Rough Sets and Higher Order Vagueness

Andrzej Skowron[1] and Roman Swiniarski[2,3]

[1] Institute of Mathematics, Warsaw University,
Banacha 2, 02-097 Warsaw, Poland
skowron@mimuw.edu.pl
[2] Institute of Computer Science, Polish Academy of Sciences,
Ordona 21, 01-237 Warsaw, Poland
[3] Department of Mathematical and Computer Sciences,
San Diego State University,
5500 Campanile Drive San Diego, CA 92182, USA
rswiniar@sciences.sdsu.edu

Abstract. We present a rough set approach to vague concept approximation within the adaptive learning framework. In particular, the role of extensions of approximation spaces in searching for concept approximation is emphasized. Boundary regions of approximated concepts within the adaptive learning framework are satisfying the higher order vagueness condition, i.e., the boundary regions of vague concepts are not crisp. There are important consequences of the presented framework for research on adaptive approximation of vague concepts and reasoning about approximated concepts. An illustrative example is included showing the application of Boolean reasoning in adaptive learning.

Keywords: vagueness, rough sets, higher order vagueness, adaptive learning.

1 Introduction

There is a long debate in philosophy on vague concepts [2]. Nowadays, computer scientists are also interested in vague (imprecise) concepts. Lotfi Zadeh [20] introduced a very successful approach to vagueness. In this approach, sets are defined by partial membership in contrast to crisp membership used in the classical definition of a set. Rough set theory [4] expresses vagueness not by means of membership but by employing the boundary region of a set. If the boundary region of a set is empty it means that a particular set is crisp, otherwise the set is rough (inexact). The non-empty boundary region of the set means that our knowledge about the set is not sufficient to define the set precisely. A discussion on vagueness in the context of fuzzy sets and rough sets can be found in [8]. In this paper some consequences on understanding of vague concepts caused by inductive extensions of approximation spaces and adaptive concept learning are outlined. This paper is an extension of [10]. In particular, we discuss a problem of adaptive learning of concept approximation assuming that learning is performed in a dynamic environment with many concepts that are linked by vague dependencies.

D. Ślęzak et al. (Eds.): RSFDGrC 2005, LNAI 3641, pp. 33–42, 2005.

2 Approximation Spaces and Their Inductive Extensions

In [4] any approximation space is defined as a pair (U, R), where U is a universe of objects and $R \subseteq U \times U$ is an indiscernibility relation defined by an attribute set.

The lower approximation, the upper approximation and the boundary region are defined as crisp sets. It means that the higher order vagueness condition is not satisfied [2]. We will return to this issue in Section 3.

We use the definition of approximation space introduced in [11]. Any approximation space is a tuple $AS = (U, I, \nu)$, where U is the universe of objects, I is an uncertainty function, and ν is a measure of inclusion called the inclusion function, generalized in rough mereology to the rough inclusion [11,13].

In this section, we consider the problem of approximation of concepts over a universe U^*, i.e., subsets of U^*. We assume that the concepts are perceived only through some subsets of U^*, called samples. This is a typical situation in machine learning, pattern recognition, or data mining [1]. In this section we explain the rough set approach to induction of concept approximations. The approach is based on inductive extension of approximation spaces.

Let $U \subseteq U^*$ be a finite sample and let $C_U = C \cap U$ for any concept $C \subseteq U^*$. Let $AS = (U, I, \nu)$ be an approximation space over the sample U. The problem we consider is how to extend the approximations of C_U defined by AS to approximation of C over U^*. We show that the problem can be described as searching for an extension $AS^* = (U^*, I^*, \nu^*)$ of the approximation space AS relevant for approximation of C. This requires showing how to induce values of the extended inclusion function to relevant subsets of U^* that are suitable for the approximation of C. Observe that for the approximation of C, it is enough to induce the necessary values of the inclusion function ν^* without knowing the exact value of $I^*(x) \subseteq U^*$ for $x \in U^*$.

We consider an example for rule-based classifiers. However, the analogous considerations for k-NN classifiers, feed-forward neural networks, and hierarchical classifiers [1]) show that their construction is based on the inductive inclusion extension [13,10].

Let AS be a given approximation space for C_U and let us consider a language L in which the neighborhood $I(x) \subseteq U$ is expressible by a formula $pat(x)$, for any $x \in U$. It means that $I(x) = \|pat(x)\|_U \subseteq U$, where $\|pat(x)\|_U$ denotes the meaning of $pat(x)$ restricted to the sample U. In the case of rule-based classifiers, patterns of the form $pat(x)$ are defined by feature value vectors.

We assume that for any new object $x \in U^* \setminus U$, we can obtain (e.g., as a result of a sensor measurement) a pattern $pat(x) \in L$ with semantics $\|pat(x)\|_{U^*} \subseteq U^*$. However, the relationships between information granules over U^*, e.g., $\|pat(x)\|_{U^*}$ and $\|pat(y)\|_{U^*}$, for different $x, y \in U^*$, are known only to a degree estimated by using relationships between the restrictions of these sets to the sample U, i.e., between sets $\|pat(x)\|_{U^*} \cap U$ and $\|pat(y)\|_{U^*} \cap U$.

The set of patterns $\{pat(x) : x \in U\}$ is usually not relevant for approximation of the concept $C \subseteq U^*$. Such patterns can be too specific or not general enough, and can directly be applied only to a very limited number of new sample elements. However, by using some generalization strategies, one can search in a family of patterns definable from $\{pat(x) : x \in U\}$ in L, for such new patterns that are relevant for approximation of concepts over U^*. Let us consider a subset $PATTERNS(AS, L, C) \subseteq L$ chosen as a set of pattern candidates for relevant approximation of a given concept C. For rule based classifiers one can search for such candidate patterns among sets definable by subsequences of feature value vectors corresponding to objects from the sample U. The set $PATTERNS(AS, L, C)$ can be selected using some quality measures evaluated on meanings (semantics) of patterns from this set restricted to the sample U (like the numbers of examples from the concept C_U and its complement that support a given pattern). Then, on the basis of properties of sets definable by these patterns over U, we induce approximate values of the inclusion function $\nu^*(X, C)$ on subsets of $X \subseteq U^*$ definable by any such pattern and the concept C. Next, we induce the value of ν^* on pairs (X, Y) where $X \subseteq U^*$ is definable by a pattern from $\{pat(x) : x \in U^*\}$ and $Y \subseteq U^*$ is definable by a pattern from $PATTERNS(AS, L, C)$.

Finally, for any object $x \in U^* \setminus U$ we induce the degree $\nu^*(\|pat(x)\|_{U^*}, C)$ applying a conflict resolution strategy $Conflict\_res$ (e.g, a voting strategy) to two families of degrees:

$$\{\nu^*(\|pat(x)\|_{U^*}, \|pat\|_{U^*}) : pat \in PATTERNS(AS, L, C)\}, \tag{1}$$

$$\{\nu^*(\|pat\|_{U^*}, C) : pat \in PATTERNS(AS, L, C)\}. \tag{2}$$

Values of the inclusion function for the remaining subsets of U^* can be chosen in any way – they do not have any impact on the approximations of C. Moreover, observe that for the approximation of C we do not need to know the exact values of uncertainty function I^* – it is enough to induce the values of the inclusion function ν^*. The defined extension ν^* of ν to some subsets of U^* makes it possible to define an approximation of the concept C in a new approximation space AS^*.

Observe, that the value $\nu^*(I^*(x), C)$ of the induced inclusion function for any object $x \in U^* - U$ is based on collected arguments *for* and *against* belonging of x to C. In this way, the approximation of concepts over U^* can be explained as a process of searching for relevant approximation spaces, in particular inducing relevant approximation spaces.

3 Approximate Reasoning About Vague Concepts Based on Adaptive Learning and Reasoning

We have recognized that for a given concept $C \subseteq U^*$ and any object $x \in U^*$, instead of crisp decision about the relationship of $I^*(x)$ and C, we can gather some arguments *for* and *against* it only. Next, it is necessary to induce from such arguments the value $\nu^*(I(x), C)$ using some strategies making it possible to resolve conflicts between those arguments [1,12]. Usually some general principles are used such as the minimal length principle [1] in searching for algorithms computing an extension $\nu^*(I(x), C)$. However, often the approximated concept over $U^* - U$ is too compound to be induced directly from $\nu(I(x), C)$. This is the reason that the existing learning methods can be not satisfactory for inducing high quality concept approximations in case of complex concepts [17]. There have been several attempts trying to omit this drawback. One of them is the incremental learning used in machine learning and also by the rough set community (see, e.g., [18]). In this case, an increasing sequence of samples $U_1 \subseteq \ldots \subseteq U_k \subseteq \ldots$ is considered and the task is to induce the extensions $\nu^{(1)}, \ldots, \nu^{(k)}, \ldots$ of inclusion functions. Still we know rather very little about relevant strategies for inducing such extensions. Some other approaches are based on hierarchical (layered) learning [14] or reinforcement learning [16]. However, there are several issues, important for learning that are not within the scope of these approaches. For example, the target concept can gradually change over time and this concept drift is a natural extension for incremental learning systems toward adaptive systems. In adaptive learning it is important not only what we learn but also how we learn, how we measure changes in a distributed environment and induce from them adaptive changes of constructed concept approximations. The adaptive learning for autonomous systems became a challenge for machine learning, robotics, complex systems, and multiagent systems. It is becoming also a very attractive research area for the rough set approach.

In general, from given information about the approximated concept C, the approximation space AS related to this information is constructed and next an extension AS^* of AS is induced. The induced approximations are only temporary, usually not matching exactly the approximated concept (even if we assume that the concept can be defined but its definition is unknown during learning). This means that the approximations will be necessary to change if some new arguments *for* and *against* will be gathered and an information or knowledge about the approximated concept will be updated. Hence, we should express a risk in prediction of decisions on the basis of the induced classification algorithms (classifiers) based on AS^* rather than exact decisions only. Such risk depends on negotiation strategies between arguments *for* and *against*, searching strategies for relevant patterns used for concept approximation, etc. This aspect is related to the higher order of vagueness [2]. Its consequence is that lower approximations, upper approximations, and boundary regions for vague concepts are not crisp.

Let us consider now some examples of adaptive concept approximation schemes.

Example 1. In Figure 1 we present an example of adaptive concept approximation scheme Sch. By $Inf(C)$ and $Inf'(C)$ we denote information about the approximated concept (e.g., decision table for C or training sample) in different (relevant) moments of time[1]. ENV denotes an environment, DS is an operation constructing an approximation space $AS_{Inf(C)}$ from a given sample $Inf(C)$. IN is an extension operation transforming the approximation space $AS_{Inf(C)}$ to an approximation space AS^* for the concept C; Q denotes a quality measure for the

[1] For simplicity, in Figure 1 we do not present time constraints.

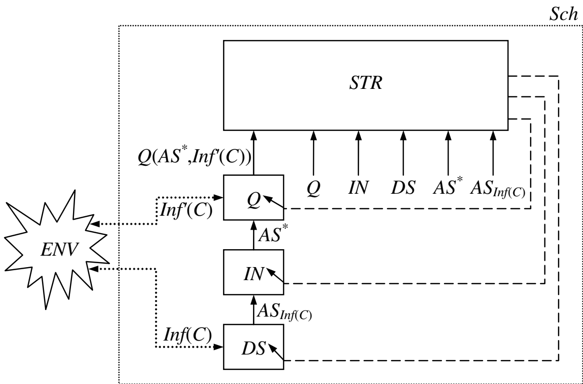

Fig. 1. An example of adaptive concept approximation scheme

induced approximation space AS^* on a new sample $Inf'(C)$. STR is a strategy that adaptively changes the approximation of C by modifying Q, IN, and DS.

The scheme Sch describes an adaptive strategy ST modifying the induced approximation space AS^* with respect to the changing information about the concept C. To explain this in more detail, let us first assume that a procedure $new_C(ENV, u)$ is given returning from the environment ENV and current information u about the concept C a new piece of information about this concept (e.g., an extension of a sample u of C). In particular, $Inf^{(0)}(C) = new_C(ENV, \emptyset)$ and $Inf^{(k+1)}(C) = new_C(ENV, Inf^{(k)}(C))$ for $k = 0, \ldots$. In Figure 1 $Inf'(C) = Inf^{(1)}(C)$. Next, assuming that operations $Q^{(0)} = Q$, $DS^{(0)} = DS$, $IN^{(0)} = IN$ are given, we define $Q^{(k+1)}$, $DS^{(k+1)}$, $IN^{(k+1)}$, $DS^{(k+1)}(Inf^{(k+1)}(C))$, and $AS^{*(k+1)}$ for $k = 0, \ldots,$ by

$$(Q^{(k+1)}, DS^{(k+1)}, IN^{(k+1)}) = \tag{3}$$

$$= STR(Q^{(k)}(AS^{*(k)}, Inf^{(k+1)}(C)), Q^{(k)}, IN^{(k)}, DS^{(k)}, AS^{*(k)}, AS^{(k)}_{Inf^{(k)}(C)})$$

$$AS^{(k+1)}_{Inf^{(k+1)}(C)} = DS^{(k+1)}(Inf^{(k+1)}(C)); \quad AS^{*(k+1)} = IN^{(k+1)}(AS^{(k+1)}_{Inf^{(k+1)}(C)}).$$

One can see that the concept of approximation space considered so far should be substituted by a more complex one represented by the scheme Sch making it possible to generate a sequence of approximation spaces $AS^{*(k)}$ for $k = 1, \ldots$ derived in an adaptive process of approximation of the concept C. One can also treat the scheme Sch as a complex information granule [12].

One can easily derive more complex adaptive schemes with metastrategies that make it possible to modify also strategies.

Example 2. In Figure 2 there is presented an idea of a scheme where a metastrategy MS can change adaptively also strategies STR_i in schemes Sch_i for $i = 1, \ldots, n$ where n is the number of schemes. The metastrategy MS can be, e.g., a fusion strategy for classifiers corresponding to different regions of the concept C. Even more compound scheme can be obtained by considering strategies

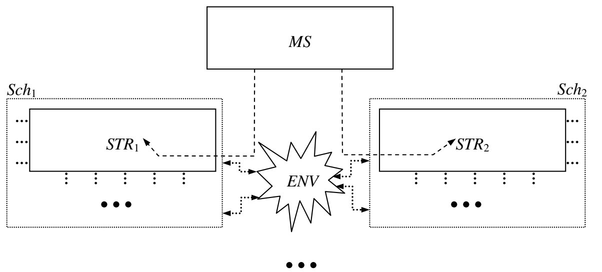

Fig. 2. An example of metastrategy in adaptive concept approximation

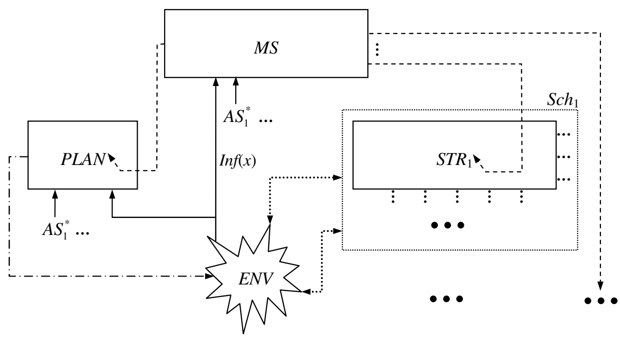

Fig. 3. An example of adaptive plan scheme

based on cooperation among the schemes for obtaining concept approximations of high quality. In Figure 3 an adaptive scheme for plan modification is presented. $PLAN$ is modified by a metastrategy MS that adaptively changes strategies in schemes Sch_i where $i = 1, \ldots, n$. This is performed on the basis of the derived approximation spaces AS_i^* induced for concepts that are guards of plan instructions and on the basis of information $Inf(x)$ about the state x of the environment ENV. The generated approximation spaces together with the plan structure are adaptively adjusted to make it possible to achieve plan goals.

The above examples are showing that the context in which sequences of approximation spaces are generated can have complex structure represented by relevant adaptive schemes.

There are some important consequences of our considerations for research on approximate reasoning about vague concepts. It is not possible to base such reasoning only on *static* models of the concepts (i.e., approximations of given concepts [4] or membership functions [20] induced from a sample available at a given moment) and on multi-valued logics widely used for reasoning about rough sets or fuzzy sets (see, e.g., [6,20,3,21]). Instead of this we need evolving systems of logics that in open and changing environments will make it possible to gradually acquire knowledge about approximated concepts and reason about them. Along this line an important research perspective arises. Among interesting topics are strategies for modeling of networks supporting such approximate reasoning (e.g., AR schemes and networks [12] can be considered as a step toward developing such strategies), strategies for adaptive revision of such networks, foundations for autonomous systems based on vague concepts.

Some recently reported results on rough sets seem to be important for developing foundations for adaptive systems. In particular, we would like to mention approximate reasoning in distributed environments based on rough mereological and granular approaches (see, e.g., [7,12]) and investigations on reasoning about changes based on rough sets and granular computing.

4 An Example: Inducing Concept Descriptions Consistent with Constraints Specified by Experts

From our considerations it follows that adaptive learning should be performed in a dynamic environment in which different vague concepts are approximated and it is necessary to preserve constraints among them. In this section we consider an illustrative example that can be treated as a starting point to further investigations on adaptive learning.

We consider together with facts, which can be represented using decision tables, some dependencies between concept approximations. These dependencies are specified by experts and represent their domain knowledge. An example of such dependency is "if road is slippery and the speed of the car is high then there is a high chance that the accident will appear". A question arises how to induce the concept approximations using together the facts represented in data tables and such dependencies. One can develop strategies for inducing decision rules preserving the dependencies between approximated concepts or for tuning the generated decision rules to preserve such dependencies. We apply another approach based on some ideas of non-monotonic reasoning. We assume that together with data tables there is given expert knowledge specified by constraints or dependencies between approximated concepts. For example, let us consider for three decisions d_1, d_2, d_3 the following constraint:

- **if** $d_1 =$*high* **and** $d_2 =$*medium* **then** $d_3 = 1$; or
- **if** *with certainty* $d_1 =$*high* **and** *one can not exclude* $d_2 =$*medium* **then** *with certainty* $d_3 = 1$.

We propose a method based on Boolean reasoning for tuning the induced from data table sets of rules (received by rough set and Boolean reasoning methods, e.g., in the form of so called minimal rules) so that the new induced concept approximations will satisfy the additional constraints specified by experts. These constraints are in the form of dependencies between approximated concepts (e.g., decision classes) or their (lower, upper) approximations or boundary regions. Let us observe that the phrase *with certainty* can be expressed by the lower approximation; and the phrase *it can not be excluded that* the upper approximation of concepts. Here, we would like to explain the main idea by example.

Example 3. Let us consider a decision table presented in Table 1. We have the following minimal decision rules of the decision table:

Table 1. Decision Table DT

	a	b	c	d
x_1	0	1	2	0
x_2	0	2	1	0
x_3	0	1	1	0
x_4	1	2	0	1
x_5	2	1	0	1
x_6	1	1	0	1

$$r_1: \ a=0 \rightarrow d=0; \quad r_2: \ c=2 \rightarrow d=0; \quad r_3: \ c=1 \rightarrow d=0;$$
$$r_4: \ c=0 \rightarrow d=1; \quad r_5: \ a=1 \rightarrow d=1; \quad r_6: \ a=2 \rightarrow d=1;$$

Let us consider the following constraint: $non(d=0 \wedge d=1)$. One can see that it is necessary to resolve conflict between left hand sides of the following pairs of rules: r_1, r_4; r_2, r_5; r_2, r_6; r_3, r_5 and r_3, r_6. These conflicts arise because conjunctions of left hand sides of listed pairs of rules are consistent (i.e., they do not include subformulas of the form $a = v \wedge a \neq v$). Hence, a new object can be matched by such rules and they will vote for different decisions 0 and 1, respectively. Let us consider the following propositional variables: $[i : a \neq v]$, $[i : a = v]$ with the intended meaning *left hand side of the rule r_i must be extended by $a \neq v$, $a = v$, respectively.* The conflicts can be encoded by the following propositional formula:

$$([1:c \neq 0] \vee [4:a \neq 0]) \wedge ([2:a \neq 1] \vee [5:c \neq 2]) \wedge ([2:a \neq 2] \vee [6:c \neq 2]) \wedge$$
$$([3:a \neq 1] \vee [5:c \neq 1]) \wedge ([3:a \neq 2] \vee [6:c \neq 1])$$

For example, the first part of the above formula describes a fact that the conflict between rules r_1, r_4 can be resolved by extending the left hand side of the rule r_1 by $c \neq 0$ or by extending the left hand side of the rule r_4 by $a \neq 0$. By computing (prime) implicants of this formula one can obtain all possible solutions, i.e., pairs of rule sets (approximating decision classes corresponding to $d = 0$ and $d = 1$) with resolved conflicts. In particular, let us consider the following implicant of the formula: $[1:c \neq 0] \wedge [2:a \neq 1] \wedge [2:a \neq 2] \wedge [3:a \neq 1] \wedge [3:a \neq 2]$. Hence, after a simplification, we obtain the following solution, i.e., a pair of rule sets:

$$r_{1'}: a=0 \wedge c=1 \rightarrow d=0; \quad r_{2'}: a=0 \wedge c=2 \rightarrow d=0; \text{ and}$$
$$r_4: \ c=0 \rightarrow d=1; \quad r_5: \ a=1 \rightarrow d=1; \quad r_6: \ a=2 \rightarrow d=1.$$

From example it follows that we can obtain different sets of rules resolving conflicts. One can look for pruning some solutions for conflict resolution using some criteria such as the rule support or descriptor occurrence frequencies on the left hand sides of the rules. Next, one can construct classifiers over such sets of rules and use them for classifying new objects using some fusion strategy. Another solution can start from generation of a sample of possible solutions (a family of sets of rules with eliminated conflicts) and next use strategies for conflict resolving between the sets of rules from the family.

A more advanced case of adaptive learning of a family of concepts is when the concepts are learned in a distributed environment consisting of distributed data tables and an additional data table with examples of "global" states, i.e., condition attribute value vectors over all data tables (see, e.g. [15]). Such vectors represent constraints for coexistence of condition attribute value vectors of data tables for different concepts. From the data table for global states one can induce rules representing constraints for local coexistence of attribute vector values from different data tables. These dependencies can be used as constraints for adaptive tuning of decision rules induced for different concepts.

For real-life data the formulas for conflict resolving can be large and efficient heuristics are necessary for solution construction. One can apply some strategies that have been developed using Boolean reasoning and rough sets [9]. Another approach can be based on decomposition of formulas using domain knowledge.

5 Conclusions

There are several conclusions from our discussion. Among them are:

1. Recognition of the importance of the inclusion function, generalized in rough mereology to rough inclusion (see, e.g., [7]). This has been used in investigations of information granule calculi, in particular those based on rough mereology (see, e.g., [12,7]) and approximation spaces based on information granules (see, e.g., [13]).
2. Observation that vague concepts cannot be approximated with satisfactory quality by *static* constructs such as induced membership inclusion functions, approximations or models derived, e.g., from a sample. Understanding of vague concepts can be only realized in a process in which the induced models are adaptively matching the concepts in a dynamically changing environment. This conclusion seems to have important consequences for further development of rough set theory in combination with fuzzy sets and other soft computing paradigms for adaptive approximate reasoning.

Acknowledgment

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. J. Friedman, T. Hastie, R. Tibshirani (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction.* Springer-Verlag, Heidelberg.
2. R. Keefe (2000). *Theories of Vagueness.* Cambridge Studies in Philosophy, Cambridge, UK.
3. J. Pavelka (1979). On Fuzzy Logic I-III. *Zeit. Math Logik Grund. Math.* **25**, 45–52, 119-134, 447-464.

4. Z. Pawlak (1991). *Rough Sets: Theoretical Aspects of Reasoning about Data.* System Theory, Knowledge Engineering and Problem Solving **9**, Kluwer Academic Publishers, Dordrecht.
5. Z. Pawlak, A. Skowron (1994). Rough Membership Functions. In R. R. Yager, M. Fedrizzi, J. Kacprzyk (eds.), *Advances in the Dempster-Schafer Theory of Evidence*, John Wiley and Sons, New York, 251–271.
6. L. Polkowski (2002). *Rough Sets: Mathematical Foundations.* Physica-Verlag, Heidelberg.
7. L. Polkowski, A. Skowron (1996). Rough Mereology: A New Paradigm for Approximate Reasoning. *International Journal of Approximate Reasoning* **15**, 333–365.
8. S. Read (1995). *Thinking about Logic. An Introduction to the Philosophy of Logic.* Oxford University Press, Oxford, New York.
9. A. Skowron (2000) Rough Sets in KDD. *16-th World Computer Congress* (IFIP'2000), Beijing, August 19-25, 2000, In: Zhongzhi Shi, Boi Faltings, Mark Musen (Eds.) Proceedings of Conference on Intelligent Information Processing (IIP2000), Publishing House of Electronic Industry, Beijing, 2000, 1–17.
10. A. Skowron (2005). Rough Sets and Vague Concepts. *Fundamenta Informaticae* **4(1-4)**, 417–431.
11. A. Skowron, J. Stepaniuk (1996). Tolerance Approximation Spaces. *Fundamenta Informaticae* **27**, 245–253.
12. A. Skowron, J. Stepaniuk (2004). Information Granules and Rough-Neural Computing. In S.K. Pal, L. Polkowski, A. Skowron (eds.), *Rough-Neural Computing: Techniques for Computing with Words*, Cognitive Technologies, Springer-Verlag, Heidelberg, 43–84.
13. A. Skowron, R. Swiniarski, P. Synak (2004). Approximation Spaces and Information Granulation. LNAI **3066**, Springer, Heidelberg, 114–123.
14. P. Stone (2000). *Layered Learning in Multi-Agent Systems: A Winning Approach to Robotic Soccer.* The MIT Press, Cambridge, MA.
15. Z. Suraj (1998). The synthesis problem of concurrent systems specified by dynamic information systems. In Polkowski, L., Skowron, A. (eds.), *Rough Sets in Knowledge Discovery 2: Applications, Case Studies and Software Systems*, Physica-Verlag, Heidelberg, 418–448.
16. R.S. Sutton, A.G. Barto (1998). *Reinforcement Learning: An Introduction.* MIT Press, Cambridge, MA.
17. V. Vapnik (1998). *Statistical Learning Theory.* John Wiley & Sons, New York, NY.
18. A. Wojna (2001). Constraint Based Incremental Learning of Classification Rules. LNAI **2005**, 428-435, Springer-Verlag, Heidelberg.
19. W. Ziarko (1993). Variable Precision Rough Set Model. *Journal of Computer and System Sciences* **46**, 39–59.
20. L.A. Zadeh (1965). Fuzzy sets. *Information and Control* **8**, 333–353.
21. L.A. Zadeh (1996). Fuzzy Logic = Computing with Words. *IEEE Transactions on Fuzzy Systems* **2**, 103–111.