Ontological Framework for Approximation

Jarosław Stepaniuk[1] and Andrzej Skowron[2]

[1] Department of Computer Science, Białystok University of Technology, Wiejska 45a, 15-351 Białystok, Poland
jstepan@ii.pb.bialystok.pl

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

Abstract. We discuss an ontological framework for approximation, i.e., to approximation of concepts and vague dependencies specified in a given ontology. The presented approach is based on different information granule calculi. We outline the rough–fuzzy approach for approximation of concepts and vague dependencies.

1 Introduction

We discuss the ontology for approximation in a granular computing framework. One of the main task in granular computing is to develop calculi of information granules [12,23,14,16]. Such calculi are aiming at constructing from elementary granules some target granules satisfying a given specification to a satisfactory degree. Hence, together with operations for construction of information granules one should define relevant measures of inclusion (to a degree) and closeness (to a degree) of granules. This idea has been presented and developed in the rough–mereological framework (see, e.g., [11,12]). Observe, that constructions of granules are often described by multilevel schemes and the final construction of the target information granule is obtained by a relevant composition of local schemes.

Vague dependencies have vague concepts in premisses and conclusions. The approach to approximation of vague dependencies based only on degrees of closeness of concepts from dependencies and their approximations (classifiers) is not satisfactory for approximate reasoning. Hence, more advanced approach should be developed. Approximation of any vague dependency is a method which allows for any object to compute the arguments "for" and "against" its membership to the dependency conclusion on the basis of the analogous arguments relative to the dependency premisses. Any argument is a compound information granule (compound pattern). Arguments are fused by local schemes (production rules) discovered from data. Further fusions are possible through composition of local schemes, called approximate reasoning schemes (AR schemes) (see, e.g., [2,12,10]). To estimate the degree to which (at least) an object belongs to concepts from ontology the arguments "for" and "against" those concepts are collected and next a conflict resolution strategy is applied to them to predict the degree.

D. Ślęzak et al. (Eds.): RSFDGrC 2005, LNAI 3641, pp. 718–727, 2005.

Several information granule calculi are involved in solving the problem of ontology approximation. Information granules in such calculi are represented by compound patterns.

By granulation of the discovered patterns to layers of vague concepts one can obtain more relevant approximations of dependencies. We outline the rough–fuzzy approach based on granulation.

The paper is organized as follows. In Section 2 we recall the basic concepts on approximation spaces. Rough information granule calculi based on the so called transducers are discussed in Section 3. In Section 4 we outline the approach to approximation of concepts and vague dependencies.

2 Approximation Spaces

In this section we recall a general definition of an approximation space [13,21]. Several known approaches to concept approximations can be covered using such spaces, e.g., the approach given in [9], approximations based on the variable precision rough set model [26] or tolerance (similarity) rough set approximations (see, e.g., [13,21] and references in [13,21]).

For every non-empty set U, let $P(U)$ denote the set of all subsets of U.

Definition 1. *[13,21] A parameterized approximation space is a system* $AS_{\#,\$} = (U, I_{\#}, \nu_{\$})$, *where*

- U *is a non-empty set of objects,*
- $I_{\#} : U \rightarrow P(U)$ *is an uncertainty function,*
- $\nu_{\$} : P(U) \times P(U) \rightarrow [0,1]$ *is a rough inclusion function,*

and $\#, \$$ *denote vectors of parameters.*

The uncertainty function defines for every object x, a set of objects described similarly. The set $I(x)$ is called the neighborhood of x (see, e.g., [9,7]). A set $X \subseteq U$ is *definable in* $AS_{\#,\$}$ if and only if it is a union of some values of the uncertainty function. The rough inclusion function defines the degree of inclusion of any $X \subseteq U$ in $Y \subseteq U$. In the simplest case it can be defined by (see, e.g., [13], [21]):

$$\nu_{SRI}(X,Y) = \begin{cases} \frac{card(X \cap Y)}{card(X)} & \text{if } X \neq \emptyset \\ 1 & \text{if } X = \emptyset. \end{cases}$$

This measure is widely used by the data mining and rough set communities. It is worth mentioning that Jan Łukasiewicz [8] was the first one who used this idea to estimate the probability of implications. However, rough inclusion can have a much more general form than inclusion of sets to a degree (see, e.g., [11,16,20]).

The lower and the upper approximations of subsets of U are defined as follows.

Definition 2. *For an approximation space* $AS_{\#,\$} = (U, I_{\#}, \nu_{\$})$ *and any subset* $X \subseteq U$, *the lower and upper approximations are defined by*

$LOW\left(AS_{\#,\$}, X\right) = \{x \in U : \nu_{\$}(I_{\#}(x), X) = 1\}$,

$UPP\left(AS_{\#,\$}, X\right) = \{x \in U : \nu_{\$}(I_{\#}(x), X) > 0\}$, *respectively.*

The lower approximation of a set X with respect to an approximation space $AS_{\#,\$}$ is the set of all objects, which can be classified with certainty as objects of X with respect to $AS_{\#,\$}$. The upper approximation of a set X with respect to an approximation space $AS_{\#,\$}$ is the set of all objects which can be possibly classified as objects of X with respect to $AS_{\#,\$}$.

The approximation spaces defined above have been generalized in [20]) to approximation spaces consisting of information granules. Approximation spaces themselves can be treated as special information granules. Granulation of approximation spaces, defined by operations of granulation and extension of relational structures, are studied, e.g., in [20]. For simplicity of considerations, we use in the paper the simple model of approximation space that has been recalled above.

3 Rough Information Granules and Transducers

Rough information granules are rough sets represented by some tuples of crisp sets (selected from lower and upper approximations, boundary regions, or complement of upper approximations). Then operations on such information granules transform rough sets into rough sets. Local schemes, called transducers [4], are used to represent operations for computing the lower and upper approximations of the target concept from the lower and upper approximations of arguments representing more elementary concepts. Multilevel schemes are representing compositions of such local schemes.

In [4], the authors study approximation transducers, devices that convert input approximate relations into output approximate ones by means of first-order theories. Different rough set techniques are applied to produce approximations of relations. In defining approximate transducers, methods of relational databases are invoked.

One can try to extend this approach to rough sets corresponding to approximations of concepts constructed using different operations such as transitive closure of relations, projections of sets, sets defined by formulas of modal logic, fixed point of some operators, etc. However, it is necessary to remember that the estimation of approximations of such concepts, obtained from approximations of concepts from which they are generated, can be of poor quality. The reason is the same as in the case of set theoretical operations, i.e., the approximation operations are not distributive with respect to such operations. Hence, the approximation quality can drop quickly with increasing of operation complexity. For example, if the only available information for approximation of the transitive closure R^* of relation R is the approximation of R then, usually, the received approximation will be of poor quality compared with the approximation that can be obtained by direct approximation of R^*, i.e., by using examples and counter examples of tuples satisfying R^*.

The conclusion is that, unfortunately, the approximation of more compound concepts has to be constructed gradually using more specific information, e.g., on patterns (information granules) discovered in construction of classifiers for some simpler concepts. In general, the high quality approximation of a concept dependent on some simpler concepts can not be derived from approximations of these simpler concepts only. In the next section we propose a step toward solution of this problem.

4 Approximation of Concepts and Dependencies from Ontology

In this section we discuss an approach to approximation of concepts and vague dependencies specified in a given concept ontology [15]. In the ontology concepts and local dependencies between them are specified. Global dependencies can be derived from local dependencies. Such derivations can be used as hints in searching for relevant compound patterns (information granules) in approximation of more compound concepts from the ontology.

The ontology approximation problem is one of the fundamental problems related to approximate reasoning in distributed environments. One should construct (in a given language that is different from the ontology specification language) not only approximations of concepts from ontology but also vague dependencies specified in the ontology. It is worthwhile to mention that an ontology approximation should be induced on the basis of incomplete information about concepts and dependencies specified in the ontology. Information granule calculi based on rough sets have been proposed as tools making it possible to solve this problem.

One can distinguish local and global dependencies between vague concepts specified in a given ontology. By a local dependency we mean a dependency consisting of concepts in premisses in a sense "close" to the concept in the conclusion so that the process of inducing of classifiers and approximation of the dependency can be performed automatically from the partial information available about the concepts. If one would like to approximate a "global" dependency in which the vague concepts on the left hand side of dependency are "far" from concept on the right hand side then one should use additional information to bound the search for relevant patterns for concept approximation and dependency approximation. The solution in this case can be based on hierarchical learning (see, e.g., [22,2,3,5]).

Any concept from the left hand side of a given vague dependency is called its premise and the dependency conclusion is the concept from the right hand side of the dependency. By approximation of a given vague dependency we understood a method which allows for any object to compute the arguments "for" and "against" its membership to the dependency conclusion on the basis of analogous arguments relative to the dependency premisses. Any argument "for" or "against" is a compound information granule (pattern) consisting of a pattern together with a degree to which (at least) this pattern is included to the concept and a degree to which (at least) the analyzed object is included to the pattern. Any local scheme (production rule) (see, e.g., [16]) or rough mereological connective (see, e.g., [12]) yields the fusion result of arguments for premisses that is next taken as the argument for the dependency conclusion. By composition of local schemes more advanced fusion schemes are obtained, called approximate reasoning schemes (AR schemes) (see, e.g., [2,16,12,19]). They show how the arguments from premisses of dependencies are fused to arguments for more compound concepts derived in a given ontology from premisses. AR schemes can correspond to different parts of complex spatio-temporal objects. Hence, there is a need for composing AR schemes for parts into AR schemes for objects composed from these parts [19].

We assume that there are distinguished some primitive concepts in a given ontology for which it is possible to derive arguments "for" and "against" from experimental data tables (e.g., with sensory attributes). To estimate the degree to which a given object belongs (at least) to a given concept C from ontology there are collected arguments "for" and "against" by using appropriate AR schemes for this concept C and next is used a conflict resolution strategy for predicting the degree.

Observe also that the discovered information granules (patterns) can be used to specify different regions of the object universe on which different "parts" of approximation of a given vague dependency can be expressed in a more relevant way.

Now, we present more details on approximation of concepts and dependencies.

Patterns for a more complex concept C can be constructed along derivation of C from C_1, C_2 using the existing dependencies in ontology. The derivation helps gradually to construct patterns for more compound concepts in the derivation from less compound concepts (closer to C_1, C_2).

Let us now consider as an example of the dependency:

$$\text{If } C_1 \text{ and } C_2 \text{ then } C, \tag{1}$$

where C_1, C_2, C are vague concepts. We assume that examples of positive and negative cases for such concepts are given. We also assume that the condition attributes for C_1, C_2 are specified. To approximate the target concept C relevant patterns should be derived. The main idea is presented in Figure 1. We assume that basic information granules (basic patterns) used for approximation of concepts C_1, C_2 can be induced for C_1, C_2. Such patterns can be defined, e.g., by left hand sides of decision rules with decisions corresponding to the concepts C_1, C_2 and to their complements. For such basic granules, one can define operations of construction of more complex patterns relevant for approximation of the target concept C. The relevant patterns can be obtained by tuning parameters of the operations. One of the most important kind of such operations is defined by the constrained sums of information systems specified by patterns [18]. These operations filter objects satisfying the constraint from objects satisfying basic patterns.

For discovered patterns the degrees of their inclusion into the considered concepts are estimated.

Next, some rules are derived that make it possible to predict the degrees of inclusion of objects to target patterns (i.e., discovered for the dependency

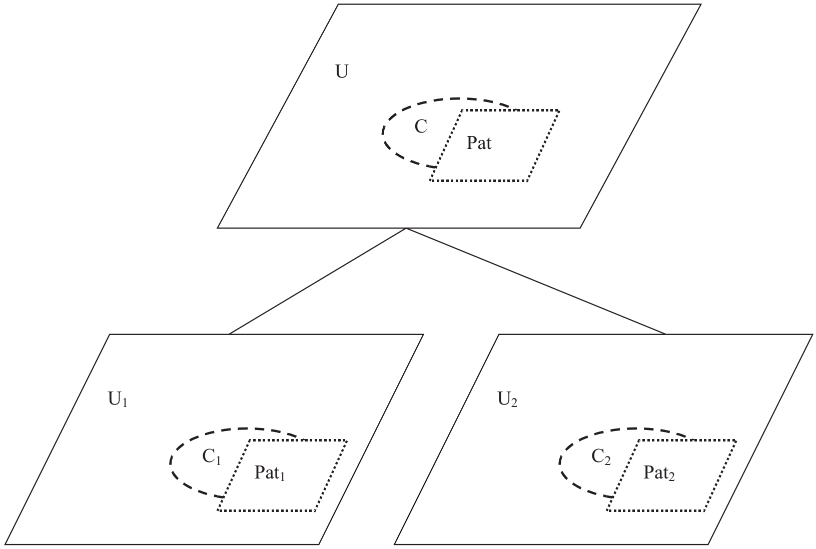

Fig. 1. Vague Concepts in U_1, U_2, U and Patterns

conclusion) from degrees of inclusion to source patterns (i.e., discovered for the dependency premisses) used for construction of the target patterns. Such rules are of the following form:

$$\text{If the degree of inclusion of } x \text{ in } Pat_i \text{ is at least } deg_i \text{ for } i = 1, 2$$

$$\text{then the degree of inclusion of } x \text{ in } Pat \text{ is at least } deg$$

where Pat is a pattern constructed from Pat_1 and Pat_2 and deg_1, deg_2, deg, are the degrees of inclusion of x in patterns Pat_1, Pat_2, Pat, respectively, and Pat_1, Pat_2, Pat are relevant patterns discovered for approximation of concepts $C_1, , C_2, C$, respectively (for the details, the reader is referred to [14]).

The discovered patterns with their degrees of inclusion into concepts are used in the construction of classifiers. The degrees of inclusion of patterns in the concepts and objects in patterns are used by a conflict resolution strategy to predict the decision, i.e., to decide if the analyzed concept belongs to a concept or not. This is the standard procedure for construction of classifiers. In our example, the procedure of a classifier construction is performed for concepts C_1, C_2, C. In such a construction of classifiers inductive reasoning is used.

One can take into account some concordance conditions between strategies for conflict resolution in the constructed classifiers for C_1, C_2 and C. By tuning such conditions one can optimize the approximation of vague dependency on different object regions. This idea is depicted in Figure 2 (see also [5,3] for an application of weights defined by rule votes in extracting relevant patterns).

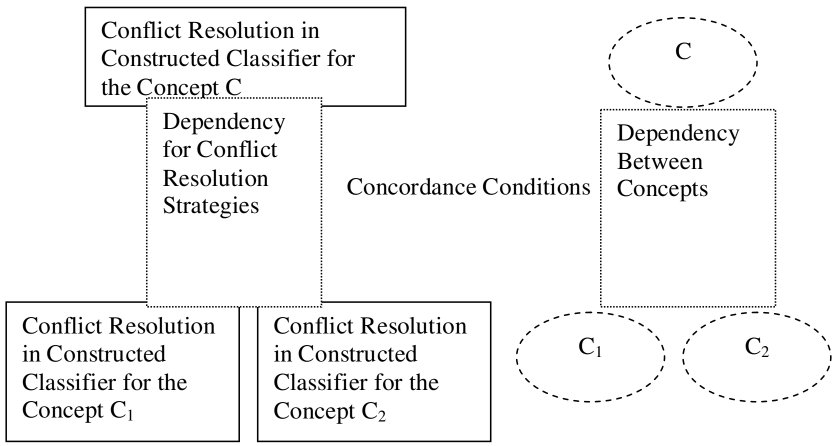

Fig. 2. Dependencies and Concordance Conditions

By granulation of discovered local schemes (production rules) more compound local schemes (production rules) can be discovered for approximation of concepts and dependencies. Such local schemes can represent dependencies between different layers of vague concepts. In this case one can use an approach based on the rough–fuzzy approach. To explain this idea we outline the approach using rough–fuzzy granules. These granules make it possible to derive a family of dependencies approximating a given dependency between vague concepts. The family consists of dependencies corresponding to different layers of the vague concepts.

Let us now discuss shortly an example of rough–fuzzy granules. Let $DT = (U, A, d)$ be a decision table with a binary decision $d : U \longrightarrow \{0, 1\}$, i.e., d is the characteristic function of some $X \subseteq U$. If the decision table is inconsistent [9], then one can define a new decision deg such that $deg(x) \in [0, 1]$ for any $x \in U$, may be interpreted as a degree to which x belongs to X [9,16]. Let us consider such a new decision table $DT' = (U, A, deg)$.

For given reals $0 < c_1 < \ldots < c_k$, where $c_i \in (0, 1]$ for $i = 1, \ldots, k$, we define c_i-cut by $X_i = \{x \in U : \nu(x) \geq c_i\}$. Assume that $X_0 = U$ and $X_{k+1} = X_{k+2} = \emptyset$. Any $B \subseteq A$ satisfying the following condition:

$$UPP(AS_B, (X_i - X_{i+1})) \subseteq (X_{i-1} - X_{i+2}), \text{ for } i = 1, \ldots, k, \quad (2)$$

is called relevant for approximation of cuts $0 < c_1 < \ldots < c_k$ in DT'.

The condition (2) expresses the fact that the boundary region of the set between any two successive cuts is included into the union of this set and two adjacent to it such sets.

The language $\mathcal{L}_{rf}$ of rough–fuzzy patterns for DT' consists of tuples $(B, c_1, \ldots, c_k)$ defining approximations of regions between cuts, i.e.,

$$(LOW(AS_B, (X_i - X_{i+1})), UPP(AS_B, (X_i - X_{i+1}))), \text{ for } i = 0, \ldots, k, \quad (3)$$

where we assume that B is relevant for approximation of cuts $0 < c_1 < \ldots < c_k$ in DT'.

Observe that searching for relevant patterns describing regions between cuts is related to tuning parameters $(B, c_1, \ldots, c_k)$ to obtain relevant patterns for the target concept approximation.

From a concept description in DT' (on a sample U) one can induce the concept approximation on an extension $U^* \supseteq U$. We consider, in a sense, richer classifiers, i.e., the classifiers that make it possible to predict different degrees to which the concept is satisfied. Such degrees can correspond to linguistic terms (e.g., low, medium, high) linearly ordered and to the boundary regions between successive degrees.

Any local scheme corresponding to a local dependency between vague concepts can be considered as a family of transducers satisfying a monotonicity property with respect to the linear order of linguistic degrees. This property can be expressed as follows:

if the granule corresponds to a linguistic membership degree *deg* of the dependency conclusion and it is constructed by the local scheme from some more elementary granules corresponding to some membership degrees $(deg_1, \ldots, deg_n)$

then this local scheme will yield a granule corresponding to *deg* at least from granules corresponding to $(deg'_1, \ldots, deg'_n)$ satisfying $deg'_i \geq deg_i$ for $i = 1, \ldots, n$, where $\geq$ denotes the linear order between linguistic degrees.

There are several problems to be solved to construct relevant layers of vague concepts for the approximation of dependencies between vague concepts. Among them are:

1. the problem of inducing relevant layers for the concepts;
2. the problem of inducing classifiers for the layers;
3. the monotonicity problem (the family of dependencies should satisfy the monotonicity property).

The discovered layers should make it possible to represent dependencies between vague concepts on the regions corresponding to the layers. We have developed strategies discovering such dependencies. Here, one can find an analogy to a neuron but a much more advanced neuron than that used in artificial neural networks [10].

We can conclude that taking fuzzy sets as models for vague concepts one can use rough sets for their constructive approximation. Then information granules we are searching for are families of approximations of different layers of concepts and dependencies between such approximated layers.

Observe that the family of dependencies discussed above may also be used in reasoning about changes. A linear order between layers corresponding to linguistic degrees of concept membership makes it possible to predict the degree of inclusion of a given object x to C on the basis of changes of degrees to which x is included in C_1, C_2.

Conclusions

We have discussed an ontological approach to the approximation problem. This approach covers the approximation of vague concepts and dependencies specified in a given ontology. Our ontology is presented in the framework of information granule calculi. The outlined methods for hierarchical construction of patterns, classifiers and vague dependency approximation create the basic step in our current project aiming at developing approximate reasoning methods in distributed or multiagent systems.

Acknowledgements

The authors gratefully acknowledge the suggestions made by Jim Peters concerning this paper. The research has been supported by the grants 3 T11C 002 26 and 4 T11C 014 25 from Ministry of Scientific Research and Information Technology of the Republic of Poland.

References

1. Barwise, J., Seligman, J.: *Information Flow: The Logic of Distributed Systems*. Cambridge University Press Tracts in Theoretical Computer Science 44, 1997.
2. Bazan, J., Skowron, A.: Classifiers based on approximate reasoning schemes. In: Dunin-Keplicz, B., Jankowski A., Skowron A., and Szczuka M., (Eds.), *Monitoring, Security, and Rescue Tasks in Multiagent Systems MSRAS*, Advances in Soft Computing. Springer, Heidelberg, 2005, 191–202.
3. Bazan, J., Nguyen, S. Hoa, Nguyen, H. Son, Skowron, A.: Rough set methods in approximation of hierarchical concepts. *Proc. of RSCTC'2004*, LNAI **3066**, Springer, Heidelberg, 2004, 346–355.
4. Doherty, P., Łukaszewicz, W., Skowron, A., Szałas, A.: Approximation transducers and trees: A technique for combining rough and crisp knowledge. In: [10], 189–218.
5. Nguyen, S. Hoa, Bazan, J., Skowron, A., Nguyen, H. Son: Layered learning for concept synthesis. *Transactions on Rough Sets I: LNCS Journal Subline*, LNCS **3100**, Springer, Heidelberg, 2004, 187–208,
6. Kloesgen, W., Żytkow, J. (eds.): *Handbook of Knowledge Discovery and Data Mining*. Oxford University Press, Oxford, 2002.
7. Lin, T.Y.: Granular computing and binary relations I & II. In: Skowron A. and Polkowski L. (Eds.), *Rough Sets in Knowledge Discovery* **1-2**, Physica-Verlag, Heidelberg, 1998, 107–120 and 121–140.
8. Łukasiewicz, J.: Die logischen Grundlagen der Wahrscheinilchkeitsrechnung, Kraków 1913. In: Borkowski, L. (ed.), *Jan Łukasiewicz - Selected Works*. North Holland, Amstardam, Polish Scientific Publishers, Warsaw, 1970.
9. Pawlak, Z.: *Rough Sets. Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publishers, Dordrecht, 1991.
10. Pal, S.K., Polkowski, L., Skowron, A. (Eds.): *Rough-Neural Computing: Techniques for Computing with Words*. Springer-Verlag, Berlin, 2004.
11. Polkowski, L., Skowron, A.: Rough mereology: A new paradigm for approximate reasoning. *Journal of Approximate Reasoning* **15**(4), 1996, 333–365.

12. Polkowski, L., Skowron, A.: Towards adaptive calculus of granules. In: [24], 201–227.
13. Skowron, A., Stepaniuk, J.: Tolerance approximation spaces. *Fundamenta Informaticae* **27**, 1996, 245–253.
14. Skowron, A., Stepaniuk, J.: Information granules: Towards foundations of granular omputing. *International Journal of Intelligent Systems* **16**(1), 2001, 57–86.
15. Staab, S., Studer, R., (Eds.): *Handbook on Ontologies.* International Handbooks on Information Systems, Springer, Heidelberg, 2004.
16. Skowron, A., Stepaniuk, J.: Information granules and rough-neural computing. In [10], 43–84.
17. Skowron, A., Stepaniuk, J., Peters, J.F.: Rough sets and infomorphisms: Towards approximation of relations in distributed environments. *Fundamenta Informaticae* **54**(1-2), 2003, 263–277.
18. Skowron, A., Stepaniuk, J.: Constrained sums of information systems. In: *Proc. RSCTC 2004*, LNCS **3066**, Springer , Heidelberg, 2004, 300–309.
19. Skowron, A., Synak, P.: Complex patterns. *Fundamenta Informaticae* **60**(1-4), 2004, 351–366.
20. Skowron, A., Swiniarski, R., Synak, P.: Approximation spaces and information granulation. *Transactions on Rough Sets III: LNCS Journal Subline*, LNCS **3400**, Springer, Heidelberg, 2005, 175–189.
21. Stepaniuk, J.: Knowledge discovery by application of rough set models. In: L. Polkowski, S. Tsumoto, T.Y. Lin (Eds.), *Rough Set Methods and Applications. New Developments in Knowledge Discovery in Information Systems*, Physica–Verlag, Heidelberg, 2000, 137–233.
22. Stone, P.: *Layered Learning in Multi-Agent Systems: A Winning Approach to Robotic Soccer.* MIT Press, Cambridge, MA, 2000.
23. Zadeh, L.A.: Toward a theory of fuzzy information granulation and its certainty in human reasoning and fuzzy logic. *Fuzzy Sets and Systems* **90** (1997) 111–127.
24. Zadeh, L.A., Kacprzyk, J. (eds.): *Computing with Words in Information/Intelligent Systems* **1-2**, Physica-Verlag, Heidelberg, 1999.
25. Zadeh, L.A.: A new direction in AI: Toward a computational theory of perceptions. *AI Magazine* **22**(1), 2001, 73–84.
26. Ziarko, W.: Variable precision rough set model, *Journal of Computer and System Sciences* **46**, 1993, 39–59.