INTERACTIVE ROUGH-GRANULAR COMPUTING IN PATTERN RECOGNITION

Andrzej Skowron
Marcin Wojnarski
Warsaw University
skowron@mimuw.edu.pl

Jan Bazan
Rzeszow University
bazan@univ.rzeszow.pl

PReMI 2009, Delhi, December 18-20, 2009

I would like to express my deepest gratitude to the Organizers of the PReMI 2009 conference, especially to
Professors Sankar K. Pal and Santanu Chaudhury
for the invitation
as well as to many colleagues and friends for discussion and/or collaboration, in particular to:
Zdzisław Pawlak (1926-2006),
Lotfi Zadeh,
Mohua Banerjee, Jan Bazan, Mihir Chakraborty, Anna Gomolinska, Andrzej Jankowski, Jiming Liu, Hung Son Nguyen, Tuan Trung Nguyen, Singh Hoa Nguyen,
Sankar K. Pal, James F. Peters, Lech Polkowski, Sheela Ramana, Dominik Slezak, Jaroslaw Stepaniuk, Zbigniew Suraj, Roman Swiniarski, Piotr Synak, Marcin Szczuka, Marcin Wojnarski, Jakub Wroblewski, YiYu Yao, Ning Zhong,
...

AGENDA

- Motivation for interactive rough-granular computation (IRGC)
- Granules and their interactions
  - elementary (atomic)
  - granules obtained by fusion of existing granules
    - relational structures (e.g., tolerance classes, approximation spaces) and their clusters
    - approximation of changes and trajectories of changes
    - rules of coexistence of local states: discovery of process models from data and domain knowledge
    - coalitions
- Interactive granules in approximation of complex concepts from data and domain knowledge
- Research topics:
  - searching for relevant interactive granules
  - adaptation in IRGC
  - discovery of interaction structures
- Software: RoughICE; TunedIT
- Conclusions: IRGC in WisTech program

Interaction is a fundamental dimension for modeling and engineering complex computational systems. More generally, interaction is a critical issue in the understanding of complex systems of any sorts: as such, it has emerged in several well-established scientific areas other than computer science, like biology, physics, social and organizational sciences.

Andrea Omicini, Alessandro Ricci, and Mirko Viroli, The Multidisciplinary Patterns of Interaction from Sciences to Computer Science. In: D. Goldin, S. Smolka, P. Wagner (eds.): Interactive computation: The new paradigm, Springer 2006

Why interactive computations on granules are needed?

[Existing] *Algorithms are metaphorically dumb and blind because they cannot adapt interactively while they compute*.

Peter Wegner: Why interaction is more powerful than algorithms. COMM ACM 40(5): (1997) 81-91

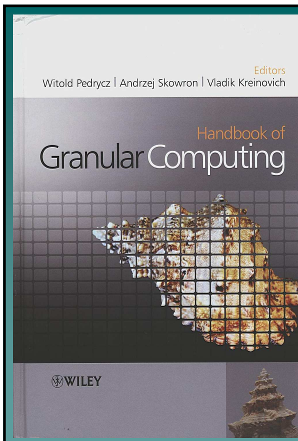
ADAPTIVE JUDGMENT

While employing IRGC, interactions and process mining we must stay in touch with the reality we are trying to model (describe) and predict.

If for some reason the decisions we are making are inconsistent with real life, we need to adapt our judgment.

COMPUTATIONS IN IRGC

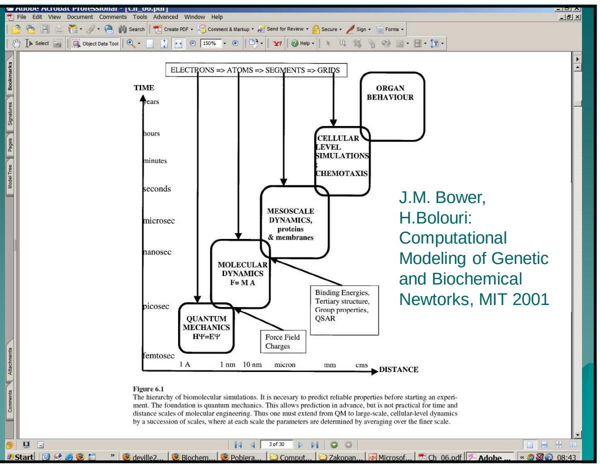
- Are performed on complex concepts called granules
  - Involve uncertainty, noise, vagueness
  - Manage parts of (descriptions and patterns for) complex concepts
- Are interactive
  - Performed by many autonomous, interacting units (agents)
  - Influenced by changes in data/knowledge and in the way co-operation goes.



Plays a key role in implementation of the strategy of divide-and-conquer in human problem-solving – Lotfi Zadeh

Over 1000 pages describing:

- Various approaches to granularity
- Foundations of GrC
- Methodologies and algorithms
- Applications
- …



J.M. Bower, H.Bolouri: Computational Modeling of Genetic and Biochemical Newtorks, MIT 2001

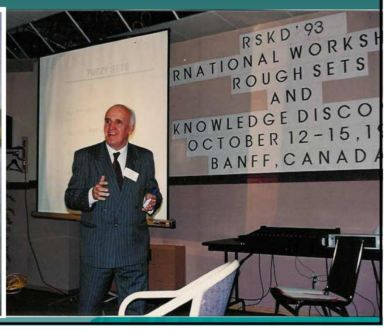
(INFORMATION) GRANULES: OBJECTS CONSTRUCTED IN THE GRANULATION AND DEGRANULATION PROCESSES

E.G., IN SEARCHING FOR COMPLEX CONCEPT APPROXIMATION

EXAMPLES OF GRANULES:

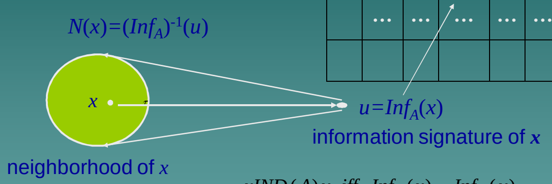
FROM
NEIGHBORHOODS OF OBJECTS TO CLUSTERS,
APPROXIMATION SPACES,
CLASSIFIERS,
ONTOLOGIES AND THEIR APPROXIMATION,
BEHAVIORAL PATTERNS,
PROCESS MODELS,
ADAPTIVE SCHEMES OF AGENTS

ROUGH SETS



INDISCERNIBILITY GRANULES

	a_1	a_2	…	a_m	d
x_1	v_1	v_2	…	v_m	1
	…	…	…	…	…



$xIND(A)y \;\; iff \;\; Inf_A(x)=Inf_A(y)$

τ

tolerance or similarity

APPROXIMATION SPACES

A. Skowron, J. Stepaniuk, Generalized Approximation Spaces 1994

$$AS = (U, N, \nu)$$

$N : U \rightarrow P(U)$ neighborhood function

$\nu : P(U) \times P(U) \rightarrow [0,1]$ rough inclusion

partial function

$x \rightarrow Inf(x) \rightarrow N(x) = Inf^{-1}(Inf(x))$

neighborhood of x

APPROXIMATION SPACE

$$AS = (U, N, \nu)$$

$$LOW(AS, X) = \{x \in U : \nu(N(x), X) = 1\}$$

$$UPP(AS, X) = \{x \in U : \nu(N(x), X) > 0\}$$

ROUGH MEREOLOGY

MEREOLOGY

St. LEŚNIEWSKI (1916)

x is_a_ part_of y

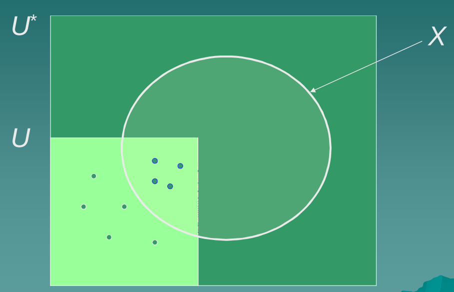
ROUGH MEREOLOGY

L. Polkowski and A. Skowron (1994-……)

x is_a_ part_of y in a degree

L. Polkowski, A. Skowron, Rough mereology, ISMIS'94, LNAI 869, Springer, 1994, 85-94

INDUCTION

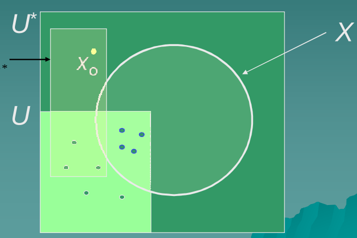


PARTITIONS → COVERINGS

$$N : U^* \rightarrow P(U^*)$$

$$Inf_A(x_o) = \{a=1, b=0, c=2\}$$

$\|a=1 \wedge b=0 \wedge c=2\|_{U^*}$



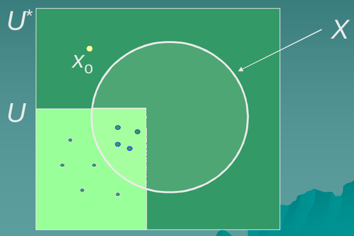
INDUCTION

What if $Inf_A(x_o)$ for $x_o \in U^* - U$ is different from any $Inf_A(x)$ for $x \in U$?

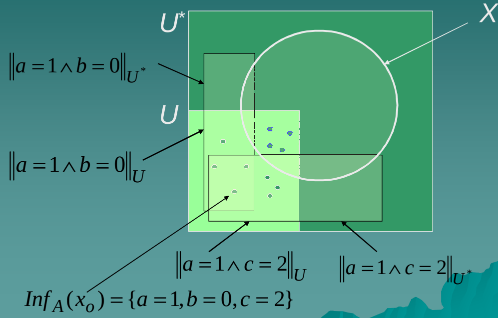
$\downarrow$

similarity of $Inf_A(x_o)$ with $Inf_A(x)$

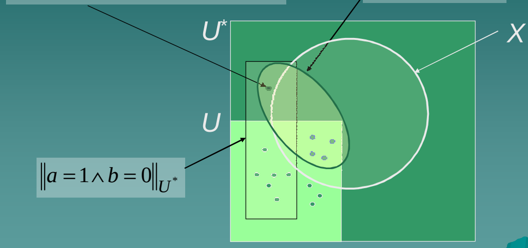
partial matching of $Inf_A(x_o)$ with $Inf_A(x)$



$$N : U^* \to P^2(U^*)$$



$$\nu(N(x_0), X) = ?$$



$$N : U \to P(U)$$

$\downarrow$ $\quad P_\omega(U^*) = \bigcup_{k \geq 1} P^k(U^*)$

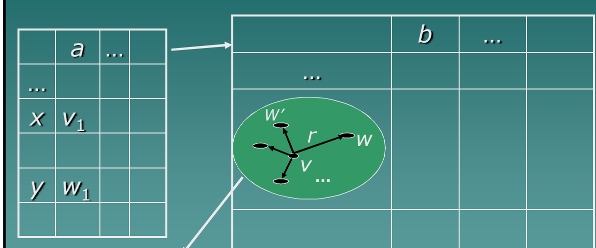
$$N : U^* \to P_\omega(U^*)$$

$$\nu : P(U) \times P(U) \to [0,1]$$

$\downarrow$ $\quad P_\omega(U^*) = \bigcup_{k \geq 1} P^k(U^*)$

$$\nu : P_\omega(U^*) \times P_\omega(U^*) \to [0,1]$$

GENERALIZATIONS OF GRANULES BY GRANULE FUSION: TOLERANCE GRANULES



$v = (v_1, ..., v_m)$; $w = (w_1, ..., w_m)$

$v\ r\ w$ iff $v_i r_i w_i$ for $i = 1, ... m$

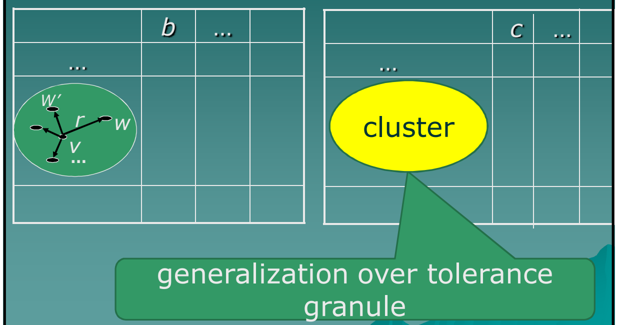
$r(v) = \{w : v\ r\ w\}$

$\|r(v)\| = \bigcup \{\|w\| : w \in r(v)\}$

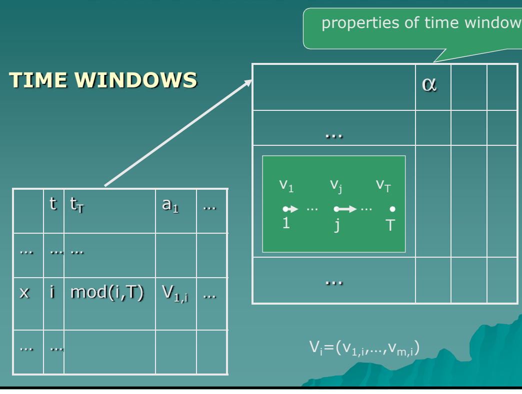
GENERALIZATION

from v to $r(v)$

GENERALIZATIONS OF TOLERANCE GRANULES
GENERALIZATION OPERATORS
by Ryszard Michalski



GRANULES REPRESENTING STRUCTURES OF OBJECTS



FUSION AND GENARALIZATION OF GRANULES

- fusion of granules:
  - tolerance granules
  - clusters over tolerance granules
  - relational structures and their clusters, e.g., approximation spaces
  - ...
  - degrees of matching: fusion, propagation
  - ...

INTERACTIVE COMPUTATIONS ON GRANULES IN DISCOVERY OF PROCESS MODELS FROM DATA AND DOMAIN KNOWLEDGE (PROCESS MINING)

Interacive processes

There are two components:

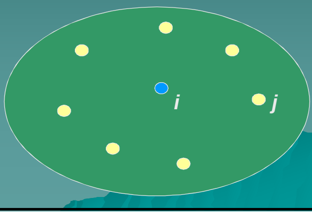
1. Process – changes of states of the system occur with time.
2. Interaction – the change of a given state in the process depends not only on time but also on exchange of information with other states.

INTERACTIVE GRANULES

- cellular automata
- differential equations
- approximation of changes
- MAS
- coexistence of local states
- interactions with experts
- reinforcement learning

INTERACTIONS REPRESENTED BY CHANGES OF LOCAL STATES. THE CHANGES ARE DEFINED BY INTERACTIONS OF LOCAL STATES IN NEIGHBORHOODS

e.g.,cellular automata



PROBLEMS

- States are complex and only uncertain information about them is available
- How to define neighborhoods?
- How to approximate changes in states as the results of interactions?
- …

DISCOVERY OF INTERACTION MODELS FROM DATA AND DOMAIN KNOWLEDGE

DEFINING INTERACTION vs INDUCING MODELS OF INTERACTION

PROCESS MODELS AND INTERACTIONS examples: coupled map lattice, oscilator

$$x_i(t+1) = f(x_i(t)) + \kappa \frac{1}{d_i} \sum_{\substack{j \\ j\sim i}} (f(x_j(t-\tau)) - f(x_i(t)))$$

neighborhood relation

$$\dot{x}_i(t) = f(x_i(t);\varepsilon) + \varepsilon\kappa \frac{1}{d_i} \sum_{\substack{j \\ j\sim i}} (x_j(t-\tau) - x_i(t))$$

Feng, J., Jost, J., Minping, Q. (eds): Network: From Biology to Theory, Springer, Berlin, 2007

?

$$\frac{ds}{dt} = G(t, s(t), e(t))$$

$$\frac{de}{dt} = H(t, s(t), e(t))$$

?

Approximation of functions G,H:

- rough, fuzzy methods
- statistical methods

DISCOVERY PROCESS MODELS FROM DATA:

METHODS FOR APPROXIMATION OF FUNCTIONS CHARACTERIZING CHANGES

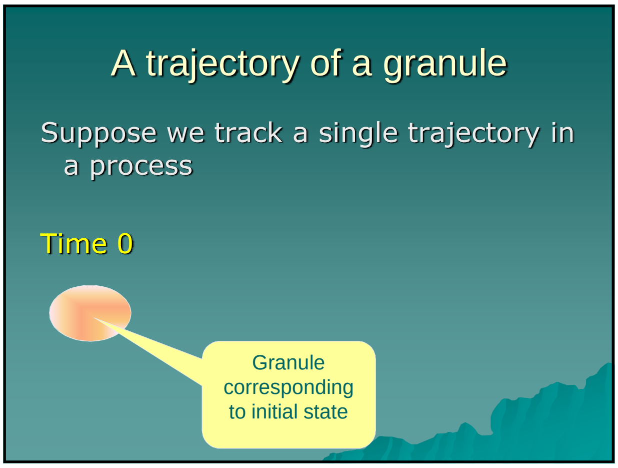
HIERARCHICAL LEARNING IS NEEDED !

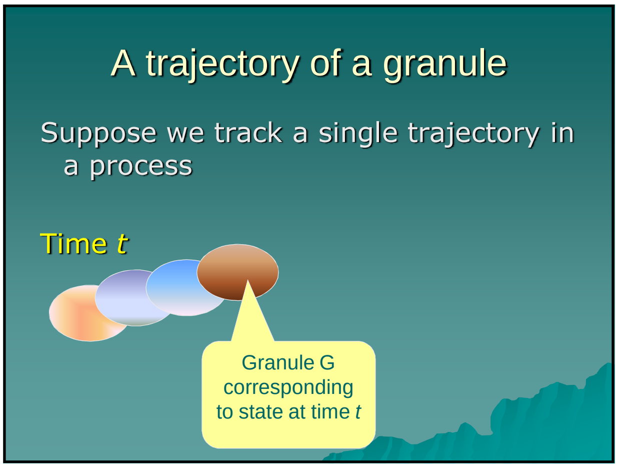
attributes relevant for characterizing changes

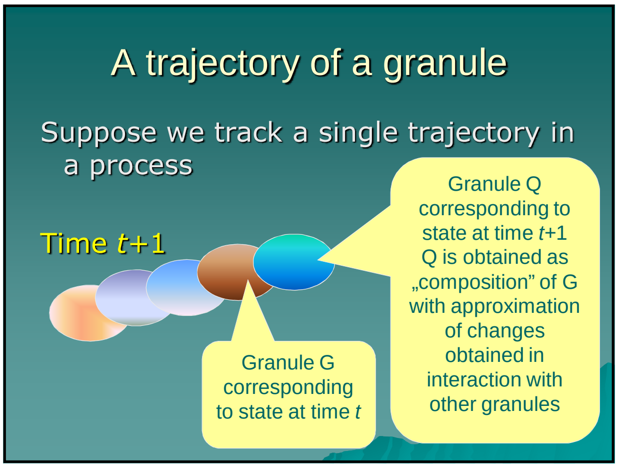
	a_1	…	a_m	d
(x,y)				
…				

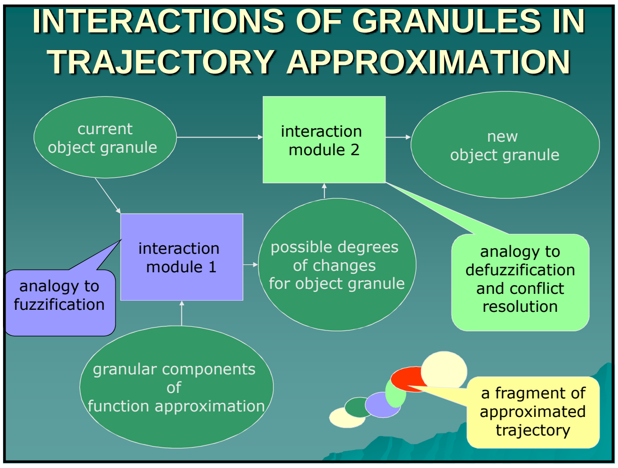
degrees of changes

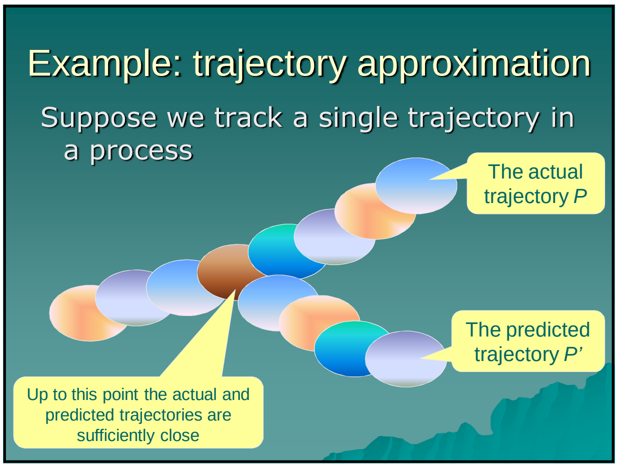
current and next configurations

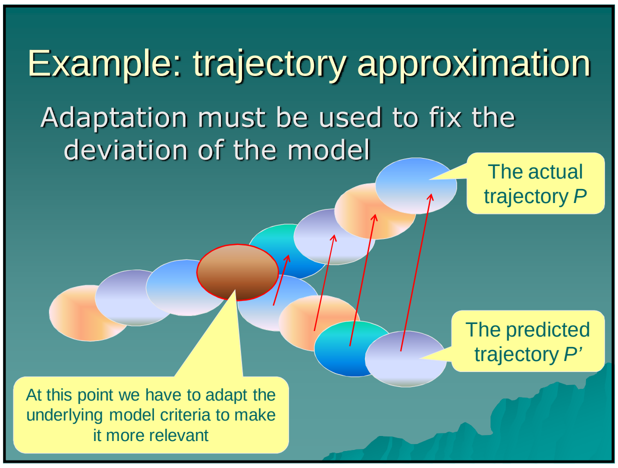

A trajectory of a granule
Suppose we track a single trajectory in a process
Time 0
Granule corresponding to initial state


A trajectory of a granule
Suppose we track a single trajectory in a process
Time *t*
Granule G corresponding to state at time *t*

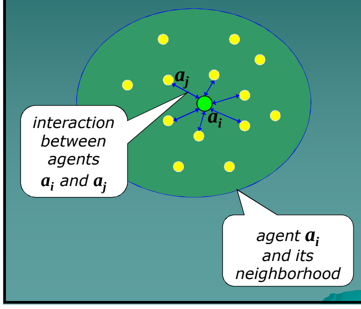

A trajectory of a granule
Suppose we track a single trajectory in a process
Time *t*+1
Granule G corresponding to state at time *t*
Granule Q corresponding to state at time *t*+1 Q is obtained as „composition" of G with approximation of changes obtained in interaction with other granules


INTERACTIONS OF GRANULES IN TRAJECTORY APPROXIMATION
current object granule
interaction module 2
new object granule
interaction module 1
analogy to fuzzification
possible degrees of changes for object granule
analogy to defuzzification and conflict resolution
granular components of function approximation
a fragment of approximated trajectory


Example: trajectory approximation
Suppose we track a single trajectory in a process
The actual trajectory *P*
The predicted trajectory *P'*
Up to this point the actual and predicted trajectories are sufficiently close


Example: trajectory approximation
Adaptation must be used to fix the deviation of the model
The actual trajectory *P*
The predicted trajectory *P'*
At this point we have to adapt the underlying model criteria to make it more relevant

INTERACTIONS IN MAS



- complex states
- partial information
- conflicts
- negotiations,
- cooperation
- coalition,
- competitions,
- intentions,
-…

$E := E^* := Perception\,(E_r);$

$G := G^* := Perception\,(G_r);$

while Property(G ,E) do

begin

$a := Select\_Action(G^*, G, E^*, E);$

$E^* := Predict\_Env(G, E, a);$

$G^* := Predict\_Gran(G, E, a);$

$G_n := I_a^{\leftarrow}(G_r, E_r);$

$E_n := I_a^{\rightarrow}(G_r, E_r);$

$G_r := G_n; E_r := E_n;$

$G := Perception\,(G_r);$

$E := Perception\,(E_r)$

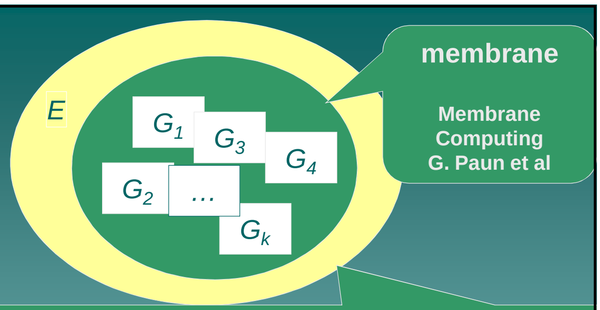
end

$cancel\,(G)$

Each granule has a scheme of interaction obtained by specifying:
- ***Property;***
- ***Select_Action;***
- ***Perception;***
- ***Predict_...***
- ***Cancel***

$I_a^{\leftarrow}$ and $I_a^{\rightarrow}$

are perceived only through *Perception.*



Granule fusion:
coalition of $G_1,\ldots G_k$

with functionalities for interaction with E and members of the coalition

---

INTERACTIONS OF GRANULES ARE BASED ON LOCAL LOGICS

- set of (high level) concepts with (approximate) rules of inference
- concepts and rules are adaptively changing

INTERACTIONS FORCED BY DEPENDENCIES OR RULES (DISCOVERED FROM DATA) PRESERVING COEXISTENCE OF LOCAL STATES IN CONCURRENT SYSTEMS

SPECIFICATION OF CONCURRENT SYSTEMS BY INFORMATION SYSTEMS

- Pawlak, Z.: Concurrent versus sequential the rough sets perspective. Bulletin of the EATCS 48 (1992) 178—190
- Skowron, A., Suraj, Z.: Rough sets and concurrency. Bull. Acad. Polon. Sci. 41(3) (1993) 237—254
- Suraj, Z.: Rough set methods for the synthesis and analysis of concurrent processes. In: L. Polkowski, S. Tsumoto, T.Y. Lin (eds), Rough Set Methods and Applications Studies in Fuzziness and Soft Computing 56, Springer/Physica Verlag (2000) 379-488

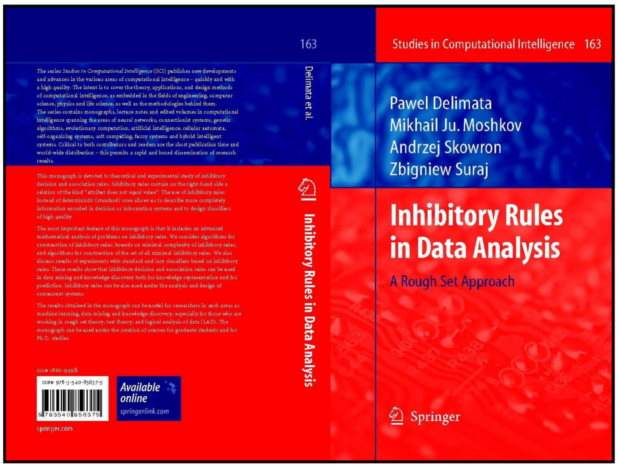
MAIN IDEA WE USE IN PROCERSS MINING

DATA A → THEORY OF DATA A Th(A) → PROCESS MODEL CONSISTENT WITH Th(A)

- DATA A: e.g., information system
- THEORY OF DATA A Th(A): e.g., set of rules defined by information system
- PROCESS MODEL CONSISTENT WITH Th(A): e.g., Petri net with reachability markings consistent with all rules in Th(A)

ADVANTAGES

- Complex Petri Nets can be generated automatically from their specification by data tables
- Petri Net can be adaptively modified with changes of data



CONTINUATION

- Which kinds of rules should be used (e.g., non-deterministic, probabilistic, temporal, spatio-temporal)?
- How to characterize the expressibility of different rule sets?
- How to extend the approach by adding information on transition relation or temporal dependencies?

Research by Z. Suraj and his team, also M. Moshkov and A. Skowron,…

DISCOVERY OF STRUCTURES OF INTERACTING PROCESSES ALONG DOMAIN ONTOLOGY

G set of paths of interaction of P_1 and P_2 → **PN** generating (consistent with) G

PN_1 generating (consistent with) G_1 ← G_1 set of paths of process P_1

PN_2 generating (consistent with) G_2 ← G_2 set of paths of process P_2

How PN is constructed from PN1 and PN2 ?

INTERACTION WITH EXPERTS

APPROXIMATION OF VAGUE COMPLEX CONCEPTS USING DOMAIN ONTOLOGY APPROXIMATION

…when you have a technical description x of the object and have some impression x about this object you have two forms of description: a formal description and a holistic description or Gestalt description. Using both descriptions during training can help to find a better decision function. This technique remains master-class learning, like musicians training in master classes. The teacher does not show exactly how to play. He talks to students and gives some images transmitting some hidden information - and this helps. So, the challenge is to create an algorithm which using additional information, will generalize better than classical algorithms.*

Vladimir Vapnik (2008): http://learningtheory.org

I believe that understanding human experience will be a driving challenge for work in AI in the years to come, and that the work that will result will profoundly impact our knowledge of how we live and interact with the world and with each other.

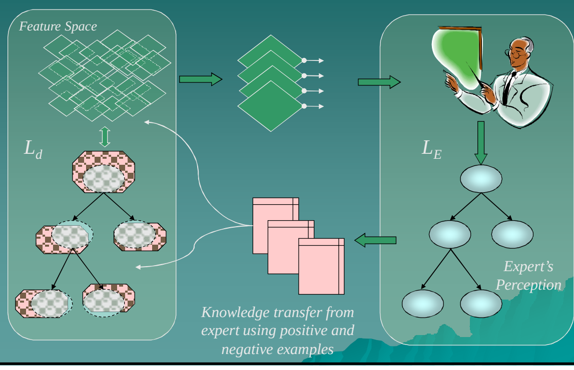
Henry Kautz (2005) Artificial Intelligence: The Next Twenty-Five Years, AI Magazine, 26(4): Winter 2005, 85–97

UNDERSTANDING THE ORGANIZATION AND PRINCIPLES OF HIGHER BRAIN FUNCTIONS: HIERARCHICAL LEARNING

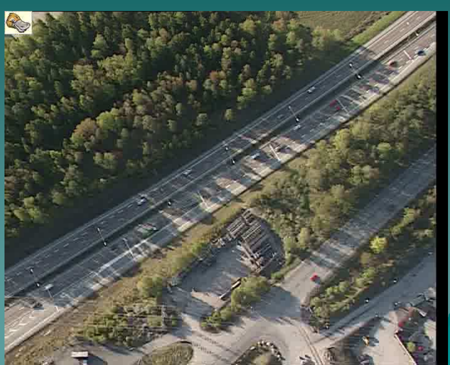
- Organization of cortex – for instance visual cortex –is strongly hierarchical.
- Hierarchical learning systems show superior performance in several engineering applications.
- This is just one of several possible connections, still to be characterized, between learning theory and the ultimate problem in natural science – the organization and the principles of higher brain functions.

T. Poggio, S. Smale: The Mathematics of Learning: Dealing with Data, Notices AMS, Vol.50, May 2003

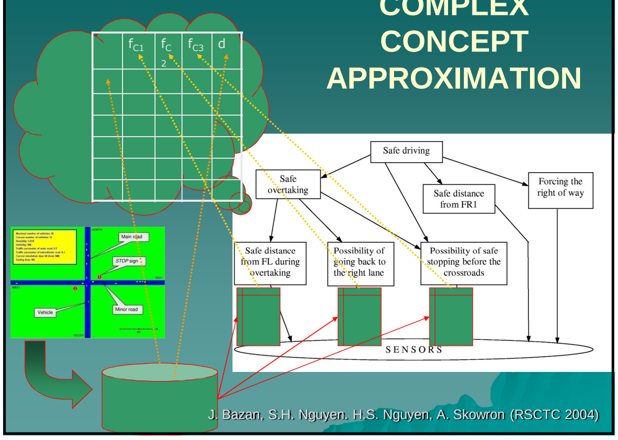
ROUGH SET BASED ONTOLOGY APPROXIMATION



UAV

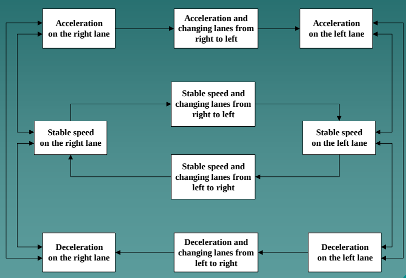


COMPLEX CONCEPT APPROXIMATION



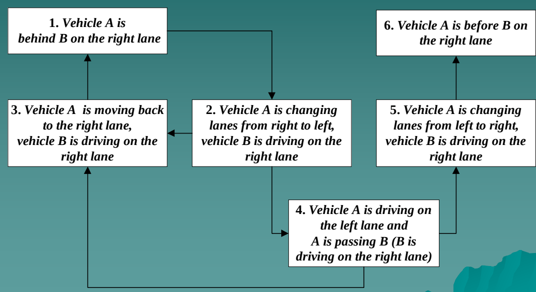
J. Bazan, S.H. Nguyen, H.S. Nguyen, A. Skowron (RSCTC 2004)

AN EXAMPLE OF BEHAVIORAL GRAPH FOR SINGLE VEHICLE



BEHAVIORAL GRAPH FOR A GROUP OF OBJECTS

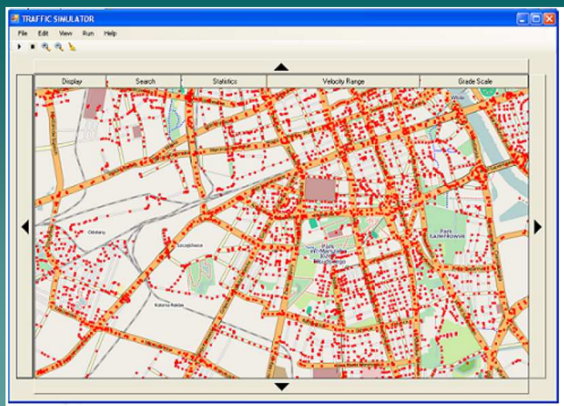
(TWO VEHICLE OF OBJECTS DURING OVERTAKING)



Results of experiments for concept: ''Is the vehicle driving safely?''

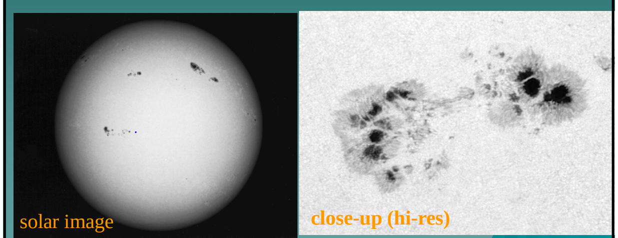
Decision class	Method	Accuracy	Coverage	Real accuracy
YES	RS1 RS2	**0.978** 0.962	0.946 **0.992**	0.925 **0.954**
NO	RS1 RS2	0.633 **0.862**	0.740 **0.890**	0.468 **0.767**
All classes (YES + NO)	RS1 RS2	**0.964** 0.958	0.935 **0.987**	0.901 **0.945**

Real accuracy = Accuracy * Coverage



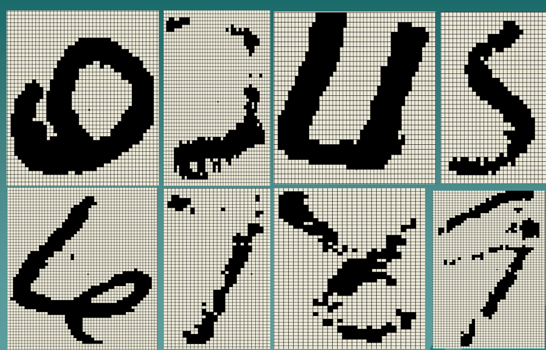
Pawel Gora

SUNSPOT CLASSIFICATION

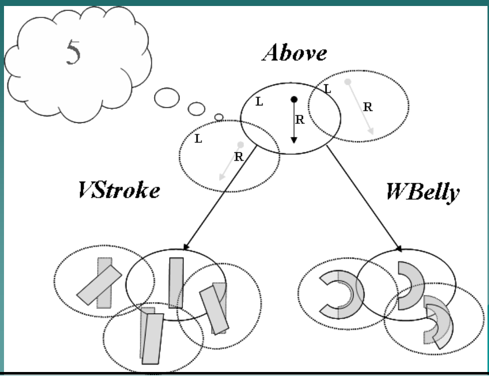


Son Nguyen et al

HARD SAMPLES



OUTLIER CLASSIFICATION



COMPLEX DYNAMIC SYSTEMS
(AUTONOMOUS MULTIAGENT SYSTEMS)

- *Systems of complex objects with the following features:*
  - objects are changing over time
  - dependencies between objects
  - cooperation between objects
  - objects able to perform flexible autonomous complex actions

- ***Examples:***
  - Complex dynamic system: *a patient (e.g., a newborn infant)*
  - Complex object: *a disease (e.g., respiratory failure)*

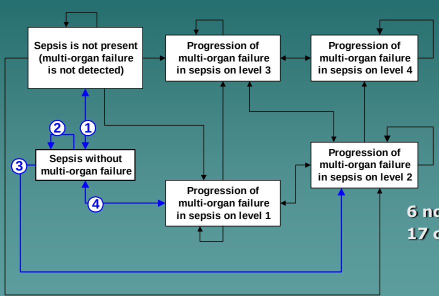
THE RESPIRATORY FAILURE

- **The respiratory failure develops when the rate of gas exchange between the atmosphere and blood is unable to match the body's metabolic demands**
- Arterial blood gas can be used to define respiratory failure – lower level of blood oxygen and accumulation of carbon dioxide
  - Clinical symptoms: increased rate of breathing, accessory respiratory muscles use, peripheral cyanosis
  - Other useful procedures: X-ray lung examination, lung biopsy, bronchoalveolar lavage, echocardiography

Data sets

- The experiments have been performed on the data sets obtained from *Neonatal Intensive Care Unit in Department of Pediatrics, Collegium Medicum, Jagiellonian University, Cracow.*
  - The data were collected between 2002 and 2004.
  - The detailed information about treatment of 340 newborns:
    - perinatal history, birth weight, gestational age, lab tests results, imagine techniques results, detailed diagnoses during hospitalization, procedures and medication.
- Train&test method has been performed to estimate accuracy, sensitivity and specificity.
  - A train data set consists of 5810 objects and a test data set consists of 5289 objects

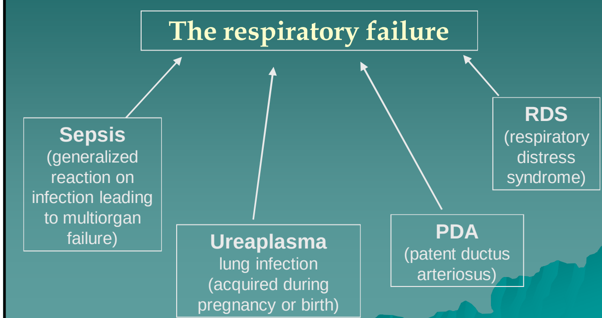
AN EXAMPLE OF BEHAVIORAL GRAPH
(the simple model of behavior for a single patient in sepsis)



6 nodes and 17 connections

Four possibilities of transition from the node: *Sepsis without multi-organ failure*

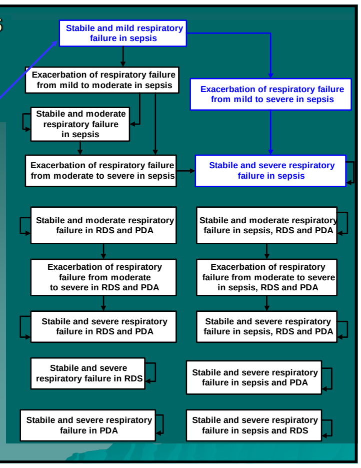
THE RESPIRATORY FAILURE AS A COMPLEX PROCESS



BEHAVIORAL GRAPH AS A BEHAVIORAL PATTERN

(the risk pattern of death due to respiratory failure)

- **The visualization of infant behavior** by a path in the behavioral graph.
- **Behavioral graph** (behavioral pattern) **as a classifier**



Results of experiments for the risk pattern of death due to respiratory failure

Decision class	Results
Yes (the high risk of death)	0.992 (sensitivity)
No (the low risk of death)	0.936 (specificity)
All classes (Yes + No)	0.956 (accuracy)

- Measures description:
  - ***sensitivity*** - the proportion those cases having a positive test result of all positive cases tested,
  - ***specificity*** - the proportion of true negatives of all the negative cases tested,
  - ***accuracy*** - the ratio of the number of all properly classified cases to the total number of tested cases.

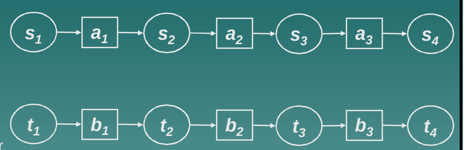
THE APPROACH WAS EXTENDED FOR AUTOMATED PLANNING OF TREATMENT OF INFANTS WITH RESPIRATORY FAILURE

- As a measure of planning success (or failure), we use the special classifier that can predict the similarity between two plans as a number between 0.0 and 1.0.
  - This classifier has been constructed on the basis of the ontology specified by human experts and clinical data sets
- The average similarity between plans for all tested situations was **0.82**

THE PROBLEM OF COMPARISON OF PLANS

Plan 1:
(e.g., proposed by human experts)

Plan 2:
(e.g., generated automatically by our computer system)

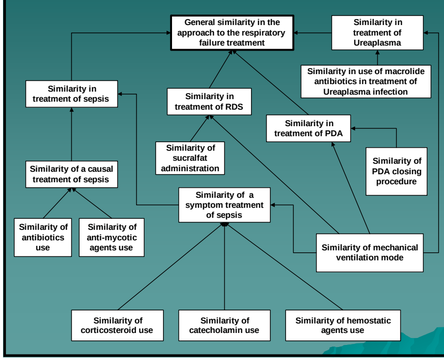


Problem: *How to compare Plan 1 and Plan 2?*

Solution: *A tool to estimate similarity between plans.*

AN EXAMPLE OF MEDICAL ONTOLOGY

TO SUPPORT THE ESTIMATION OF SIMILARITY BETWEEN PLANS OF THE TREATMENT OF NEWBORN INFANTS WITH THE RESPIRATORY FAILURE



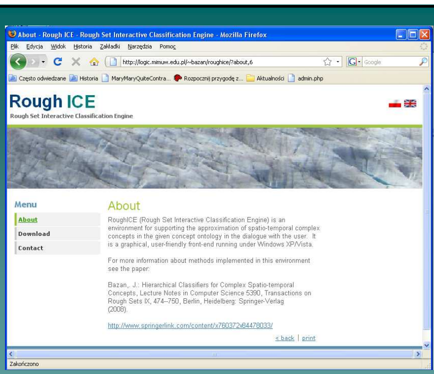
16 concepts and 18 connections

Any concept represents different aspect of similarity between medical plans

RESULTS OF EXPERIMENTS

FOR THE AUTOMATED PLANNING OF TREATMENT OF INFANTS WITH RESPIRATORY FAILURE

- As a measure of planning success (or failure) in our experiments, we use the special classifier that can predict the similarity between two plans as a number between 0.0 and 1.0
  - The classifier has been constructed on the basis of the ontology specified by human experts and data sets
- The average similarity between plans for all tested situations was **0.82**



On the WWW

http://logic.mimuw.edu.pl/~bazan/roughice/

TUNEDIT: www.tunedit.org

Automated evaluation of machine-learning and data-mining algorithms

Generation of reproducible experimental results
→ for high-quality research papers

Collaboration between researchers: sharing of algorithms, datasets, experimental results and other resources; project: IRGC in discovery of new features

Benchmarks of algorithms: currently stores performance data for nearly 100 algorithms tested on several tens of datasets. Included: Weka, Rseslib algorithms, UCI datasets

CONCLUSIONS
Wisdom Technology (WisTech) Program

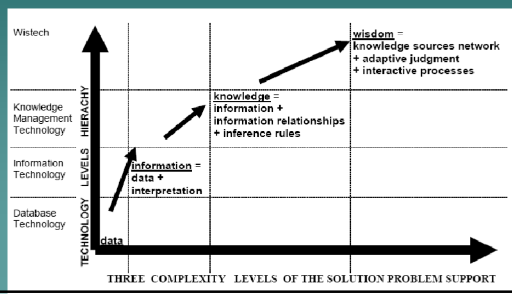
The basic meta-equation of WisTech

wisdom = network of knowledge sources + adaptive judgment + interactive processes

IRGC = systems based on interactive computations on granules with use of domain (expert) knowledge, process mining and concept learning

CONCLUSIONS

We discussed some issues of WisTech in the framework of ROUGH GRANULAR COMPUTING. In our further study we plan to develop foundations for WisTech based on RGC.



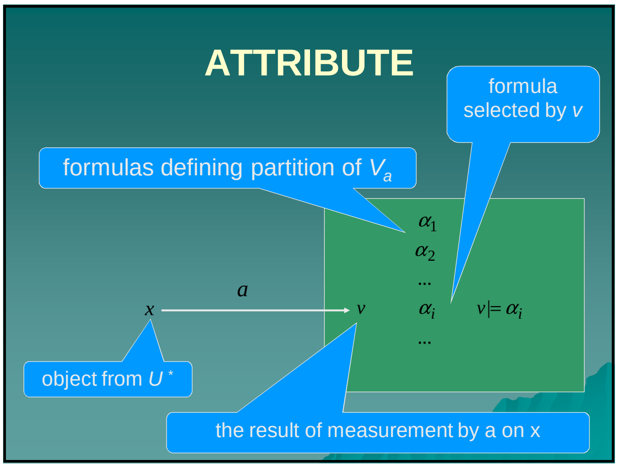
Why WisTech?

Aristotle's man of practical wisdom, the *phronimos*, … is observant of principles and, at the same time, open to their modification. He begins with *nomoi* – established law - and employs **practical wisdom** to determine how it should be applied in particular situations and when departures are warranted. Rules provide the guideposts for inquiry and critical reflection.

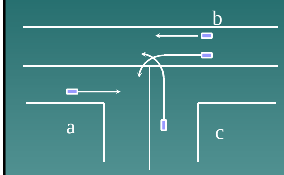
L. P. Thiele. The Heart of Judgment: Practical Wisdom, Neuroscience, and Narrative. Cambridge Univ. Press, 2006. p.5.

http://logic.mimuw.edu.pl/

THANK YOU !



EXAMPLE: LIGHT CONTROL



Local processes

	a	b	c
u_1	1	1	0
u_2	0	2	0
u_3	0	0	2

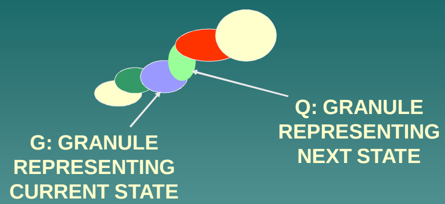
Global states

u_1, u_2, u_3 - states

a, b, c - movement directions

$0, 1, 2$ - light color (red, green, green arrow)

TRAJECTORY APPROXIMATION



Q= „COMPOSITION" of G with approximation of changes

- In some cases hints for adaptation can be acquired from experts but quite often they will be expressed in natural language and complex vague concepts will be involved in them. Such hints with vague complex concepts should be approximated.