FROM INFORMATION SYSTEMS TO INTERACTIVE INFORMATION SYSTEMS

Andrzej Skowron
Institute of Mathematics UW
&
IBS PAN

FedCSIS'16 Plenary Panel on the Legacy of Professor Zdzisław Pawlak, Gdańsk, September 11-14, 2016

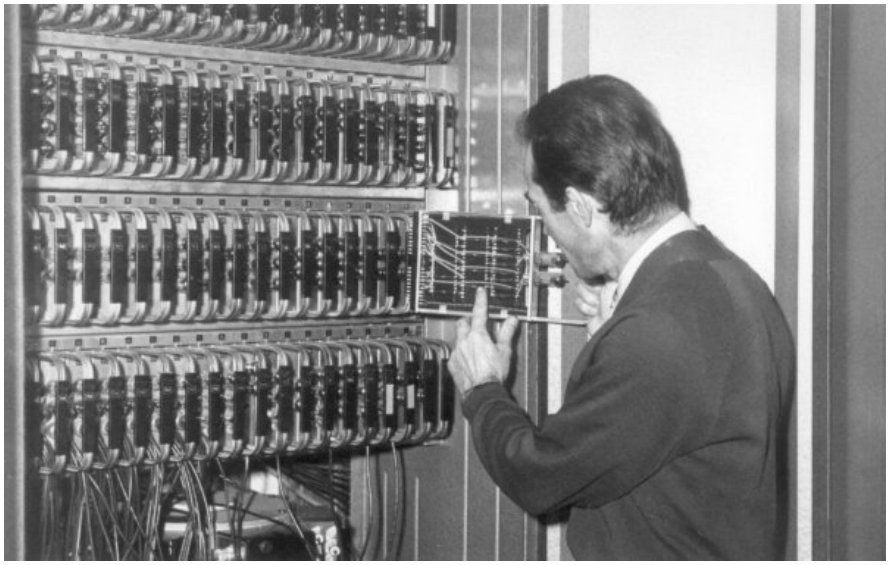
ZDZISŁAW PAWLAK (1926-2006)

Pawlak, Z.: Rough sets. International Journal of Computer and Information Sciences 11 (1982)

http://rsds.univ.rzeszow.pl/home: more than 20 000 articles

Google scholar returns more than 2.9 mln references to the frase *rough set* with more tan 240 000 references quoted in the last 3 years

UMC1: 1961



The original arithmetic for the UMC1 computer system with base “-2” was due to Pawlak.

Everybody associates his name with
rough sets, but it is very little known
that he is one of the pioneers of today
molecular computing, by his
chapter on genetics in his book in Polish
"Grammar and Mathematics"
published in the sixties.

Solomon Marcus
(Romanian Academy of Science)

CONFLICT ANALYSIS

issues

a – autonomous Palestinian state on the West Bank and Gaza
b – Israeli military outpost along the Jordan River
c – Israeli retains East Jerusalem
d – Israeli military outposts on the Golan Heights
e – Arab countries grant citizenship to Palestinians who choose to remain within their borders

agents

1 – Israel
2 – Egypt
3 – Palestinians
4 – Jordan
5 – Syria
6 – Saudi Arabia

U	a	b	c	d	e
1	$-$	$+$	$+$	$+$	$+$
2	$+$	0	$-$	$-$	$-$
3	$+$	$-$	$-$	$-$	0
4	0	$-$	$-$	0	$-$
5	$+$	$-$	$-$	$-$	$-$
6	0	$+$	$-$	0	$+$

Kiekrz 1992

RASIOWA - PAWLAK SCHOOL

AGENDA

INFORMATION SYSTEMS AND INFORMATION RETRIEVAL

INFORMATION SYSTEMS AND ROUGH SETS

INFORMATION SYSTEMS AND CONCURRENT SYSTEMS

INFORMATION SYSTEMS IN GRANULAR COMPUTING (GrC)
(e.g., HIERARCHICAL MODELING AND LEARNIG)

WHAT NEXT?
INTERACTIVE INFORMATION SYSTEMS
IN INTERACTIVE GRANULAR COMPUTING (IGrC)

INFORMATION SYSTEMS

- deterministic
- non-deterministic
- with missing values
- stochastic (probabilistic)
- distributed
- incremental
- …

INFORMATION SYSTEMS

	Age	LEMS
x1	16-30	50
x2	16-30	0
x3	31-45	1-25
x4	31-45	1-25
x5	46-60	26-49
x6	16-30	26-49
x7	46-60	26-49

- *IS* is a pair (*U*, *A*)
- *U* is a non-empty finite set of objects.
- *A* is a non-empty finite set of attributes such that $a : U \to V_a$ for every $a \in A$.
- V_a is called the value set of *a*.

INFORMATION SYSTEMS AND INFORMATION RETRIEVAL

PRACE CO PAN • CC PAS REPORTS

Zdzisław Pawlak

Mathematical foundations of information retrieval

101
1973
WARSZAWA

CENTRUM OBLICZENIOWE POLSKIEJ AKADEMII NAUK
COMPUTATION CENTRE POLISH ACADEMY OF SCIENCES
WARSAW, PKIN, P. O. Box 22, POLAND

This note contains a simple mathematical formulation of basic ideas concerning information retrieval and its computer implementation. The presented theory is based on the results given in [1], [2] and [3].

1. Descriptive systems

By a descriptive system we mean triplet $D = \langle A_D, X_D, \sigma_D \rangle$ (or briefly $D = \langle A, X, \sigma \rangle$), where

A - is a (finite or infinite) set; elements of A are called objects of D,

X - is a finite set of symbols; elements of X are referred to as elementary descriptors of D,

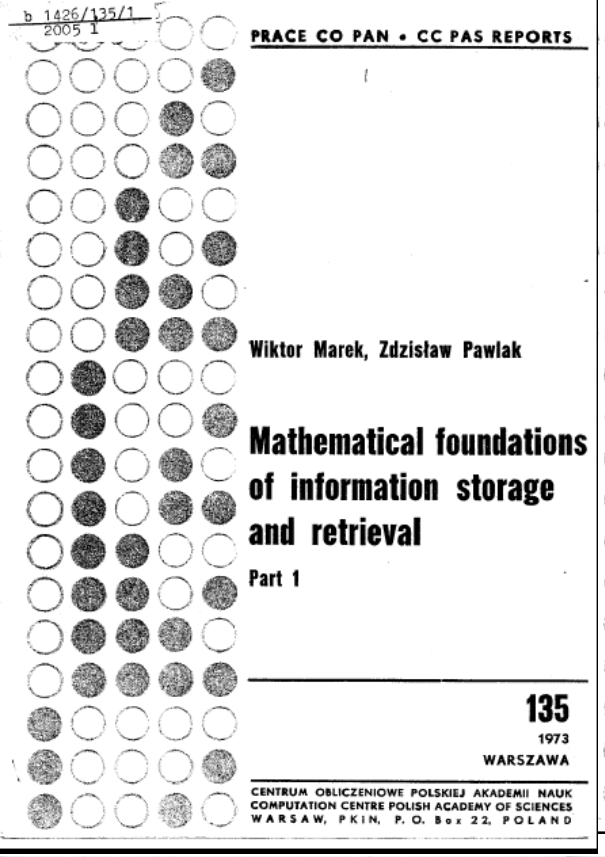
$\sigma \subseteq A \times X$ - is a binary relation, called description relation (or description) in D.

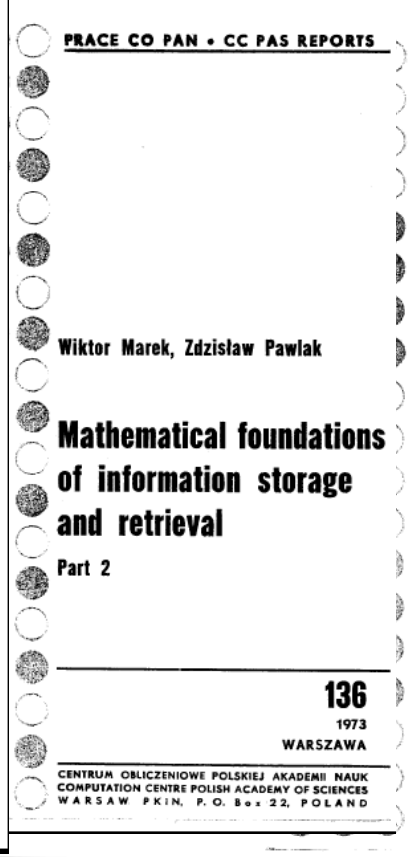
Relation σ may be replaced by the function:

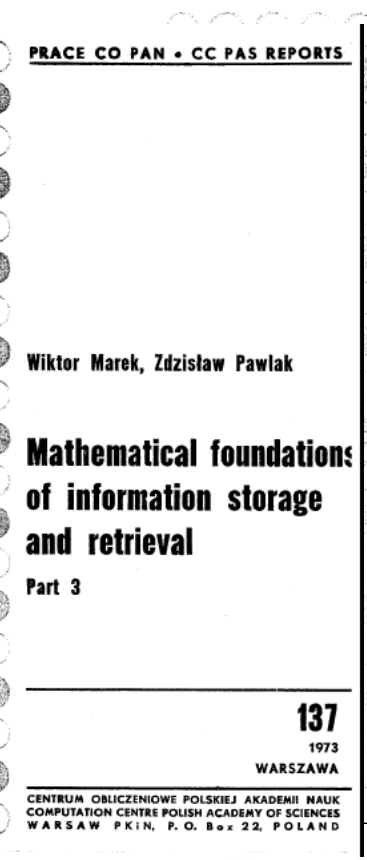
$$\Psi : X \to 2^A$$

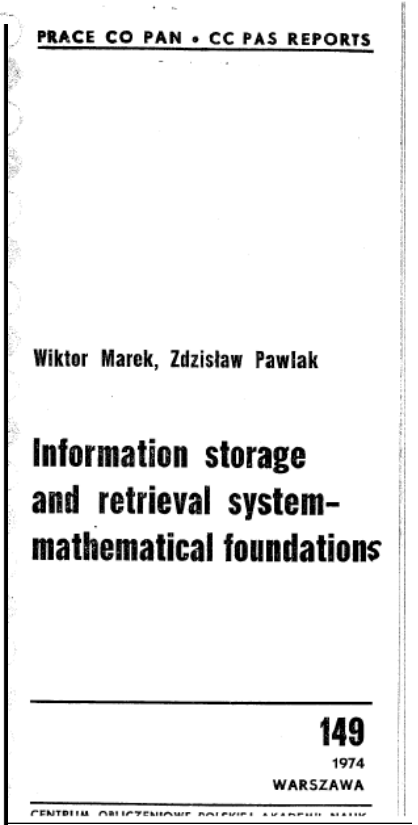
such that:

$$\Psi(x) = \{a \in A ; \sigma(a,x)\}.$$









Annales Societatis Mathematicae Polonae
Series IV: Fundamenta Informaticae VIII.1(1985)

ROUGH SETS AND INFORMATION SYSTEMS

Wiktor Marek
University of Warsaw
Zdzisław Pawlak
Institute of Computer Science PAS

Received October 25, 1982
AMS Categories: 68H05

Abstract: We apply rough sets to characterize definable subsets of the universe of the information system.

BULLETIN DE L'ACADÉMIE POLONAISE DES SCIENCES
Série des sciences math., astr. et phys. — Vol. XXII, No. 4, 1974

MATHEMATICS (COMPUTER SC.

On the Foundations of Information Retrieval

by

W. MAREK and Z. PAWLAK

Presented by A. MOSTOWSKI on October 5, 1973

Theoretical Computer Science 1 (1976) 331–354.

INFORMATION STORAGE AND RETRIEVAL SYSTEMS: MATHEMATICAL FOUNDATIONS

Wiktor MAREK and Zdzisław PAWLAK

Institute of Mathematics, Polish Academy of Sciences, Warsaw, Poland
Computation Center, Polish Academy of Sciences, Warsaw, Poland

Communicated by E. Engeler
Received April 1974
Revised January 1975

PRACE IPI PAN • ICS PAS REPORTS

Zdzisław Pawlak

Distributed information systems

370
1979
WARSZAWA

INSTYTUT PODSTAW INFORMATYKI POLSKIEJ AKADEMII NAUK
INSTITUTE OF COMPUTER SCIENCE POLISH ACADEMY OF SCIENCES
00-901 WARSAW, P. O. Box 22, POLAND

PRACE IPI PAN • ICS PAS REPORTS

Zdzisław Pawlak

Toward the theory of information systems

I. The notion of an information system

419
JUNE 1980
WARSZAWA

INSTYTUT PODSTAW INFORMATYKI POLSKIEJ AKADEMII NAUK
INSTITUTE OF COMPUTER SCIENCE POLISH ACADEMY OF SCIENCES
00-901 WARSAW, P.O. Box 22, POLAND

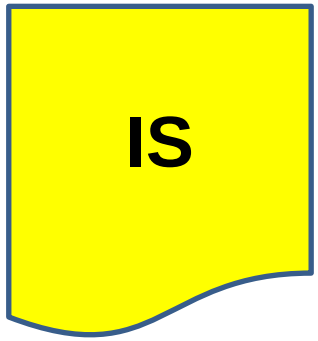
Zdzisław Pawlak

SYSTEMY INFORMACYJNE
Podstawy teoretyczne

WYDAWNICTWA NAUKOWO-TECHNICZNE • WARSZAWA 1983

Z. Pawlak, Information Systems. Theoretical Foundations, PWM Warsaw 1983 (in Polish)

INFORMATION RETRIEVAL



Boolean queries:

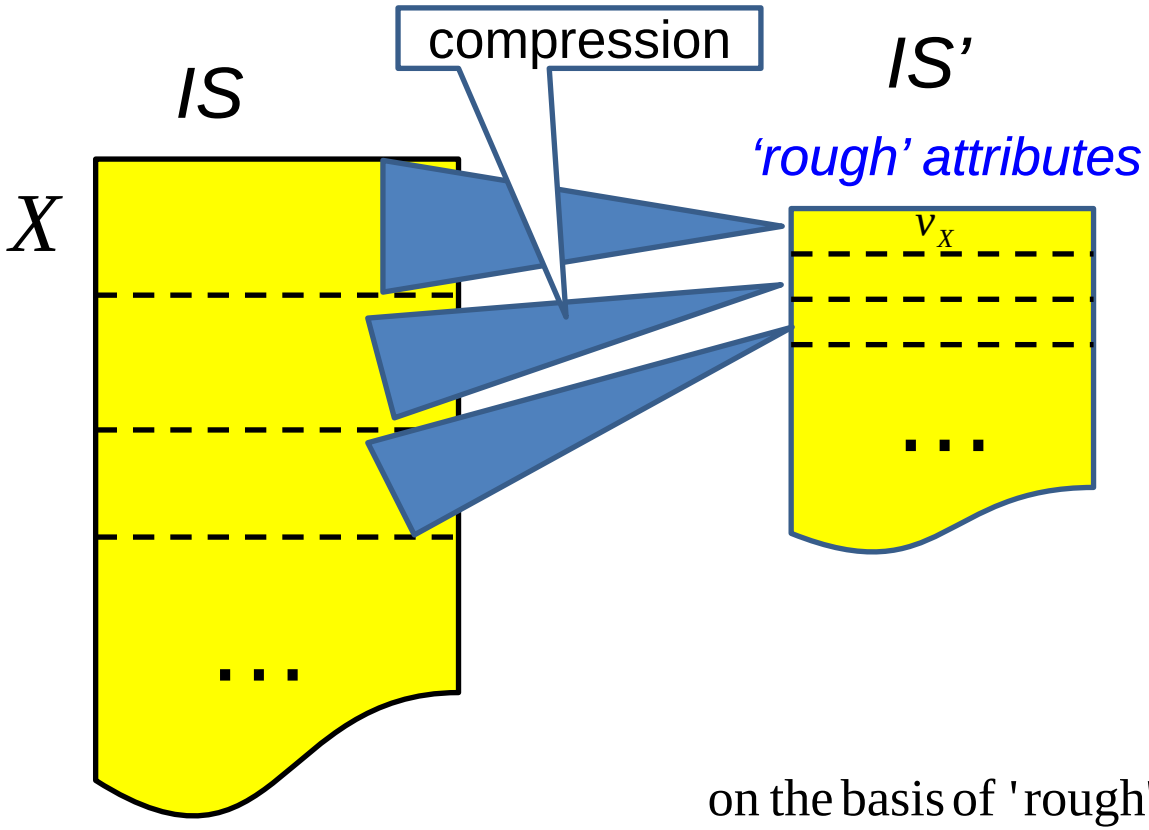
α – *Boolean combination of descriptors*

$$\|\alpha\|_{IS} = ?$$

$$\|\alpha\|_{IS} \neq \varnothing$$

$$x \in \|\alpha\|_{IS}$$

INFORMATION RETRIEVAL

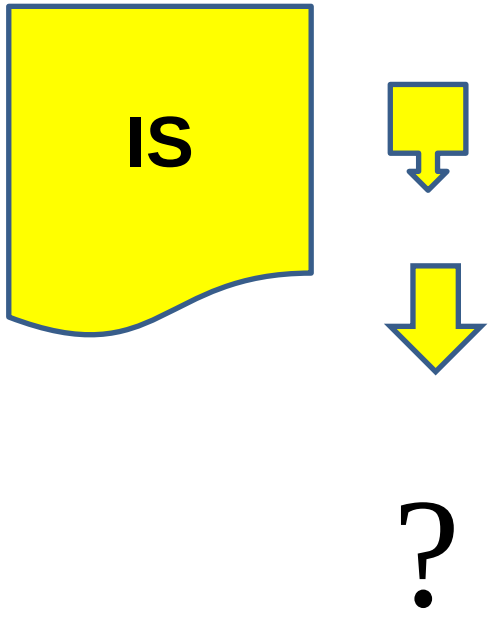


www.infobright.com/

$card(\|\alpha\|_{IS}) = ?$

on the basis of 'rough' information v_X about X
it is often possible easy to check that $X \subseteq \|\alpha\|_{IS}$ or
$X \subseteq U \setminus \|\alpha\|_{IS}$

INFORMATION RETRIEVAL



queries in natural language:

α

JUDEA PEARL- TURING AWARD 2011

for fundamental contributions to artificial intelligence through the development of a calculus for probabilistic and causal reasoning.

Traditional statistics is strong in devising ways of describing data and inferring distributional parameters from sample. Causal inference requires two additional ingredients:

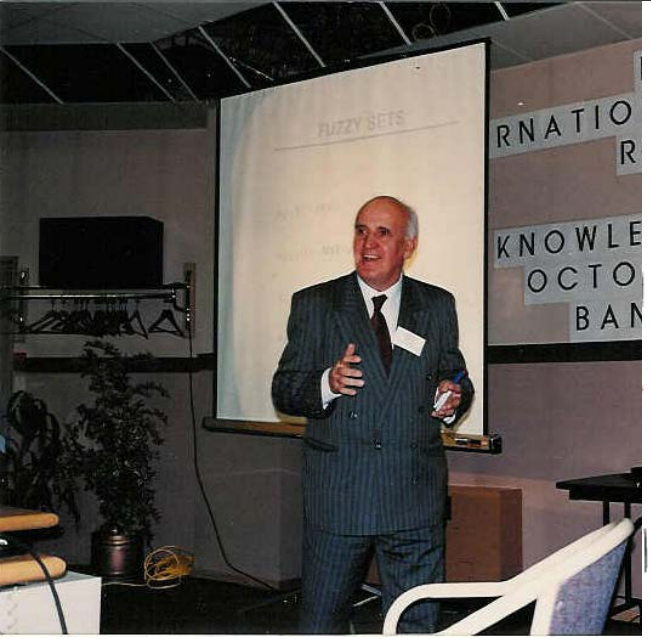
- *a science-friendly language for articulating causal knowledge,*

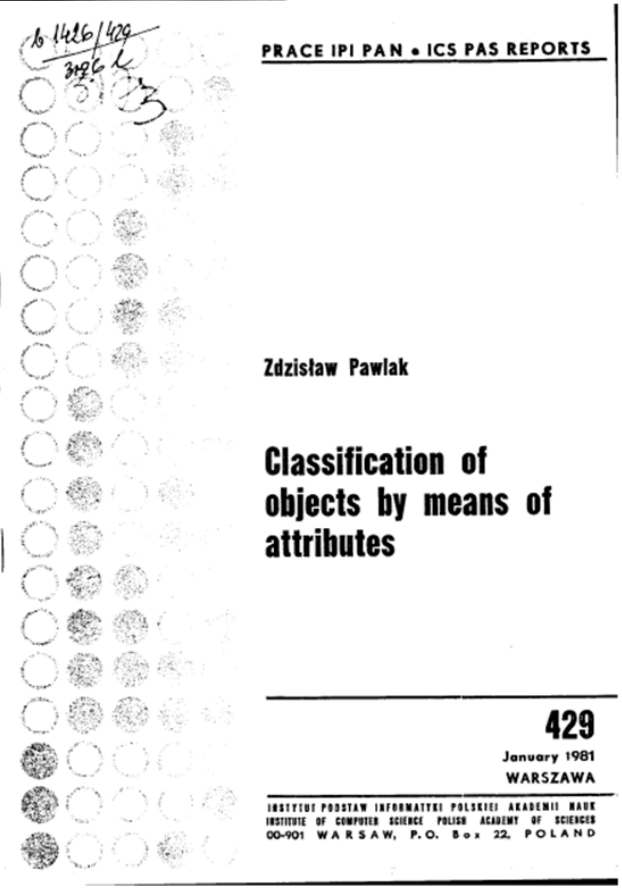
and

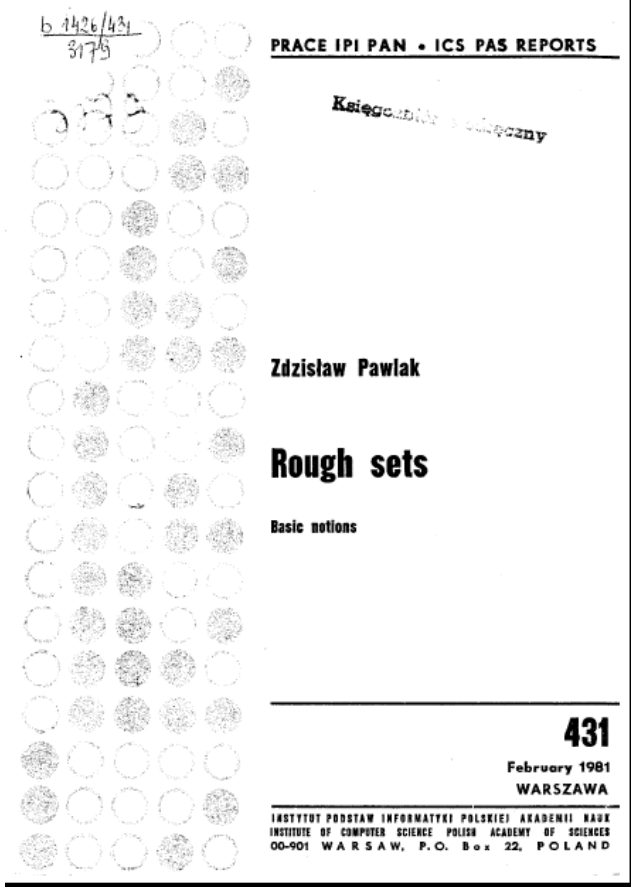
- *a mathematical machinery for processing that knowledge, combining it with data and drawing new causal conclusions about a phenomenon.*

Judea Pearl: Causal inference in statistics: An overview. Statistics Surveys 3, 96-146 (200)

INFORMATION SYSTEMS AND ROUGH SETS







International Journal of Computer & Information Sciences

October 1982, Volume 11, Issue 5, pp 341–356

Rough sets

Zdzisław Pawlak

Abstract

We investigate in this paper approximate operations on sets, approximate equality of sets, and approximate inclusion of sets. The presented approach may be considered as an alternative to fuzzy sets theory and tolerance theory. Some applications are outlined.

PRACE IPI PAN • ICS PAS REPORTS

Ewa Orłowska, Zdzisław Pawlak

Expressive power of knowledge representation systems

432
April 1981
WARSZAWA

INSTYTUT PODSTAW INFORMATYKI POLSKIEJ AKADEMII NAUK
INSTITUTE OF COMPUTER SCIENCE POLISH ACADEMY OF SCIENCES
00-901 WARSAW, P.O. Box 22, POLAND

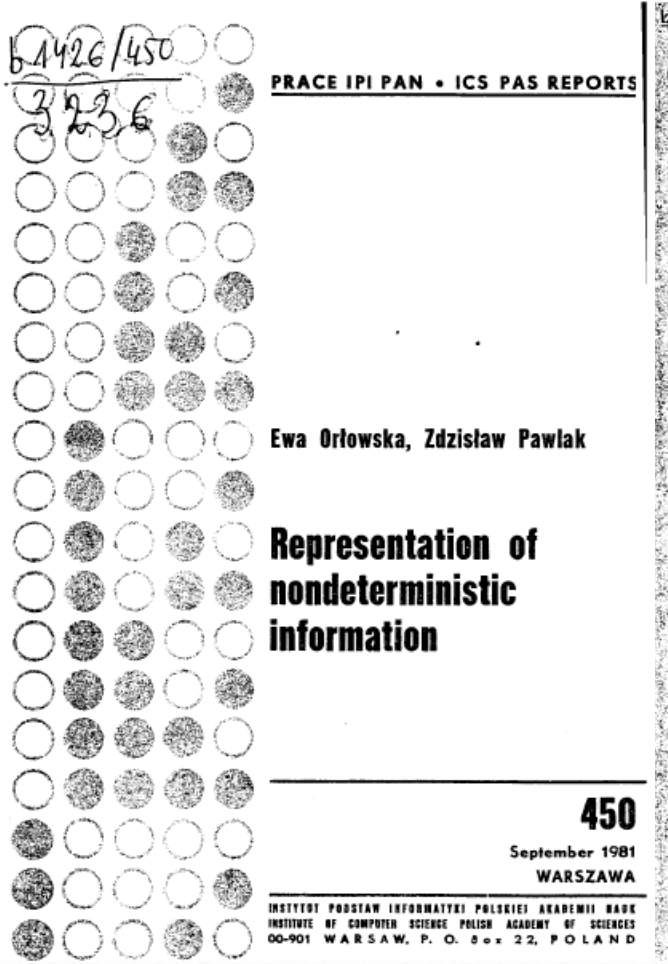
PRACE IPI PAN • ICS PAS REPORTS

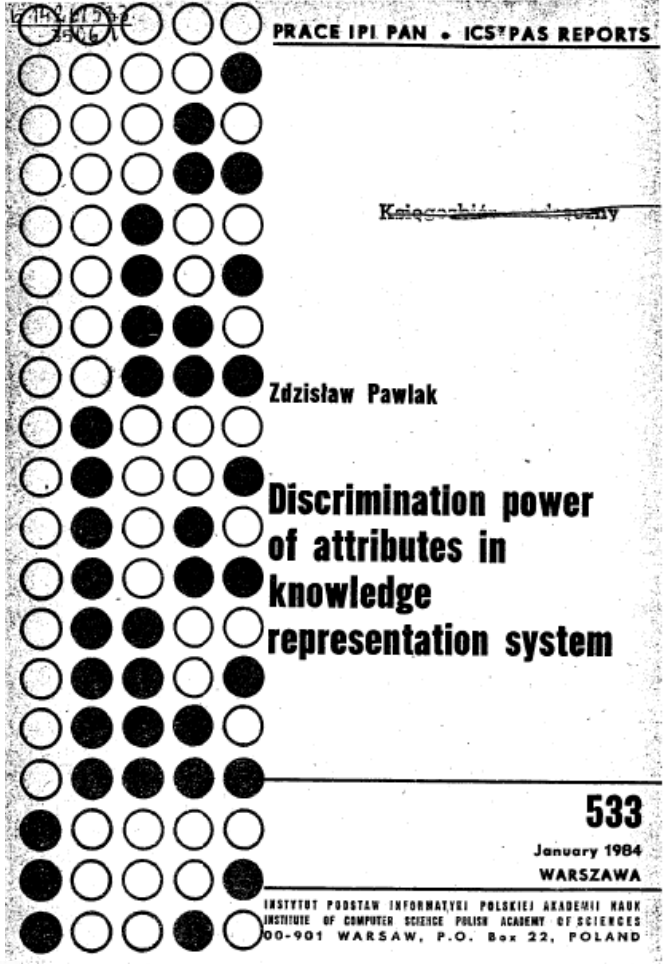
Wiktor Marek, Zdzisław Pawlak

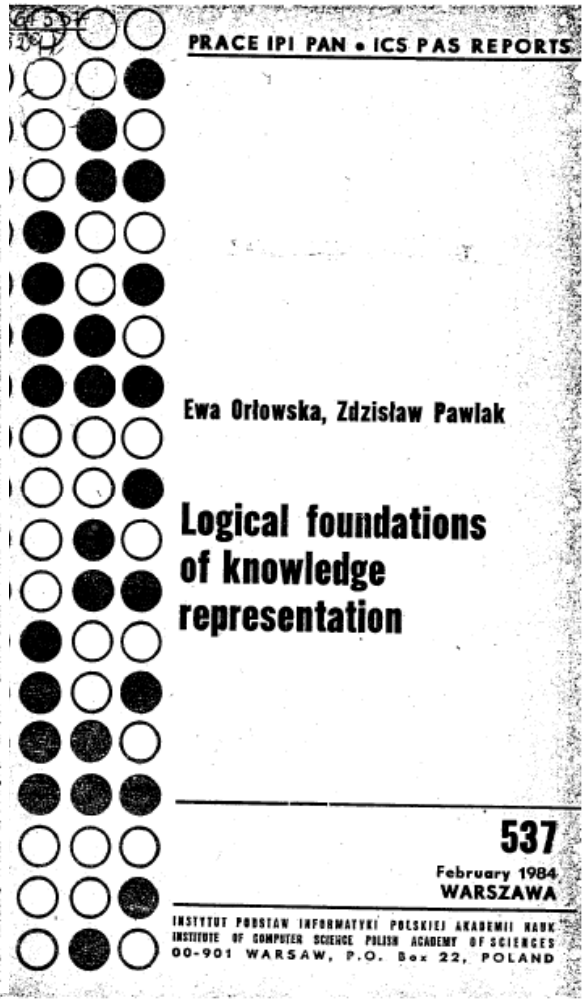
Rough sets and information systems

441
July 1981
WARSZAWA

INSTYTUT PODSTAW INFORMATYKI POLSKIEJ AKADEMII NAUK
INSTITUTE OF COMPUTER SCIENCE POLISH ACADEMY OF SCIENCES
00-901 WARSAW, P.O. Box 22, POLAND







Theoretical Computer Science 29 (1984) 27–39
North-Holland

REPRESENTATION OF NONDETERMINISTIC INFORMATION

Ewa ORŁOWSKA and Zdzisław PAWLAK
Institute of Computer Science, Polish Academy of Sciences, 00-901 Warsaw, PKiN, Poland

Communicated by E. Engeler
Received October 1982
Revised March 1983

Abstract. In this paper we develop a method of dealing with nondeterministic information. We introduce the concept of knowledge representation system of nondeterministic information and we define a language providing a means for defining nondeterministic information. We also develop deduction methods for the language.

PODSTAWY STEROWANIA
18, 3-4(1988), 175-200

Rough sets and information systems

Zdzisław PAWLAK

Department of Complex Control Systems, Polish Academy of Sciences
44-100 Gliwice ul. Bałtycka 5,

Received on 88.03.15

1. Introduction

In this paper we are going to give some basic ideas underlying the concept of a rough set, introduced by the author in [13] in order to deal with the vague and imprecise data.

The most interesting case is when data is arranged in the form of an information system (see [12]). The application of rough sets to the analysis of information systems is shown and discussed here.

The proposed approach has ben applied successfully in many areas (see e.g. [1, 11] and [15]).

The rough set concept can be vieved as an alternative to the fuzzy sets (see [18]). Comparison of these two concepts can be found in [2, 14] and [17].

More properties concerning rough sets and information systems are published in [2–10] and [16].

Annales Societatis Mathematicae Polonae
Series IV: Fundamenta Informaticae VI.3-4 (1983)

ON A REPRESENTATION OF ROUGH SETS BY MEANS OF INFORMATION SYSTEMS

Miroslav N o v o t n ý
Czechoslovak Academy of Sciences, Branch Brno
Zdzisław P a w l a k
Polish Academy of Sciences

Received October 7, 1982
AMS Categories: 68H05

ARCHIWUM INFORMATYKI
TEORETYCZNEJ I STOSOWANEJ
Tom 2 (1990), z. 3-4
pp. 139-166

Information systems and decision tables a rough set perspective

ZDZISŁAW PAWLAK

Institute of Theoretical and Applied Informatics, Polish Academy of Sciences
ul. Baltycka 5, 44-100 Gliwice

(*Received 1989. 07. 15*)

Abstract. In this paper we are going to show how the concept of a rough set can be employed as a theoretical basis of information systems and decision tables. It turns out that many problems, in particular in AI, like machine learning, expert systems, pattern recognition decision support systems and others can be reduced to the proposed schemes. In fact the approach has found many real life applications in medicine [46, 47], cement kiln control algorithms [19], aircraft pilots performance evaluation [10] – and others.

Inform. Systems Vol. 6, No. 3, pp. 205–218, 1981
Printed in Great Britain.

INFORMATION SYSTEMS THEORETICAL FOUNDATIONS

Z. PAWLAK
Institute of Computer Science, Polish Academy of Sciences. P.O. Box 22. 00-901 Warsaw PKiN. Poland

(*Received* 14 *March* 1980; *in revised form* 9 *December* 1980)

Abstract—Some basic concepts concerning information systems are defined and investigated. With every information system a query language is associated and its syntax and semantics is formally defined. Some elementary properties of the query language are stated. The presented approach leads to a new information systems organization. The presented idea was implemented and the implementation shows many advantages compared with other methods.

UNCERTAINTY IN OBJECT PERCEPTION
INDISCERNIBILITY RELATIONS

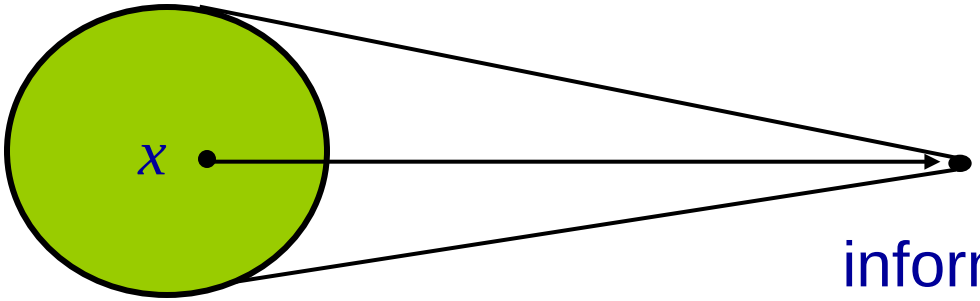
information system (data table)

$IS = (U, A)$

$U = \{x_1, ..., x_n\},\ A = \{a_1, ..., a_m\}$

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

$N(x) = (Inf_A)^{-1}(u)$



neighborhood of x

$u = Inf_A(x)$

information signature of $\boldsymbol{x}$

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

τ

tolerance or similarity

$IND(B) \ \ for \ \ B \subseteq A$

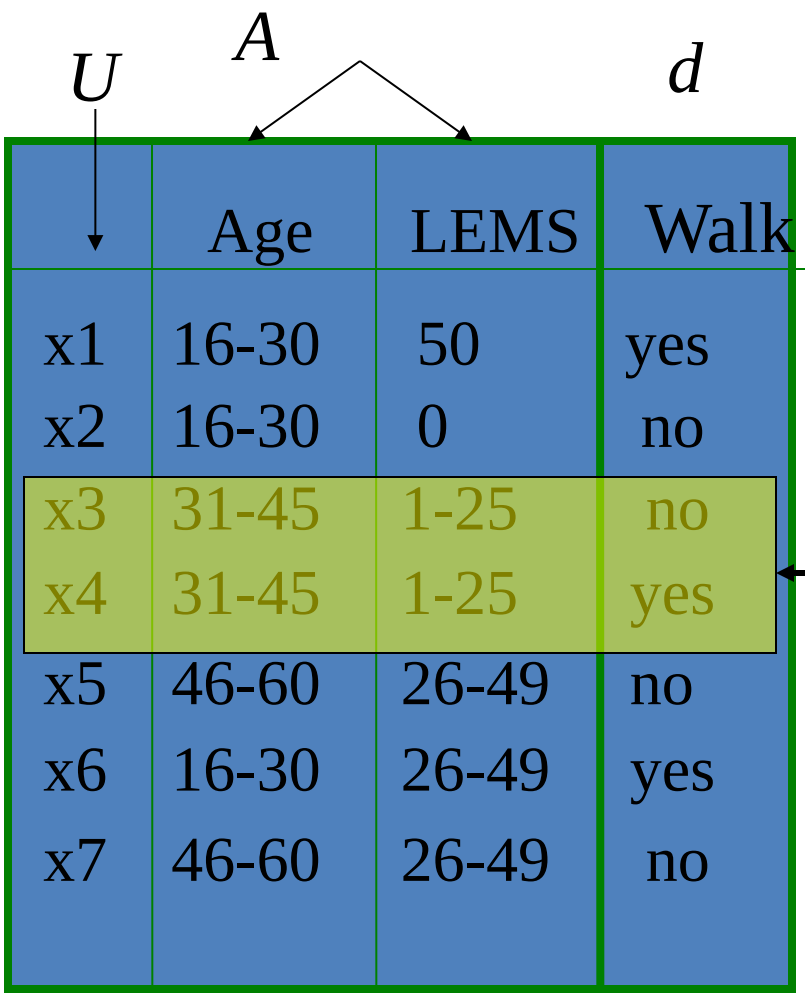
$[x]_{IND(B)} = [x]_B = \{y \in U : xIND(B)y\}$

$U / B = \{[x]_B : x \in U\}$

UNCERTAINTY
IN SIGNATURES OF OBJECTS

- missing values – different interpretations
- uncertainty in attribute value measurement
- noise
- …

DECISION SYSTEMS



$$DT = (U, A, d) \quad d \notin A$$

condition attributes

decision attribute

$$d : U \to V_d$$

decision classes

$$X_i = \{x \in U : d(x) = i\} \ \ for \ \ i \in V_d$$

inconsistency

Generalized decision:

$$\partial_B : U \to P(V_d) \ \ where \ \ B \subseteq A$$

$$\partial_B(x) = \{v' : \exists x'(xIND(B)x' \wedge d(x') = v')\} = d([x]_B)$$

Remark. Possible generalization for many decisions.

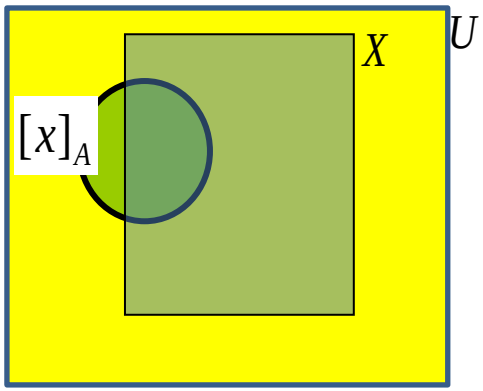
UNCERTAINTY IN OBJECT PERCEPTION
APPROXIMATION OF DECISION CLASSES

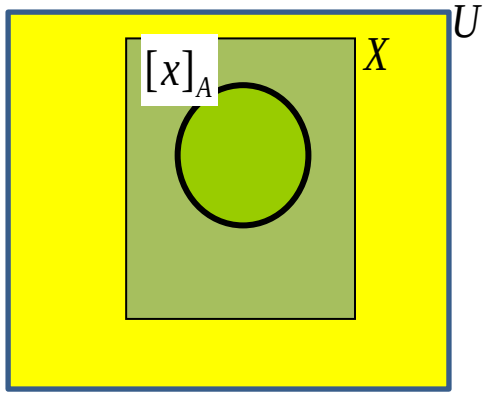
$X = \{x \in U : d(x) = 1\}$

decision system (data table)

$DT = (U, A, d)$

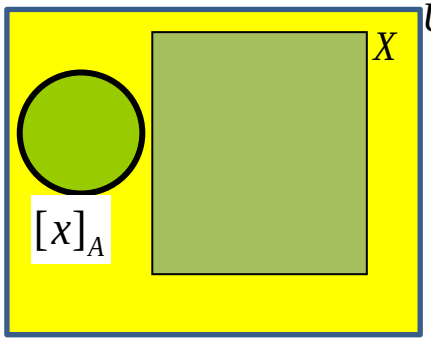
	a_1	a_2	…	a_m	d
x_1	v_1	v_2	…	v_m	1
	…	…	…	…	…





$[x]_{IND(A)} = [x]_A =$

$\{y \in U : xIND(A)y\}$



A-definable sets: unions of indiscernibility classes

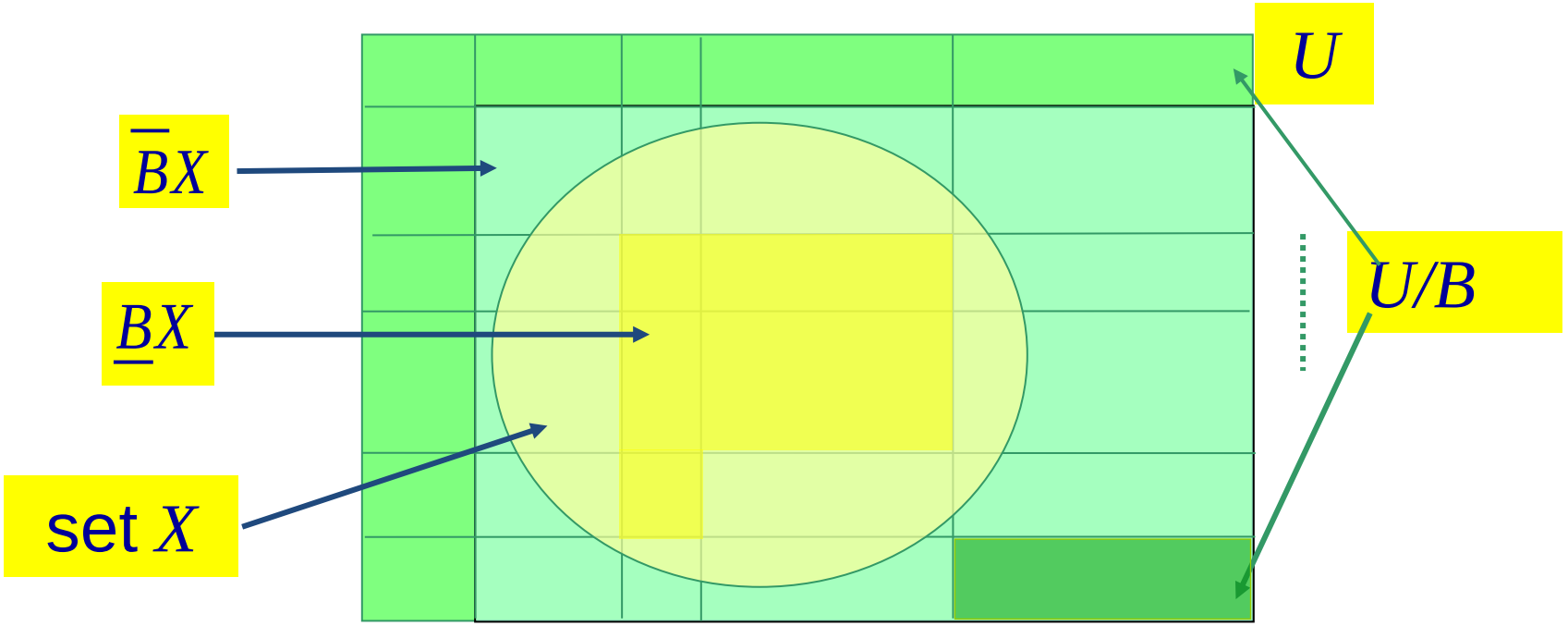
PROBLEM: Is a given decision class definable (relative to A)?

LOWER AND UPPER APROXIMATION

$$X \subseteq U,\ B \subseteq A$$

$$\underline{B}X = \bigcup \{Y \in U / B : Y \subseteq X\}$$

$$\overline{B}X = \bigcup \{Y \in U / B : Y \cap X \neq 0\}$$



BOUNDARY REGION

$$BN_B(X) = \overline{B}X \setminus \underline{B}X$$

ROUGH SETS

BOUNDARY REGION

$$BN_B(X) = \overline{B}X \setminus \underline{B}X$$

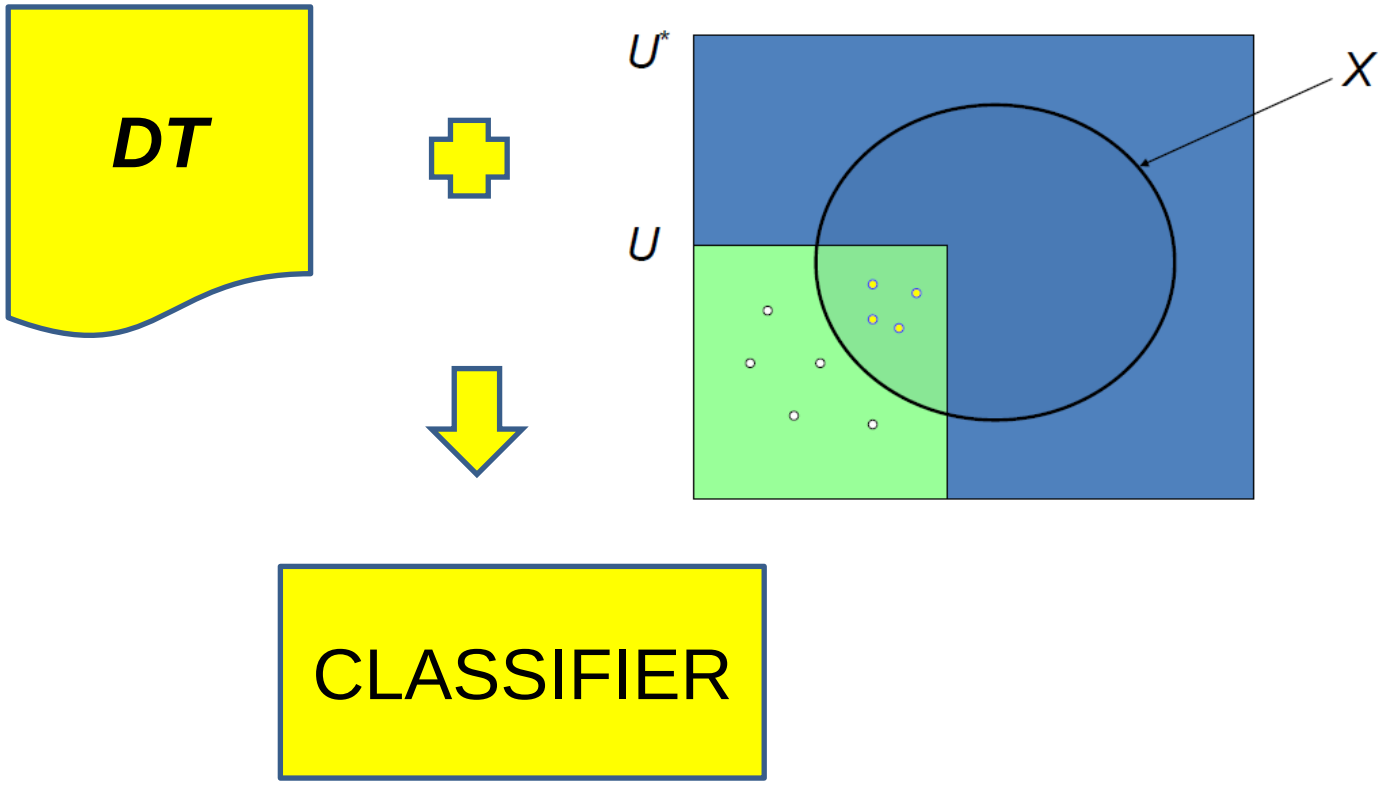
CRISP SET

$$BN_B(X) = \emptyset$$

ROUGH SET

$$BN_B(X) \neq \emptyset$$

INFORMATION SYSTEMS AND INDUCTION



TRANSFORMATIONS OF INFORMATION SYSTEMS WITH PRESERVING RELEVANT PROPERTIES

REDUCTION OF INFORMATION (DECISION) SYSTEMS

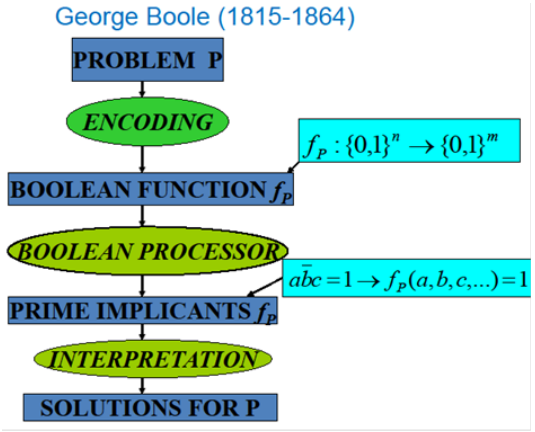
$(V_a, =)$

$(V_a, \leq)$

$(P(V_a), \cup, \cap, 0, 1)$

...

new features defined by formulas interpreted over relational structures: Boolean reasoning used in searching for relevant patterns (discretization, symbolic value grouping, association rules, dominance RS, …)

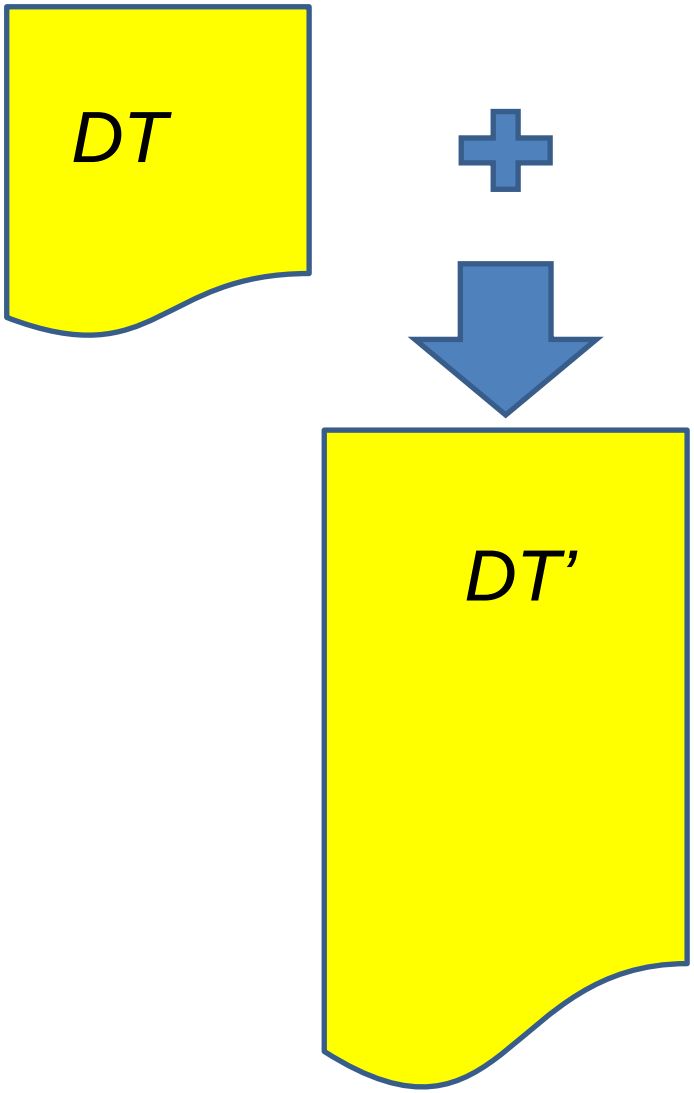


DT

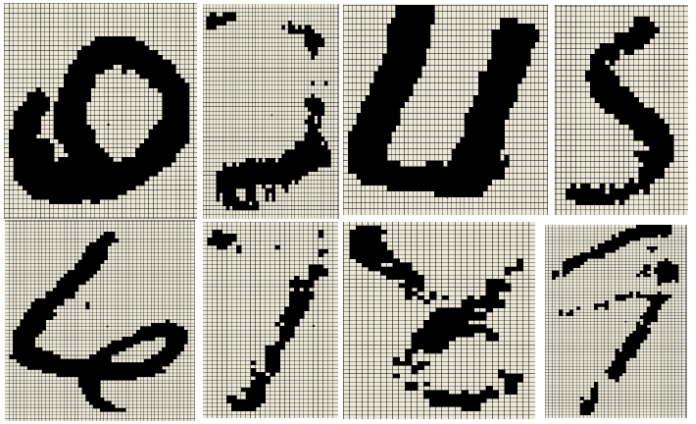
DT'

reduction of the size with preserving (e.g., discernibility, quality of classification approximation)
MINIMUM DESCRIPTION LENGTH PRINCIPLE (MDL)

EXTENSION OF INFORMATION (DECISION) SYSTEMS



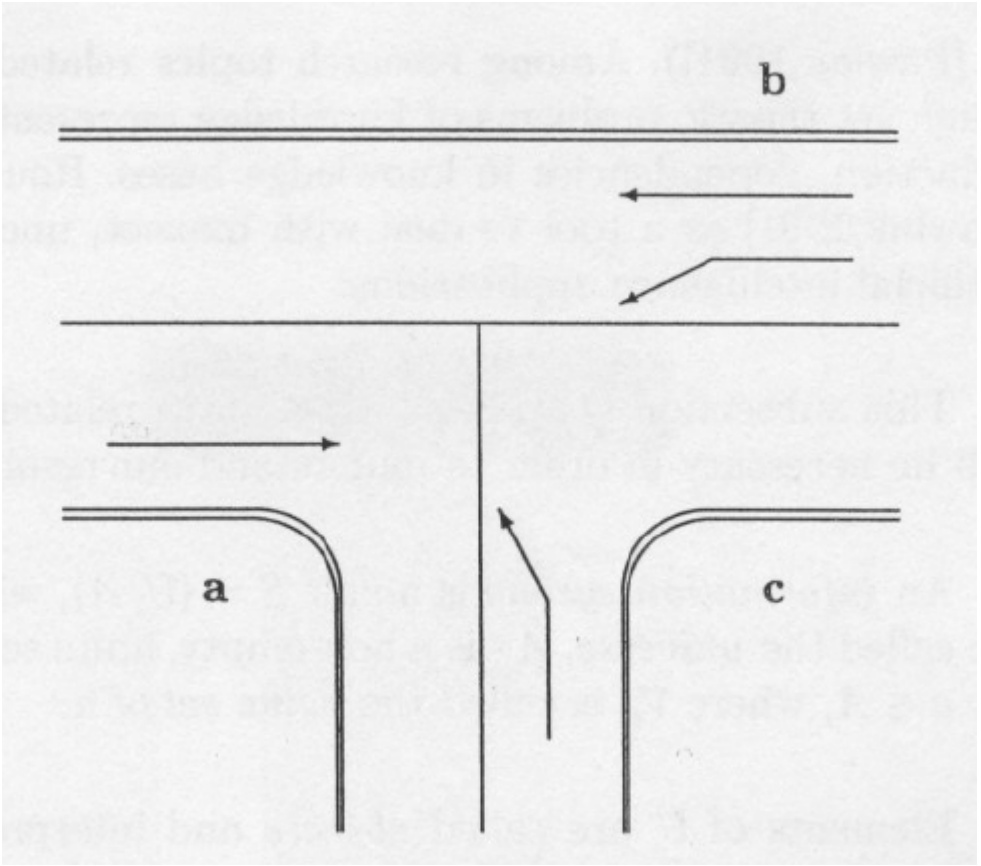
set of transformations preserving classification



INFORMATION SYSTEMS AND CONCURRENT SYSTEMS

Zdzislaw Pawlak. Concurrent versus sequential - the rough set perspective. Bulletin of the European Association for Theoretical Computer Science (EATCS), 48:178-190, 1992.

U/A	a	b	c
u_1	1	1	0
u_2	0	2	0
u_3	0	0	2



In this case we assume that attributes a, b, and c denote the traffic signals, objects labeled by u_1, u_2, u_3 denote the possible states of the observed system, whereas entries of the table 0, 1 and 2 denote colours of the traffic lights, red, green and green arrow, respectively.

INFORMATION SYSTEMS AND CONCURRENT SYSTEMS

$$IS$$

$$Dec\_Rules : set\ of\ rules$$

$$Petri\ Net$$

consistent with the maximal extension of set of states consistent with *Dec_Rules*

CONFLICT ANALYSIS

issues

a – autonomous Palestinian state on the West Bank and Gaza
b – Israeli military outpost along the Jordan River
c – Israeli retains East Jerusalem
d – Israeli military outposts on the Golan Heights
e – Arab countries grant citizenship to Palestinians who choose to remain within their borders

agents

1 – Israel
2 – Egypt
3 – Palestinians
4 – Jordan
5 – Syria
6 – Saudi Arabia

U	a	b	c	d	e
1	$-$	$+$	$+$	$+$	$+$
2	$+$	0	$-$	$-$	$-$
3	$+$	$-$	$-$	$-$	0
4	0	$-$	$-$	0	$-$
5	$+$	$-$	$-$	$-$	$-$
6	0	$+$	$-$	0	$+$

PRACE IPI PAN • ICS PAS REPORTS

Zdzisław Pawlak

Distributed information systems

370
1979
WARSZAWA

INSTYTUT PODSTAW INFORMATYKI POLSKIEJ AKADEMII NAUK
INSTITUTE OF COMPUTER SCIENCE POLISH ACADEMY OF SCIENCES
00-901 WARSAW, P. O. Box 22, POLAND

DISTRIBUTED INFORMATION SYSTEMS

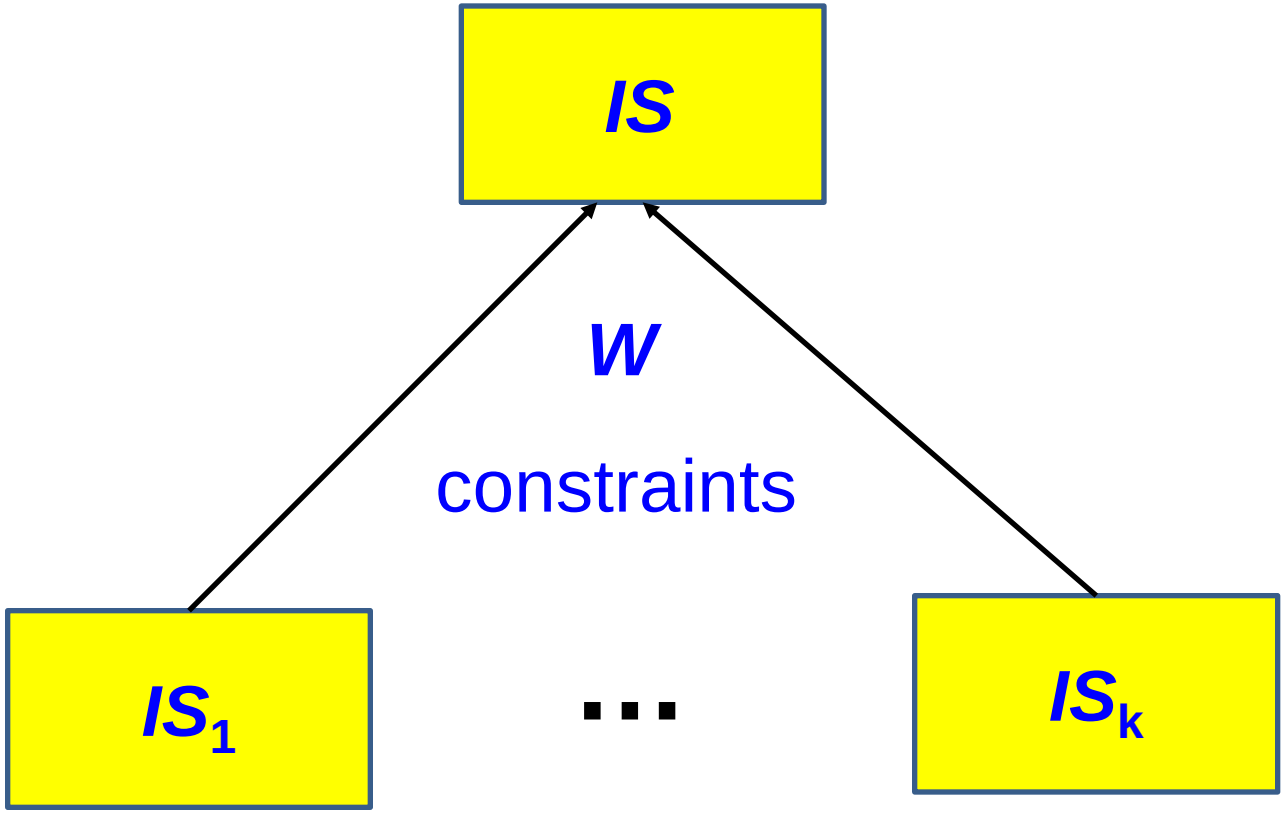
DISTRIBUTED INFORMATION SYSTEMS

- relationships with the information flow by Bairwise
- hierarchical learning
- multi-agent systems
- decomposition and synthesis
- …

STRUCTURAL OBJECTS

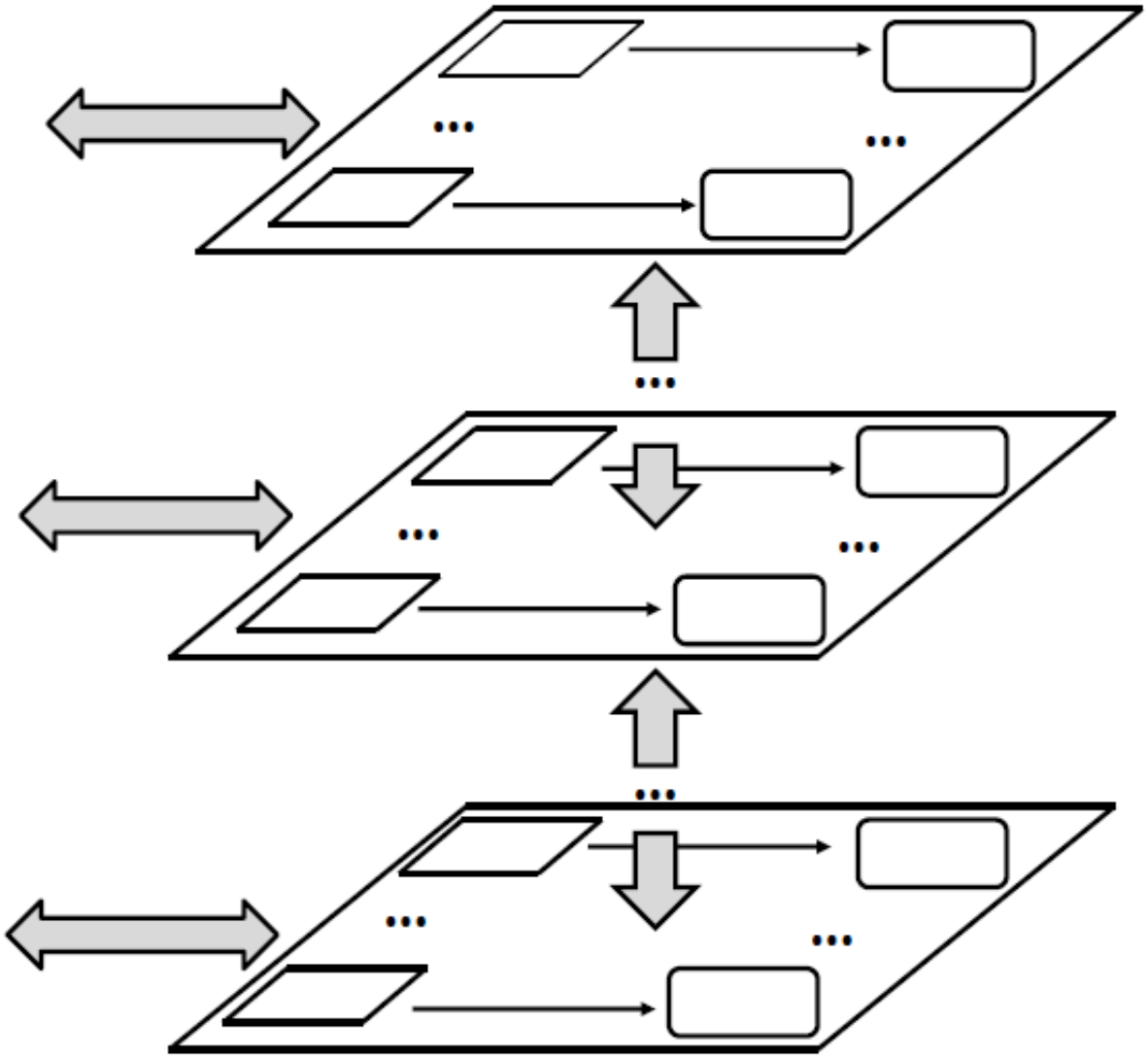
SEARCHING FOR RELEVANT FEATURES

JOIN WITH CONSTRAINTS

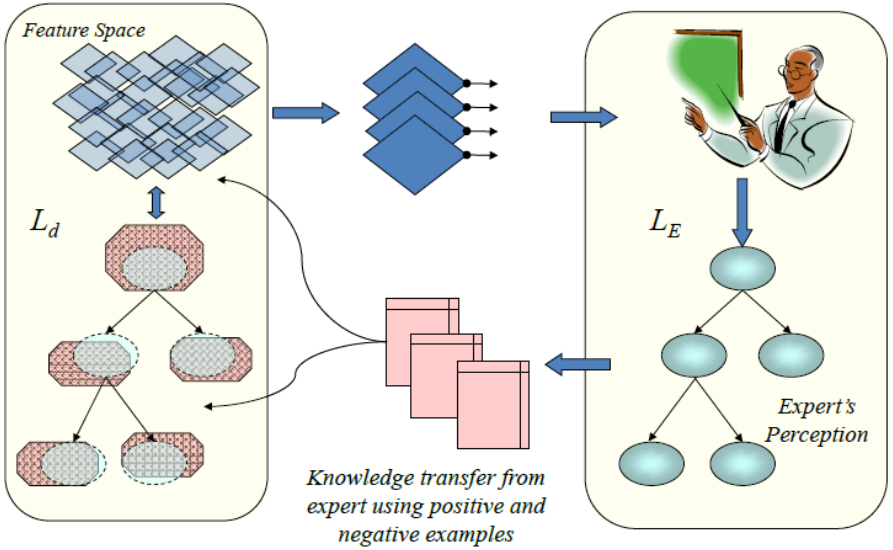


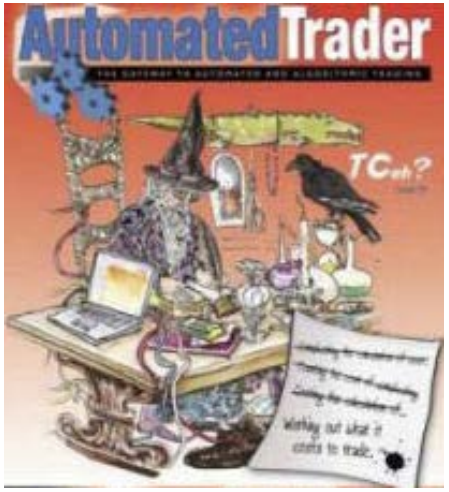
Objects (granules) in *IS* are composed out of attribute value vectors from $IS_1 \ldots IS_k$ satisfying *W*

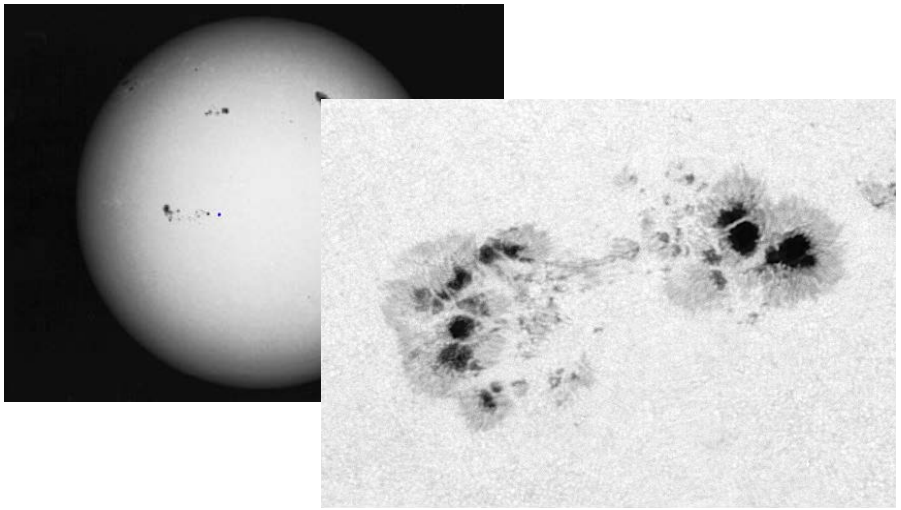
INTERACTIVE HIERARCHICAL STRUCTURES



HIERARCHICAL LEARNING:
RS BASED ONTOLOGY APPROXIMATION



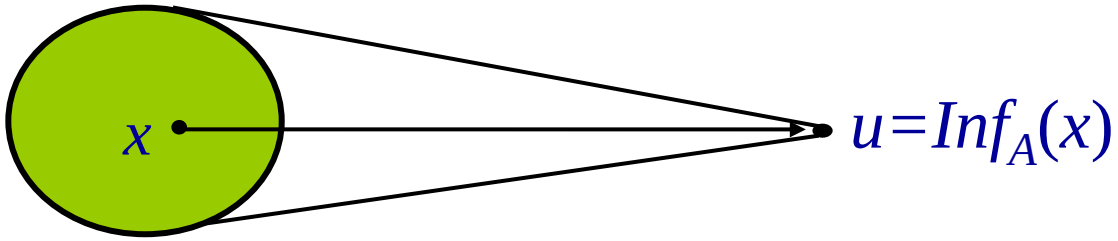




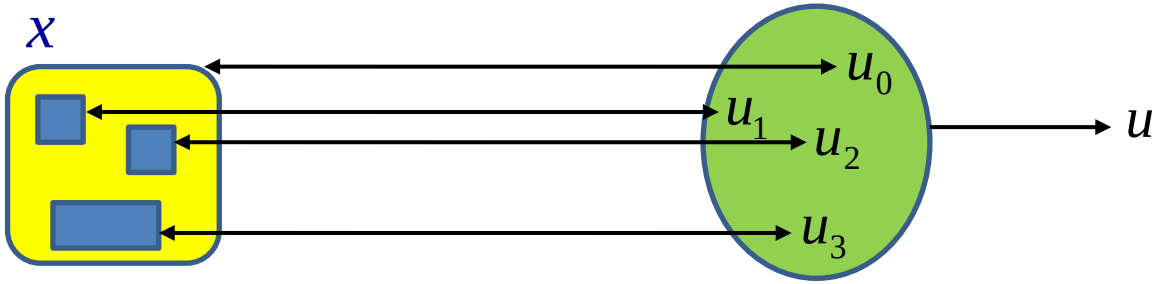
WHAT NEXT ?

INTERACTIVE INFORMATION SYSTEMS

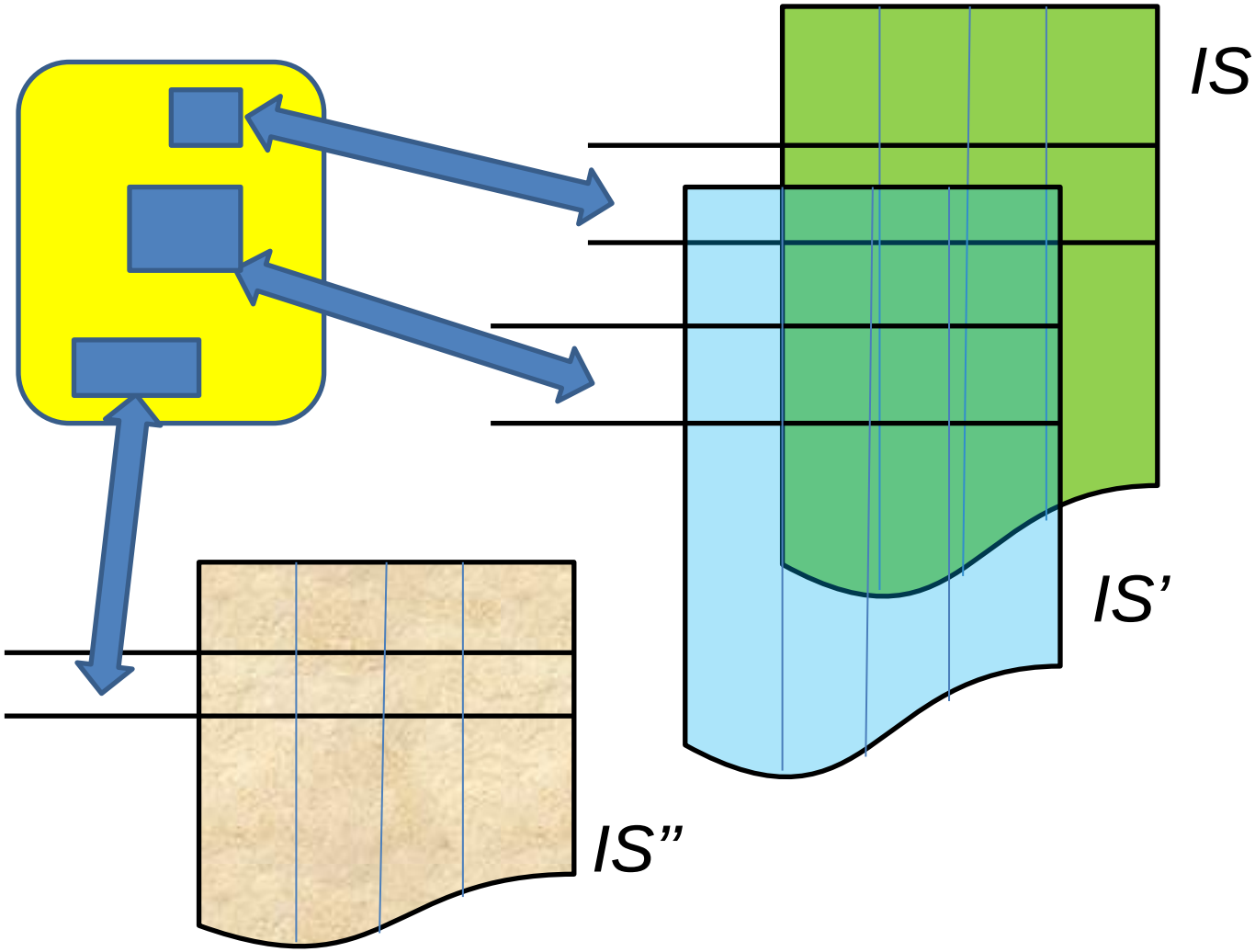
INTERACTIVE INFORMATION SYSTEMS



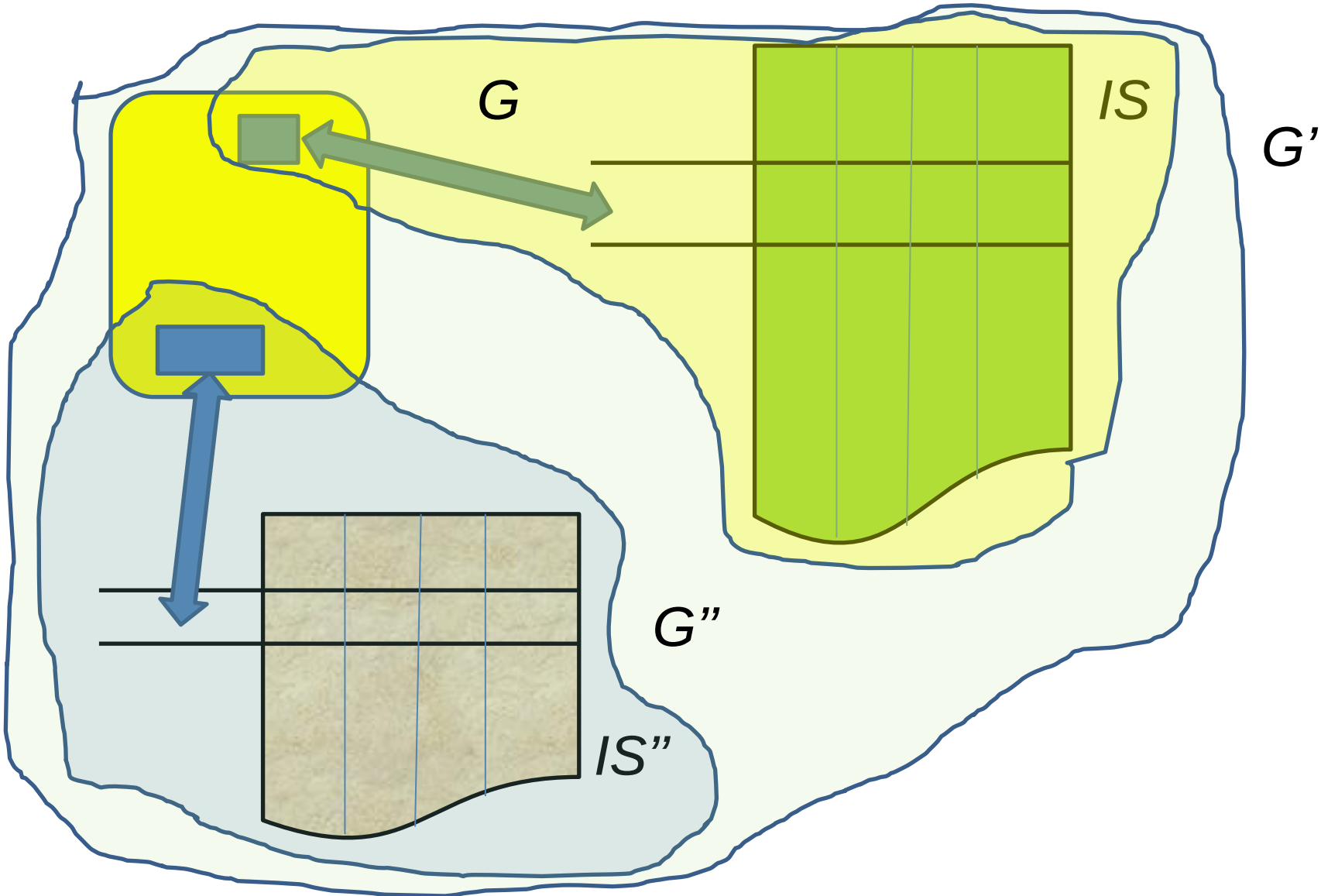
for complex physical objects we need to model interaction with them



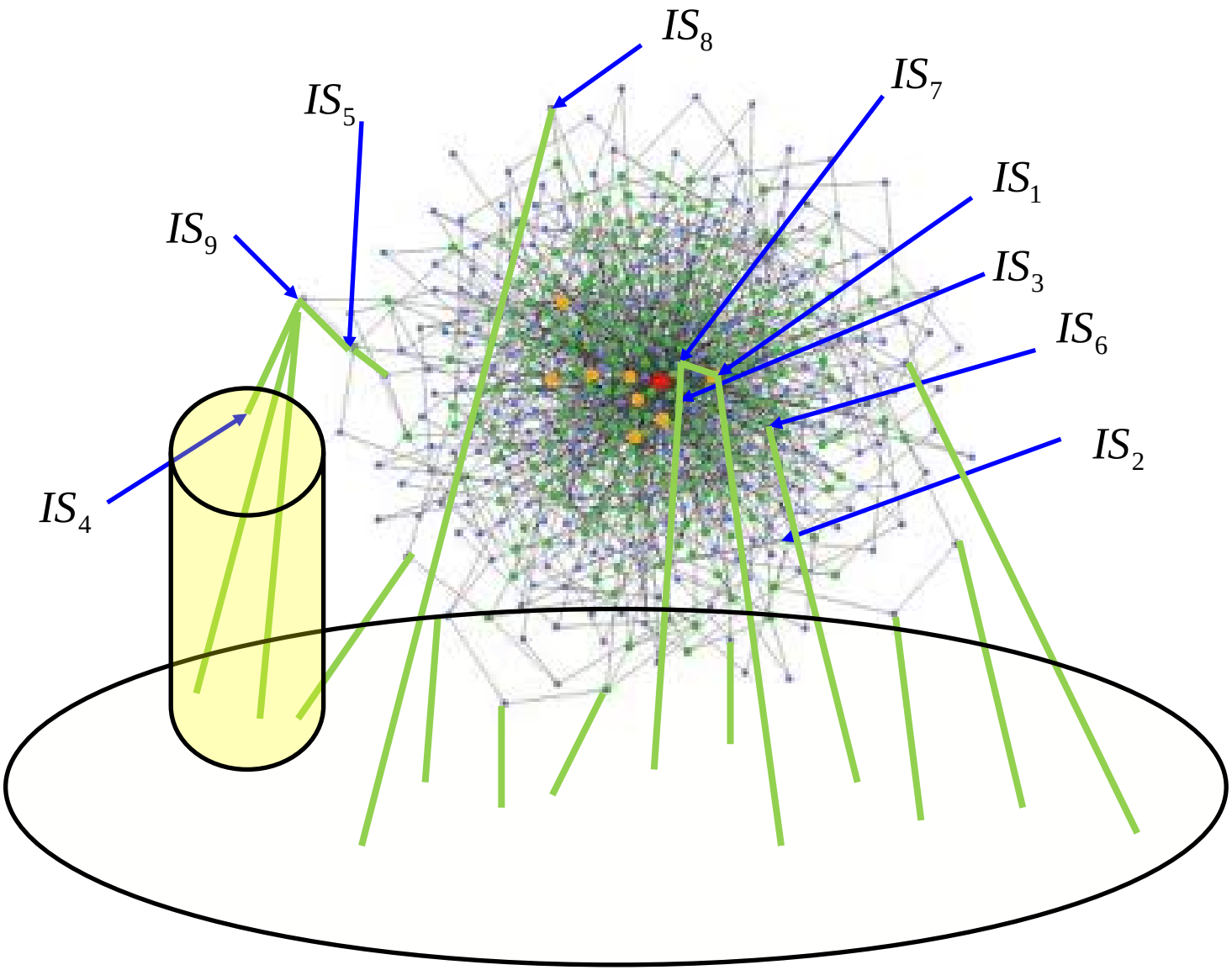
INTERACTIVE INFORMATION SYSTEMS ARE LINKED WITH PHYSICAL OBJECTS BY COMPLEX GRANULES (c-granules)



INTERACTIVE INFORMATION SYSTEMS ARE LINKED WITH PHYSICAL OBJECTS BY COMPLEX GRANULES (c-granules)

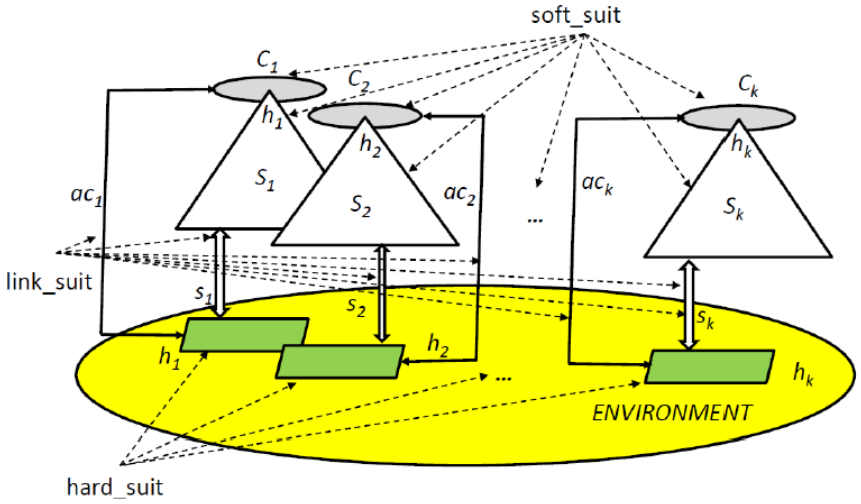


C-GRANULES



WHAT NEXT ?

COMPLEX GRANULES

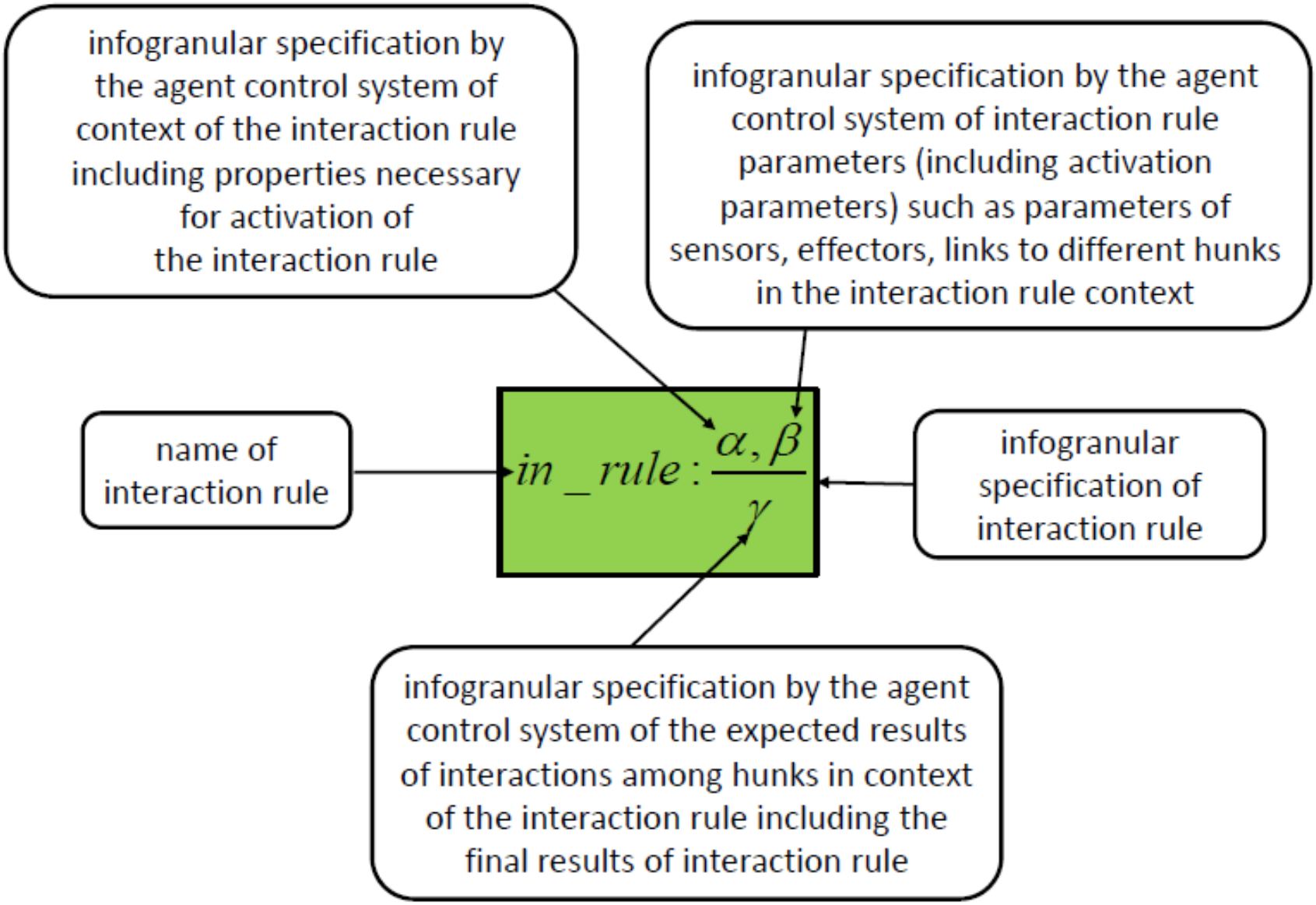


THREATS AND VULNERABILITIES

COMPLEX GRANULES IN DEALING WITH PROBLEMS BEYOND ONTOLOGIES

- EVOLVING LANGUAGES FOR PERCEIVING, REASONING AND ACTING TOWARD ACHIEVING GOALS
- RISK MANAGEMENT BASED ON JUDGMENT ON COMPUTATIONS OVER COMPLEX GRANULES

INTERACTION RULE



ECORITHMS

The algorithms I discuss in this book are special. Unlike most algorithms, they can be run in environments unknown to the designer, and they learn by interacting with the environment how to act effectively in it. After sufficient interaction they will have expertise not provided by the designer, but extracted from the environment. I call these algorithms **ecorithm**.

Valiant, L.: Probably Approximately Correct. Nature's Algorithms for Learning and Prospering in a Complex World. Basic Books, A Member of the Perseus Books Group, New York (2013)

BEYOND THE TURING TEST
&
JUDGMENT

The Turing test, as originally conceived, focused on language and reasoning; **problems of perception and action were conspicuously absent**. The proposed tests will provide an opportunity to bring four important areas of AI research (language, reasoning, perception, and action) back into sync after each has regrettably diverged into a fairly independent area of research.

C. L. Ortitz Jr. Why we need a physically embodied Turing test and what it might look like.
AI Magazine 37 (2016) 55–62.

International Rough Set Society http://www.roughsets.org

Group at Warsaw University:

http://logic.mimuw.edu.pl

RSES: http://logic.mimuw.edu.pl/~rses/

Rough Set Database System:

http://rsds.univ.rzeszow.pl/

RoughSets: Data Analysis Using Rough Set and Fuzzy Rough Set Theories (package in R)

https://cran.r-project.org/web/packages/RoughSets/index.html

Journal: Transactions on Rough Sets

http://roughsets.home.pl/www/index.php?option=com_content&task=view&id=14&Itemid=32

http://scholar.google.com/citations?user=fYu9ryIAAAAJ&hl=en&oi=ao

http://scholar.google.com/citations?user=zVpMZBkAAAAJ&hl=en&oi=ao

THANK YOU !