (dangerous) location of two vehicles. Such patterns can be formulated by an expert or learnt from past observations of the road. They can be defined as a directed graph, where

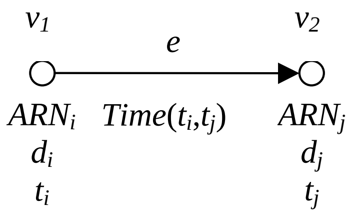

Figure 4. AR network for a pattern describing behaviour of a complex object

vertices are labelled by AR networks $ARN_1, \ldots, ARN_n$, corresponding to states of a complex object, and edges are labelled by spatio-temporal relations specifying, e.g., constraints on time that takes to move from one state (vertex) to another. Such a graph can be cyclic, e.g., from a situation that vehicle A goes after B, we can move to the case when A starts overtaking, and then back to the state when A follows B (because A was not able to overtake). Edges may have many different interpretations: spatial and/or temporal, e.g., that moving from one situation to another does not take more than Δt, or that it can take as long as possible. Observe that for any pattern represented by AR networks, it should be possible to estimate the degree to which such a network is matched by a given process (e.g., history of situations on the road), represented by another network close to measurements. In this way patterns for the complex concepts, like overtaking, are defined by means of patterns for simpler concepts and spatio-temporal relations between them. Patterns represented by AR networks can be used for prediction.

In Figure 4 a simple example of a fragment of AR network is presented. Two vertices v_1, v_2 of this network are connected by an edge $e = \langle v_1, v_2 \rangle$. The vertices v_1, v_2 are labelled by AR networks and the edge e by a time constraint $Time(t_i, t_j)$. Using the degree to which a given tested object is matching a given pattern (AR network) labelling v_1 and the degree of satisfiability of the temporal constraint labelling e, we can estimate the degree to which a pattern related to v_2 is satisfied (at time specified by the temporal constraint). Such graphs corresponding to local patterns can be transformed to AR networks, where nodes are labelled by other AR networks and edges by temporal relations.

One can see that we deal with two kinds of information in constructing of AR networks. Some parts of such networks can be identified as AR schemes describing spatial (static) properties of objects, related to a given moment (period) of time and different AR schemes are related by some time constraints. In our example the former parts describe spatial features of location of vehicles, e.g., “A after B”, while the latter describe soft temporal properties of objects, e.g., in a short period of time.

6.4. Matching complex concepts by AR networks

Let us recall that the constructed AR networks form a spatio-temporal knowledge base. They are a kind of patterns for approximation of complex concepts. With each concept, there are associated AR networks matching it to a satisfactory degree. An important problem is to select all AR networks matched by the currently observed situation. The size of spatio-temporal knowledge base can be huge and, thus, the process of choosing the most relevant AR network can be complex. From the practical point of view (e.g., in the case of real-time systems), it cannot be possible to search sequentially through the whole space of networks. One approach to this problem is based on constructing of a dedicated hardware for parallel selecting AR networks matched by a given situation. Another possibility is to perform relevant indexing of the stored AR networks on the basis of patterns close to measurements (e.g., sensory patterns).

7. Conclusions

In the paper we have discussed some problems related to discovery of spatio-temporal patterns from data and soft domain knowledge. Such patterns are related to spatial structure of objects as well as to their changes in time. The spatial patterns can be represented by AR schemes. To represent spatio-temporal patterns we have proposed AR networks. A collection of such AR networks create a spatio-temporal knowledge base for approximate spatio-temporal reasoning about a given task. Several aspects of AR networks were discussed. In particular,

- representation of spatio-temporal patterns by AR networks;
- learning of AR networks from data and soft knowledge bases;
- learning of protocols controlling behaviour of parts of complex objects for preserving global invariants.

There are more complex spatio-temporal patterns which are important to be discovered from data and domain knowledge. They can be related to changes in time of rules in AR networks or the whole AR networks.

In our further work we would like to develop multi-layered approximate reasoning system across layers specified by the knowledge base of AR networks. The main task will be to develop methods of learning (from data and domain knowledge) a multi-layered structure of spatio-temporal knowledge bases. In particular, it will be necessary to extend methods for AR schemes synthesis (see, e.g., [28, 13]) 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] Barsalou, L. W.: Perceptual Symbol Systems, *Behavioral and Brain Sciences*, **22**, 1999, 577–660.

[2] Barwise, J., Seligman, J., Eds.: *Information Flow: The Logic of Distributed Systems*, vol. 44 of *Tracts in Theoretical Computer Science*, Cambridge University Press, Cambridge, UK, 1997.

[3] Bennett, B., Cohn, A. G., Wolter, F., Zakharyaschev, M.: Multi-Dimensional Modal Logic as a Framework for Spatio-Temporal Reasoning, *Applied Intelligence*, **17**(3), 2002, 239–251.

[4] Breiman, L.: Statistical Modeling: The Two Cultures, *Statistical Science*, **16**(3), 2001, 199–231.

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

[6] Fahle, M., Poggio, T.: *Perceptual Learning*, MIT Press, 2002.

[7] Friedman, J. H., Hastie, T., Tibshirani, R.: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer-Verlag, Heidelberg, 2001.

[8] Harnad, S.: *Categorical Perception: The Groundwork of Cognition*, Cambridge University Press, New York, NY, 1987.

[9] Kloesgen, W., Żytkow, J., Eds.: *Handbook of Knowledge Discovery and Data Mining*, Oxford University Press, 2002.

[10] McCarthy, J.: Notes on Formalizing Context, *Thirteenth International Joint Conference on Artificial Intelligence IJCAI*, Morgan Kaufmann, Chambéry, France, 1993.

[11] McCarthy, J., Hayes, P.: Some Philosophical Problems from the Standpoint of Artificial Intelligence, *Machine Intelligence*, **4**, 1969, 463–502.

[12] Mitchell, T. M.: *Machine Learning*, Mc Graw Hill, 1997.

[13] Pal, S. K., Polkowski, L., Skowron, A., Eds.: *Rough-Neural Computing: Techniques for Computing with Words*, Cognitive Technologies, Springer-Verlag, Heidelberg, 2004.

[14] Pal, S. K., Skowron, A., Eds.: *Rough Fuzzy Hybridization: A New Trend in Decision-Making*, Springer-Verlag, Singapore, 1999.

[15] Pawlak, Z.: *Rough Sets: Theoretical Aspects of Reasoning about Data*, vol. 9 of *System Theory, Knowledge Engineering and Problem Solving*, Kluwer Academic Publishers, Dordrecht, 1991.

[16] Peters, J. F., Skowron, A., Synak, P., Ramanna, S.: Rough Sets and Information Granulation, *Tenth International Fuzzy Systems Association World Congress IFSA* (T. Bilgic, D. Baets, O. Kaynak, Eds.), *Lecture Notes in Artificial Intelligence*, **2715**, Springer-Verlag, Heidelberg, 2003, 370–377.

[17] Poggio, T., Smale, S.: The Mathematics of Learning: Dealing with Data, *Notices of the AMS*, **50**(5), 2003, 537–544.

[18] Polkowski, L., Skowron, A.: Rough Mereological Approach to Knowledge-Based Distributed AI, *Third World Congress on Expert Systems* (J. K. Lee, J. Liebowitz, J. M. Chae, Eds.), Cognizant Communication Corporation, Soeul, Korea, February 5-9, 1996, 774–781.

[19] Polkowski, L., Skowron, A.: Rough Mereology: A New Paradigm for Approximate Reasoning, *International Journal of Approximate Reasoning*, **15**(4), 1996, 333–365.

[20] Polkowski, L., Skowron, A.: Towards Adaptive Calculus of Granules, *Computing with Words in Information/Intelligent Systems* (L. A. Zadeh, J. Kacprzyk, Eds.), Physica-Verlag, Heidelberg, 1999, 201–227.

[21] Polkowski, L., Skowron, A.: Rough Mereology in Information Systems. A Case Study: Qualitative Spatial Reasoning, in: *Rough Set Methods and Applications: New Developments in Knowledge Discovery in Information Systems* (L. Polkowski, T. Y. Lin, S. Tsumoto, Eds.), *Studies in Fuzziness and Soft Computing*, **56**, Springer-Verlag, Heidelberg, 2000, 89–135.

[22] Polkowski, L., Skowron, A.: Rough Mereological Calculi of Granules: A Rough Set Approach to Computation, *Computational Intelligence*, **17**(3), 2001, 472–492.

[23] Roddick, J. F., Hornsby, K., Spiliopoulou, M.: An Updated Bibliography of Temporal, Spatial and Spatio-Temporal Data Mining Research, *Post-Workshop Proceedings of the International Workshop on Temporal, Spatial and Spatio-Temporal Data Mining TSDM* (J. F. Roddick, K. Hornsby, Eds.), *Lecture Notes in Artificial Intelligence*, **2007**, Springer-Verlag, Berlin, 2001, 147–163.

[24] Roddick, J. F., Hornsby, K., Spiliopoulou, M.: YABTSSTDMR - Yet Another Bibliography of Temporal, Spatial and Spatio-Temporal Data Mining Research, *SIGKDD Temporal Data Mining Workshop* (K. P. Unnikrishnan, R. Uthurusamy, Eds.), ACM Press, San Francisco, CA, 2001, 167–175.

[25] Sandewall, E., Ed.: *Features and Fluents: The Representation of Knowledge About Dynamical Systems*, vol. 1, Oxford University Press, 1994.

[26] Skowron, A.: Toward Intelligent Systems: Calculi of Information Granules, *Bulletin of the International Rough Set Society*, **5**(1-2), 2001, 9–30.

[27] Skowron, A., Stepaniuk, J.: Tolerance Approximation Spaces, *Fundamenta Informaticae*, **27(2-3)**, 1996, 245–253.

[28] Skowron, A., Stepaniuk, J.: Information Granule Decomposition, *Fundamenta Informaticae*, **47**(3-4), 2001, 337–350.

[29] Skowron, A., Stepaniuk, J.: Information Granules: Towards Foundations of Granular Computing, *International Journal of Intelligent Systems*, **16**(1), 2001, 57–86.

[30] Skowron, A., Stepaniuk, J.: Information Granules and Rough-Neural Computing, in: Pal et al. [13], 43–84.

[31] Skowron, A., Stepaniuk, J., Peters, J.: Rough Sets and Infomorphisms: Towards Approximation of Relations in Distributed Environment, *Fundamenta Informaticae*, **54**(2-3), 2003, 263–277.

[32] Skowron, A., Synak, P.: Patterns in Information Maps, *Third International Conference on Rough Sets and Current Trends in Computing RSCTC* (J. J. Alpigini, J. F. Peters, A. Skowron, N. Zhong, Eds.), *Lecture Notes in Artificial Intelligence*, **2475**, Springer-Verlag, Heidelberg, 2002, 453–460.

[33] Skowron, A., Synak, P.: Reasoning in Information Maps, *Fundamenta Informaticae*, **59**(2-3), 2004, 241–259.

[34] SPACENET: Project Web Site, WWW.
URL `http://agora.scs.leeds.ac.uk/spacenet/`

[35] Stone, P.: *Layered Learning in Multi-Agent Systems: A Winning Approach to Robotic Soccer*, The MIT Press, Cambridge, MA, 2000.

[36] Świniarski, R., Skowron, A.: Rough Set Methods in Feature Selection and Recognition, *Pattern Recognition Letters*, **24**(6), 2003, 833–849.

[37] Vapnik, V.: *Statistical Learning Theory*, John Wiley & Sons, New York, NY, 1998.

[38] WITAS: Project Web Site, WWW.
URL `http://www.ida.liu.se/ext/witas/eng.html`

[39] Zadeh, L. A.: Fuzzy Sets, *Information and Control*, **8**(3), 1965, 338–353.

[40] Zadeh, L. A.: A New Direction in AI: Toward a Computational Theory of Perceptions, *AI Magazine*, **22**(1), 2001, 73–84.