Dominance-based Rough Set Approach to Multiple Criteria Decision Support

Salvatore Greco
Università di Catania, Italy

Benedetto Matarazzo
Università di Catania, Italy

Roman Słowiński
Politechnika Poznańska, Poland

Plan

- Knowledge discovery from data
- Inconsistencies in data – Rough Set Theory
- Dominance-based Rough Set Approach (DRSA)
  - Dominance principle as monotonicity principle
  - Granular computing with dominance cones
  - Induction of decision rules from dominance-based rough approximations
- Decision rules
  - Attractiveness measures of decision rules
  - Knowledge representation and prediction
  - Bayesian confirmation measures
  - Effectiveness of intervention
- Multiple criteria decision support with DRSA
- Examples of application
- Other extensions of DRSA
- Conclusions

Knowledge Discovery from Data

Knowledge discovery from data

- The gap between data generation and data comprehension grows up
- *Knowledge Discovery* techniques try to bridge this gap
- Knowledge discovery is an inductive process aiming at identification of:
  - true,
  - non-trivial,
  - useful,
  - directly comprehensible

  patterns in data
- Pattern = rule, trend, phenomenon, regularity, anomaly, hypothesis, function etc.
- The patterns are useful for explanation of situations described by data, for prediction of future situations and for buiding a strategy of intervention

What form of a pattern: real-valued function ?

- Description of complex phenomena by recursive estimation techniques applied on historical data (*Int. J. Environment and Pollution, vol.12, no.2/3, 1999*)
- The pattern shows the dependence of the size of the mouth of a river in month k, represented by the relative tidal energy (RTE_k), from RTE_{k-1}, the river flow (F_{k-1}), the onshore wind (W_{k-1}) and the crude monthly count of storm events (S_k) (Elford et al. 1999; Murray Mouth, Australia):

$$\boxed{RTE_k = A_1 RTE_{k-1} + A_2 \frac{(F_{k-1} - 200)^{2.4}}{8RTE_{k-1} + 1} + A_3 \frac{W_{k-1}}{8RTE_{k-1} + 1} + A_4 S_k + \varepsilon_k}$$

where the exponent 2.4 was tuned by „trial and error”, coefficients A_1, A_2, A_3, A_4 were computed using a recursive least squares (RLS) approach, and ε_k is the model error

- The pattern is used to produce a strategy for the opening of barrages that will control the river flow, and thus, the size of the mouth

What form of a pattern: logical statements, rules ?

- Description of complex phenomena by recursive estimation techniques applied on historical data (*Int. J. Environment and Pollution, vol.12, no.2/3, 1999*)
- The pattern shows the impact of urban stormwater on the quality of the receiving water (Rossi, Słowiński, Susmaga 1999; Lausanne and Genève).
- Polluants: solid particles, organic matter, nitrogen and phosphorus, bacteria, viruses, lead and hydrocarbons, petroleum residues, pesticides etc.
- Example of rule induced from empirical observation of some sites:

If *the site is of type 2 (residential),* **and** *total rainfall is low (up to 8 mm),* **and** *max intensity of rain is between 2.7 and 11.2 mm/h,* **then** *total mass of suspended solids is < 2.2 kg/ha*

- The pattern involves heterogeneous data: nominal, qualitative and quantitative

Example of technical diagnostics

- 176 buses (objects)
- 8 symptoms (attributes)
- Decision = technical state:

  3 – good state (in use)

  2 – minor repair

  1 – major repair (out of use)
- Discover patterns = find relationships between symptoms and the technical state
- Patterns explain expert's decisions and support diagnosis of new buses

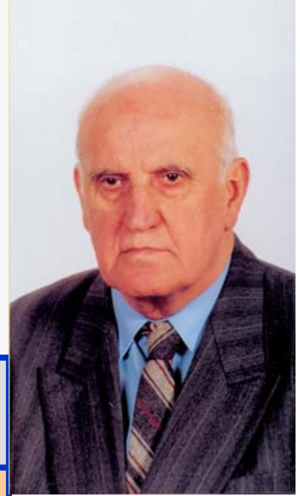
Examples:

	MaxSpeed	ComprPressure	Blacking	Torque	SummerCons	WinterCons	OilCons	HorsePower	State
1.	90	2	38	481	21	26	0	145	3
2.	76	2	70	420	22	25	2	110	1
3.	63	1	82	400	22	24	3	101	1
4.	90	2	49	477	21	25	1	138	3
5.	85	2	52	460	21	25	1	130	2
6.	72	2	73	425	23	27	2	112	1
7.	88	2	50	480	21	24	1	140	3
8.	87	2	56	465	22	27	1	135	3
9.	90	2	16	486	26	27	0	150	3
10.	60	1	95	400	23	24	4	96	1
11.	80	2	60	451	21	26	1	125	1
12.	78	2	63	448	21	26	1	120	2
13.	90	2	26	482	22	24	0	148	3
14.	62	1	93	400	22	28	3	100	1
15.	82	2	54	461	22	26	1	132	2
16.	65	2	67	402	22	23	2	103	1
17.	90	2	51	468	22	26	1	138	3
18.	90	2	15	488	20	23	0	150	3
19.	76	2	65	428	27	33	2	116	1
20.	85	2	50	454	21	26	1	129	2
21.	85	2	58	450	22	25	1	126	2
22.	88	2	48	458	22	25	1	130	3
23.	60	1	90	400	24	28	4	95	1
24.	64	2	71	420	23	25	2	105	1
25.	75	2	64	432	22	25	1	114	2
26.	74	2	64	420	21	25	1	110	2
27.	68	2	70	400	22	26	2	100	1

Attributes: 9 of 10 | Examples: 76 | Decision: State | Missing Values: No

Inconsistencies in Data – Rough Set Theory

Inconsistencies in data – Rough Set Theory



- Zdzisław Pawlak (1926 – 2006)

Student	Mathematics	Physics	Literature	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**medium**	**medium**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**medium**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

Inconsistencies in data – Rough Set Theory

- Objects with the same description are indiscernible and create blocks

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**medium**	**medium**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**medium**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

Inconsistencies in data – Rough Set Theory

- Objects with the same description are indiscernible and create blocks

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**medium**	**medium**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**medium**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

Inconsistencies in data – Rough Set Theory

- Objects with the same description are indiscernible and create blocks

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**medium**	**medium**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**medium**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

Inconsistencies in data – Rough Set Theory

- Objects with the same description are indiscernible and create blocks

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**medium**	**medium**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**medium**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

Inconsistencies in data – Rough Set Theory

- Objects with the same description are indiscernible and create blocks

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**medium**	**medium**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**medium**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

Inconsistencies in data – Rough Set Theory

- Objects with the same description are indiscernible and create granules

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**medium**	**medium**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**medium**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

Inconsistencies in data – Rough Set Theory

- Another information assigns objects to some classes (sets, concepts)

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**medium**	**medium**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**medium**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

Inconsistencies in data – Rough Set Theory

- Another information assigns objects to some classes (sets, concepts)

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**medium**	**medium**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**medium**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

Inconsistencies in data – Rough Set Theory

- Another information assigns objects to some classes (sets, concepts)

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**medium**	**medium**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**medium**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

Inconsistencies in data – Rough Set Theory

- The granules of indiscernible objects are used to approximate classes

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**medium**	**medium**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**medium**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

Inconsistencies in data – Rough Set Theory

- Lower approximation of class „good"

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**medium**	**medium**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**medium**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

Lower Approximation

Inconsistencies in data – Rough Set Theory

- Lower and upper approximation of class „good”

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**medium**	**medium**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**medium**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

Upper Approximation

Lower Approximation

CRSA – decision rules induced from rough approximations

- Certain decision rule supported by objects from lower approximation of class „good” (discriminant rule)

  If Lit=good, *then* Student is certainly good
  {S5,S6}

- Possible decision rule supported by objects from upper approximation of class „good” (partly discriminant rule)

  If Phys=medium & Lit=medium, *then* Student is possibly good
  {S3,S4}

- Approximate decision rule supported by objects from the boundary of class „medium” and „good”

  If Phys=medium & Lit=medium, *then* Student is medium or good
  {S3,S4}

Classical Rough Set Approach (CRSA)

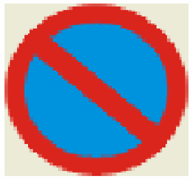
- Let U be a finite universe of discourse composed of objects (actions) described by a finite set of attributes
- Sets of objects indiscernible w.r.t. attributes create granules of knowledge (elementary sets)
- Any subset $X \subseteq U$ may be expressed in terms of these granules:
  - either precisely – as a union of the granules
  - or roughly – by two ordinary sets, called *lower* and *upper approximations*
- The lower approximation of X consists of all the granules included in X
- The upper approximation of X consists of all the granules having non-empty intersection with X

Classical Rough Set Approach (CRSA)

Example

- Classification of basic traffic signs
- There exist three main classes of traffic signs corresponding to:
  - warning (W),
  - interdiction (I),
  - order (O).
- These classes may be distinguished by such attributes as the shape (S) and the principal color (PC) of the sign
- Finally, we give few examples of traffic signs

CRSA – example of traffic signs

Traffic sign	Shape (S)	Primary Color (PC)	Class
a) 	triangle	yellow	W
b) 	circle	white	I
c) 	circle	blue	I
d) 	circle	blue	O

CRSA – example of traffic signs

Traffic sign	Shape (S)	Primary Color (PC)	Class
a) 	triangle	yellow	W
b) 	circle	white	I
c) 	circle	blue	I
d) 	circle	blue	O

- Granules of knowledge:

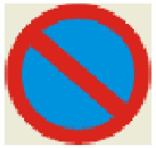
  $W=\{a\}_{\text{Class}}$, $I=\{b,c\}_{\text{Class}}$, $O=\{d\}_{\text{Class}}$

  $\{a\}_{S,PC}$, $\{b\}_{S,PC}$, $\{c,d\}_{S,PC}$

- Explanation of classification in terms of granules generated by S and PC
  - class W includes sign *a* certainly and no other sign possibly
  - class I includes sign *b* certainly and signs *b*, *c* and *d* possibly
  - class O includes no sign certainly and signs *c* and *d* possibly
- *Lower* and *upper approximation* of the classes by attributes S and PC:
  - $\text{lower\_appx.}_{S,PC}(W)=\{a\}$, $\text{upper\_appx.}_{S,PC}(W)=\{a\}$
  - $\text{lower\_appx.}_{S,PC}(I)=\{b\}$, $\text{upper\_appx.}_{S,PC}(I)=\{b,c,d\}$
  - $\text{lower\_appx.}_{S,PC}(O)=\varnothing$, $\text{upper\_appx.}_{S,PC}(O)=\{c,d\}$
  - $\text{boundary}_{S,PC}(I)=\text{upper\_appx.}_{S,PC}(I) - \text{lower\_appx.}_{S,PC}(I)=\{c,d\}$
  - $\text{boundary}_{S,PC}(O)=\text{upper\_appx.}_{S,PC}(O) - \text{lower\_appx.}_{S,PC}(O)=\{c,d\}$
- The *quality of approximation*: 2/4

CRSA – example of traffic signs

- To increase the quality of approximation (decrease the ambiguity) we add a new attribute – secondary color (SC)

Traffic sign	Shape (S)	Primary Color (PC)	Secondary color (SC)	Class
a) 	triangle	yellow	red	W
b) 	circle	white	red	I
c) 	circle	blue	red	I
d) 	circle	blue	white	O

- The granules: $\{a\}_{S,PC,SC}$, $\{b\}_{S,PC,SC}$, $\{c\}_{S,PC,SC}$, $\{d\}_{S,PC,SC}$
- Quality of approximation: 4/4=1

CRSA – example of traffic signs

- Are all three attributes necessary to characterize precisely the classes W, I, O ?

Traffic sign	~~Shape (S)~~	Primary Color (PC)	Secondary color (SC)	Class
a) 	~~triangle~~	yellow	red	W
b) 	~~circle~~	white	red	I
c) 	~~circle~~	blue	red	I
d) 	~~circle~~	blue	white	O

- The granules: $\{a\}_{PC,SC}$, $\{b\}_{PC,SC}$, $\{c\}_{PC,SC}$, $\{d\}_{PC,SC}$
- Quality of approximation: 4/4=1

CRSA – example of traffic signs

Traffic sign		Shape (S)	~~Primary Color (PC)~~	Secondary color (SC)	Class
a)		triangle	~~yellow~~	red	W
b)		circle	~~white~~	red	I
c)		circle	~~blue~~	red	I
d)		circle	~~blue~~	white	O

- The granules: $\{a\}_{S,SC}$, $\{b,c\}_{S,SC}$, $\{d\}_{S,SC}$
- *Reducts* of the set of attributes: {PC, SC} and {S, SC}
- Intersection of reducts is the *core*: {SC}

CRSA – example of traffic signs

- The minimal representation of knowledge contained in the Table – *decision rules*

Traffic sign		Shape (S)	~~Primary Color (PC)~~	Secondary color (SC)	Class
a)		triangle	~~yellow~~	~~red~~	W
b)		circle	~~white~~	red	I
c)		circle	~~blue~~	red	I
d)		~~circle~~	~~blue~~	white	O

rule #1: *if* S=triangle, *then* Class=W $\{a\}$
rule #2: *if* S=circle *and* SC=red, *then* Class=I $\{b,c\}$
rule #3: *if* SC=white, *then* Class=O $\{d\}$

- Decision rules are *classification patterns* discovered from data contained in the table

CRSA – example of traffic signs

- Alternative set of decision rules

Traffic sign	~~Shape (S)~~	Primary Color (PC)	Secondary color (SC)	Class
a) 	~~triangle~~	yellow	~~red~~	W
b) 	~~circle~~	white	~~red~~	I
c) 	~~circle~~	blue	red	I
d) 	~~circle~~	~~blue~~	white	O

rule #1’: *if* PC=yellow, *then* Class=W {*a*}
rule #2’: *if* PC=white, *then* Class=I {*b* }
rule #3’: *if* PC=blue *and* SC=red, *then* Class=I {*c*}
rule #4’: *if* SC=white, *then* Class=O {*d*}

CRSA – example of traffic signs

- Decision rules induced from the original table

Traffic sign	Shape (S)	Primary Color (PC)	Class
a) 	triangle	~~yellow~~	W
b) 	~~circle~~	white	I
c) 	~~circle~~	blue	I
d) 	~~circle~~	blue	O

rule #1”: *if* S=triangle, *then* Class=W {*a*}
rule #2”: *if* PC=white, *then* Class=I {*b* }
rule #3”: *if* PC=blue, *then* Class=I or O {*c*,*d*}

- Rules #1” & #2” – certain rules induced from lower approximations of W and I
- Rule #3” – approximate rule induced from the boundary of I and O

CRSA – example of traffic signs

- Useful results:
  - a characterization of decision classes (even in case of inconsistency) in terms of chosen attributes by lower and upper approximation,
  - a measure of the quality of approximation indicating how good the chosen set of attributes is for approximation of the classification,
  - reduction of knowledge contained in the table to the description by relevant attributes belonging to reducts,
  - the core of attributes indicating indispensable attributes,
  - decision rules induced from lower and upper approximations of decision classes show classification patterns existing in data.

CRSA – formal definitions

- Approximation space

  U = finite set of objects (universe)

  C = set of condition attributes

  D = set of decision attributes

  $C \cap D = \emptyset$

$$X_C = \prod_{q=1}^{|C|} X_q \text{ – condition attribute space}$$

$$X_D = \prod_{q=1}^{|D|} X_q \text{ – decision attribute space}$$

CRSA – formal definitions

- Indiscernibility relation in the approximation space

  x* is indiscernible with *y by $P \subseteq C$ in X_P iff $x_q = y_q$ for all $q \in P$

  x* is indiscernible with *y by $R \subseteq D$ in X_D iff $x_q = y_q$ for all $q \in R$

  $I_P(x)$, $I_R(x)$ – equivalence classes including x

  I_D makes a partition of U into decision classes $\boldsymbol{Cl} = \{Cl_t, t = 1,\dots,\mathrm{m}\}$

- Granules of knowledge are bounded sets:

  $I_P(x)$ in X_P and $I_R(x)$ in X_R ($P \subseteq C$ and $R \subseteq D$)

- Classification patterns to be discovered are functions representing granules $I_R(x)$ by granules $I_P(x)$

CRSA – illustration of formal definitions

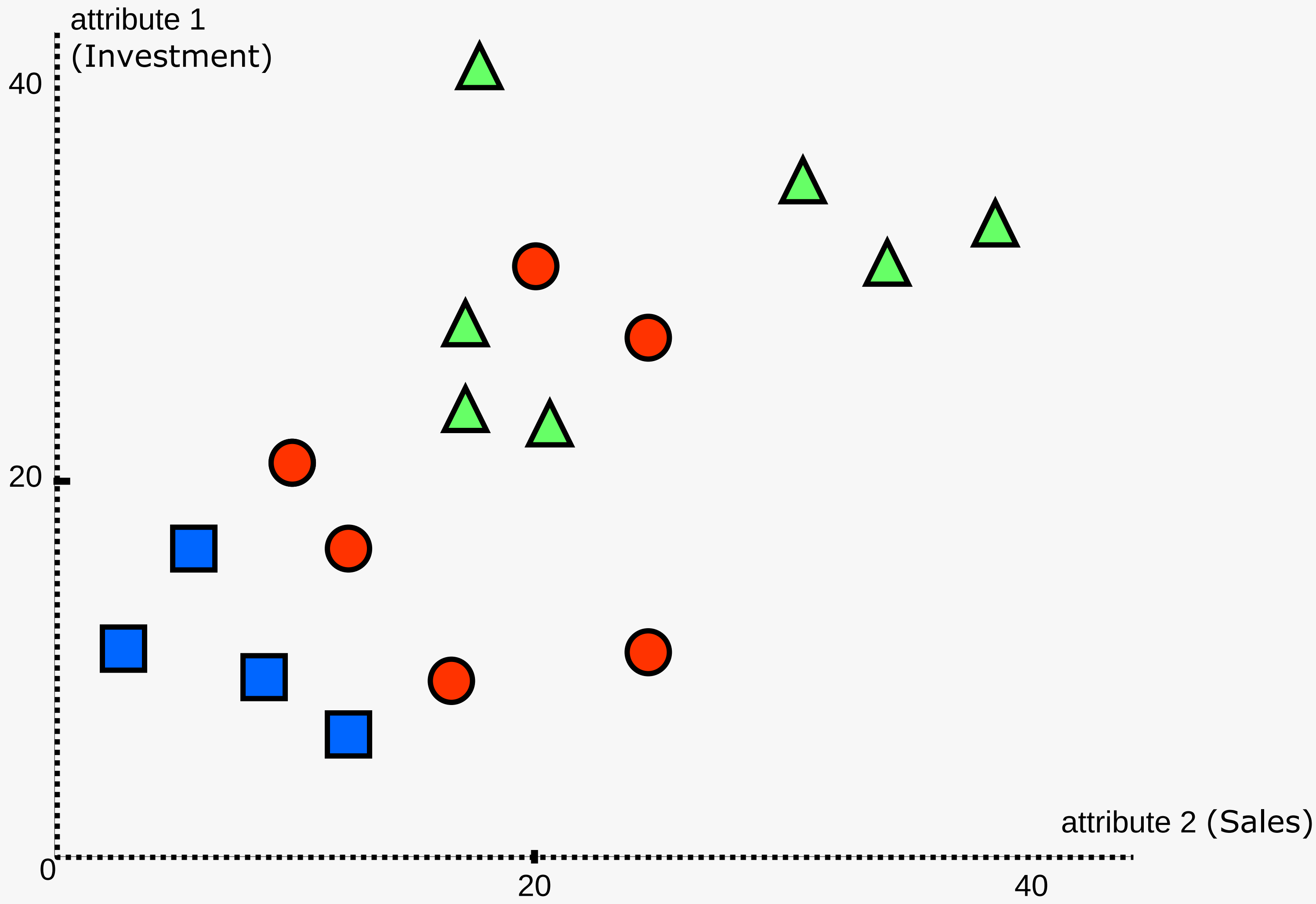
- Example

Objects = firms

Investments	Sales	Effectiveness
40	17,8	High
35	30	High
32.5	39	High
31	35	High
27.5	17.5	High
24	17.5	High
22.5	20	High
30.8	19	Medium
27	25	Medium
21	9.5	Medium
18	12.5	Medium
10.5	25.5	Medium
9.75	17	Medium
17.5	5	Low
11	2	Low
10	9	Low
5	13	Low

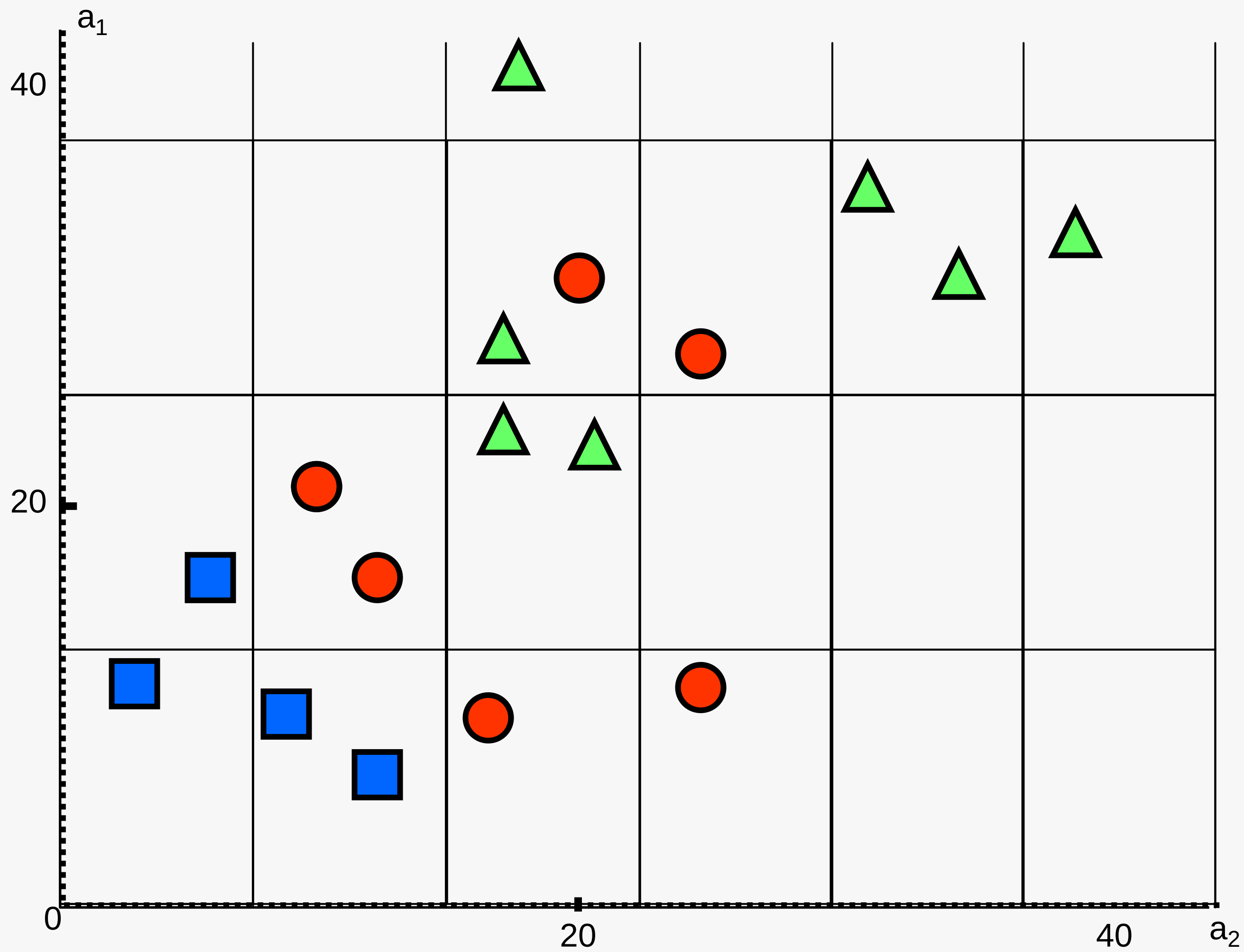
CRSA – illustration of formal definitions

Objects in condition attribute space



CRSA – illustration of formal definitions

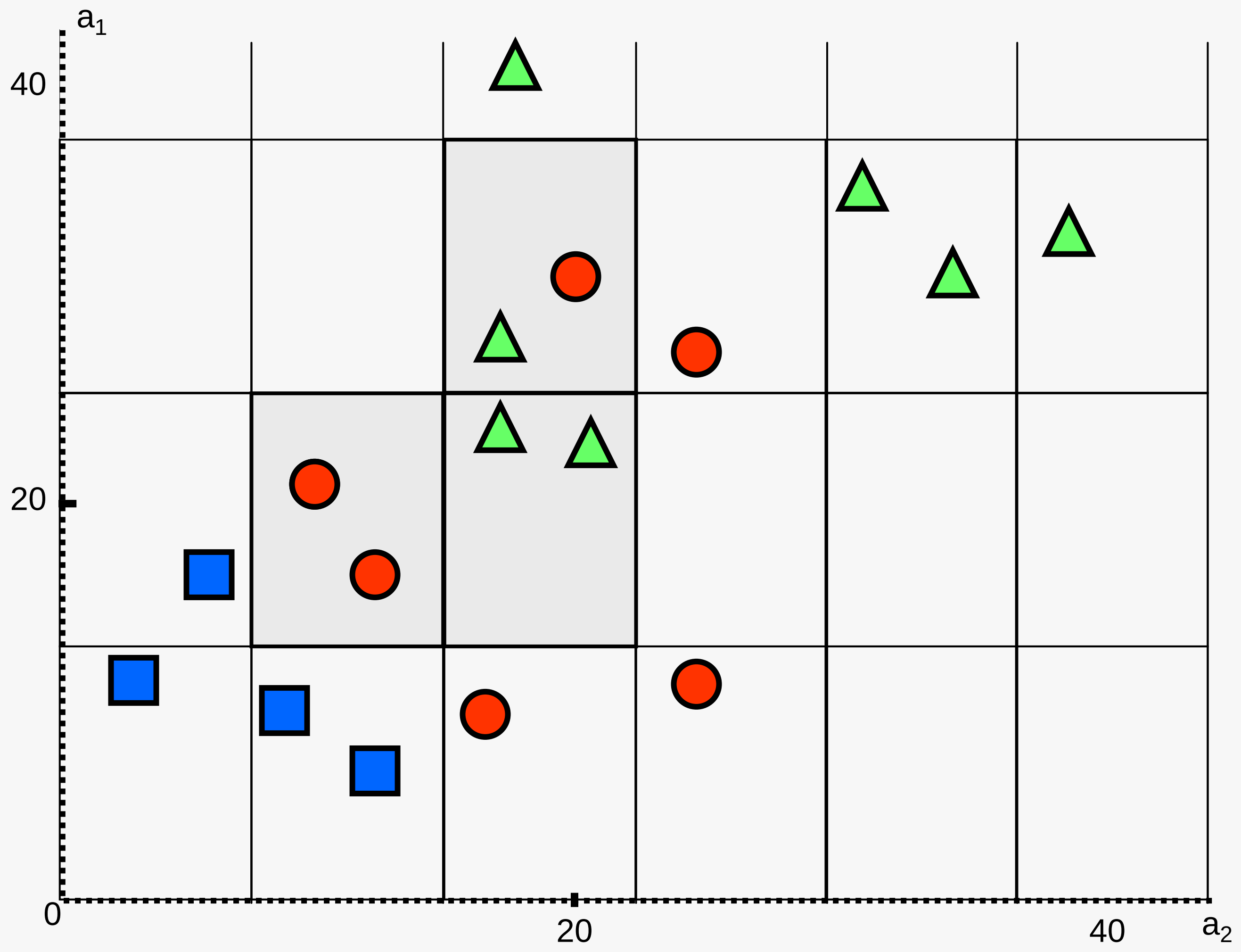
Indiscernibility sets



Quantitative attributes are discretized according to perception of the user

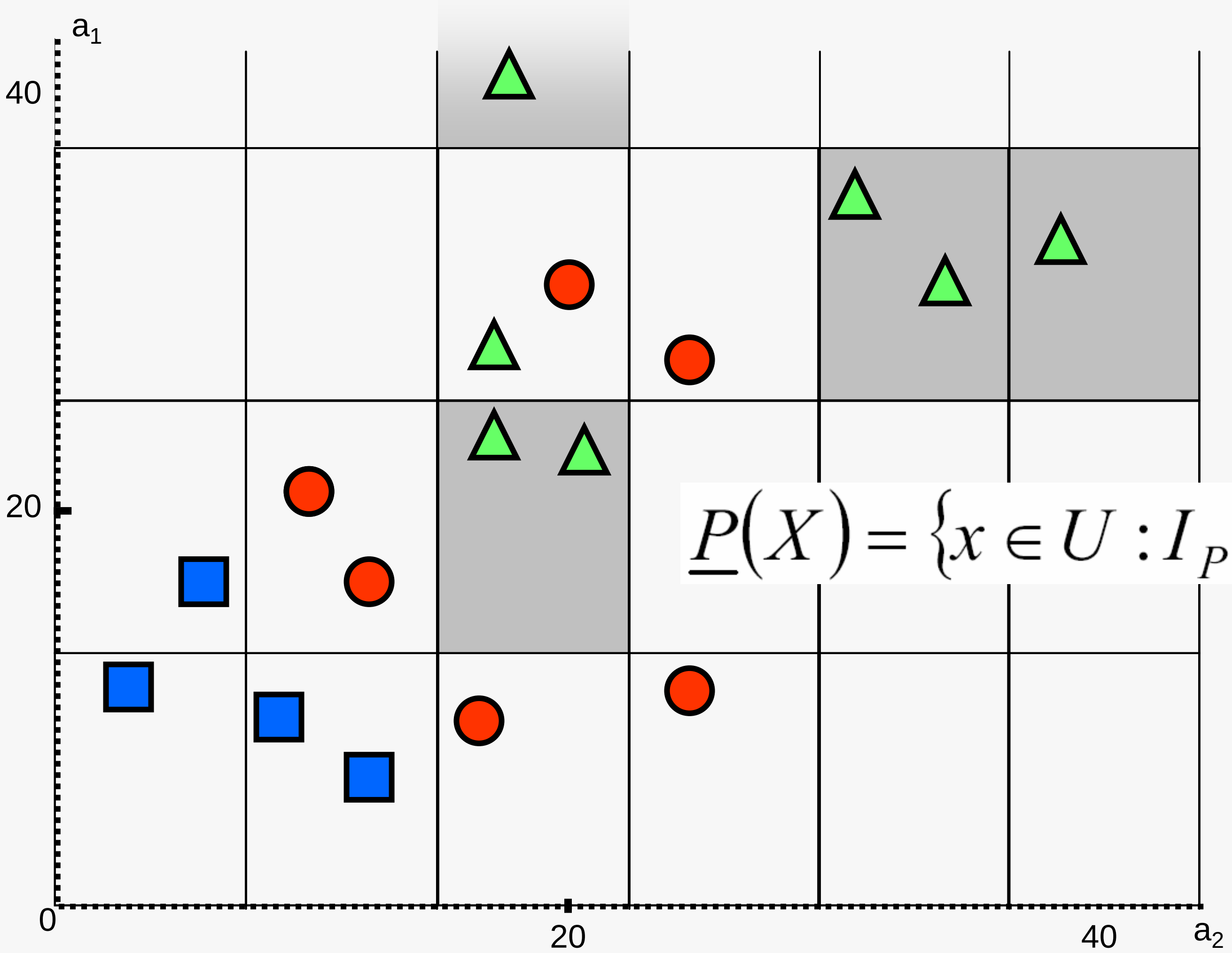
CRSA – illustration of formal definitions

Granules of knowlegde are bounded sets $I_P(x)$



CRSA – illustration of formal definitions

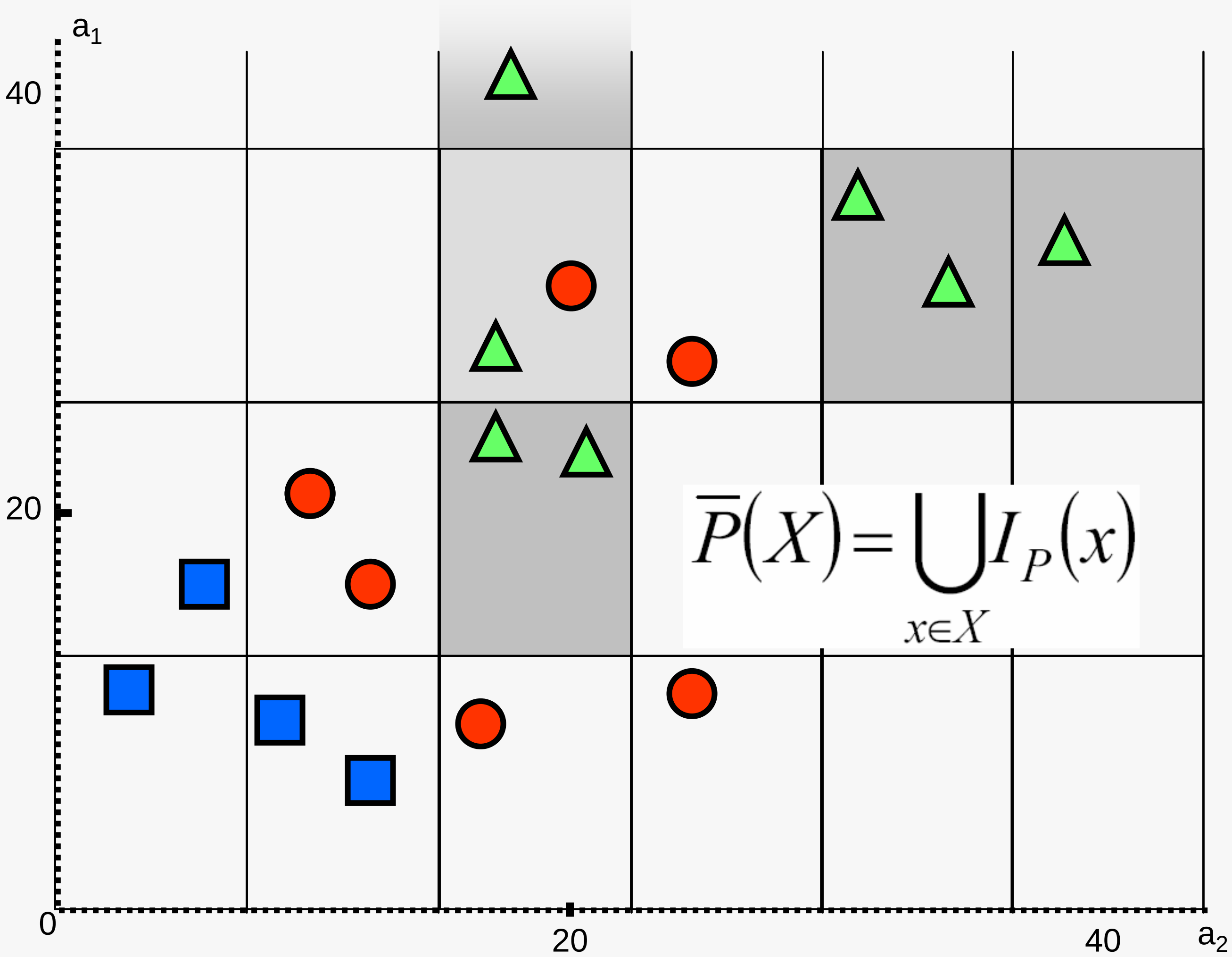
Lower approximation of class High △



$$\underline{P}(X) = \{x \in U : I_P(x) \subseteq X\}$$

CRSA – illustration of formal definitions

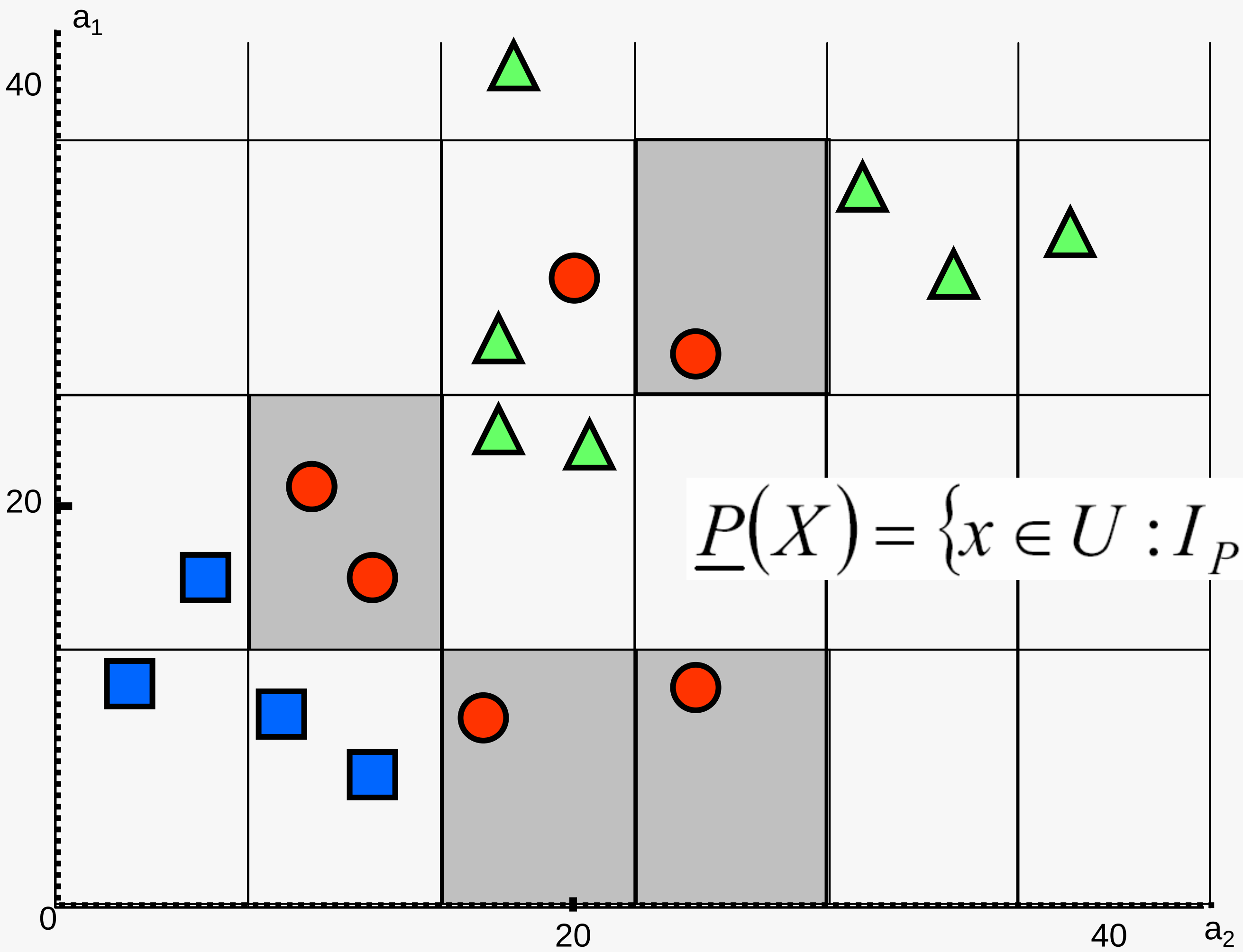
Upper approximation of class High



$$\overline{P}(X)=\bigcup_{x\in X} I_P(x)$$

CRSA – illustration of formal definitions

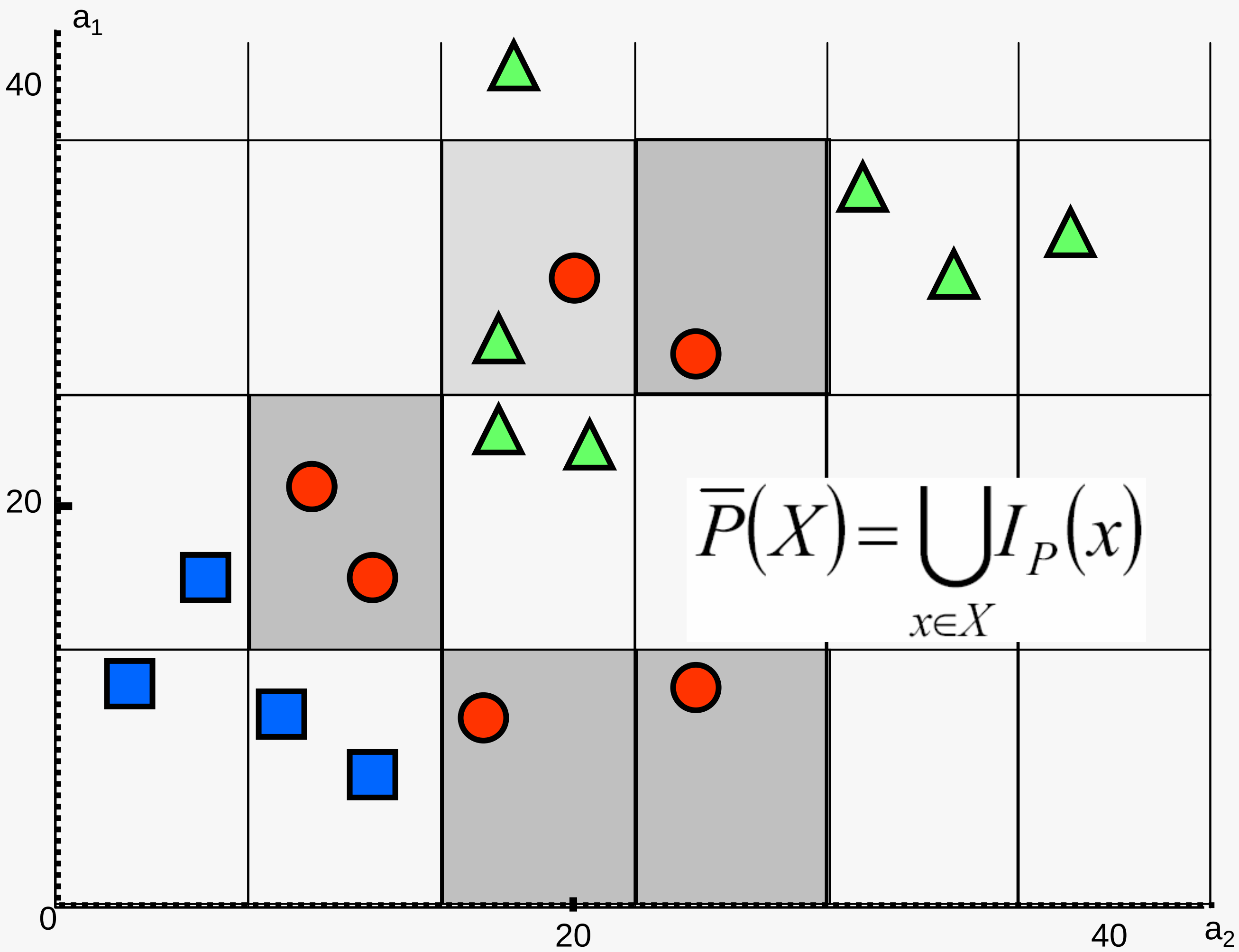
Lower approximation of class Medium



$$\underline{P}(X) = \{x \in U : I_P(x) \subseteq X\}$$

CRSA – illustration of formal definitions

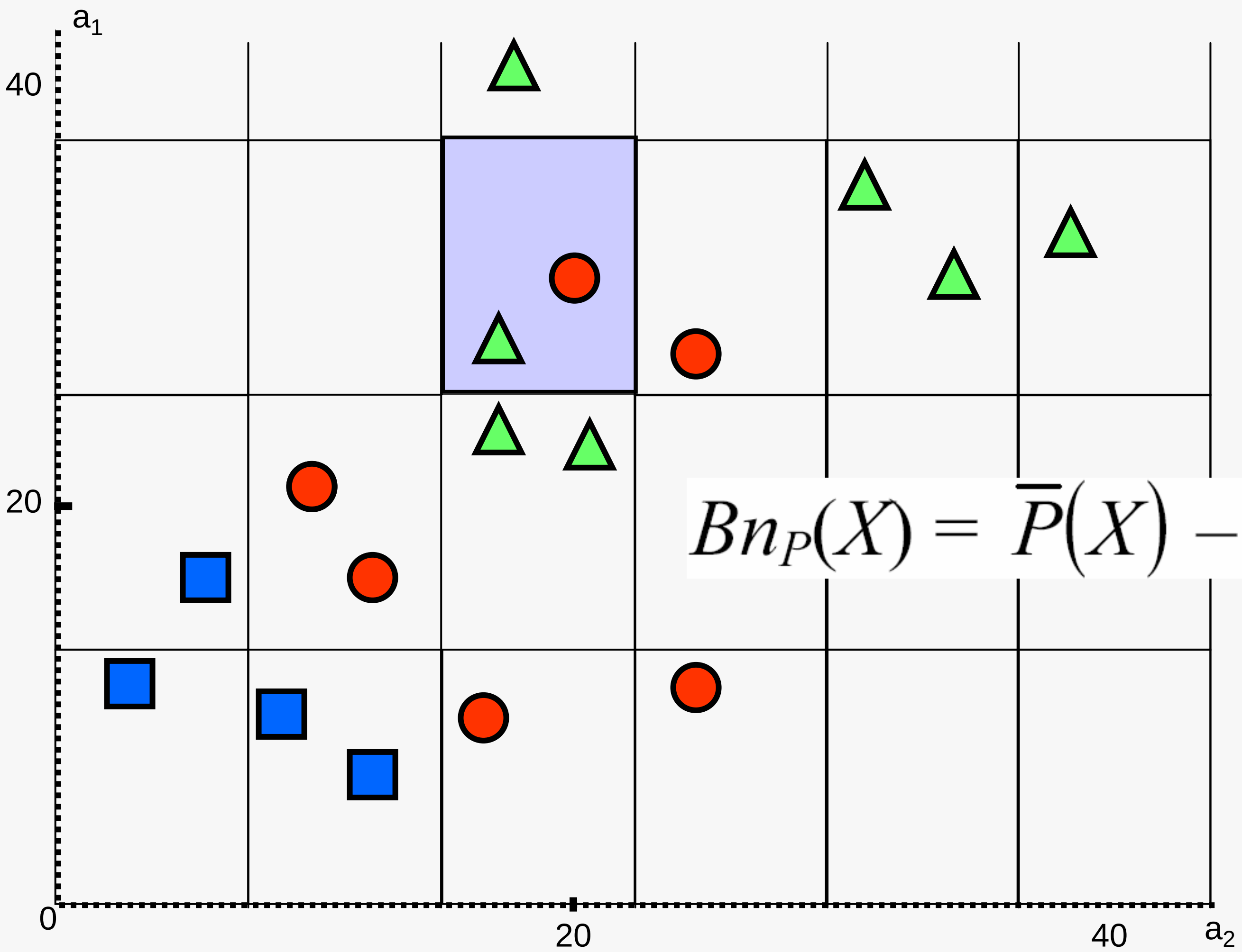
Upper approximation of class Medium



$$\overline{P}(X) = \bigcup_{x \in X} I_P(x)$$

CRSA – illustration of formal definitions

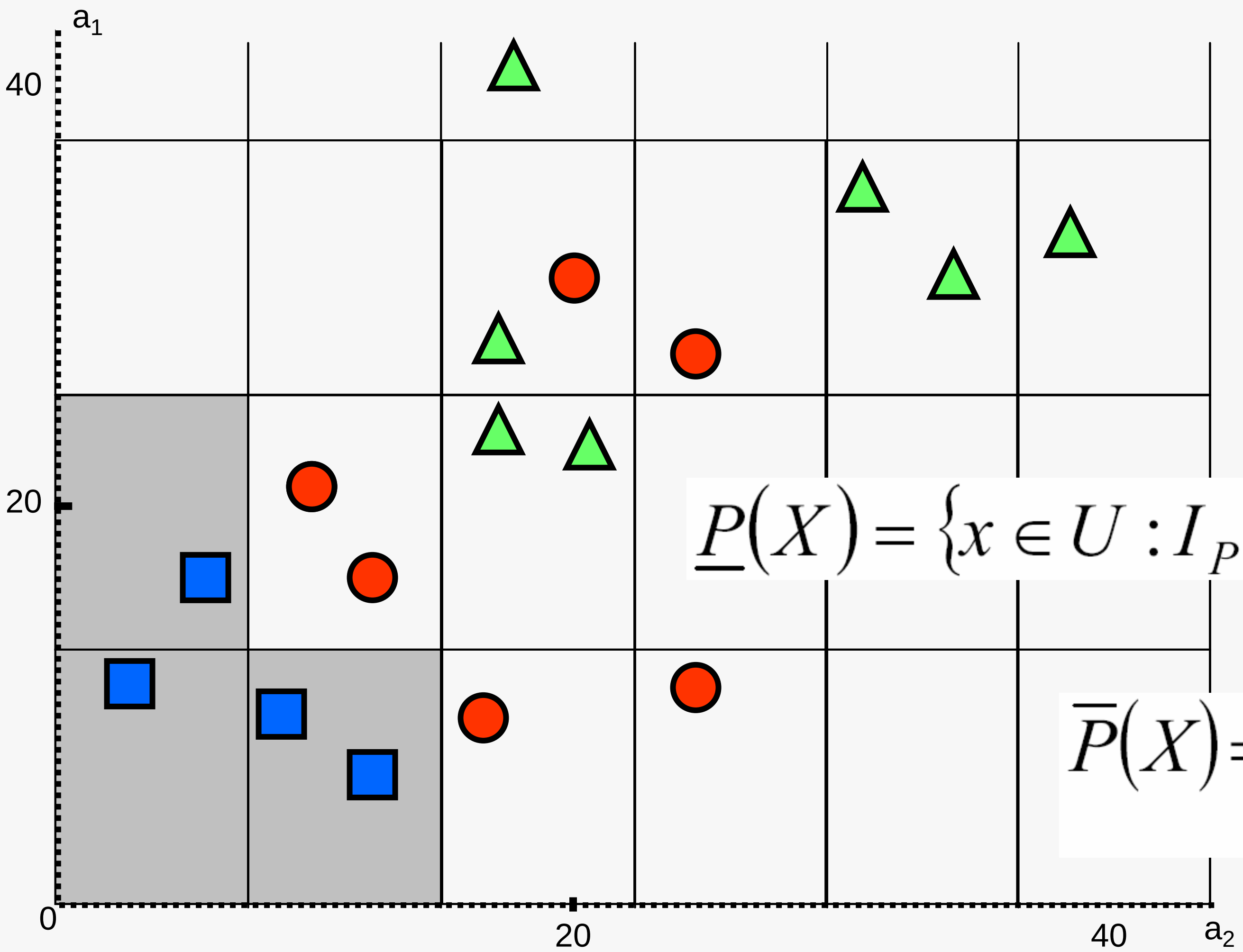
Boundary set of classes High and Medium



$$Bn_P(X) = \overline{P}(X) - \underline{P}(X)$$

CRSA – illustration of formal definitions

Lower = Upper approximation of class Low ■



$$\underline{P}(X) = \{x \in U : I_P(x) \subseteq X\}$$

$$\overline{P}(X) = \bigcup_{x \in X} I_P(x)$$

CRSA – formal definitions

- Basic properies of rough approximations

$$\underline{P}(X) \subseteq X \subseteq \overline{P}(X) \qquad\qquad \underline{P}(X) = U - \overline{P}(U - X)$$

- Accuracy measures
  - Accuracy and quality of approximation of $X \subset U$ by attributes $P \subseteq C$

$$\alpha_P(X) = card(\underline{P}(X))/card(\overline{P}(X)) \qquad \gamma_P(X) = card(\underline{P}(X))/card(X)$$

  - Quality of approximation *of classification* $\boldsymbol{Cl}=\{Cl_t,\ t=1,...m\}$ by attributes $P \subseteq C$

$$\gamma_P(\boldsymbol{Cl}) = \frac{\sum_{t=1}^{m} card(\underline{P}(Cl_t))}{card(U)}$$

  - Rough membership of $x \in U$ to $X \subset U$, given $P \subseteq C$

$$\mu_X^P(x) = \frac{card(X \cap I_P(x))}{card(I_P(x))}$$

CRSA – formal definitions

- ***Cl*-reduct** of $P \subseteq C$, denoted by $RED_{\boldsymbol{Cl}}(P)$, is a minimal subset P' of P which keeps the quality of classification ***Cl*** unchanged, i.e.

$$\gamma_{P'}(\boldsymbol{Cl}) = \gamma_P(\boldsymbol{Cl})$$

- ***Cl*-core** is the intersection of all the ***Cl***-reducts of P:

$$CORE_{\boldsymbol{Cl}}(P) = \bigcap RED_{\boldsymbol{Cl}}(P)$$

R.Słowiński, D.Vanderpooten: A generalized definition of rough approximations based on similarity. *IEEE Transactions on Data and Knowledge Engineering*, 12 (2000) no. 2, 331-336

CRSA – decision rules induced from rough approximations

- Certain decision rule supported by objects from lower approximation of Cl_t (discriminant rule)

$$if \;\; x_{q_1} = r_{q_1} \;\; and \;\; x_{q_2} = r_{q_2} \;\; and \ldots x_{q_p} = r_{q_p}, \;\; then \;\; x \in Cl_t$$

- Possible decision rule supported by objects from upper approximation of Cl_t (partly discriminant rule)

$$if \;\; x_{q_1} = r_{q_1} \;\; and \;\; x_{q_2} = r_{q_2} \;\; and \ldots x_{q_p} = r_{q_p}, \;\; then \;\; x \in Cl_t$$

- Approximate decision rule supported by objects from the boundary of Cl_t

$$if \;\; x_{q_1} = r_{q_1} \;\; and \;\; x_{q_2} = r_{q_2} \;\; and \ldots x_{q_p} = r_{q_p}, \;\; then \;\; x \in Cl_t \; or \; Cl_s \; or \ldots Cl_u$$

where $\{q_1, q_2, \ldots, q_p\} \subseteq C, \; (r_{q_1}, r_{q_2}, \ldots, r_{q_p}) \in V_{q_1} \times V_{q_2} \times \ldots \times V_{q_p}$

$Cl_t, Cl_s, \ldots, Cl_u$ are classes to which belong inconsistent objects supporting this rule

Dominance-based Rough Set Approach (DRSA)

Classical Rough Set Theory

⇓

Indiscernibility principle

If *x* and *y* are indiscernible with respect to all relevant **attributes**,

then *x* should classified to the same class as *y*

Dominace-based Rough Set Theory

⇓

Dominance principle

If *x* is at least as good as *y* with respect to all relevant **criteria**,

then *x* should be classified at least as good as *y*

S.Greco, B.Matarazzo, R.Słowiński: Rough sets theory for multicriteria decision analysis. *European J. of Operational Research*, 129 (2001) no.1, 1-47

What is a criterion ?

- Criterion is a real-valued function g_i defined on U, reflecting a value of each action from a particular point of view, such that in order to compare any two actions $a,b \in U$ from this point of view it is sufficient to compare two values: $g_i(a)$ and $g_i(b)$
- Scales of criteria:
  - Ordinal scale – only the order of values matters; a distance in ordinal scale has no meaning of intensity, so one cannot compare differences of evaluations (e.g. school marks, customer satisfaction, earthquake scales)
  - Cardinal scales – a distance in ordinal scale has a meaning of intensity:
    - **Interval scale** – „zero" in this scale has no absolute meaning, but one can compare differences of evaluations (e.g. Fahrenheit scale)
    - **Ratio scale** – „zero" in this scale has an absolute meaning, so a ratio of evaluations has a meaning (e.g. weight, Kelvin scale)

Dominance principle as monotonicity principle

- Interpretation of the dominance principle

> The better the evaluation of x with respect to considered criteria,
>
> the better its comprehensive evaluation

- Many other relationships of this type, e.g.:
  - The faster the car, the more expensive it is
  - The higher the inflation, the higher the interest rate
  - The larger the mass and the smaller the distance, the larger the gravity
  - The colder the weather, the greater the energy consumption

- The Dominance-based Rough Set Approach does not only permit representation and analysis of decision problems but, more generally, representation and analysis of all phenomena involving monotonicity

Monotonicity: general idea

- Monotonicity concerns relationship between different aspects of a phenomenon described by data

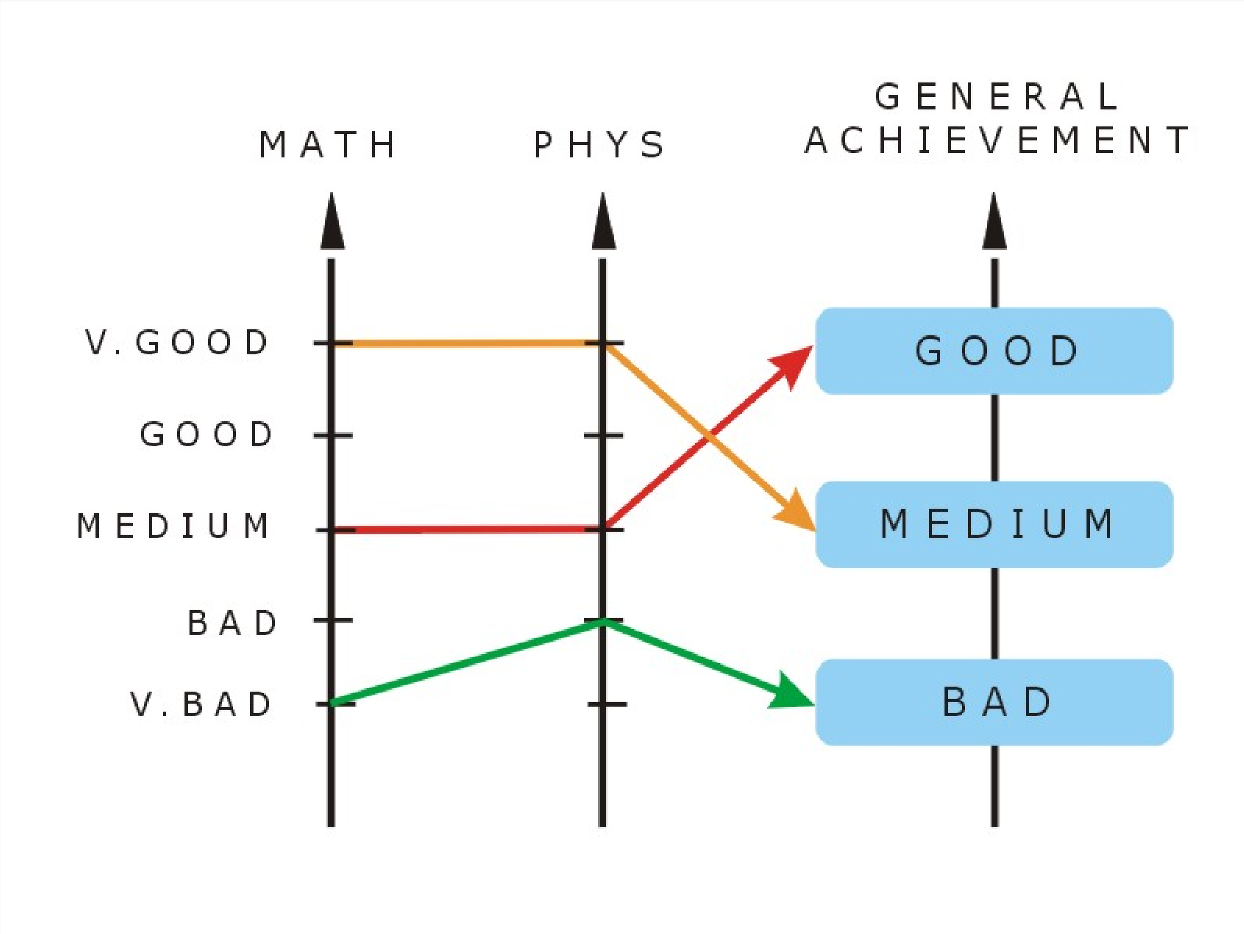
- Whenever we discover a relationship between different aspects of a phenomenon, this relationship can be represented by monotonicity with respect to some specific measures or perceptions of the considered aspects

E.g. „*the more a tomato is red, the more it is ripe*"

R.Słowiński, S.Greco, B.Matarazzo: Rough set based decision support. Chapter 16 [in]: E.K.Burke and G.Kendall (eds.), *Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques*, Springer-Verlag, New York, 2005, pp. 475-527

Why Classical Rough Set Approach has to be adapted to MCDM?

- Ordinal classification with monotonicity constraints: inconsistency w.r.t. dominance principle (Pareto principle)



Why Classical Rough Set Approach has to be adapted to MCDM?

- Classical rough set approach does not detect inconsistency w.r.t. dominance (Pareto principle)

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**medium**	**medium**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**bad**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

Monotonicity, induction and data analysis: between Wittgenstein and Mill

- *„The process of induction is the process of assuming the simplest law that can be made to harmonize with our experience"* (Wittgenstein 1922)

- This simplest law is just monotonicity and, therefore, data analysis can be seen as a specific way of dealing with monotonicity

- Considering monotonicity in data mining means to search for positive or negative relations between magnitudes of considered variables and this is concordant with the method of concomitant variation (Mill 1843)

Monotonicity, induction and data analysis: between Wittgenstein and Mill

- *„Whatever phenomenon varies in any manner whenever another phenomenon varies in some particular manner, is either a cause or an effect of that phenomenon, or it is connected with it through some causation”* (Mill 1843)

- *„The one canon, which receives the least attention [in data mining] is that of concomitant variation, and it is this which is believed to have the greatest potential for the discovery of knowledge, in such areas as biology and biomedicine, as it addresses parameters which are forever present and inseparable, but do change”* (Cornish & Elliman 1995)

Dominance-based Rough Set Approach (DRSA)

- Sets of condition (C) and decision (D) criteria are monotonically dependent
- $\succeq_q$ – weak preference relation (outranking) on U w.r.t. criterion $q \in \{C \cup D\}$ (complete preorder)
- $x_q \succeq_q y_q$: "x_q is at least as good as y_q on criterion q"
- xD_Py : ***x* dominates *y*** with respect to $P \subseteq C$ in condition space X_P if $x_q \succeq_q y_q$ for all criteria $q \in P$
- $D_P = \bigcap_{q \in P} \succeq_q$ is a partial preorder
- Analogically, we define xD_Ry in decision space X_R, $R \subseteq D$

Dominance-based Rough Set Approach (DRSA)

- For simplicity : $D=\{d\}$
- I_d makes a partition of U into decision classes $\boldsymbol{Cl}=\{Cl_t,\ t=1,...,m\}$
- $[x \in Cl_r,\ y \in Cl_s,\ r>s] \Rightarrow x \succ y$ ($x \succeq y$ and *not* $y \succeq x$)
- In order to handle monotonic dependency between condition and decision criteria:

$Cl_t^{\geq} = \bigcup_{s \geq t} Cl_s$ – upward union of classes, t=2,...,m („*at least*" class Cl_t)

$Cl_t^{\leq} = \bigcup_{s \leq t} Cl_s$ – downward union of classes, t=1,...,m-1 („*at most*" class Cl_t)

- $Cl_t^{\geq}$ and $Cl_t^{\leq}$ are positive and negative dominance cones in X_D, with D reduced to single dimension d

Granular computing with dominance cones

- Granules of knowledge are dominance cones in condition space X_P ($P \subseteq C$)

  $D_P^+(x) = \{y \in U: yD_Px\}$: *P-dominating set*

  $D_P^-(x) = \{y \in U: xD_Py\}$: *P-dominated set*

- *P*-dominating and *P*-dominated sets are positive and negative dominance cones in X_P

- Classification patterns (preference model) to be discovered are functions representing granules $Cl_t^{\geq}$, $Cl_t^{\leq}$, by granules $D_P^+(x)$, $D_P^-(x)$

DRSA – illustration of formal definitions

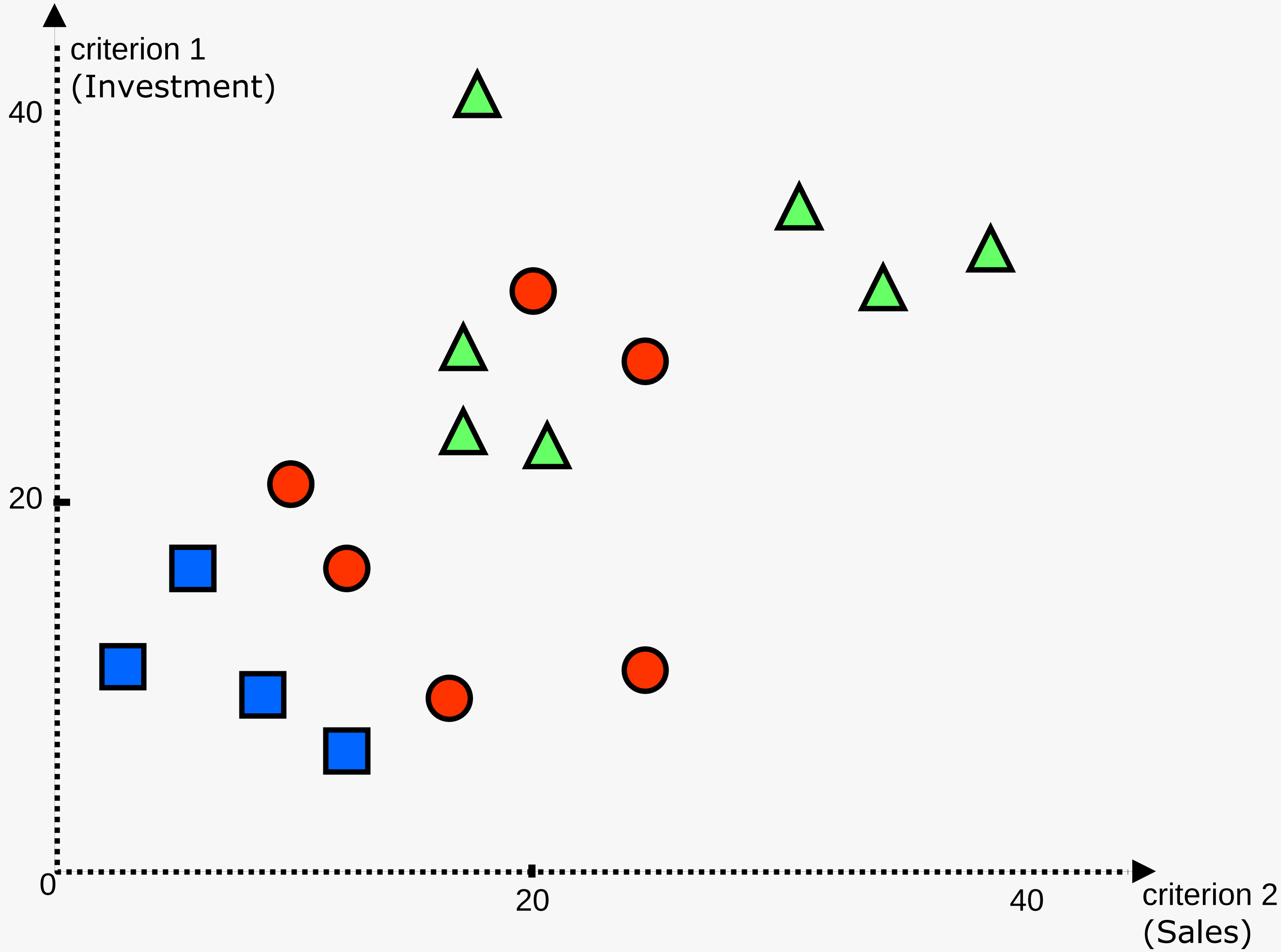
- Example

△ $\succ$ ○ $\succ$ □

Investments ↑	Sales ↑	Effectiveness ↑
40	17,8	△ High
35	30	△ High
32.5	39	△ High
31	35	△ High
27.5	17.5	△ High
24	17.5	△ High
22.5	20	△ High
30.8	19	○ Medium
27	25	○ Medium
21	9.5	○ Medium
18	12.5	○ Medium
10.5	25.5	○ Medium
9.75	17	○ Medium
17.5	5	□ Low
11	2	□ Low
10	9	□ Low
5	13	□ Low

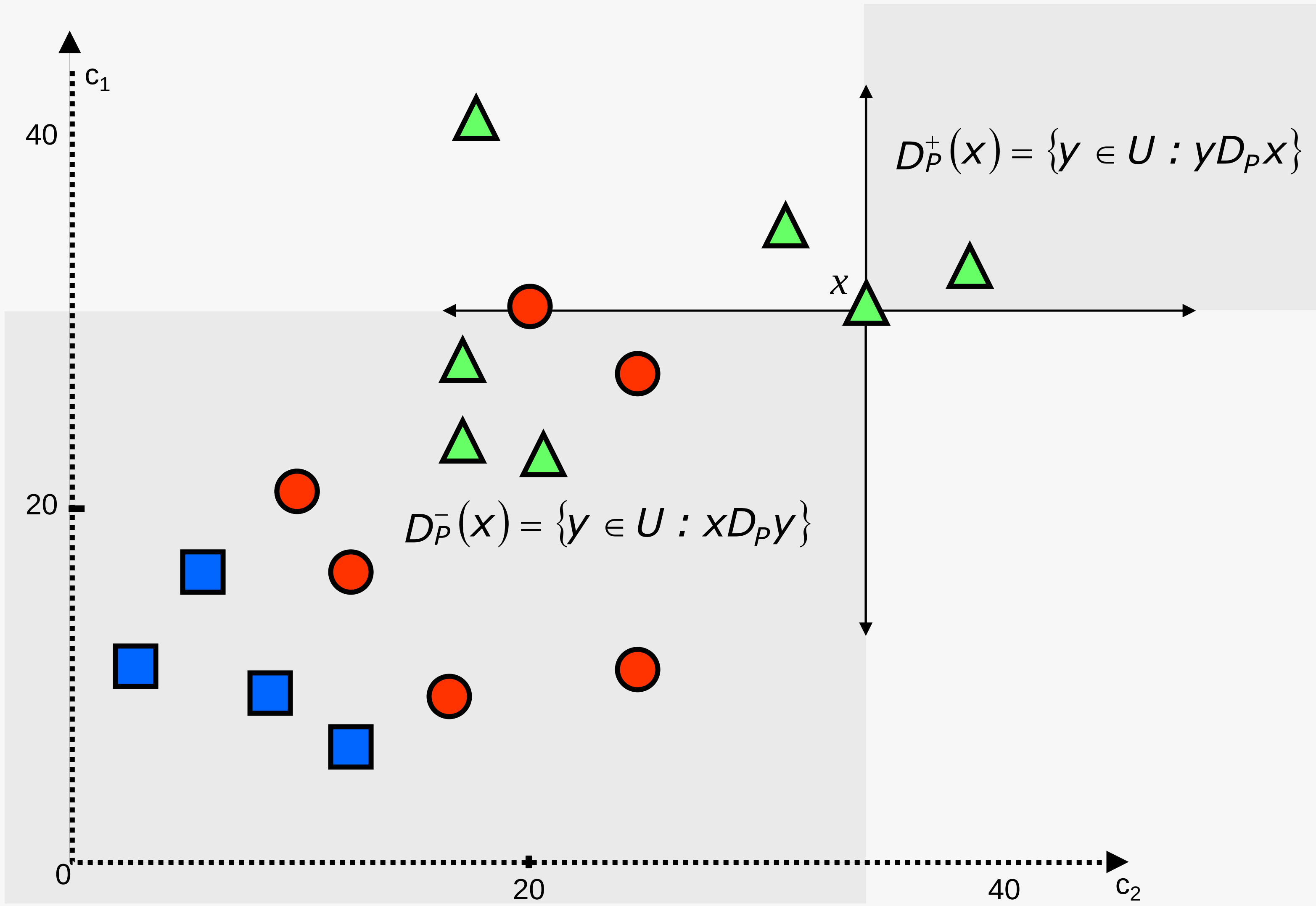
DRSA – illustration of formal definitions

Objects in condition criteria space



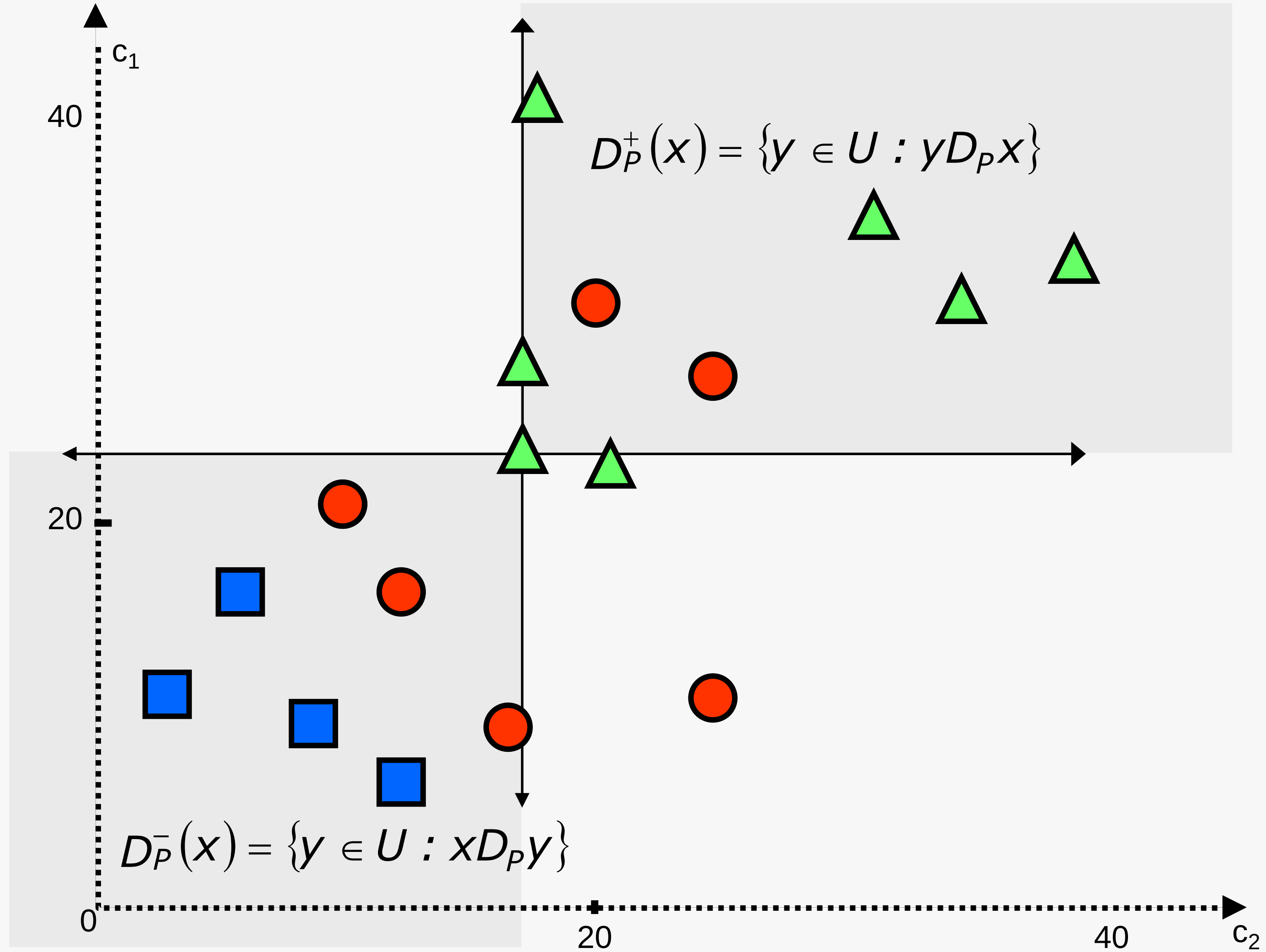
DRSA – illustration of formal definitions

Granular computing with dominance cones



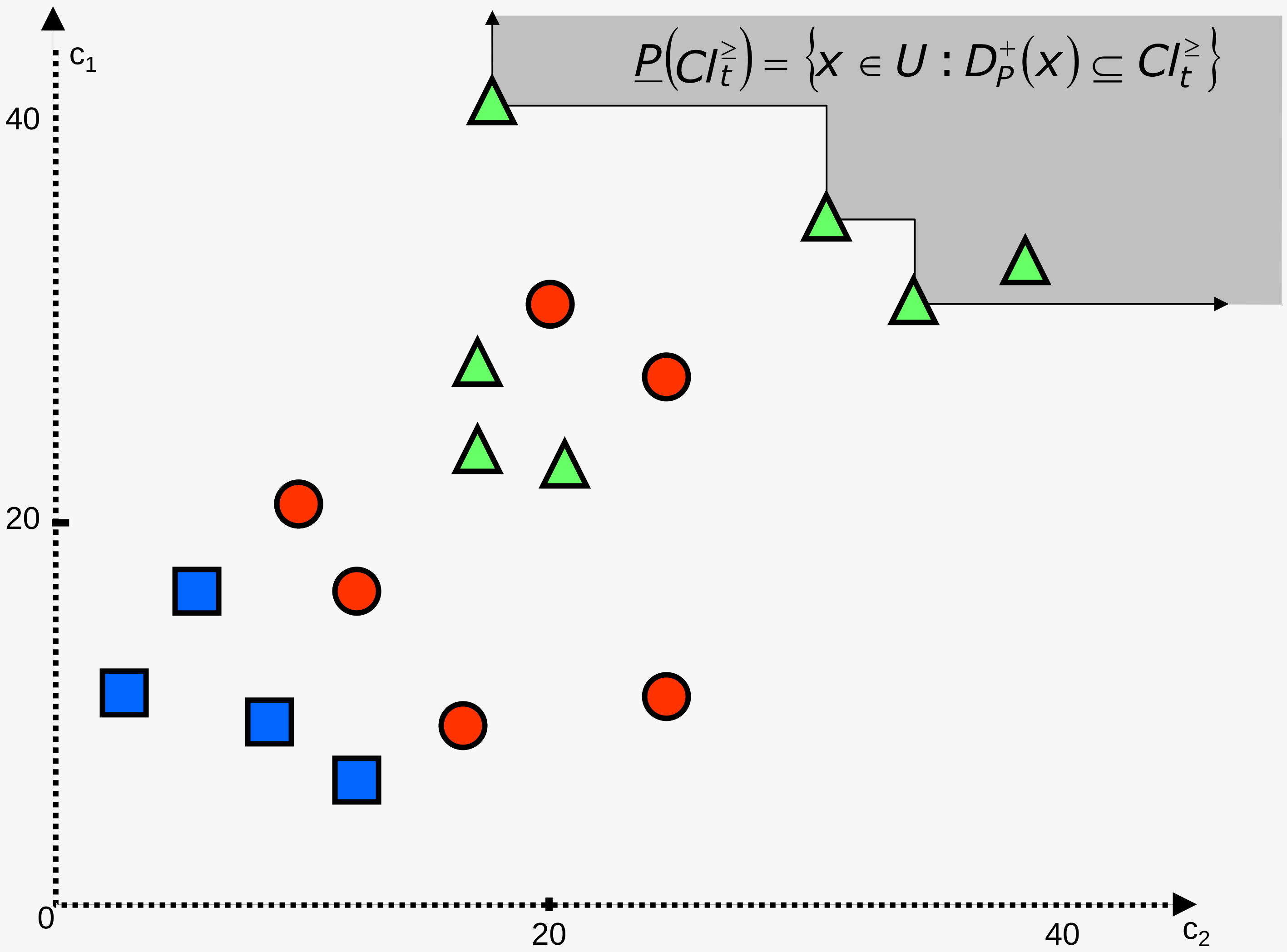
DRSA – illustration of formal definitions

Granular computing with dominance cones



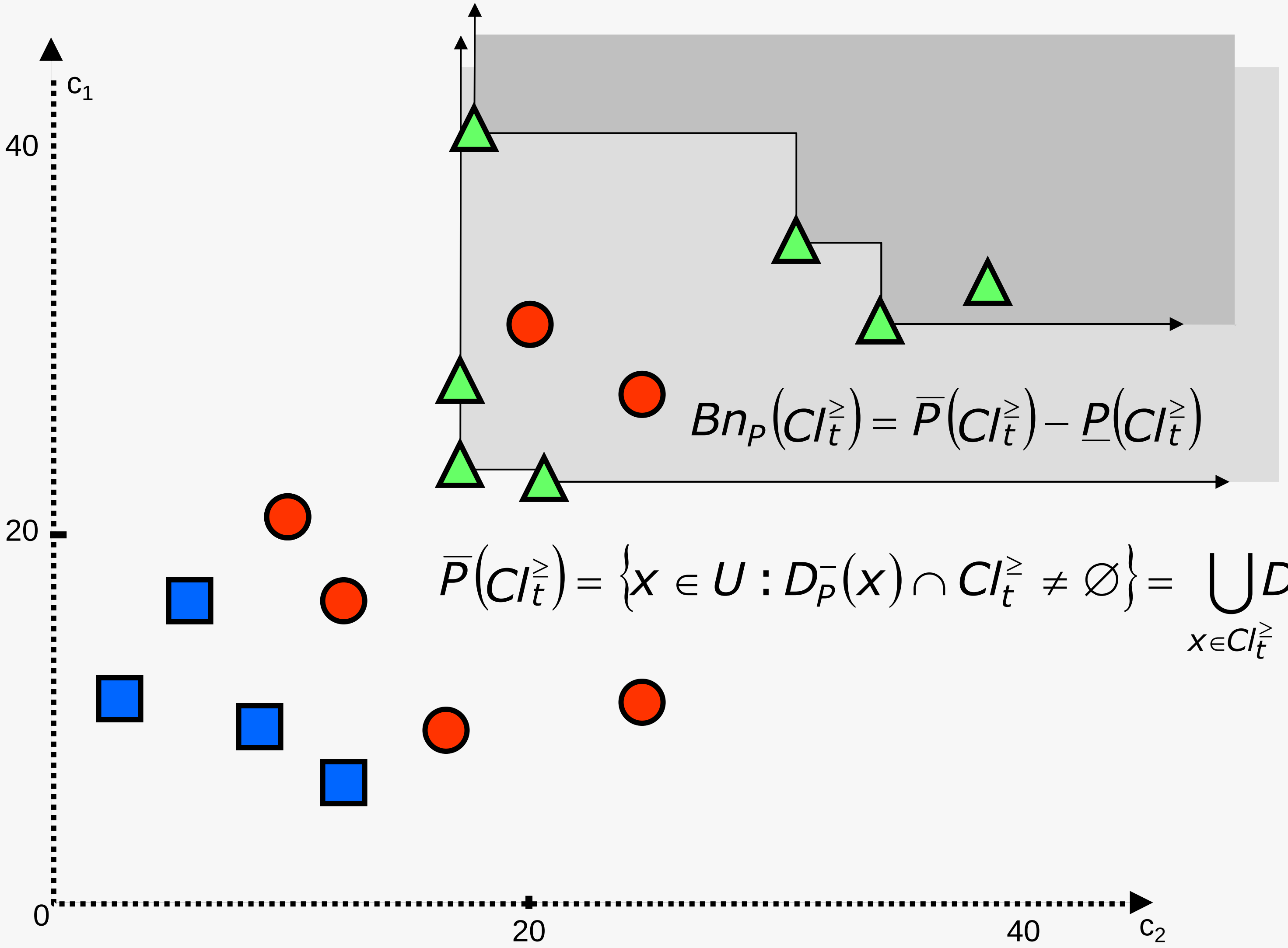
DRSA – illustration of formal definitions

Lower approximation of upward union of class High



DRSA – illustration of formal definitions

Upper approximation and the boundary of upward union of class High

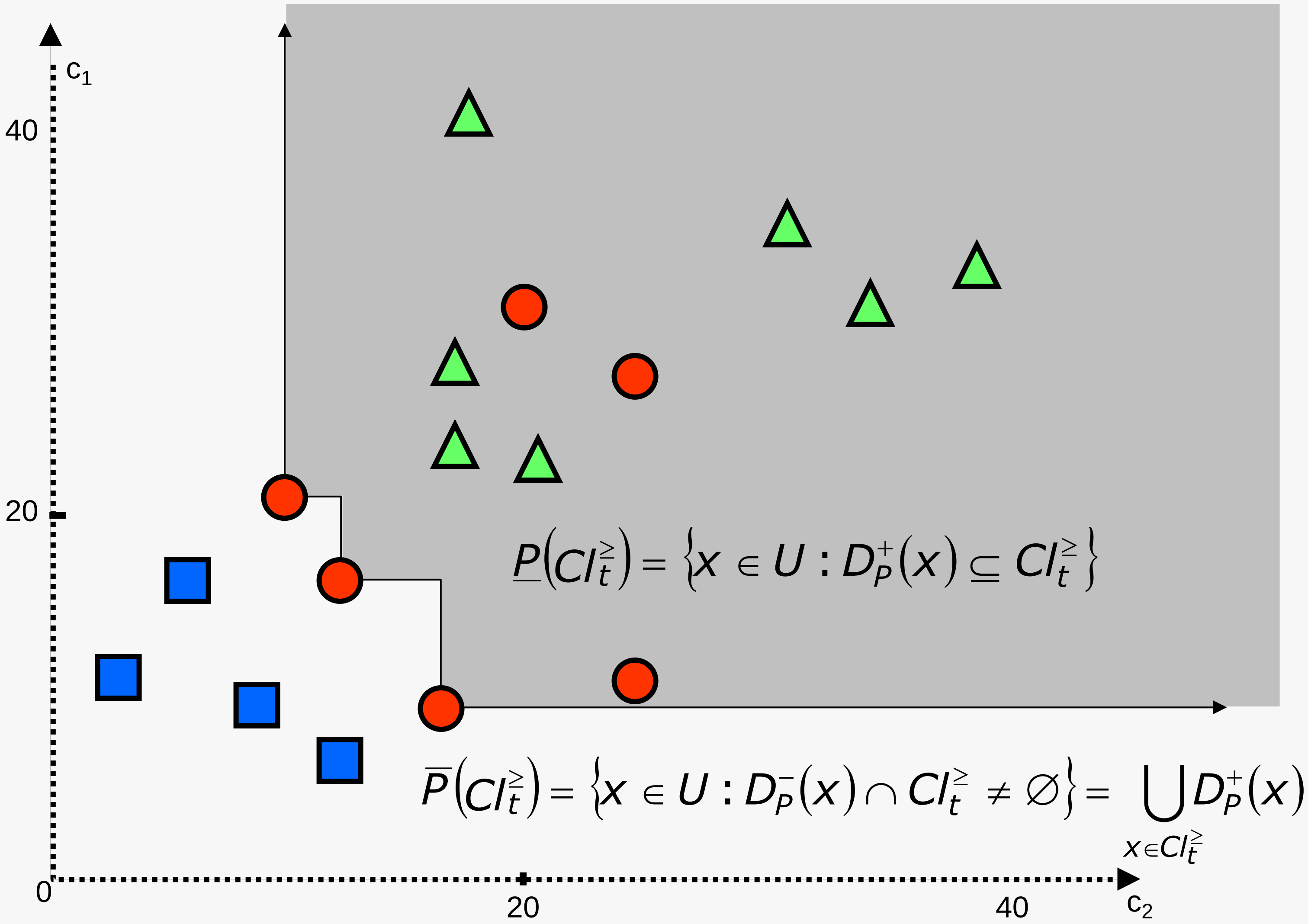


$$Bn_P\left(Cl_t^{\geq}\right) = \overline{P}\left(Cl_t^{\geq}\right) - \underline{P}\left(Cl_t^{\geq}\right)$$

$$\overline{P}\left(Cl_t^{\geq}\right) = \left\{x \in U : D_P^{-}(x) \cap Cl_t^{\geq} \neq \varnothing\right\} = \bigcup_{x \in Cl_t^{\geq}} D_P^{+}(x)$$

DRSA – illustration of formal definitions

Lower = Upper approximation of <u>upward</u> union of class Medium

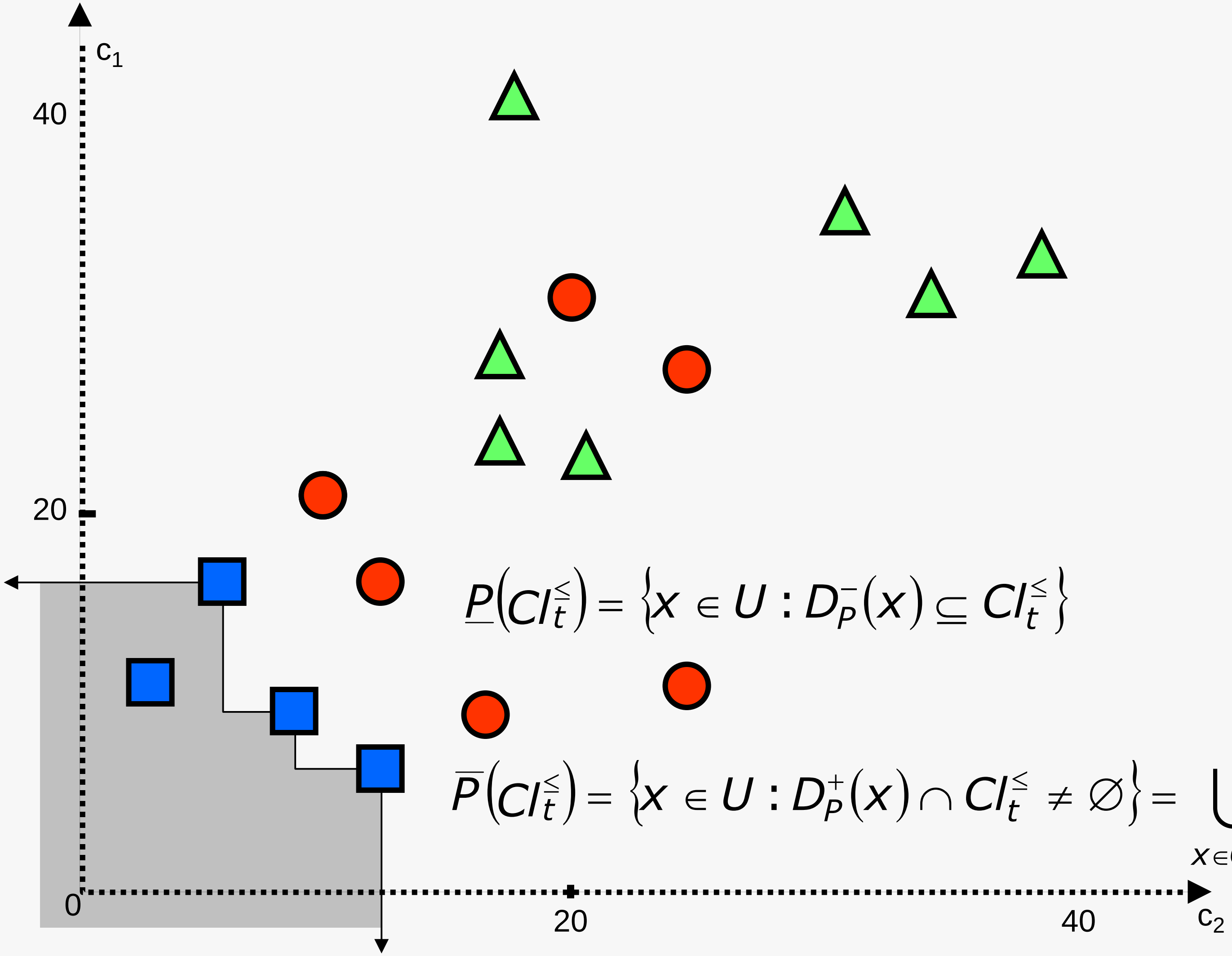


$$\underline{P}\left(Cl_t^{\geq}\right) = \left\{x \in U : D_P^+(x) \subseteq Cl_t^{\geq}\right\}$$

$$\overline{P}\left(Cl_t^{\geq}\right) = \left\{x \in U : D_P^-(x) \cap Cl_t^{\geq} \neq \emptyset\right\} = \bigcup_{x \in Cl_t^{\geq}} D_P^+(x)$$

DRSA – illustration of formal definitions

Lower = upper approximation of <u>downward</u> union of class Low

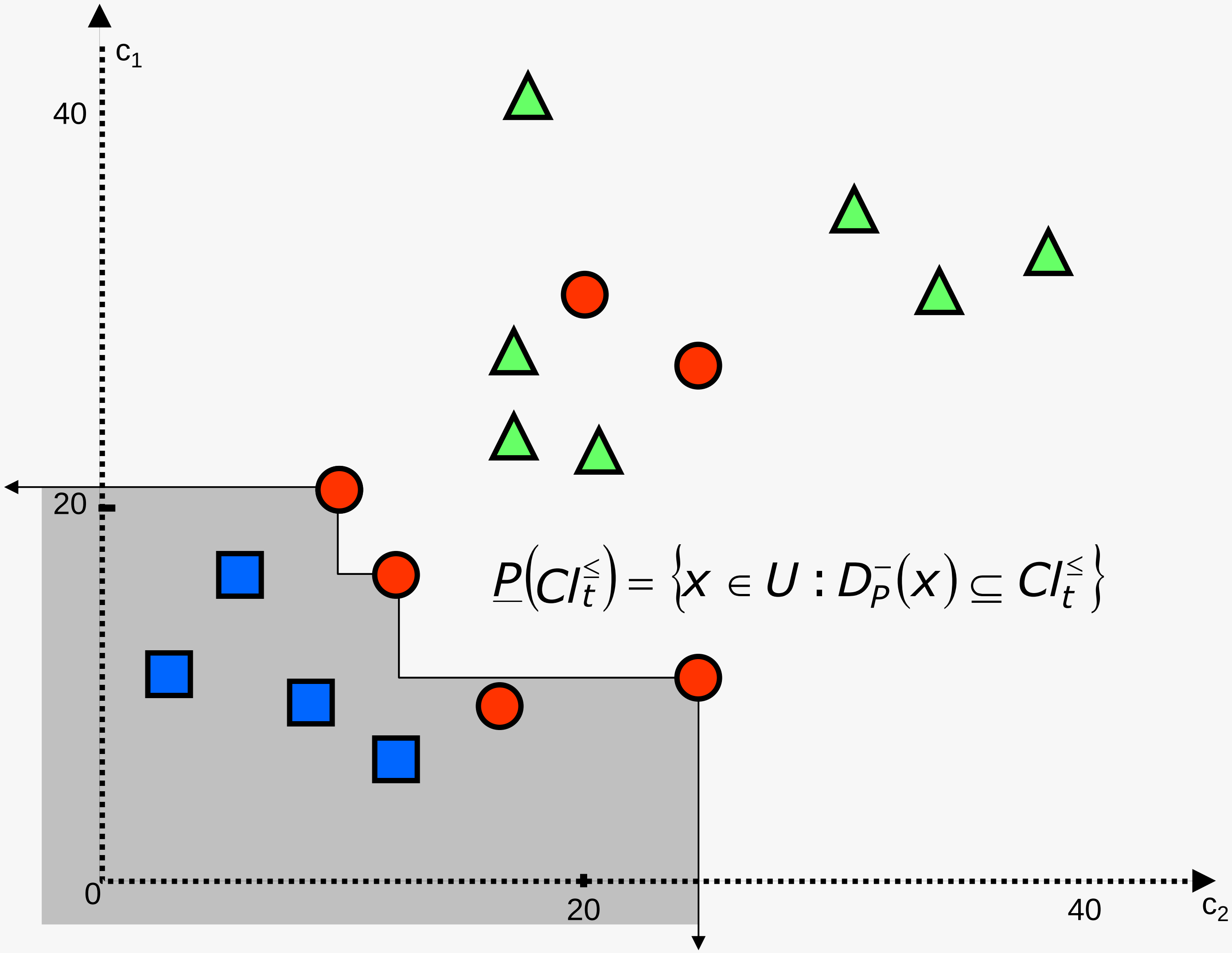


$$\underline{P}\left(Cl_t^{\le}\right) = \left\{x \in U : D_P^-(x) \subseteq Cl_t^{\le}\right\}$$

$$\overline{P}\left(Cl_t^{\le}\right) = \left\{x \in U : D_P^+(x) \cap Cl_t^{\le} \neq \emptyset\right\} = \bigcup_{x \in Cl_t^{\le}} D_P^-(x)$$

DRSA – illustration of formal definitions

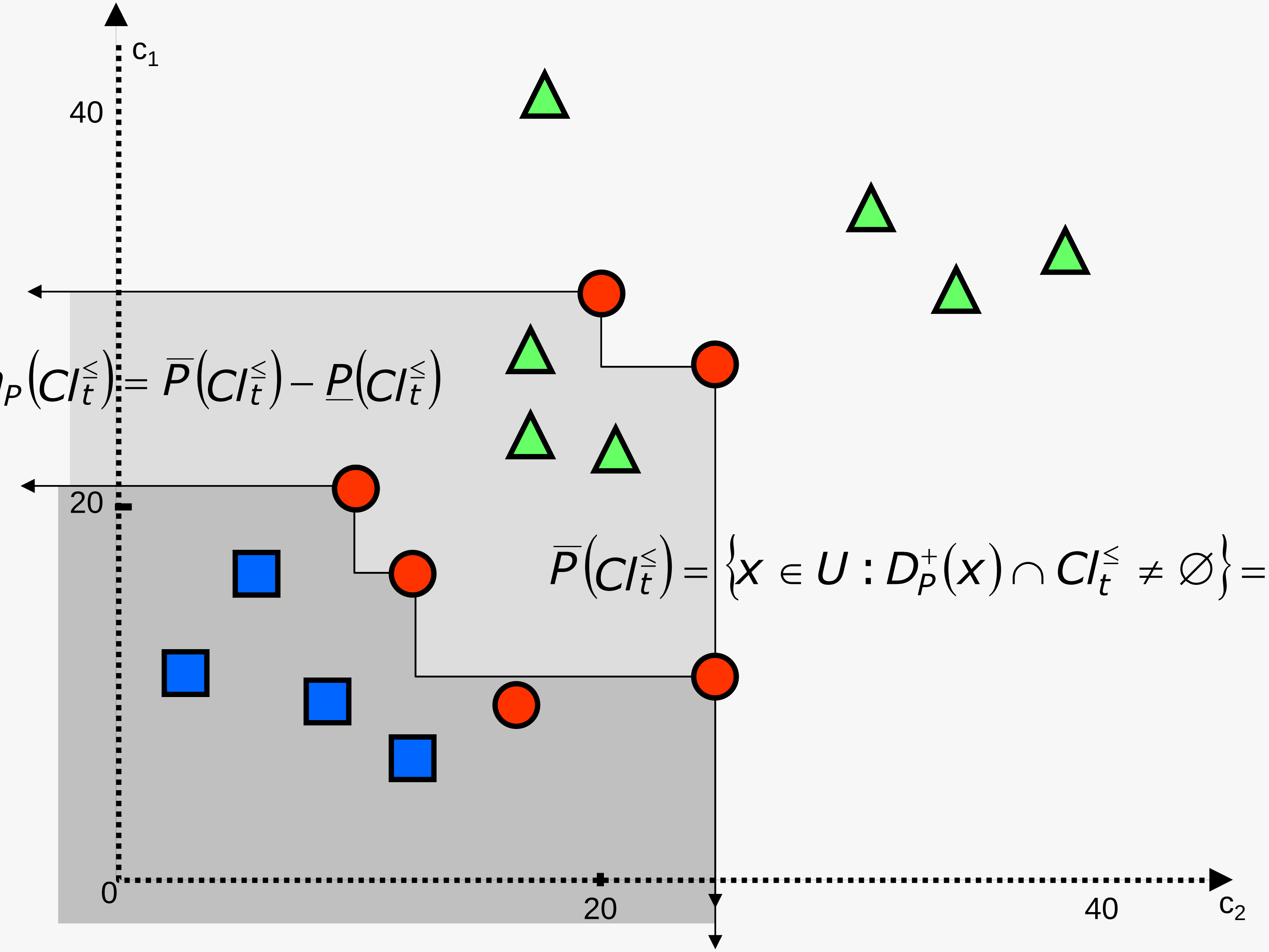
Lower approximation of downward union of class Medium



$$\underline{P}\left(Cl_t^{\le}\right) = \left\{x \in U : D_P^{-}(x) \subseteq Cl_t^{\le}\right\}$$

DRSA – illustration of formal definitions

Upper approximation and the boundary of downward union of class Medium



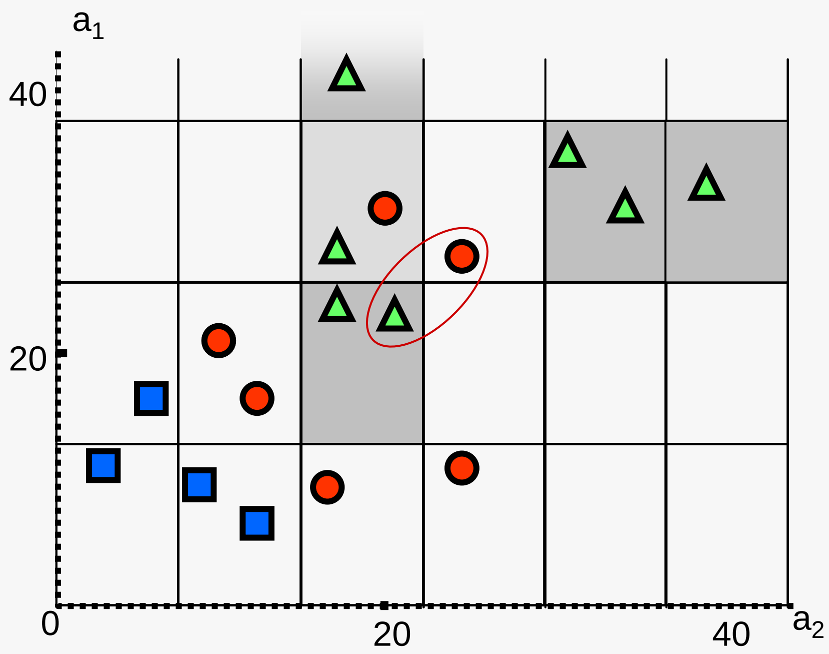
$$Bn_P\left(Cl_t^{\leq}\right) = \overline{P}\left(Cl_t^{\leq}\right) - \underline{P}\left(Cl_t^{\leq}\right)$$

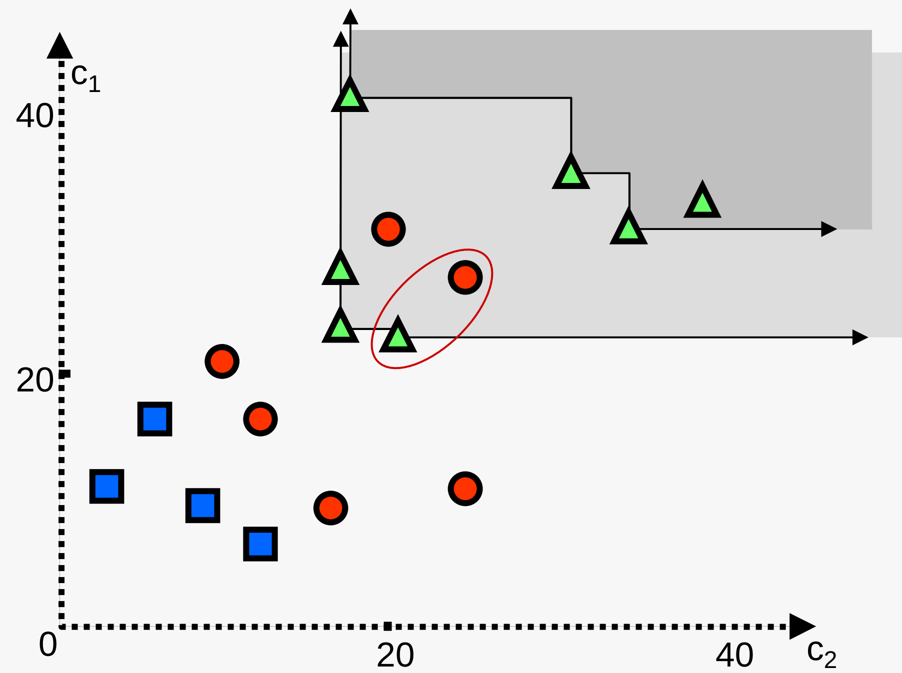
$$\overline{P}\left(Cl_t^{\leq}\right) = \left\{x \in U : D_P^{+}(x) \cap Cl_t^{\leq} \neq \emptyset\right\} = \bigcup_{x \in Cl_t^{\leq}} D_P^{-}(x)$$

Dominance-based Rough Set Approach vs. Classical RSA

Comparison of CRSA and DRSA

Classes: $\triangle \succ \bullet \succ \blacksquare$





$$\underline{P}(X) = \{x \in U : I_P(x) \subseteq X\}$$

$$\overline{P}(X) = \bigcup_{x \in X} I_P(x)$$

$$\underline{P}\left(Cl_t^{\geq}\right) = \left\{x \in U : D_P^{+}(x) \subseteq Cl_t^{\geq}\right\}$$

$$\overline{P}\left(Cl_t^{\geq}\right) = \bigcup_{x \in Cl_t^{\geq}} D_P^{+}(x)$$

Rough Set approach to multiple-criteria sorting

- Example of preference information about students:

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**good**	**good**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**bad**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

- Examples of classification of S1 and S2 are inconsistent

S.Greco, B.Matarazzo, R.Słowiński: Decision rule approach. Chapter 13 [in]: J.Figueira, S.Greco and M.Ehrgott (eds.), *Multiple Criteria Decision Analysis: State of the Art Surveys*, Springer-Verlag, New York, 2005, pp. 507-562

Rough Set approach to multiple-criteria sorting

- If we eliminate Literature, then more inconsistencies appear:

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**good**	**good**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**bad**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

- Examples of classification of S1, S2, S3 and S5 are inconsistent

Rough Set approach to multiple-criteria sorting

- Elimination of Mathematics does not increase inconsistencies:

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**good**	**good**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**bad**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

- Subset of criteria {Ph,L} is a reduct of {M,Ph,L}

Rough Set approach to multiple-criteria sorting

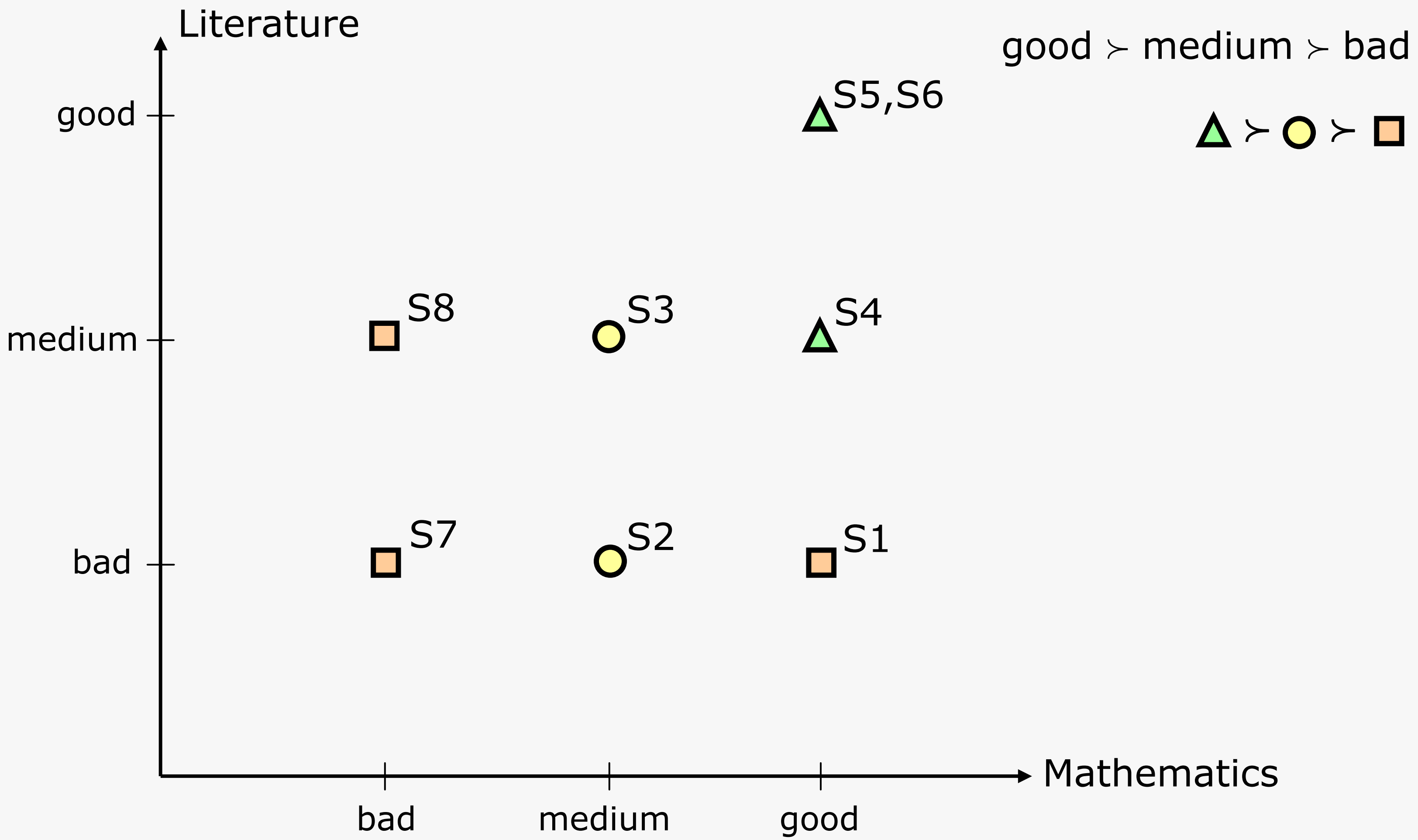
- Elimination of Physics also does not increase inconsistencies:

Student	Mathematics (**M**)	Physics (**Ph**)	Literature (**L**)	Overall class
S1	**good**	**medium**	**bad**	**bad**
S2	**medium**	**medium**	**bad**	**medium**
S3	**medium**	**medium**	**medium**	**medium**
S4	**good**	**good**	**medium**	**good**
S5	**good**	**medium**	**good**	**good**
S6	**good**	**good**	**good**	**good**
S7	**bad**	**bad**	**bad**	**bad**
S8	**bad**	**bad**	**medium**	**bad**

- Subset of criteria {M,L} is a reduct of {M,Ph,L}
- Intersection of reducts {M,L} and {Ph,L} gives the core {L}

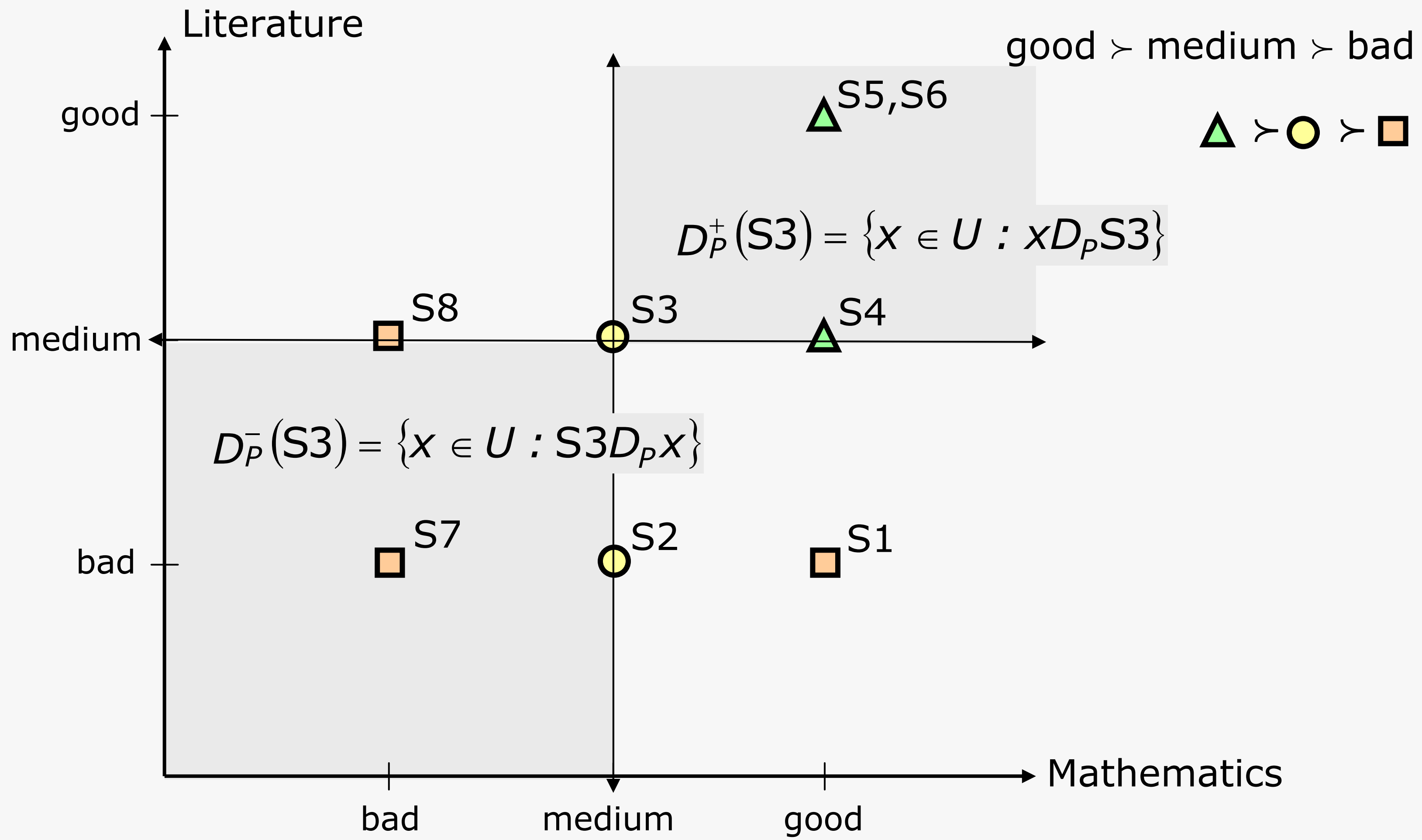
Rough Set approach to multiple-criteria sorting

- Let us represent the students in condition space {M,L} :



Rough Set approach to multiple-criteria sorting

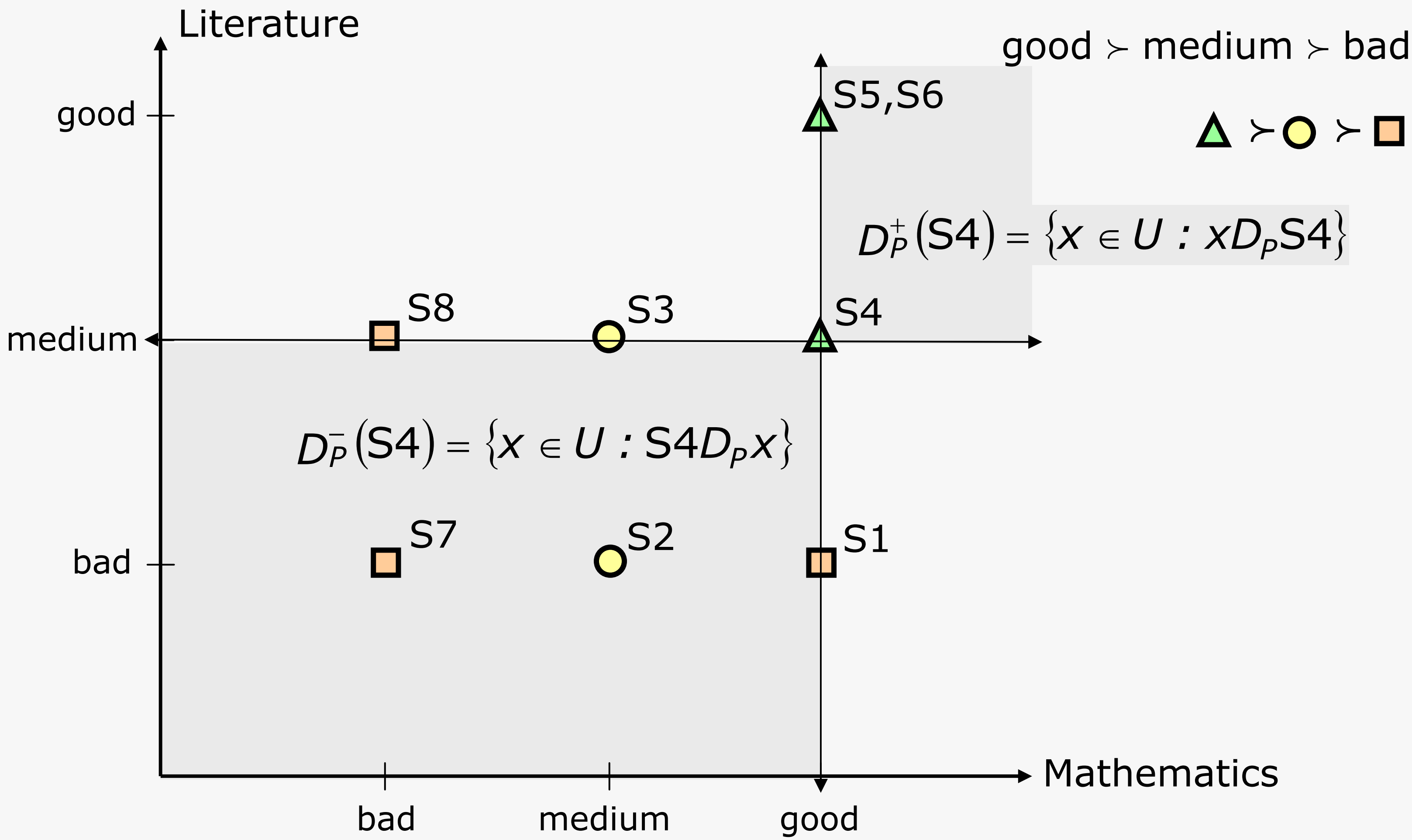
- Dominance cones in condition space {M,L} :



- P={M,L}

Rough Set approach to multiple-criteria sorting

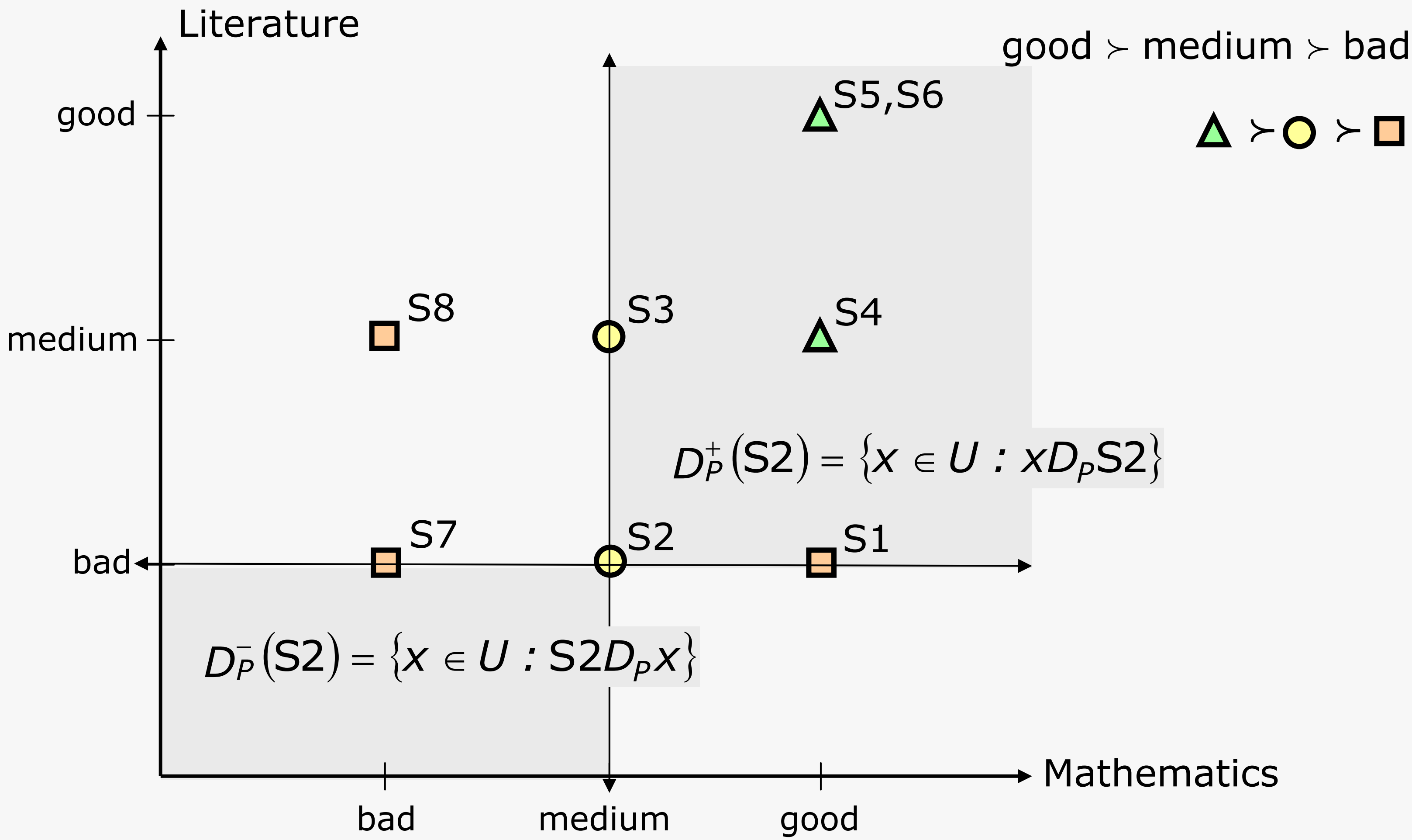
- Dominance cones in condition space {M,L} :



- P={M,L}

Rough Set approach to multiple-criteria sorting

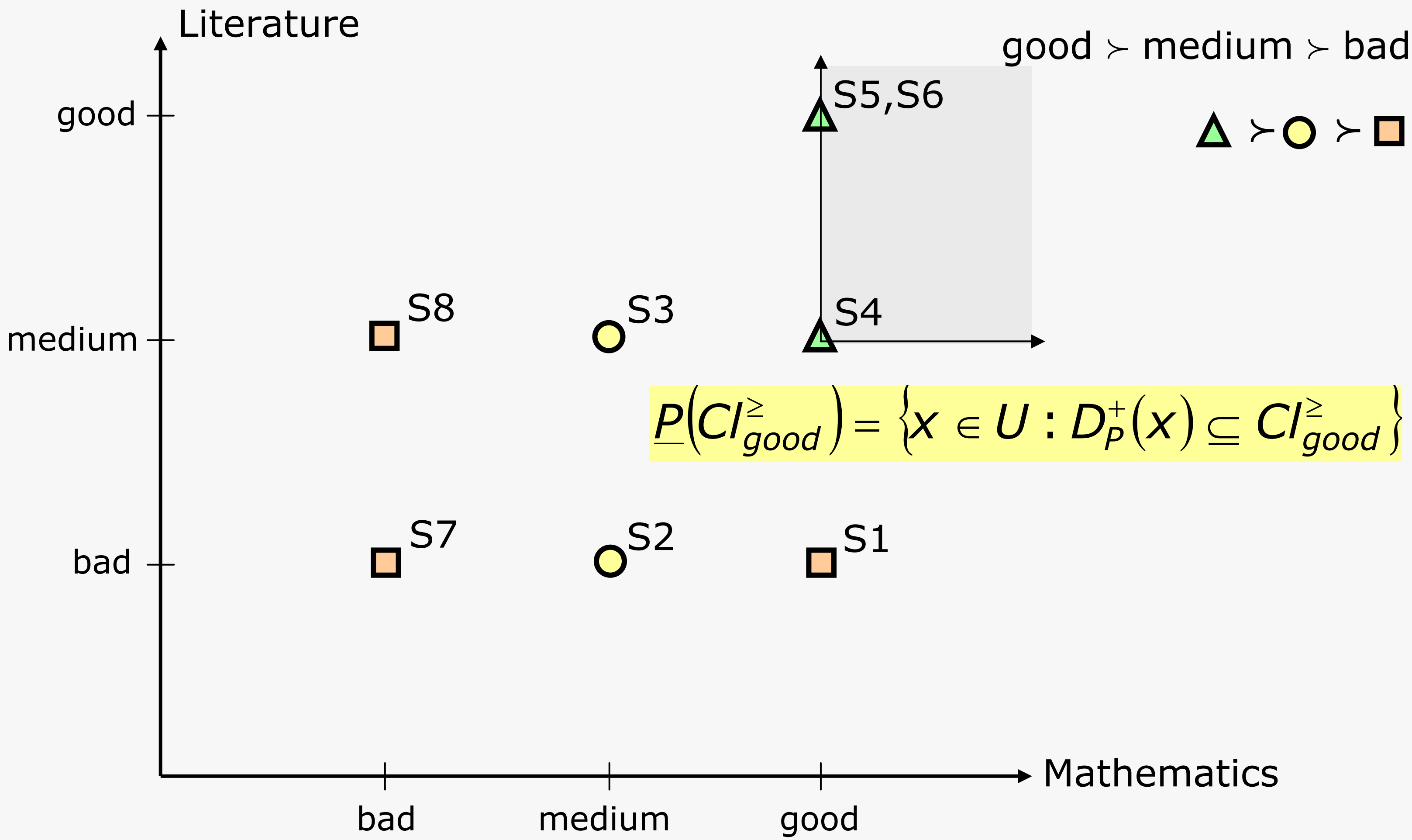
- Dominance cones in condition space {M,L} :



- P={M,L}

Rough Set approach to multiple-criteria sorting

- Lower approximation of at least good students:

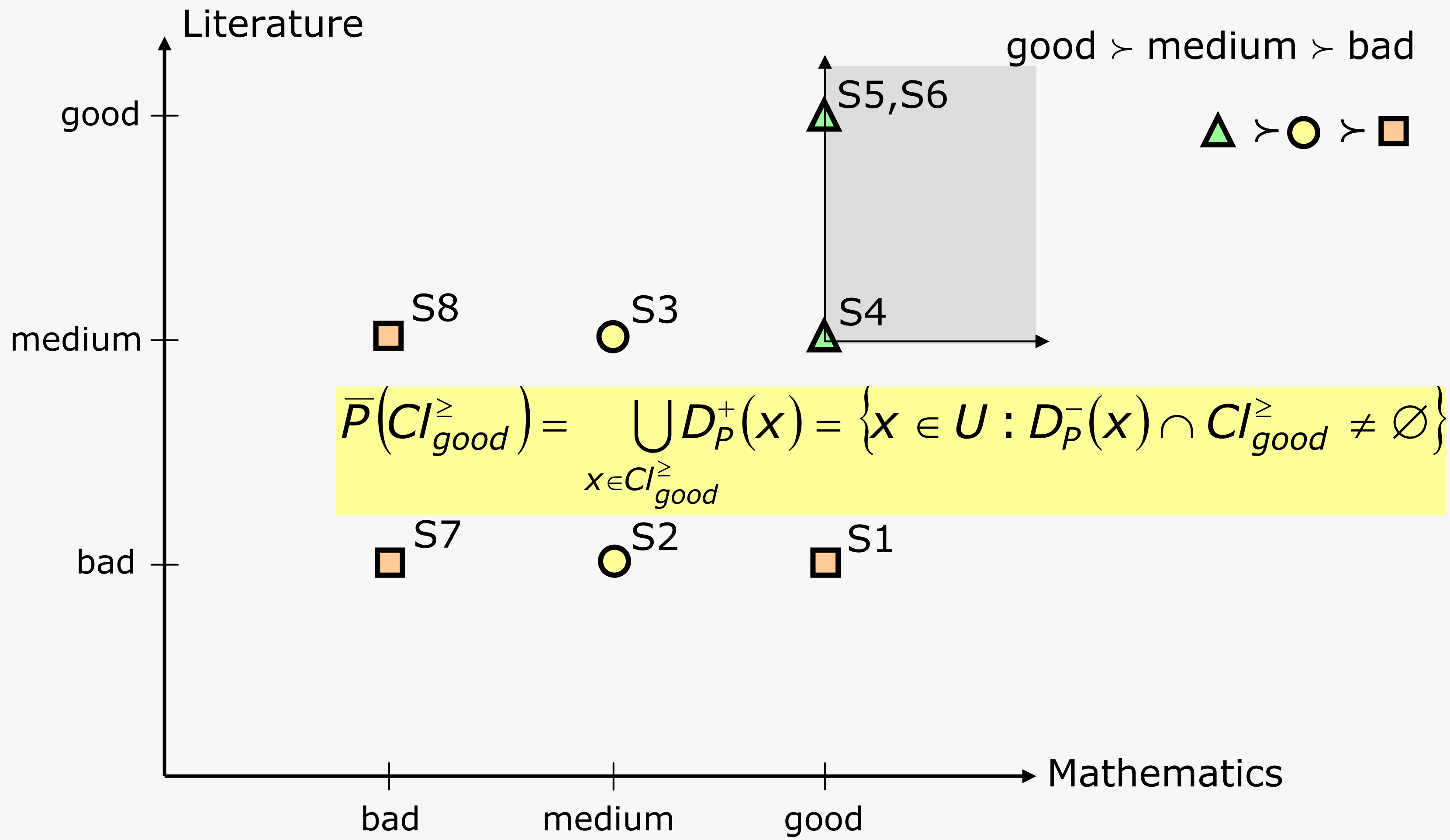


$$\underline{P}\left(Cl_{good}^{\geq}\right)=\left\{x\in U: D_{P}^{+}(x)\subseteq Cl_{good}^{\geq}\right\}$$

- P={M,L}

Rough Set approach to multiple-criteria sorting

- Upper approximation of <u>at least</u> good students:

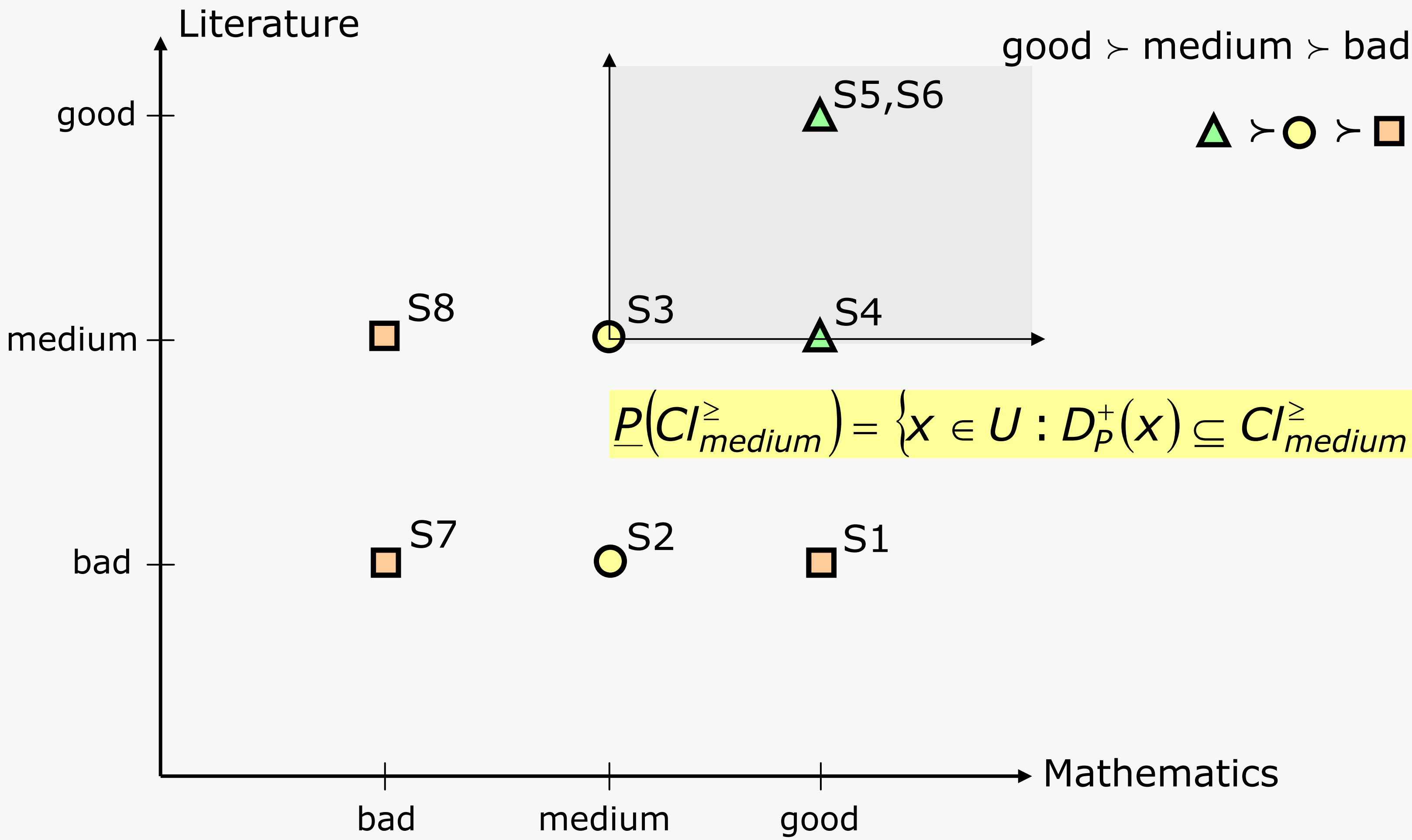


$$\overline{P}\left(Cl_{good}^{\geq}\right)=\bigcup_{x\in Cl_{good}^{\geq}}D_P^{+}(x)=\left\{x\in U: D_P^{-}(x)\cap Cl_{good}^{\geq}\neq\emptyset\right\}$$

- P={M,L}

Rough Set approach to multiple-criteria sorting

- Lower approximation of at least medium students:

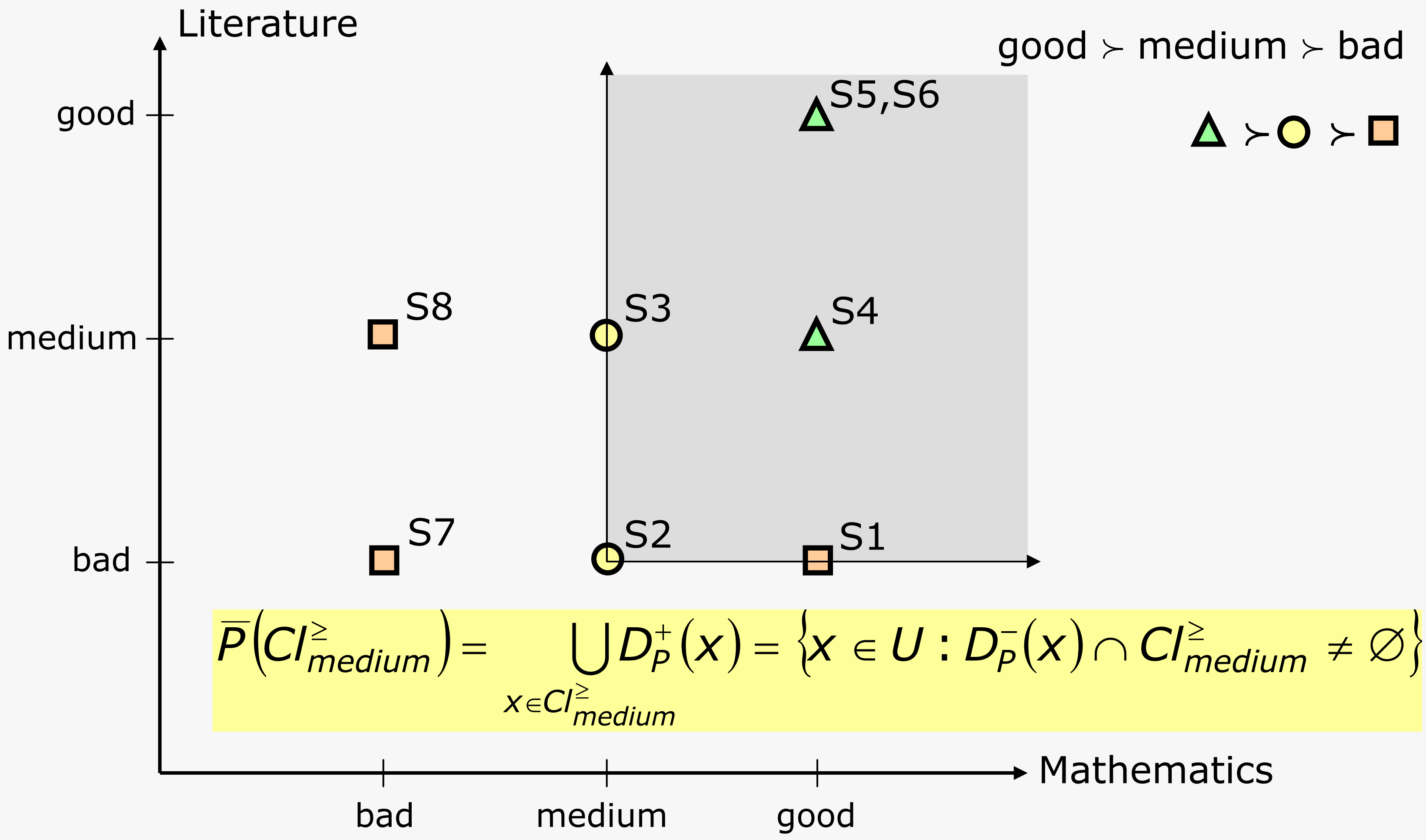


$$\underline{P}\left(Cl^{\geq}_{medium}\right) = \left\{x \in U : D_P^{+}(x) \subseteq Cl^{\geq}_{medium}\right\}$$

- P={M,L}

Rough Set approach to multiple-criteria sorting

- Upper approximation of <u>at least</u> medium students:

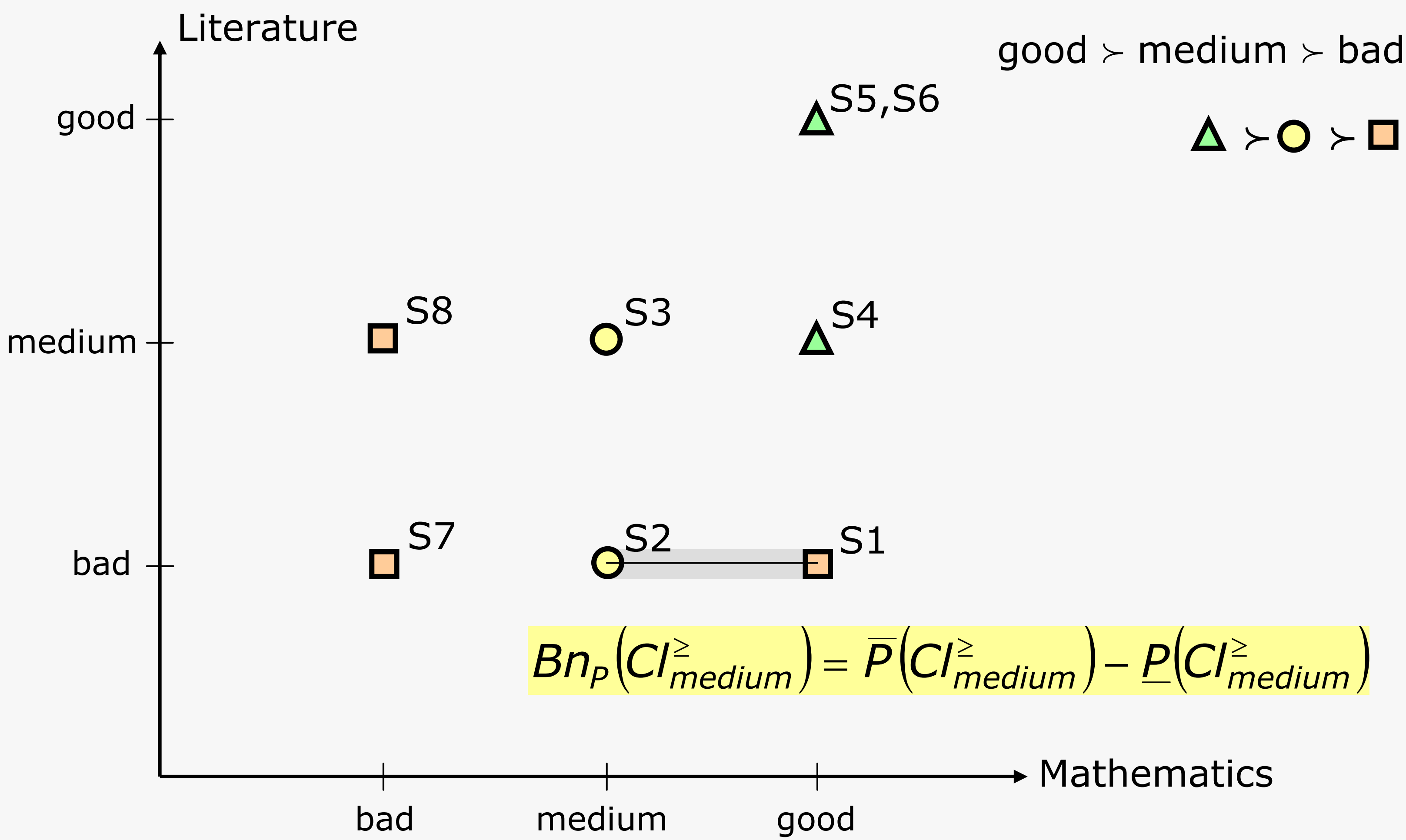


$$\overline{P}\left(Cl^{\geq}_{medium}\right) = \bigcup_{x \in Cl^{\geq}_{medium}} D_P^{+}(x) = \left\{x \in U : D_P^{-}(x) \cap Cl^{\geq}_{medium} \neq \emptyset\right\}$$

- P={M,L}

Rough Set approach to multiple-criteria sorting

- Boundary region of <u>at least</u> medium students:

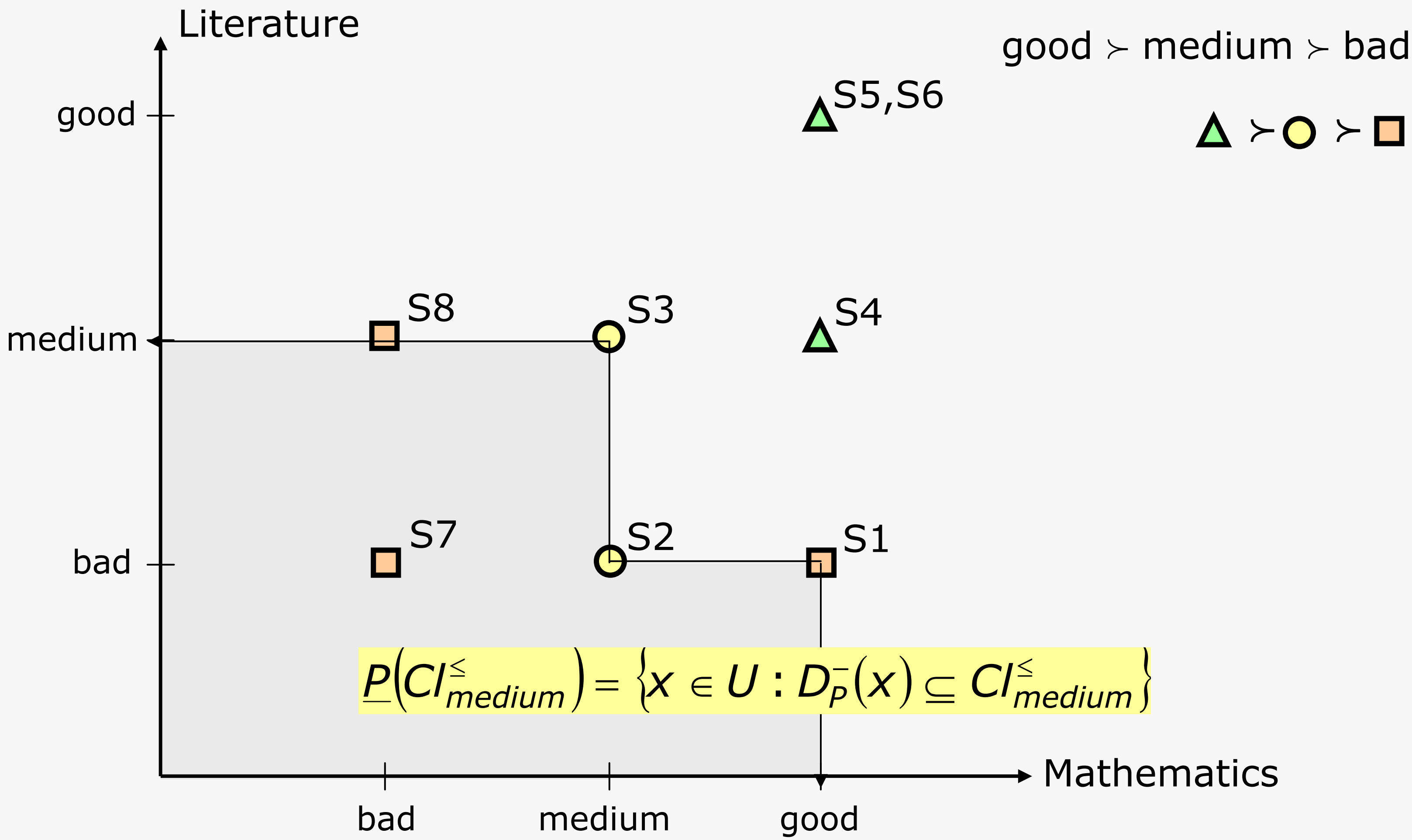


$$Bn_P\left(Cl^{\geq}_{medium}\right) = \overline{P}\left(Cl^{\geq}_{medium}\right) - \underline{P}\left(Cl^{\geq}_{medium}\right)$$

- P={M,L}

Rough Set approach to multiple-criteria sorting

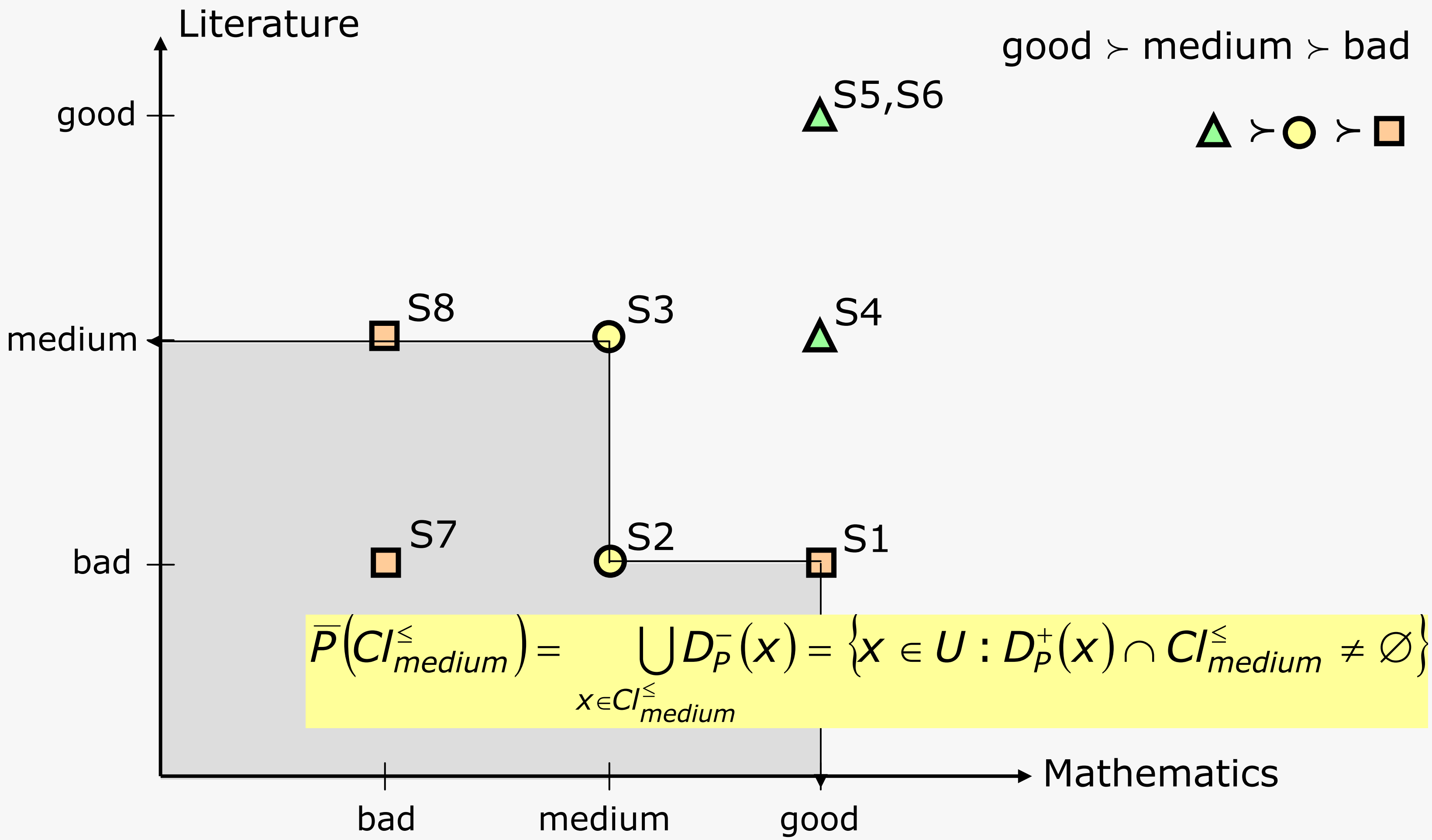
- Lower approximation of <u>at most</u> medium students:



- P={M,L}

Rough Set approach to multiple-criteria sorting

- Upper approximation of at most medium students:

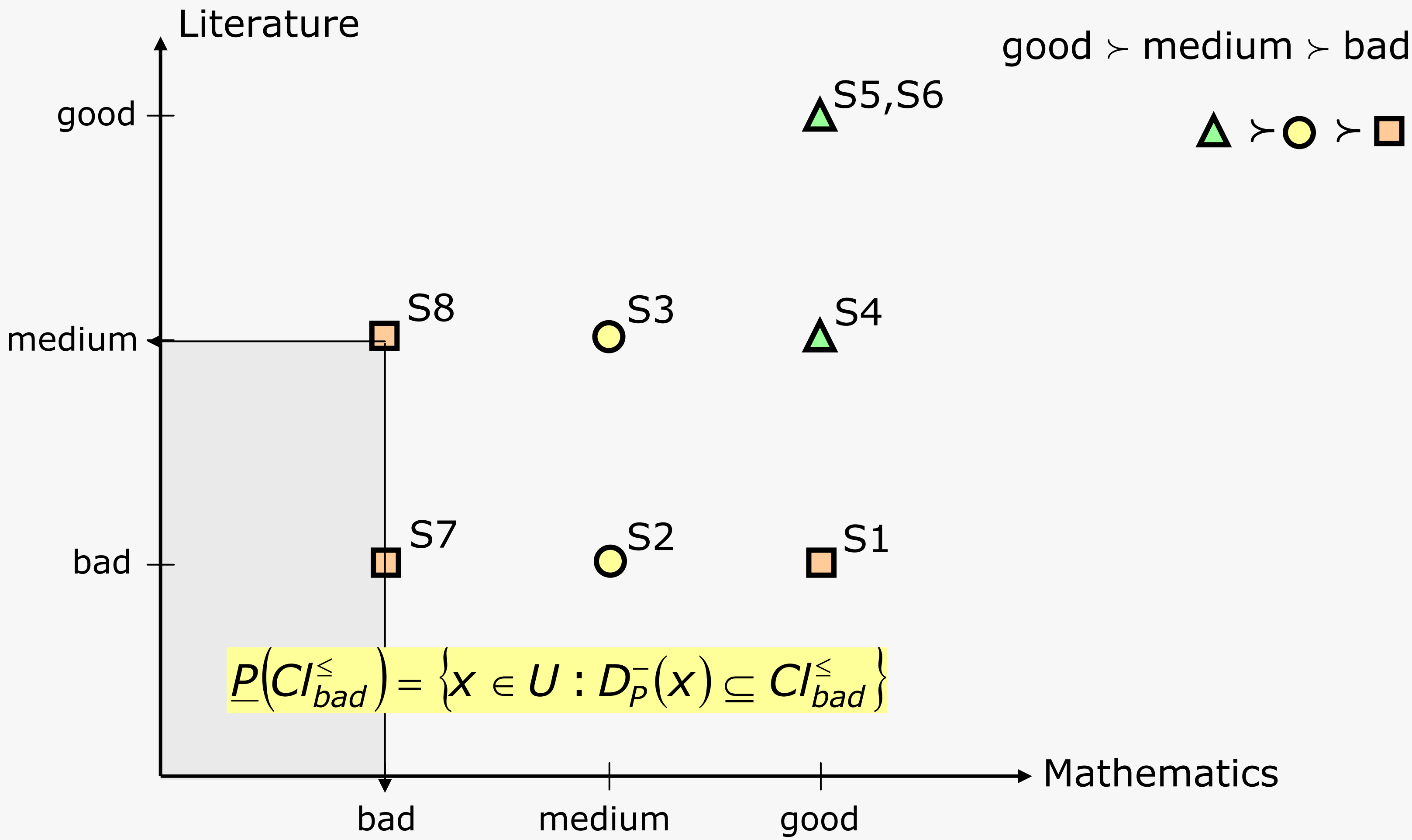


$$\overline{P}\left(Cl^{\le}_{medium}\right) = \bigcup_{x\in Cl^{\le}_{medium}} D_P^-(x) = \left\{x \in U : D_P^+(x) \cap Cl^{\le}_{medium} \neq \emptyset\right\}$$

- P={M,L}

Rough Set approach to multiple-criteria sorting

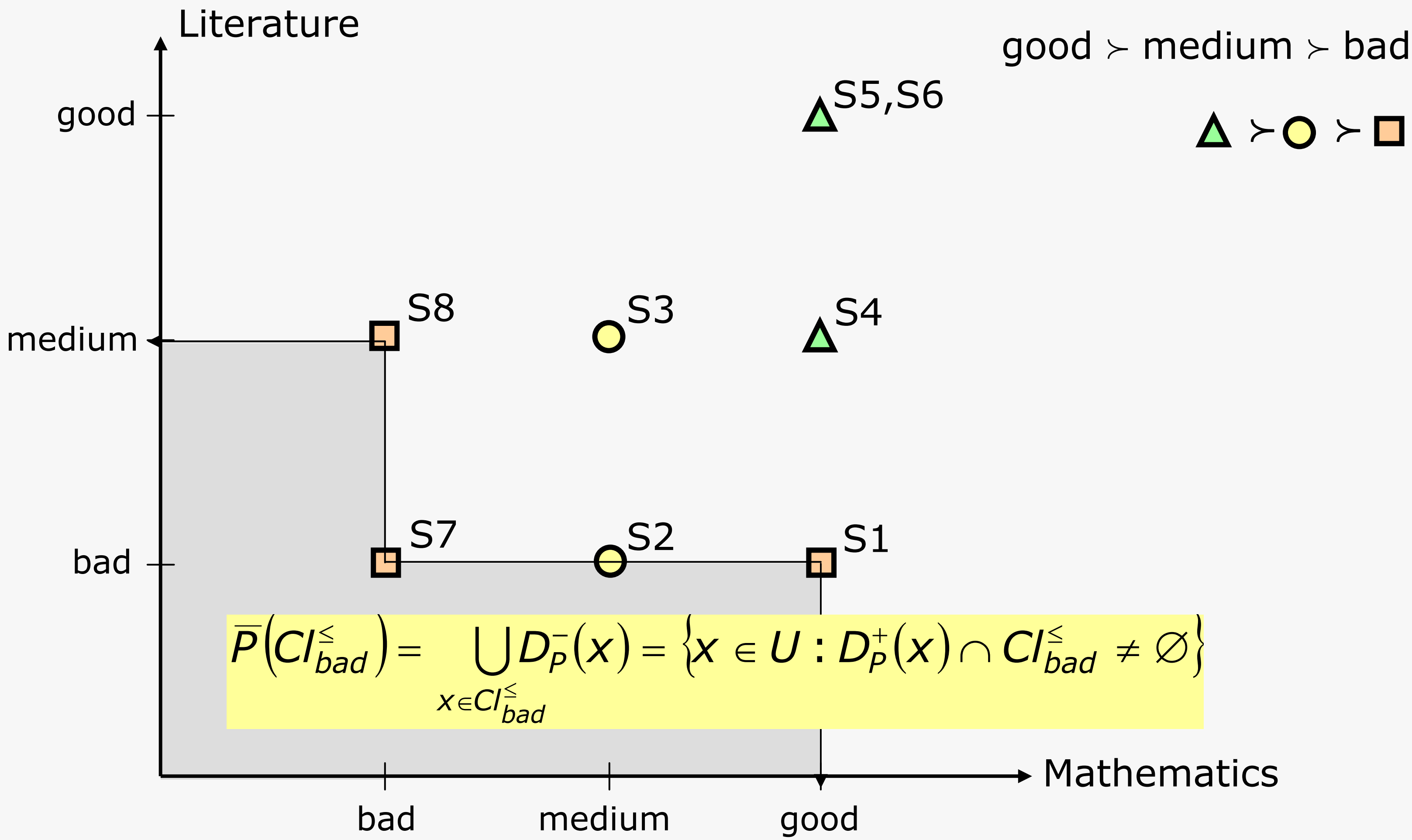
- Lower approximation of at most bad students:



- P={M,L}

Rough Set approach to multiple-criteria sorting

- Upper approximation of at most bad students:

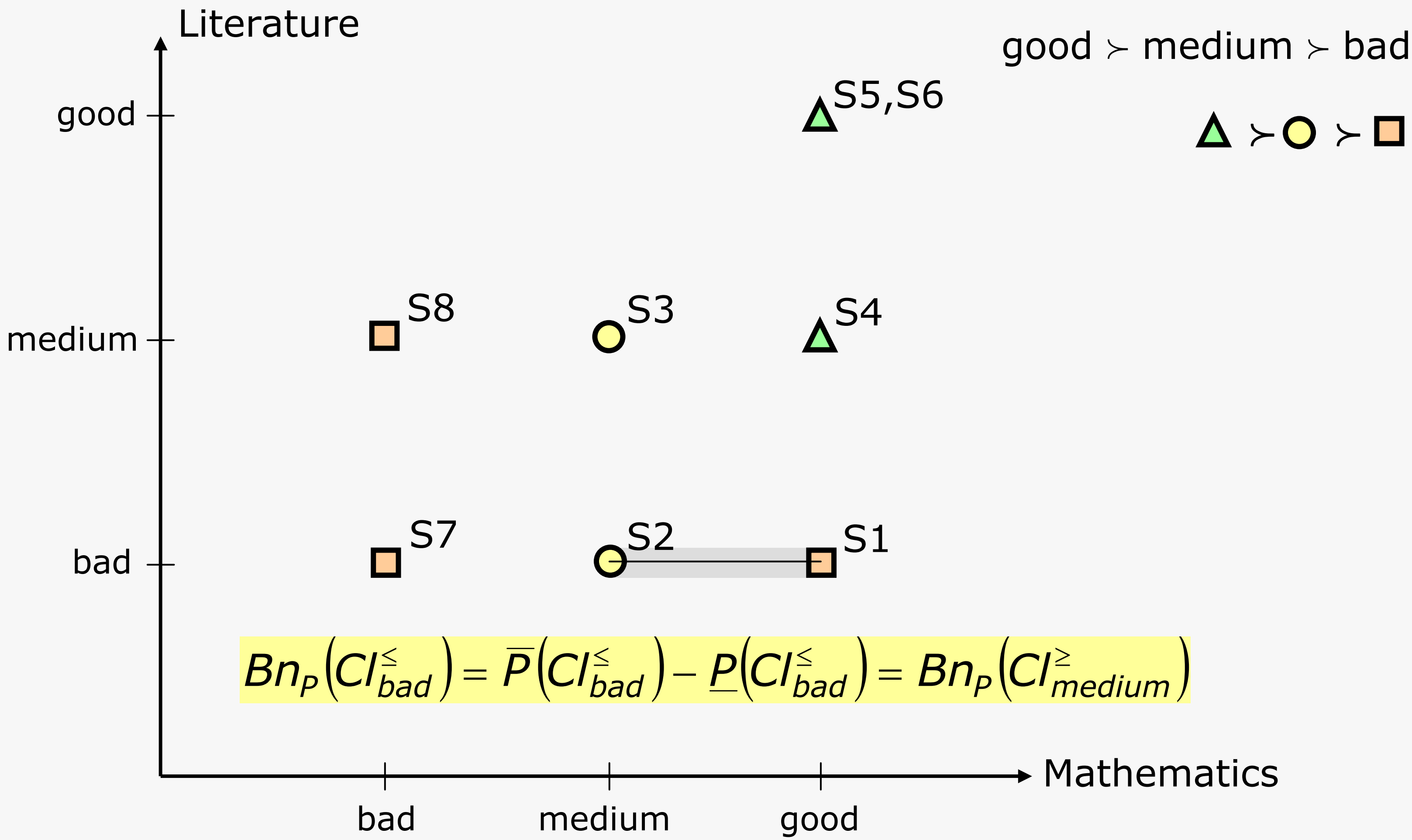


$$\overline{P}\left(Cl_{bad}^{\leq}\right) = \bigcup_{x \in Cl_{bad}^{\leq}} D_P^{-}(x) = \left\{x \in U : D_P^{+}(x) \cap Cl_{bad}^{\leq} \neq \varnothing\right\}$$

- P={M,L}

Rough Set approach to multiple-criteria sorting

- Boundary region of <u>at most</u> bad students:



$$Bn_P\left(Cl_{bad}^{\leq}\right) = \overline{P}\left(Cl_{bad}^{\leq}\right) - \underline{P}\left(Cl_{bad}^{\leq}\right) = Bn_P\left(Cl_{medium}^{\geq}\right)$$

- P={M,L}

DRSA – formal definitions

- Basic properies of rough approximations

$$\underline{P}\left(Cl_t^{\geq}\right) \subseteq Cl_t^{\geq} \subseteq \overline{P}\left(Cl_t^{\geq}\right) \qquad \underline{P}\left(Cl_t^{\leq}\right) \subseteq Cl_t^{\leq} \subseteq \overline{P}\left(Cl_t^{\leq}\right)$$

$$\underline{P}\left(Cl_t^{\geq}\right) = U - \overline{P}\left(Cl_{t-1}^{\leq}\right), \text{ for } t=2,\dots,m$$

- Identity of boundaries $Bn_P\left(Cl_t^{\geq}\right) = Bn_P\left(Cl_{t-1}^{\leq}\right)$, for $t=2,\dots,m$

- Quality of approximation of sorting $\boldsymbol{Cl}=\{Cl_t,\ t=1,\dots m\}$ by criteria $P\subseteq C$

$$\gamma_P(\boldsymbol{Cl}) = \frac{card\left(U - \bigcup_{t\in\{2,\dots,m\}} Bn_P\left(Cl_t^{\geq}\right)\right)}{card(U)}$$

- ***Cl***-reducts and ***Cl***-core of $P\subseteq C$

$$CORE_{\boldsymbol{Cl}}(P) = \bigcap RED_{\boldsymbol{Cl}}(P)$$

DRSA – induction of decision rules from rough approximations

- Induction of decision rules from rough approximations
  - *certain* D$_{\geq}$-*decision rules*, supported by objects $\in Cl_t^{\geq}$ without ambiguity:

    if $x_{q1} \succeq_{q1} r_{q1}$ *and* $x_{q2} \succeq_{q2} r_{q2}$ *and* … $x_{qp} \succeq_{qp} r_{qp}$, *then* $x \in Cl_t^{\geq}$

  - *possible* D$_{\geq}$-*decision rules*, supported by objects $\in Cl_t^{\geq}$ with or without any ambiguity:

    if $x_{q1} \succeq_{q1} r_{q1}$ *and* $x_{q2} \succeq_{q2} r_{q2}$ *and* … $x_{qp} \succeq_{qp} r_{qp}$, *then* x possibly $\in Cl_t^{\geq}$

DRSA – induction of decision rules from rough approximations

- Induction of decision rules from rough approximations
  - *certain* $D_{\leq}$*-decision rules*, supported by objects $\in Cl_t^{\leq}$ without ambiguity:

    if $x_{q1} \preceq_{q1} r_{q1}$ *and* $x_{q2} \preceq_{q2} r_{q2}$ *and* … $x_{qp} \preceq_{qp} r_{qp}$, *then* $x \in Cl_t^{\leq}$

  - *possible* $D_{\leq}$*-decision rules*, supported by objects $\in Cl_t^{\leq}$ with or without any ambiguity:

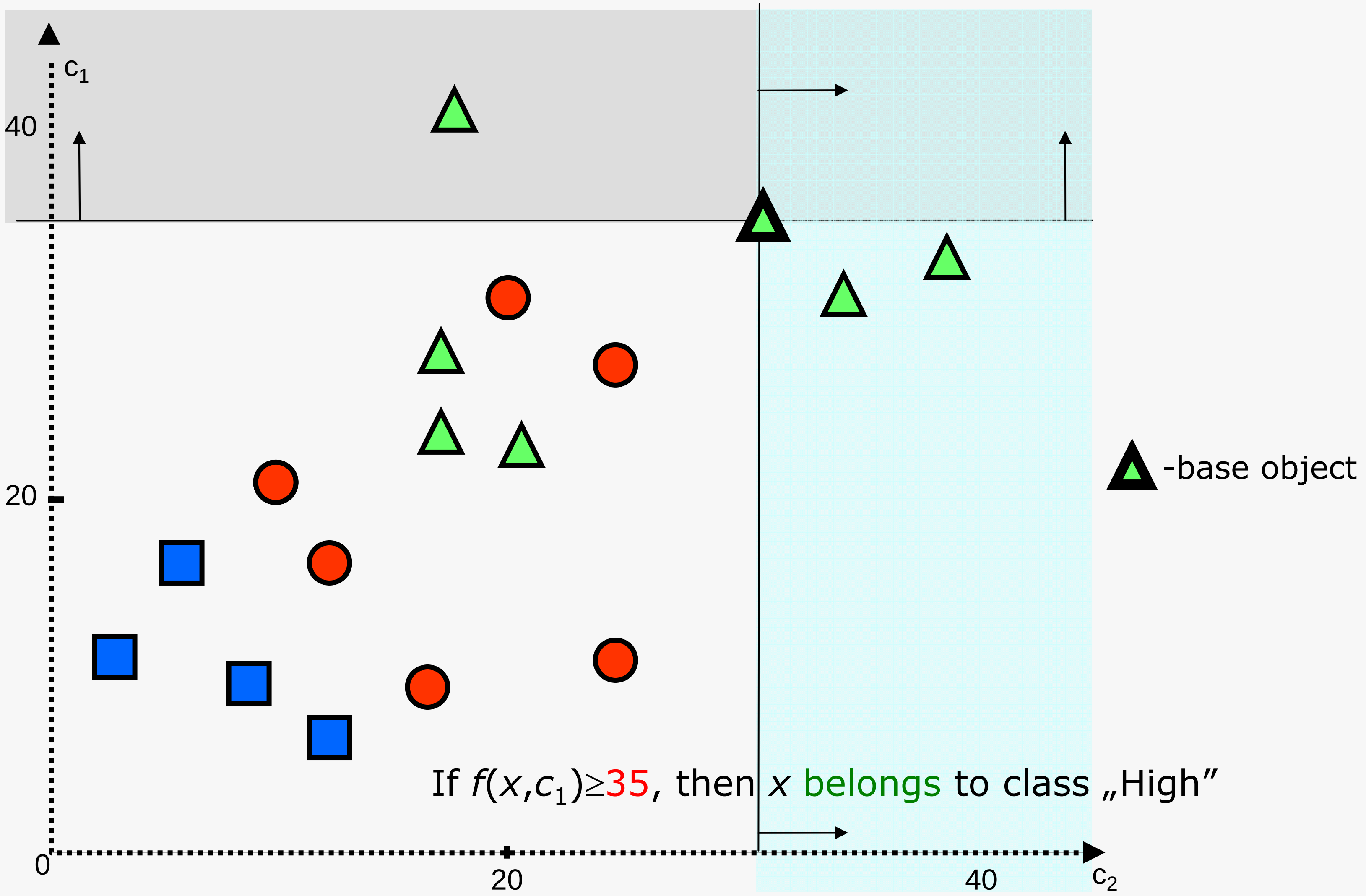
    if $x_{q1} \preceq_{q1} r_{q1}$ *and* $x_{q2} \preceq_{q2} r_{q2}$ *and* … $x_{qp} \preceq_{qp} r_{qp}$, *then* x possibly $\in Cl_t^{\leq}$

  - *approximate* $D_{\geq\leq}$*-decision rules*, supported by objects $\in Cl_s \cup Cl_{s+1} \cup \ldots \cup Cl_t$ without possibility of discerning to which class:

    if $x_{q1} \succeq_{q1} r_{q1}$ *and*… $x_{qk} \succeq_{qk} r_{qk}$ *and* $x_{qk+1} \preceq_{qk+1} r_{qk+1}$ *and* … $x_{qp} \preceq_{qp} r_{qp}$, *then* $x \in Cl_s \cup Cl_{s+1} \cup \ldots \cup Cl_t$.

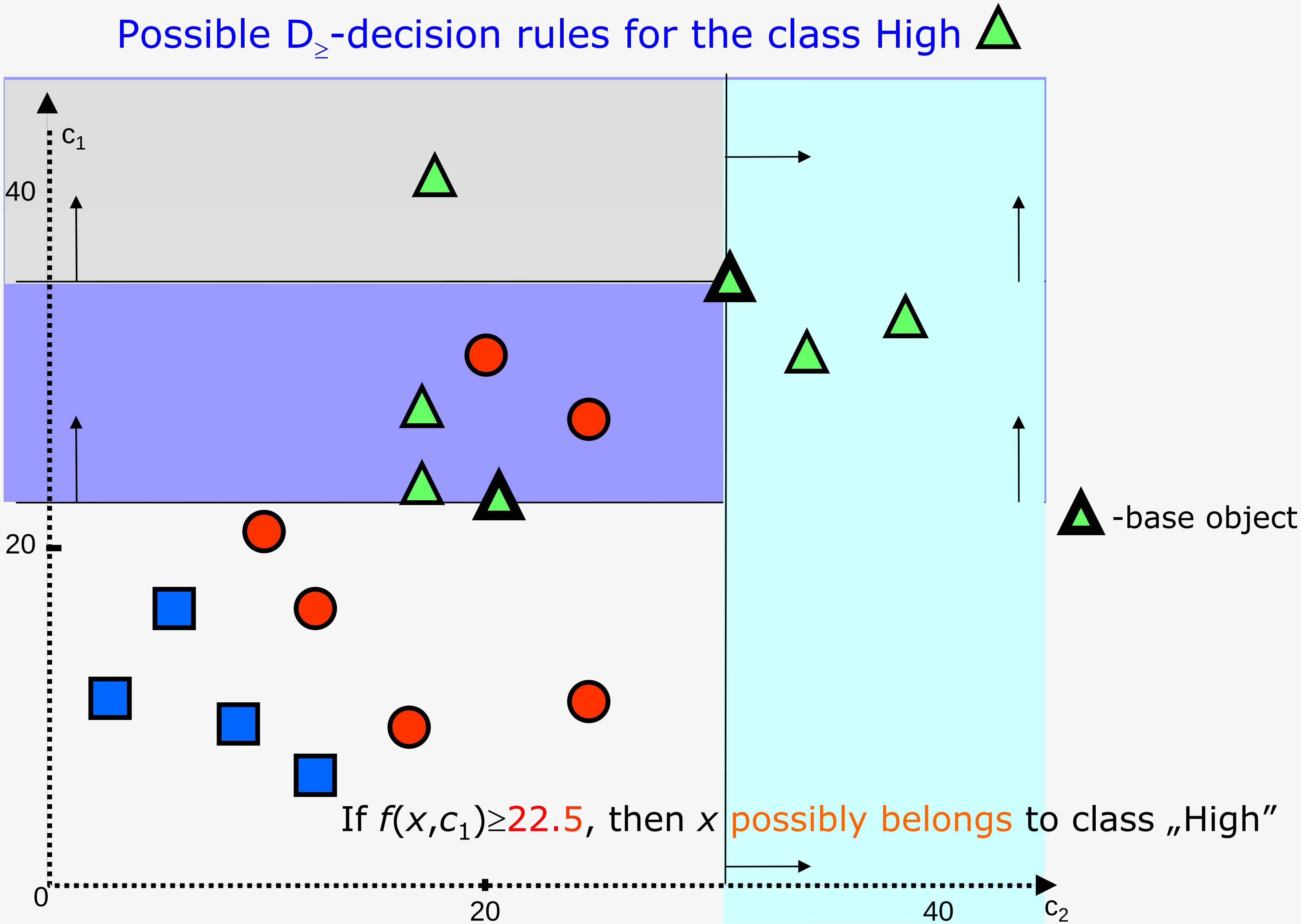
DRSA – decision rules

Certain $D_{\geq}$-decision rules for the class High



DRSA – decision rules

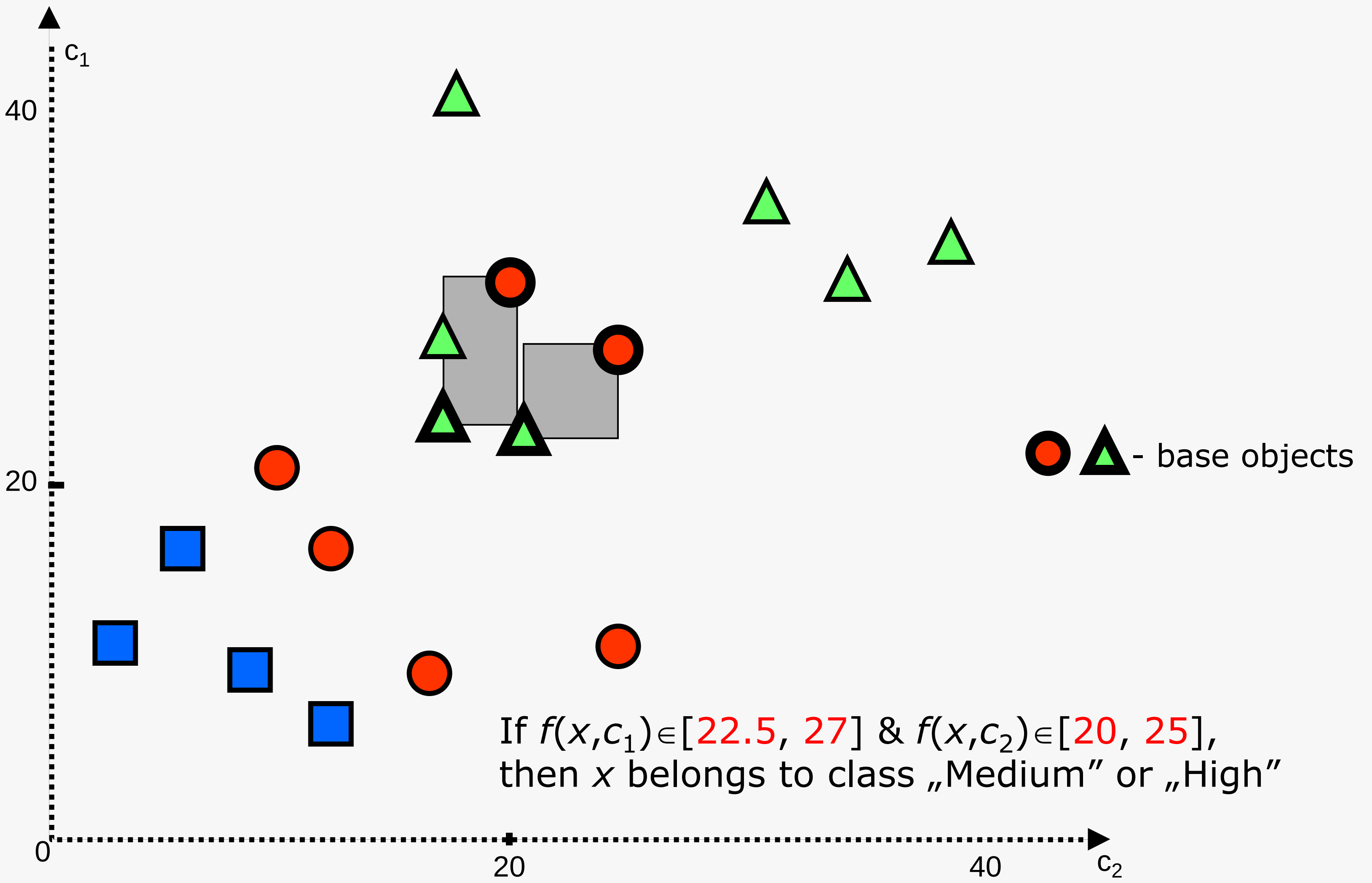
Possible $D_{\geq}$-decision rules for the class High



If $f(x,c_1)\geq 22.5$, then x possibly belongs to class „High"

DRSA – decision rules

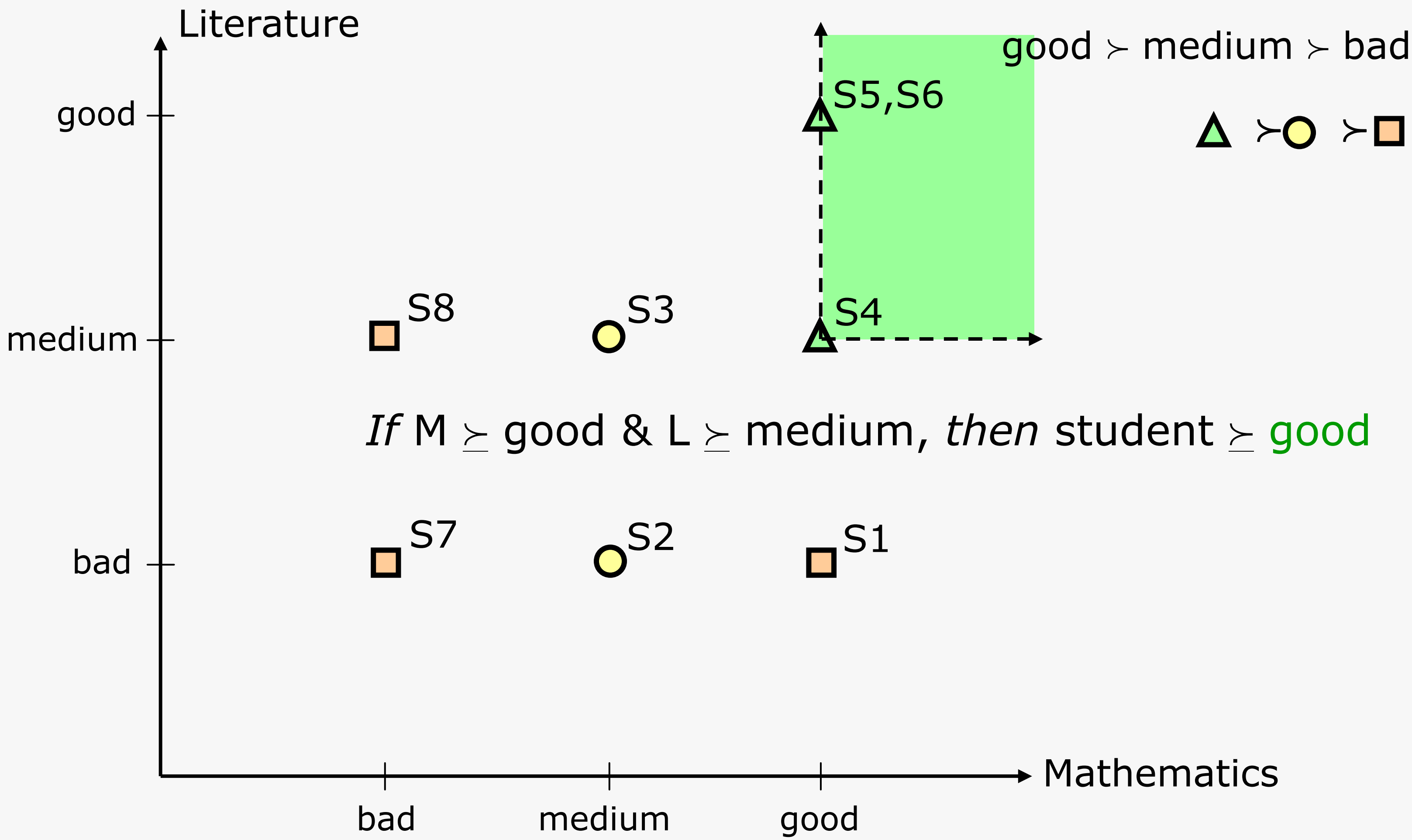
Approximate $D_{\geq\leq}$-decision rules for the class Medium or High



If $f(x,c_1)\in[22.5, 27]$ & $f(x,c_2)\in[20, 25]$,
then x belongs to class „Medium" or „High"

Rough Set approach to multiple-criteria sorting

- Decision rules in terms of {M,L} :

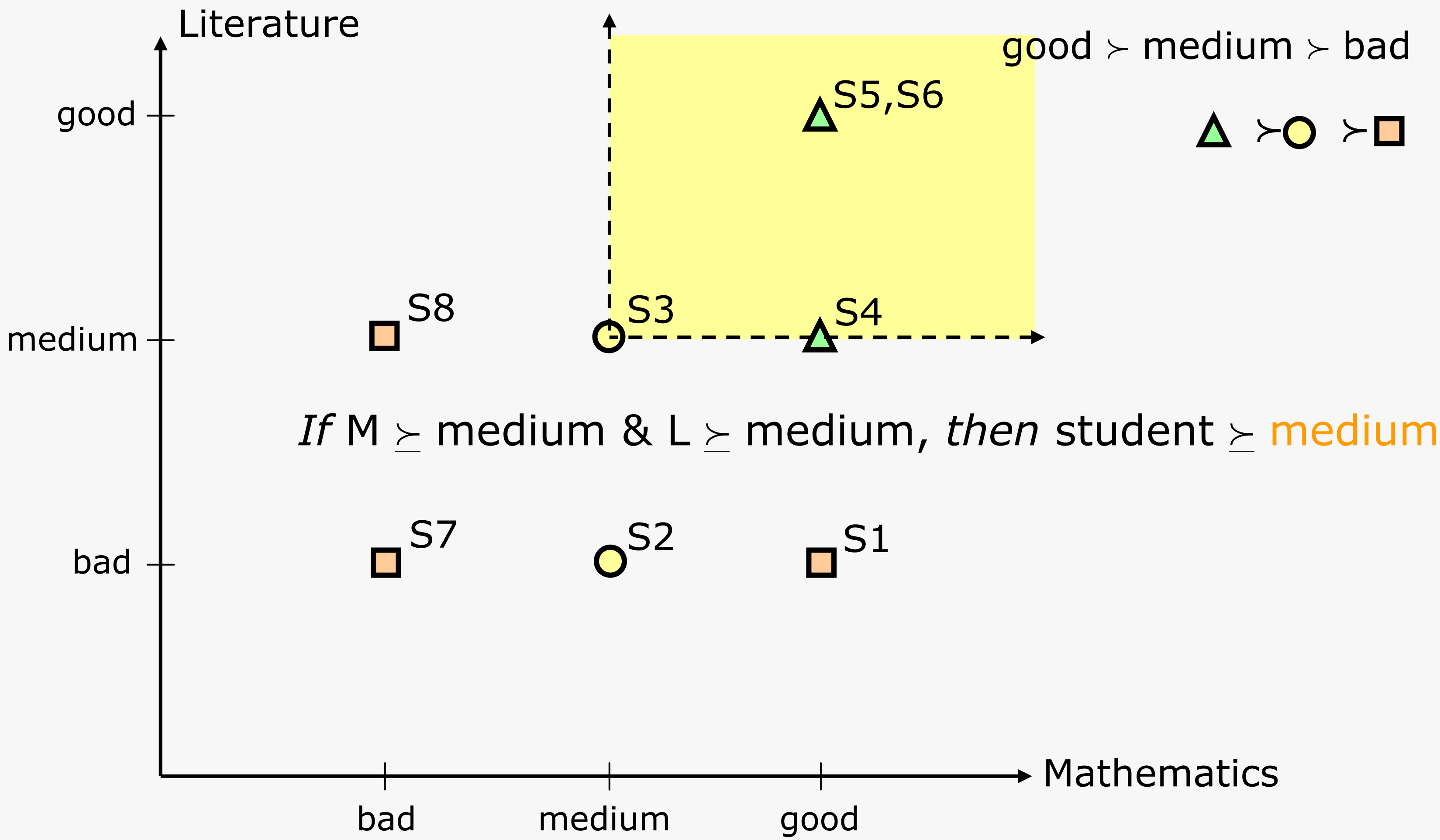


If M $\succeq$ good & L $\succeq$ medium, *then* student $\succeq$ good

- $D_{\geq}$ - certain rule

Rough Set approach to multiple-criteria sorting

- Decision rules in terms of {M,L} :

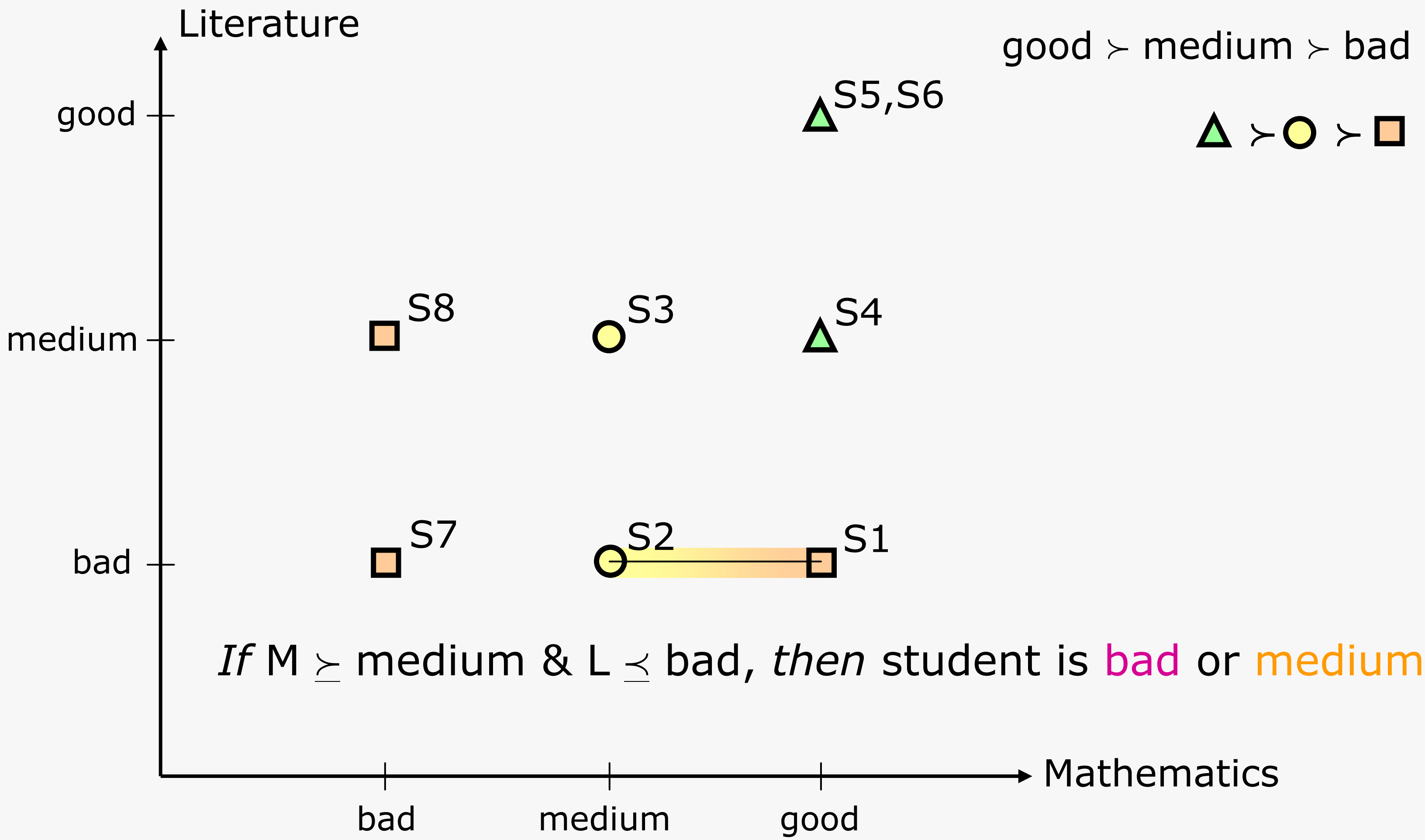


If M $\succeq$ medium & L $\succeq$ medium, *then* student $\succeq$ medium

- $D_{\geq}$ - certain rule

Rough Set approach to multiple-criteria sorting

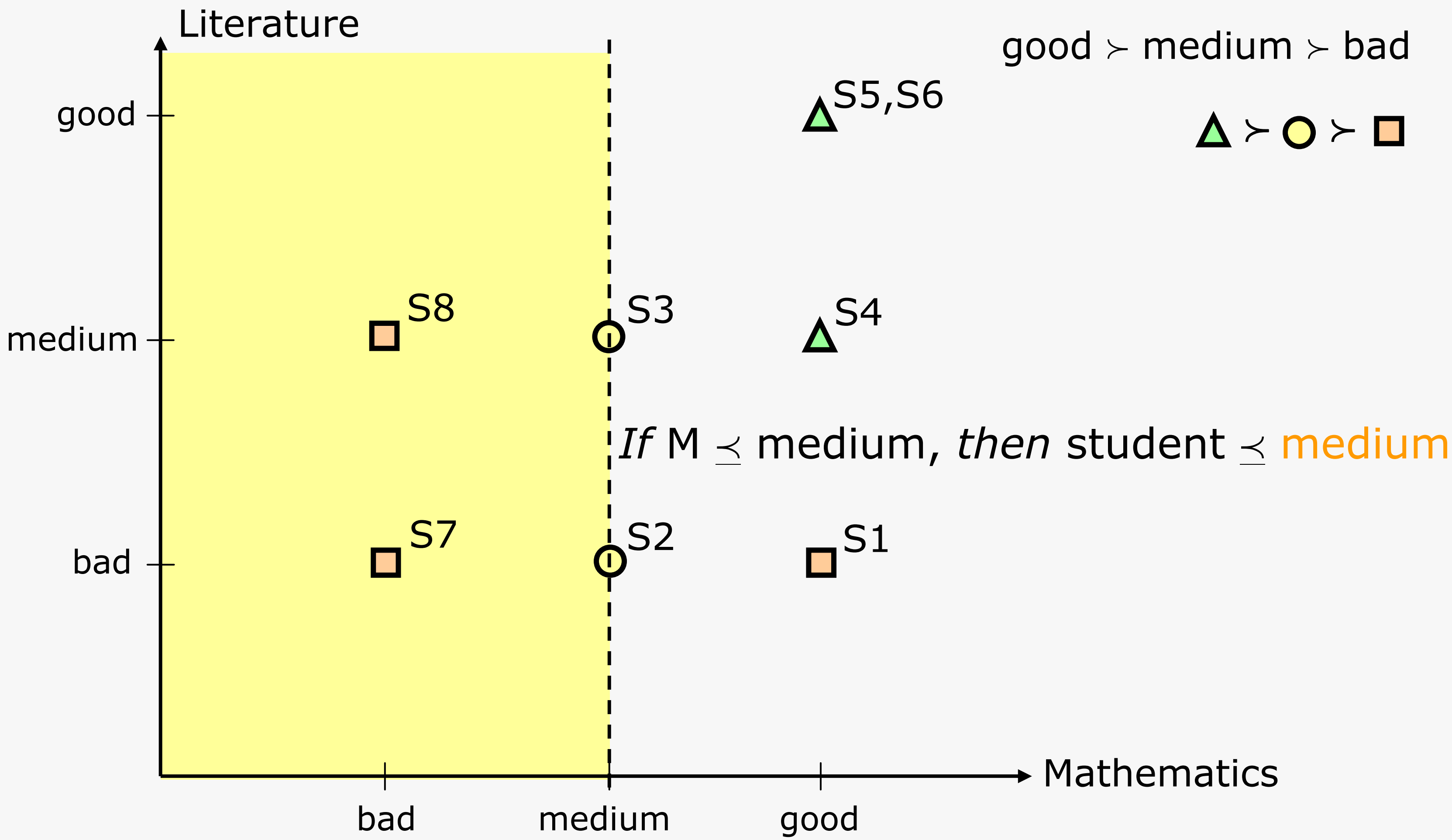
- Decision rules in terms of {M,L} :



- $D_{\geq\leq}$ - approximate rule

Rough Set approach to multiple-criteria sorting

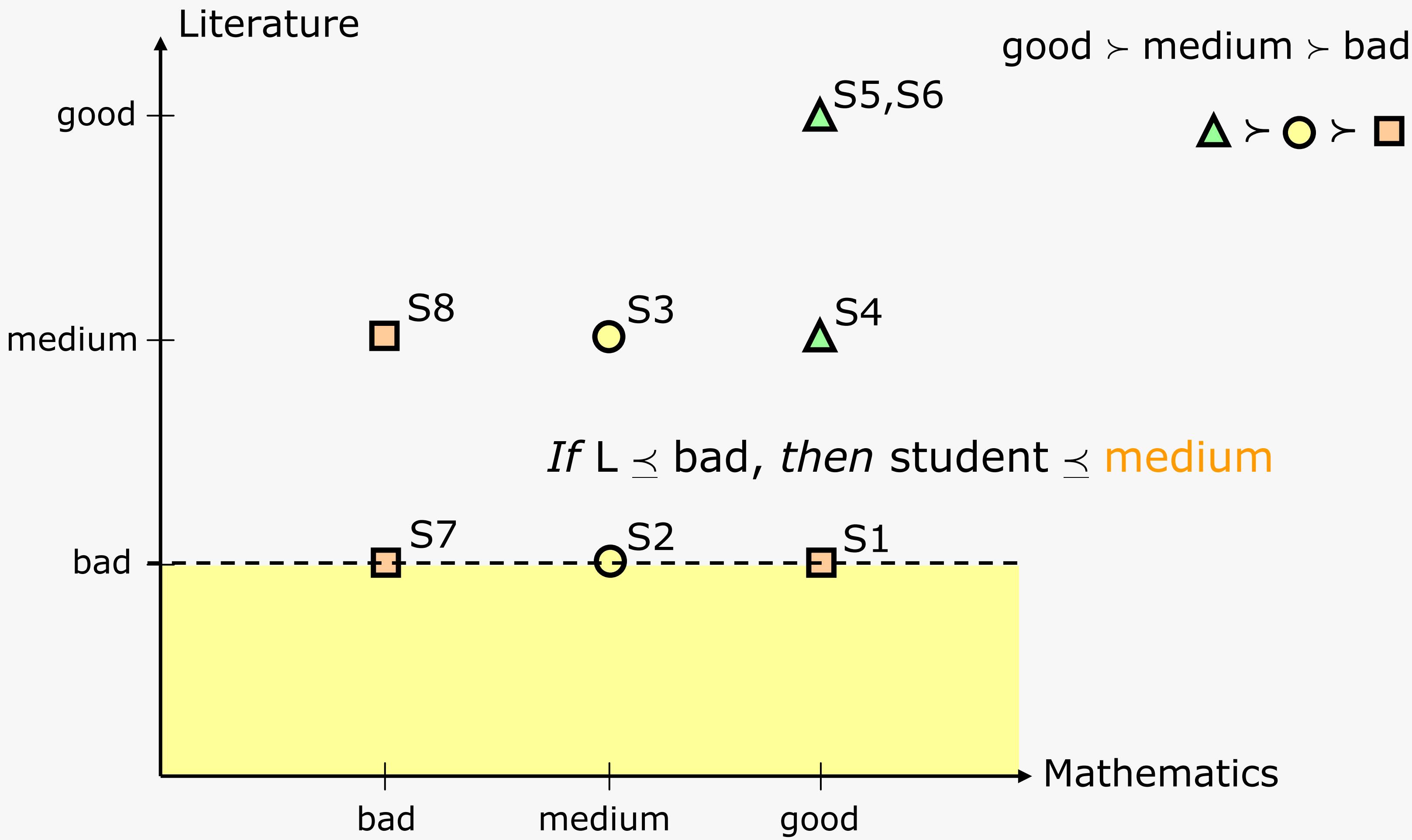
- Decision rules in terms of {M,L} :



- $D_{\leq}$ - certain rule

Rough Set approach to multiple-criteria sorting

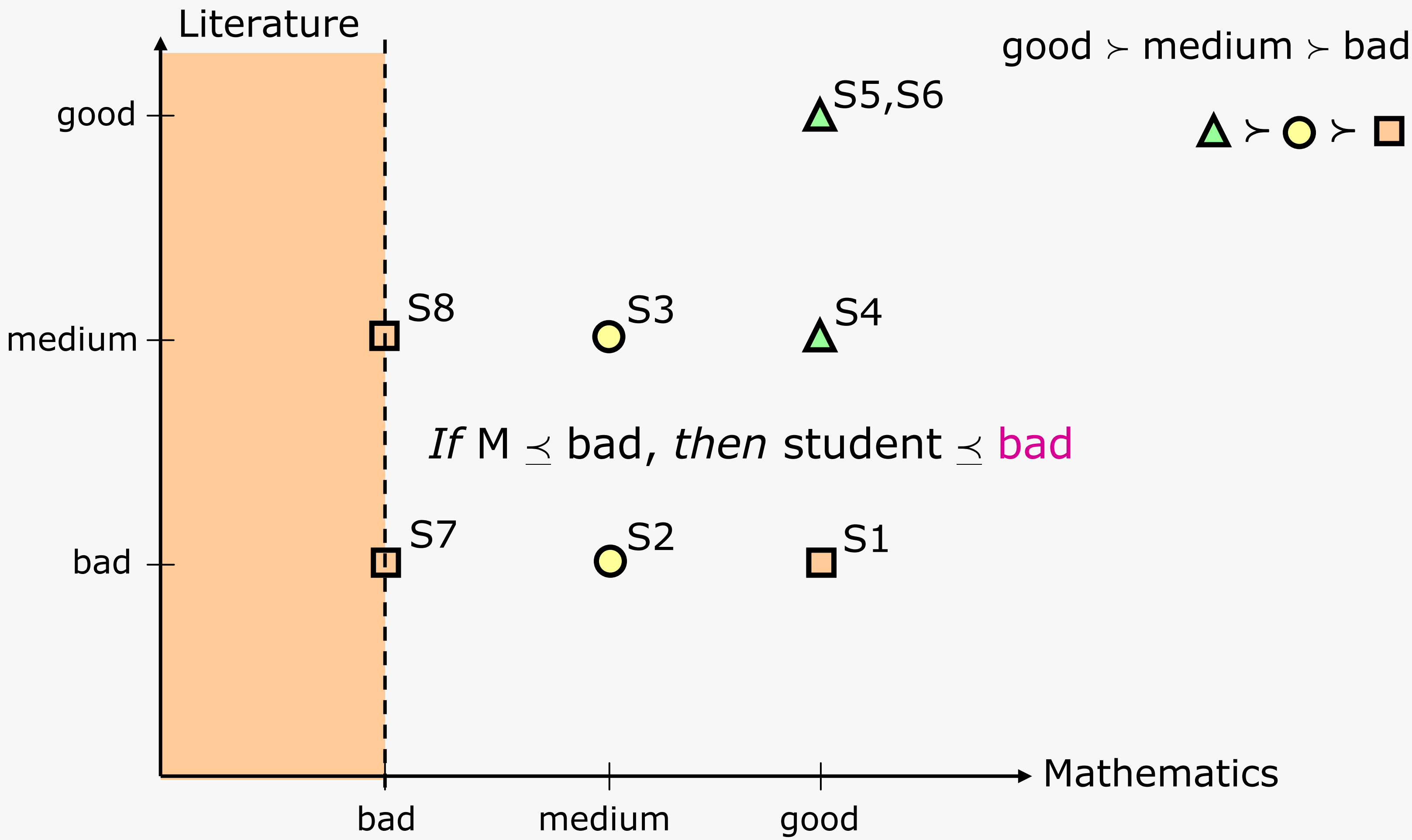
- Decision rules in terms of {M,L} :



- $D_{\leq}$ - certain rule

Rough Set approach to multiple-criteria sorting

- Decision rules in terms of {M,L} :



- $D_{\leq}$ - certain rule

Rough Set approach to multiple-criteria sorting

- Set of decision rules in terms of {M, L} representing preferences:

If M $\succeq$ good & L $\succeq$ medium, *then* student $\succeq$ good {S4,S5,S6}

If M $\succeq$ medium & L $\succeq$ medium, *then* student $\succeq$ medium {S3,S4,S5,S6}

If M $\succeq$ medium & L $\preceq$ bad, *then* student is bad or medium {S1,S2}

If M $\preceq$ medium, *then* student $\preceq$ medium {S2,S3,S7,S8}

If L $\preceq$ bad, *then* student $\preceq$ medium {S1,S2,S7}

If M $\preceq$ bad, *then* student $\preceq$ bad {S7,S8}

Rough Set approach to multiple-criteria sorting

- Set of decision rules in terms of {M,Ph,L} representing preferences:

If M $\succeq$ good & L $\succeq$ medium, *then* student $\succeq$ good {S4,S5,S6}

If M $\succeq$ medium & L $\succeq$ medium, *then* student $\succeq$ medium {S3,S4,S5,S6}

If M $\succeq$ medium & L $\preceq$ bad, *then* student is bad or medium {S1,S2}

If Ph $\preceq$ medium & L $\preceq$ medium *then* student $\preceq$ medium {S1,S2,S3,S7,S8}

If M $\preceq$ bad, *then* student $\preceq$ bad {S7,S8}

- The preference model involving all three criteria is more concise

Rough Set approach to multiple-criteria sorting

- Importance and interaction among criteria
- Quality of approximation of sorting $\gamma_P(\boldsymbol{Cl})$ ($P \subseteq C$) is a fuzzy measure with the property of Choquet capacity
  ($\gamma_\emptyset(\boldsymbol{Cl})=0$, $\gamma_C(\boldsymbol{Cl})=r$ and $\gamma_R(\boldsymbol{Cl}) \leq \gamma_P(\boldsymbol{Cl}) \leq r$ for any $R \subseteq P \subseteq C$)
- Such measure can be used to calculate Shapley value or Benzhaf index, i.e. an average „contribution" of criterion q in all coalitions of criteria, $q \in \{1,\ldots,m\}$
- Fuzzy measure theory permits, moreover, to calculate interaction indices (Murofushi & Soneda, Grabisch or Roubens) for pairs (or larger subsets) of criteria,
  i.e. an average „added value" resulting from putting together q and q' in all coalitions of criteria, $q,q' \in \{1,\ldots,m\}$

Rough Set approach to multiple-criteria sorting

- Quality of approximation of sorting students

$$\gamma_C(\boldsymbol{Cl}) = [8\text{-card}(\{S1,S2\})]/8 = 0.75$$

Set of criteria P	Ambiguous objects	Non-ambiguous objects	Quality of classification	Shapley value
{Mathematics}	S1,S2,S3,S4,S5,S6	S7,S8	0.25	0.167
{Physics}	S1,S2,S3,S5	S4,S6,S7,S8	0.5	0.292
{Literature}	S1,S2,S3,S4,S7,S8	S5,S6	0.25	0.292
{Mathematics, Physics}	S1,S2,S3,S5	S4,S6,S7,S8	0.5	-0.375
{Mathematics, Literature}	S1,S2	S3,S4,S5,S6,S7,S8	0.75	0.125
{Physics, Literature}	S1,S2	S3,S4,S5,S6,S7,S8	0.75	-0.125
{Mathematics, Physics, Literature}	S1,S2	S3,S4,S5,S6,S7,S8	0.75	-0.125

Preference modeling

- Three families of **preference models**:
  - Function, e.g. utility (value) function

$$U(a)=\sum_{i=1}^{n} k_i g_i(a), \qquad U(a)=\sum_{i=1}^{n} u_i[g_i(a)]$$

  - Relational system, e.g. outranking relation S or fuzzy relation

aSb = "a is at least as good as b"

  - Set of decision rules,

e.g. "If $g_i(a)\geq r_i$ & $g_j(a)\geq r_j$ & ... $g_h(a)\geq r_h$, then $a \rightarrow Class\ t$ or higher"

"If $\Delta_i(a,b)\geq s_i$ & $\Delta_j(a,b)\geq s_j$ & ... $\Delta_h(a,b)\geq s_h$, then aSb"

- The rule model is the most general of all three

Greco, S., Matarazzo, B., Słowiński, R.: Axiomatic characterization of a general utility function and its particular cases in terms of conjoint measurement and rough-set decision rules. *European J. of Operational Research*, 158 (2004) no. 2, 271-292

DRSA – preference modeling by decision rules

- A set of ($D_{\geq}$ $D_{\leq}$ $D_{\geq\leq}$)-rules induced from rough approximations represents a preference model of a Decision Maker

- Traditional preference models:
  - utility function (e.g. additive, multiplicative, associative, Choquet integral, Sugeno integral),
  - binary relation (e.g. outranking relation, fuzzy relation)

- Decision rule model is the most general model of preferences: a general utility function, Sugeno or Choquet inegral, or outranking relation exists <u>if and only if</u> there exists the decision rule model

Słowiński, R., Greco, S., Matarazzo, B.: "Axiomatization of utility, outranking and decision-rule preference models for multiple-criteria classification problems under partial inconsistency with the dominance principle", *Control and Cybernetics*, 31 (2002) no.4, 1005-1035

DRSA – preference modeling by decision rules

- Representation axiom (cancellation property): for every dimension i=1,…,n, for every evaluation $x_i, y_i \in X_i$ and $a_{-i}, b_{-i} \in X_{-i}$, and for every pair of decision classes $Cl_r, Cl_s \in \{Cl_1,…,Cl_m\}$:

$$\{x_i a_{-i} \in Cl_r \text{ and } y_i b_{-i} \in Cl_s\} \Rightarrow \{y_i a_{-i} \in Cl_r^{\geq} \text{ or } x_i b_{-i} \in Cl_s^{\geq}\}$$

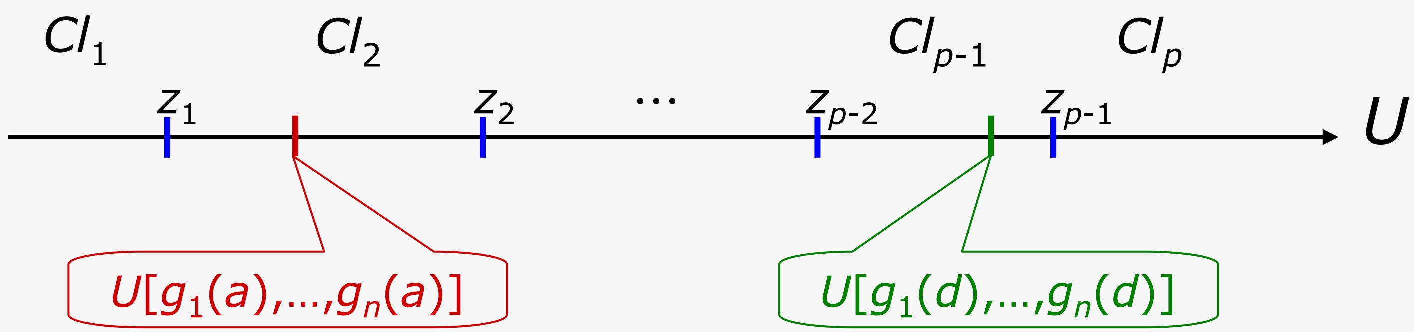
- The above axiom constitutes a minimal condition that makes the weak preference relation $\succeq_i$ a complete preorder

- This axiom does not require pre-definition of criteria scales g_i, nor the dominance relation, in order to derive 3 preference models: general utility function, outranking relation, set of decision rules $D_{\geq}$ or $D_{\leq}$

Greco, S., Matarazzo, B., Słowiński, R.: Conjoint measurement and rough set approach for multicriteria sorting problems in presence of ordinal criteria. [In]: A.Colorni, M.Paruccini, B.Roy (eds.), *A-MCD-A: Aide Multi Critère à la Décision – Multiple Criteria Decision Aiding*, European Commission Report EUR 19808 EN, Joint Research Centre, Ispra, 2001, pp. 117-144

Comparison of decision rule preference model and utility function

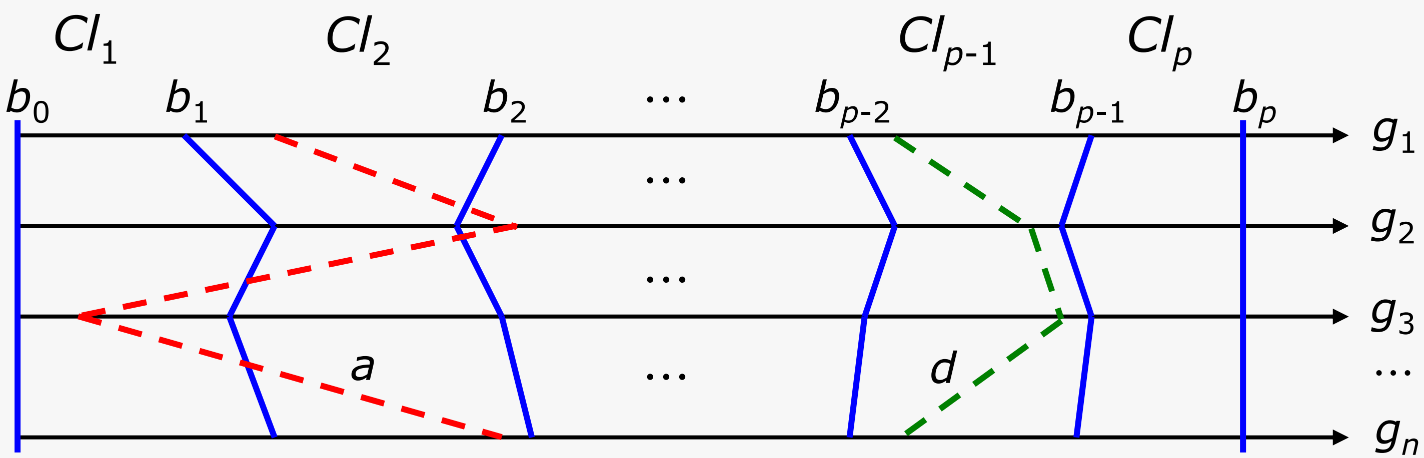
- Value-driven methods
- The preference model is a utility function U and a set of thresholds z_t, t=1,…,p-1, on U, separating the decision classes Cl_t, t=0,1,…,p



- A value of utility function U is calculated for each action $a \in A$
- e.g. $a \rightarrow Cl_2$, $d \rightarrow Cl_{p-1}$

Comparison of decision rule preference model and outranking relation

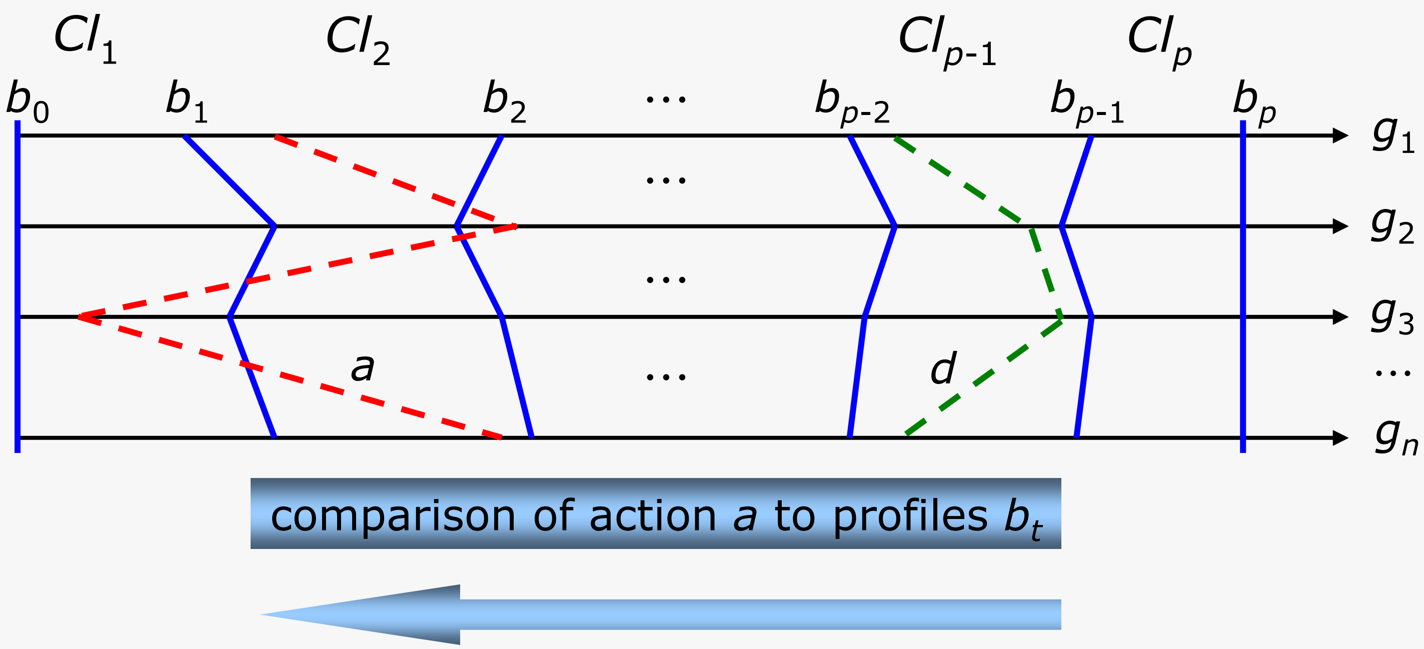
- ELECTRE TRI
- Decision classes Cl_t are caracterized by limit profiles b_t, t=0,1,…,p



- The preference model, i.e. outranking relation S, is constructed for each couple (a, b_t), for every $a \in A$ and b_t, t=0,1,…,p

Comparison of decision rule preference model and outranking relation

- ELECTRE TRI
- Decision classes Cl_t are caracterized by limit profiles b_t, t=0,1,…,p

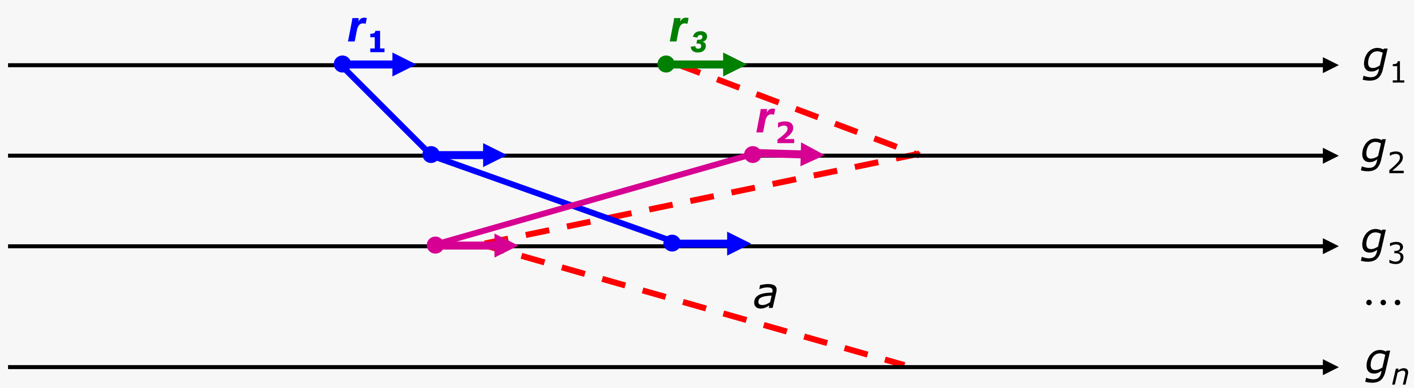


- Compare action a successively to each profile b_t, $t=p$-1,…,1,0;
  if b_t is the first profile such that aSb_t, then $a \rightarrow Cl_{t+1}$
- e.g. $a \rightarrow Cl_1$, $d \rightarrow Cl_{p-1}$

Comparison of decision rule preference model and outranking relation

- Rule-based classification
- The preference model is a set of decision rules for unions $Cl_t^{\geq}$, t=2,…,p

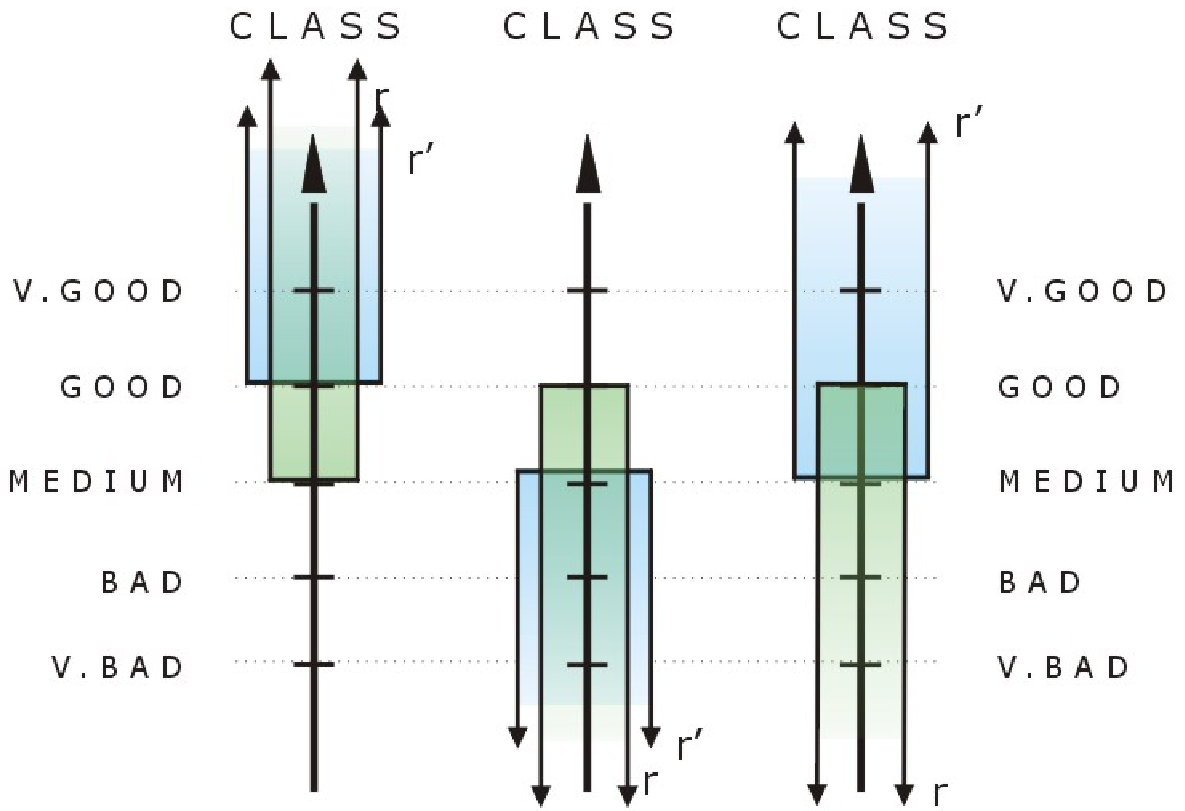
e.g. for $Cl_2^{\geq}$



- A decision rule compares an action profile to a partial profile using a dominance relation
- e.g. $a \rightarrow Cl_2^{\geq}$, because profile of a dominates partial profiles of r_2 and r_3

DRSA – application of decision rules

- Application of decision rules: „intersection” of rules matching object *x*



Final assignment: $x \in G$ $x \in M$ $x \in \langle M, G \rangle$

Decision Rules

Decision rules

- Discovering rules from data is the domain of inductive reasoning (IR)
- IR uses data about a sample of larger reality to start inference
- $S=\langle U, A\rangle$ – *data table*, where U and A are finite, non-empty sets
  U – *universe*; A – set of *attributes*
- $S=\langle U, C, D\rangle$ – *decision table*, where C – set of *condition attributes*,
  D – set of *decision attributes*, $C \cap D=\varnothing$

e.g.

Characterization of nationalities

U	*Height*	*Hair*	*Eyes*	*Nationality*	Support
1	*tall*	*blond*	*blue*	*Swede*	270
2	*medium*	*dark*	*hazel*	*German*	90
3	*medium*	*blond*	*blue*	*Swede*	90
4	*tall*	*blond*	*blue*	*German*	360
5	*short*	*red*	*blue*	*German*	45
6	*medium*	*dark*	*hazel*	*Swede*	45

$\underbrace{\textit{Height} \quad \textit{Hair} \quad \textit{Eyes}}_{C}$ $\underbrace{\textit{Nationality}}_{D}$

Decision rules

- With every subset of attributes $P \subseteq A$, one can associate a formal language of formulas ***L***, called *decision language*

- Formulas are built from attribute-value pairs (q,v), where $q \in P$ and $v \in V_q$ (domain of q), using logical connectives $\wedge$, $\vee$, $\neg$

- All formulas in ***L*** are partitioned into *condition* and *decision formulas*

- *Decision rule* or *association rule* induced from *S* is a *consequence relation*: $\Phi \rightarrow \Psi$ *read as* ***if Φ, then Ψ***

  where Φ and Ψ are condition and decision formulas expressed in ***L***

Decision rules

- $\|\Phi\|_S$ is the set of all objects from *U*, having property Φ in *S*
- $\|\Psi\|_S$ is the set of all objects from *U*, having property Ψ in *S*

- In the *Rough Set approach*, $\|\Psi\|_S$ is:
  - *C*-lower approximation, or
  - *C*-upper approximation, or
  - *C*-boundary of formula Ψ in S,

  giving thus a *certain*, or *possible*, or *approximate* rule $\Phi \rightarrow \Psi$, resp.

- Basic quantitative characteristics of rules

Measures characterizing decision rules in system $S=\langle U, C, D\rangle$

- *Support* of decision rule $\Phi \rightarrow \Psi$ in S:

$$supp_S(\Phi,\Psi) = card(\|\Phi \wedge \Psi\|_S)$$

- *Strength* of decision rule $\Phi \rightarrow \Psi$ in S:

$$str_S(\Phi,\Psi) = \frac{card(\|\Phi \wedge \Psi\|_S)}{card(U)}$$

- *Certainty factor* for decision rule $\Phi \rightarrow \Psi$ in S (Łukasiewicz, 1913):
  (called also *confidence*)

$$cer_S(\Phi,\Psi) = \frac{card(\|\Phi \wedge \Psi\|_S)}{card(\|\Phi\|_S)}$$

- *Coverage factor* for decision rule $\Phi \rightarrow \Psi$ in S:

$$cov_S(\Phi,\Psi) = \frac{card(\|\Phi \wedge \Psi\|_S)}{card(\|\Psi\|_S)}$$

Measures characterizing decision rules in system $S=\langle U, C, D\rangle$

- Certainty and coverage factors refer to *Bayes' theorem*

$$cer_S(\Phi,\Psi)=Pr(\Psi|\Phi)=\frac{Pr(\Psi\wedge\Phi)}{Pr(\Phi)},\quad cov_S(\Phi,\Psi)=Pr(\Phi|\Psi)=\frac{Pr(\Phi\wedge\Psi)}{Pr(\Psi)}$$

- Given a decision table S, the probability (*frequency*) is calculated as:

$$Pr(\Phi)=\frac{card(\|\Phi\|_S)}{card(U)},\quad Pr(\Psi)=\frac{card(\|\Psi\|_S)}{card(U)},\quad Pr(\Phi\wedge\Psi)=\frac{card(\|\Phi\wedge\Psi\|_S)}{card(U)}$$

- In fact, *without referring to prior and posterior probability*:

$$cer_S(\Phi,\Psi)\times card(\|\Phi\|_S)=cov_S(\Phi,\Psi)\times card(\|\Psi\|_S)$$

- What is the certainty factor for $\Phi\rightarrow\Psi$ is the coverage factor for $\Psi\rightarrow\Phi$
- This underlines a directional character of the statement *if Φ, then Ψ*
  (e.g. „if x is a square, then x is a rectangle")

Decision rules

- E.g. decision rules induced from „characterization of nationalities”:
  1) If (*Height*, *tall*), then (*Nationality*, *Swede*)
  2) If (*Height*, *medium*) and (*Hair*, *dark*), then (*Nationality*, *German*)
  3) If (*Height*, *medium*) and (*Hair*, *blond*), then (*Nationality*, *Swede*)
  4) If (*Height*, *tall*), then (*Nationality*, *German*)
  5) If (*Height*, *short*), then (*Nationality*, *German*)
  6) If (*Height*, *medium*) and (*Hair*, *dark*), then (*Nationality*, Swede)

Certainty and coverage factors

Rule number	Certainty	Coverage	Support	Strength
1	0.43	0.67	270	0.3
2	0.67	0.18	90	0.1
3	1.00	0.22	90	0.1
4	0.57	0.73	360	0.4
5	1.00	0.09	45	0.05
6	0.33	0.11	45	0.05

43% *tall* people are *Swede*

67% *Swede* are *tall*

certain rules

Decision rules

- Decision rules $\Phi \rightarrow \Psi$ have a double utility:
  - they **represent knowledge** about the universe
    in terms of laws relating some properties Φ with properties Ψ,
  - they can be used for **prospective decisions**.

- The use of rules for prospective decisions can be understood in two ways:
  - **matching up the rules to new objects** with property Φ
    in view of predicting property Ψ,
  - **building a strategy of intervention** based on discovered rules
    in view of transforming the universe in a desired way.

Decision rules

- For example, rules mined from medical data are useful to:
  - *represent relationships* between symptoms and diseases
    - *if* test α=P & test β=N, *then* no disease *d*
  - *diagnose new patients*
    - for patient *x*: test α=P & test β=N $\Rightarrow$ *x* is not sick of *d*
- Moreover, rules can be seen as *general laws to be considered for an intervention*:
    - for all patients with:
      - α=N & β=N
      - α=N & β=P
      - α=P & β=P

    apply a therapy aiming at getting α=P & β=N in order to get out from disease *d*

Decision rules – attractiveness measures

- In all practical applications, like *medical practice*, *market basket*, *customer satisfaction or risk analysis*, it is crucial to know how good the rules are for:
  - knowledge representation & prediction (how strong is the law $\Phi \rightarrow \Psi$, and what is the chance of getting Ψ when Φ holds ?)
  - efficient intervention (how efficient will be the action based on a rule discovered in U, and taken in U' ?)
- "How good" is a question about *attractiveness measures* of rules
- Review of literature shows that there is no single measure which would be the best for applications in all possible perspectives
  (e.g. Bayardo and Agrawal 1999, Greco, Pawlak & Slowinski 2004, Yao & Zhong 1999, Hilderman and Hamilton 2001)

Decision rules – knowledge representation and prediction

- $\Phi \rightarrow \Psi$ are laws „naturally" characterized by:
  - number of cases from U supporting them, i.e. strength

$$str_S(\Phi,\Psi) = \frac{card\left(\|\Phi \wedge \Psi\|_S\right)}{card(U)}$$

  - probability of getting decision Ψ when condition Φ holds, i.e. certainty

$$cer_S(\Phi,\Psi) = \frac{card\left(\|\Phi \wedge \Psi\|_S\right)}{card\left(\|\Phi\|_S\right)}$$

- Why not other statistical interestingness measures, like lift, conviction, laplace, piatetsky-shapiro, kamber-shingal, gini, chi-squared value… ?
- Because for a given hypothesis (fixed Ψ), the Pareto set of rules with respect to strength and certainty includes all rules that are best according to any of these measures (Bayardo and Agrawal 1999)

Decision rules – knowledge representation and prediction

Let $a = supp_S(\Phi, \Psi)$ - the number of objects in U for which Φ and Ψ hold together…

$b = supp_S(\neg\Phi, \Psi)$

$c = supp_S(\Phi, \neg\Psi)$

$d = supp_S(\neg\Phi, \neg\Psi)$

$u = card(U)$

$f = card(\|\Phi\|_S)$, $f' = card(\|\neg\Phi\|_S)$

$p = card(\|\Psi\|_S)$, $p' = card(\|\neg\Psi\|_S)$

$lift = ua / p$

$conviction = f / uc$

$laplace = (a+1)/(f+k)$, where k - number of classes

$piatetsky\text{-}shapiro = a\text{-}fp/u$

$kamber\text{-}shingal = a(1\text{-}b/d)/c$

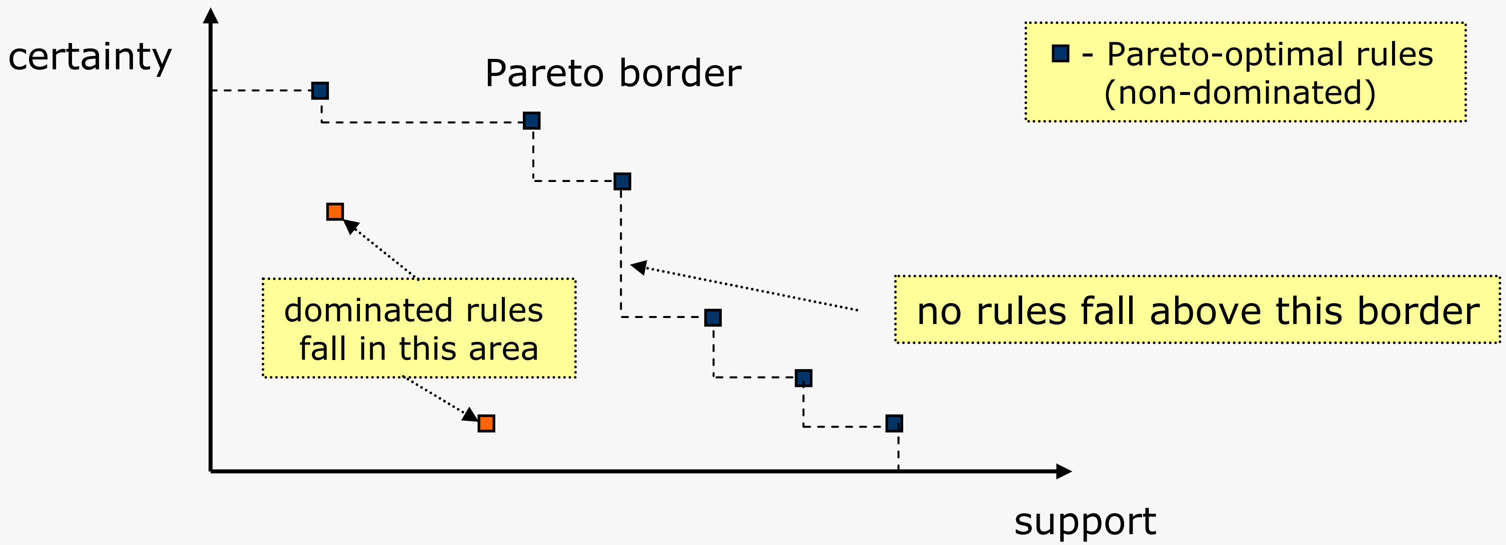
$gray\text{-}orlowska = \left[(au/fp)^h - 1\right](fp/u^2)^m$, with, e.g., $h = m = 1$

$$gini = \left[1 - \left((p/u)^2 + (u-p)^2/u^2\right)\right] - \left[(u/f)\left(1 - \left((a/f)^2 + (f-a)^2/f^2\right)\right)\right] - \\ - \left[(f'/u)\left(1 - \left((b/f')^2 + (f'-b)^2/f'^2\right)\right)\right]$$

$$chi^2 = \frac{f(a/f - p/u)^2 - f'(b/f' - p/u)^2}{p/u} + \\ + \frac{f((f-a)/f - p'/u)^2 - f'((f'-b)/f' - p'/u)^2}{p'/u}$$

Support-certainty Pareto border

- Support-certainty Pareto border is the set of non-dominated, Pareto-optimal rules with respect to both rule support and certainty



- Mining the border identifies rules optimal with respect to measures such as: *lift*, *gain*, *conviction*, *piatetsky-shapiro*,…

Support-certainty Pareto border – example

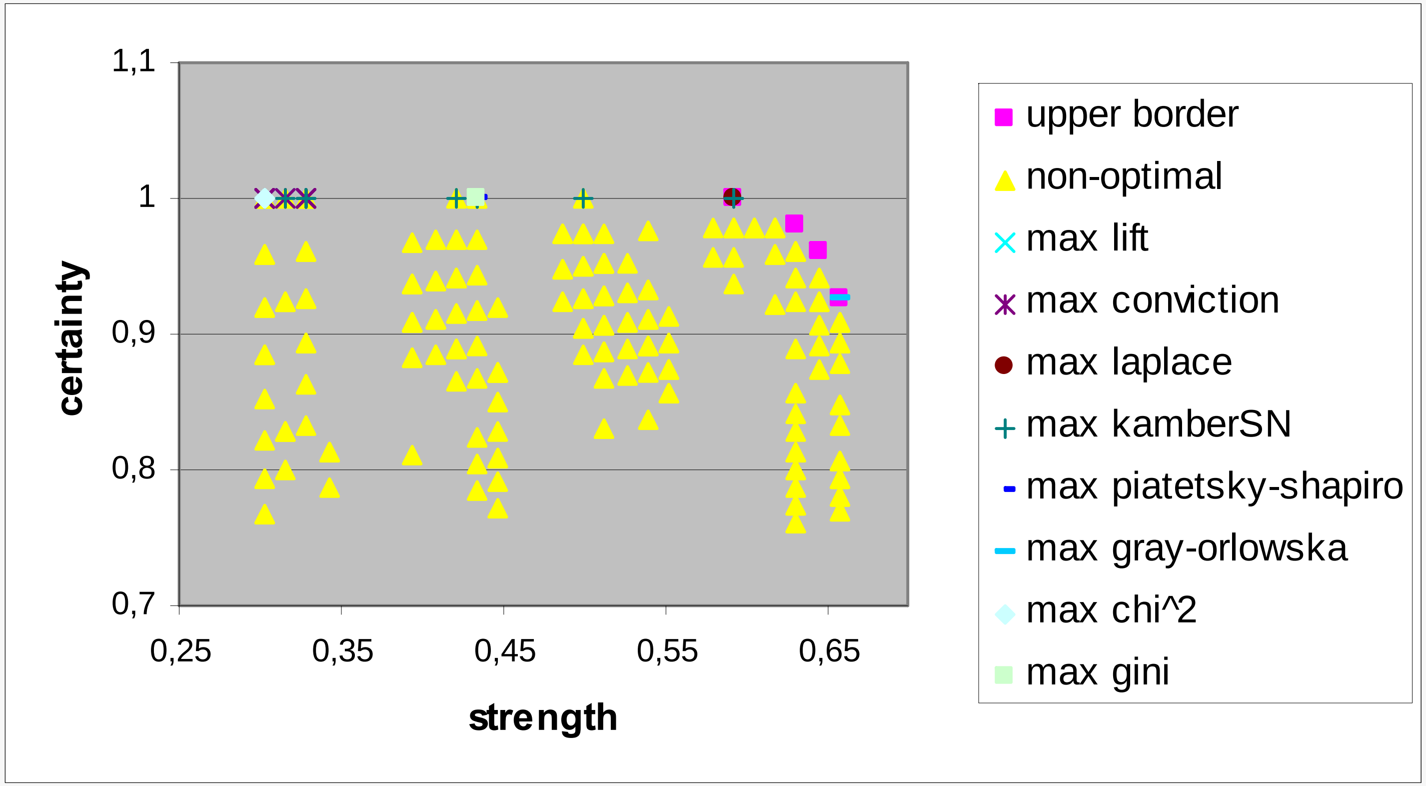
- ***Buses***

  187: (MaxSpeed ≥ 74) & (Blacking ≤ 65) & (SummerCons ≤ 26) => (State ≥ 2)

  Certainty = 0.96, Strength = 0.63

  Positive support: 1, 4, 5, 7, 8, 12, 13, 15, 17, 18, 20, 21, 22, 25, 26, 29, 30, 31, 32, 33, 35, 36, 37, 39, 41, 42, 43, 44, 49, 51, 52, 53, 54, 55, 56, 57, 59, 61, 64, 65, 66, 70, 71, 72, 73, 74, 75, 76

  Negative support: 11, 58



Decision rules – knowledge representation and prediction

- In statistics, measures of confirmation quantify the degree to which a piece of evidence Φ provides support for or against hypothesis Ψ (Fitelson 2001):

$$c(\Phi,\Psi)\begin{cases} >0 & \text{if } Pr(\Psi|\Phi) > Pr(\Psi) \\ =0 & \text{if } Pr(\Psi|\Phi) = Pr(\Psi) \\ <0 & \text{if } Pr(\Psi|\Phi) < Pr(\Psi) \end{cases}$$

- Its meaning is different from a simple statistics of co-occurrence of properties Φ and Ψ in universe U

S.Greco, Z.Pawlak, R.Słowiński: Can Bayesian confirmation measures be useful for rough set decision rules? *Engineering Applications of Artificial Intelligence*, 17 (2004) no.4, 345-361

Bayesian confirmation measure

- The most well-known measures of confirmation

$$d(\Phi,\Psi) = Pr(\Psi|\Phi) - Pr(\Psi)$$

Earman (1992), Eells (1982), Gillies (1986), Jeffrey (1992), Rosenkrantz (1994)

$$r(\Phi,\Psi) = log\left[\frac{Pr(\Psi|\Phi)}{Pr(\Psi)}\right]$$

Horwich (1982), Keynes (1921), Mackie (1969), Milne (1995, 1996), Schlesinger (1995), Pollard (1999)

$$s(\Phi,\Psi) = Pr(\Psi|\Phi) - Pr(\Psi|\neg\Phi)$$

Christensen (1999), Joyce (1999)

$$b(\Phi,\Psi) = Pr(\Psi \wedge \Phi) - Pr(\Psi)Pr(\Phi)$$

Carnap (1962)

$$l(\Phi,\Psi) = log\left[\frac{Pr(\Phi|\Psi)}{Pr(\Phi|\neg\Psi)}\right]$$

$$f(\Phi,\Psi) = \frac{Pr(\Phi|\Psi) - Pr(\Phi|\neg\Psi)}{Pr(\Phi|\Psi) + Pr(\Phi|\neg\Psi)}$$

Kemeny & Oppenheim (1952), Good (1984), Heckerman (1988), Schumm (1994), Horvitz & Heckerman (1986), Pearl (1988),

Fitelson (2001)

Why the certainty measure is not sufficient?

- Example (Popper, 1959)

  Consider the possible result of rolling a die: 1,2,3,4,5,6.

  Ψ = "the result is **6**" $\neg\Psi$ = "the result is ***not*** **6**"

  Φ = "the result is an **even** number (i.e. 2 or 4 or 6)"

  - $\Phi \rightarrow \Psi$, $cer_S(\Phi, \Psi) = 1/3$
  - $\Phi \rightarrow \neg\Psi$, $cer_S(\Phi, \neg\Psi) = 2/3$

- Probability that the result is **6** is 1/6,

  while the probability that the result is ***not*** **6** is 5/6

- Information Φ increases the probability of Ψ from 1/6 to 1/3, and decreases the probability of $\neg\Psi$ from 5/6 to 2/3

- In conclusion: $\boldsymbol{\Phi}$ confirms $\boldsymbol{\Psi}$ and disconfirms $\neg\boldsymbol{\Psi}$,

  <u>independently</u> of the fact that $cer_S(\Phi, \Psi) < cer_S(\Phi, \neg\Psi)$

Bayesian confirmation measure

- Given a decision rule $\Phi \to \Psi$, the Bayesian confirmation measure gives the credibility of the proposition:

Ψ is satisfied more frequently when Φ is satisfied rather than when Φ is not satisfied

Bayesian confirmation measure

- $c(\Phi, \Psi) > 0$ means that property Ψ is satisfied more frequently when Φ is satisfied (then, this frequency is $cer_S(\Phi, \Psi)$), rather than generically (frequency is $Fr_S(\Psi)$),

- $c(\Phi, \Psi) = 0$ means that property Ψ is satisfied with the same frequency whether Φ is satisfied or not

- $c(\Phi, \Psi) < 0$ means that property Ψ is satisfied less frequently when Φ is satisfied, rather than generically

Bayesian confirmation measure for decision rules

- Assuming $Fr_S(\Psi) = \dfrac{card(\|\Psi\|_S)}{card(U)}$:

$$c(\Phi,\Psi)\begin{cases} >0 & \text{if } \; Pr(\Psi|\Phi) > Pr(\Psi) \\ =0 & \text{if } \; Pr(\Psi|\Phi) = Pr(\Psi) \\ <0 & \text{if } \; Pr(\Psi|\Phi) < Pr(\Psi) \end{cases}$$

$$\Downarrow$$

$$c(\Phi,\Psi)\begin{cases} >0 & if \;\; cer_S(\Phi,\Psi) > Fr_S(\Psi) \\ =0 & if \;\; cer_S(\Phi,\Psi) = Fr_S(\Psi) \\ <0 & if \;\; cer_S(\Phi,\Psi) < Fr_S(\Psi) \end{cases}$$

Bayesian confirmation measure for decision rules

- The most well-known measures of confirmation

$$d(\Phi,\Psi) = cer_S(\Phi,\Psi) - Fr_S(\Psi)$$

$$r(\Phi,\Psi) = log\left[\frac{cer_S(\Phi,\Psi)}{Fr_S(\Psi)}\right]$$

$$s(\Phi,\Psi) = cer_S(\Phi,\Psi) - cer_S(\neg\Phi,\Psi)$$

$$b(\Phi,\Psi) = str_S(\Phi,\Psi) - Fr_S(\Psi)\,Fr_S(\Phi)$$

$$l(\Phi,\Psi) = log\left[\frac{cer_S(\Psi,\Phi)}{cer_S(\neg\Psi,\Phi)}\right]$$

$$f(\Phi,\Psi) = \frac{cer_S(\Psi,\Phi) - cer_S(\neg\Psi,\Phi)}{cer_S(\Psi,\Phi) + cer_S(\neg\Psi,\Phi)}$$

Bayesian confirmation measure

- Desirable properties of $c(\Phi,\Psi)$:
  - *hypothesis symmetry* (Eells, Fitelson 2002): $c(\Phi,\Psi) = -c(\Phi,\neg\Psi)$
  - *monotonicity property* (M) (Greco, Pawlak, Słowiński 2004):

    $a=supp_S(\Phi,\Psi)$, $b=supp_S(\neg\Phi,\Psi)$, $c=supp_S(\Phi,\neg\Psi)$, $d=supp_S(\neg\Phi,\neg\Psi)$

    $c(\Phi,\Psi) = F(a,b,c,d)$, where F is a function non-decreasing with respect to a and d and non-increasing with respect to b and c

- Among all popular confirmation measures, the only ones that satisfy both properties are (Greco, Pawlak, Słowiński 2004):

$$f(\Phi,\Psi) = \frac{cer_S(\Psi,\Phi) - cer_S(\neg\Psi,\Phi)}{cer_S(\Psi,\Phi) + cer_S(\neg\Psi,\Phi)}$$

$$l(\Phi,\Psi) = log\left[\frac{cer_S(\Psi,\Phi)}{cer_S(\neg\Psi,\Phi)}\right]$$

- $l(\Phi,\Psi)$ and $f(\Phi,\Psi)$ are *ordinally equivalent* (Fitelson 2001)

Interpretation of the monotonicity property M

- E.g. (Hempel) consider rule $\phi \rightarrow \psi$:

  if x is a raven then x is black

- ϕ is the property *to be a raven*, ψ is the property *to be black*

  - *a* – the number of objects in *U* which are black ravens

    (the more black ravens we observe, the more credible becomes the rule)

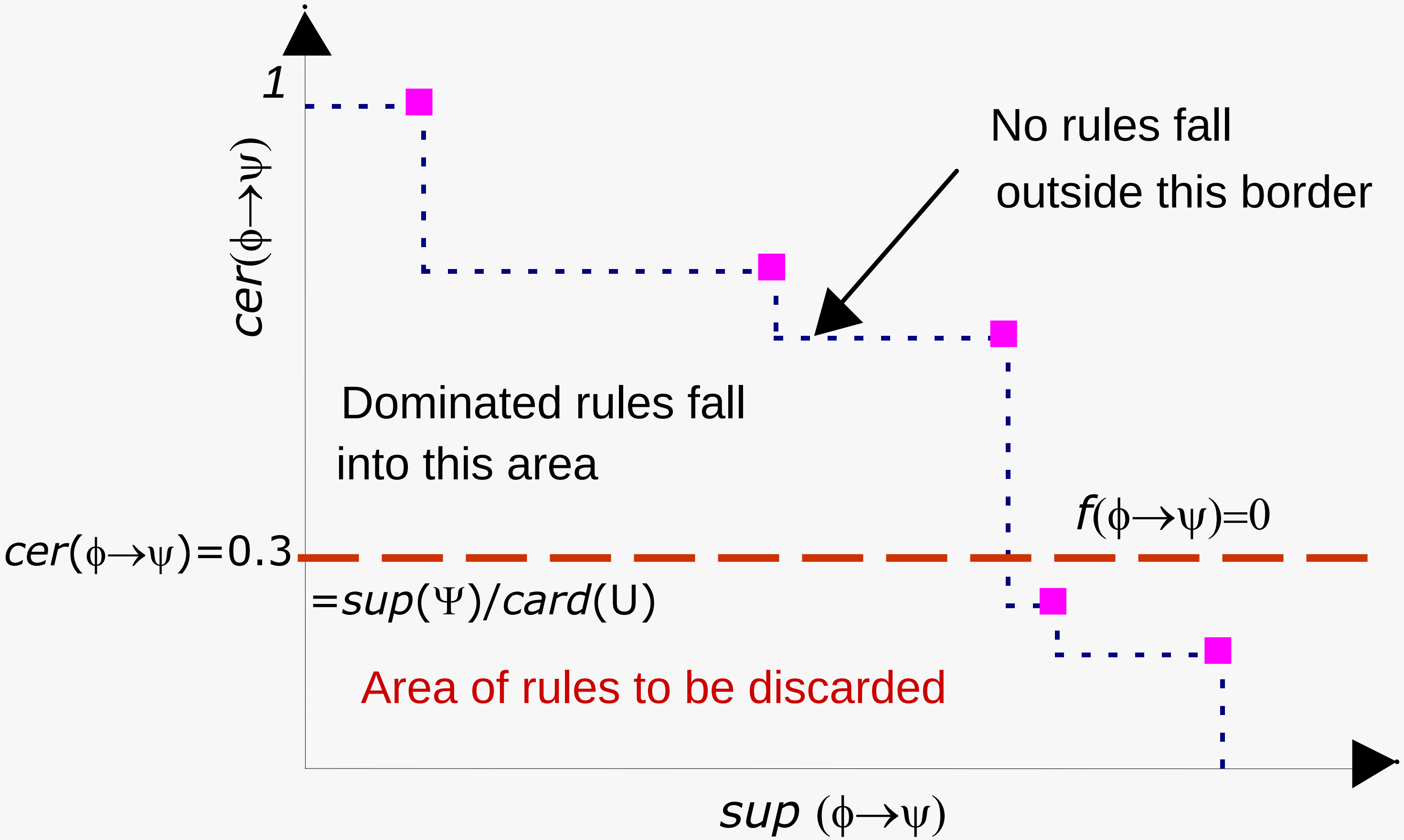
  - *b* – the no. of objects in *U* which are black non-ravens

  - *c* – the no. of objects in *U* which are non-black ravens

  - *d* – the no. of objects in *U* which are non-black non-ravens

Support-certainty vs. support-confirmation Pareto border

The set of rules located on the **support-certainty** Pareto border is exactly the same as on the **support-*f*** Pareto border (Greco, Brzezińska, Słowiński 2006)



The **support-*f*** Pareto border **is more meaningful** than the support-certainty Pareto border

Confirmation perspective on support-confidence space

- Is there a curve separating rules with negative value of any measure with the confirmation property in the support-confidence space?

- *Theorem:*

  Rules lying above a constant:

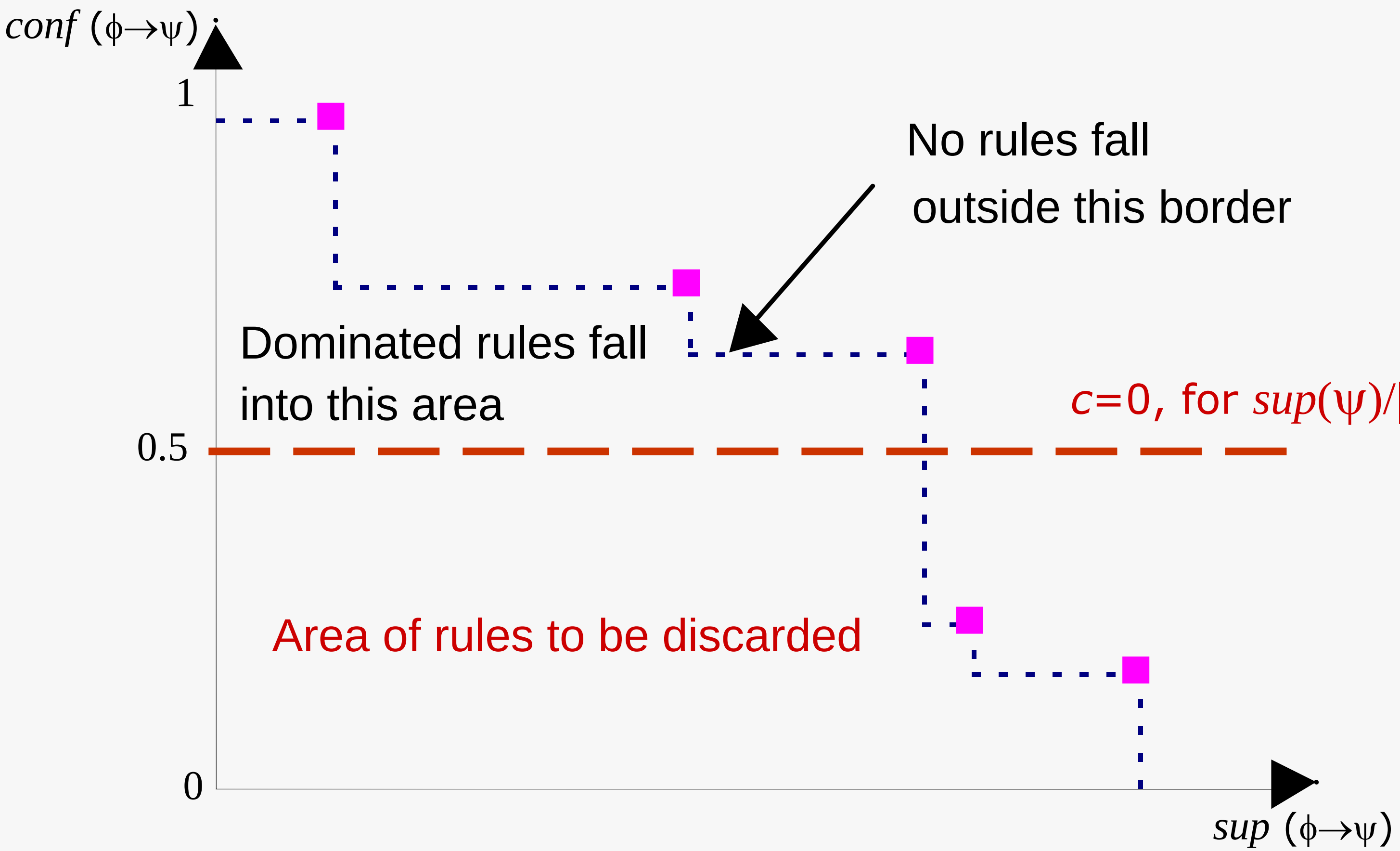
  $$sup(\psi)/|U|$$

  have a negative value of any confirmation measure.

  For those rules, the premise only disconfirms the conclusion!

Słowiński R., Szczęch I., Greco S.: Mining Association Ruleswith respect to Support and Anti–support - experimental results.

Confirmation perspective on support-confidence space



For rules lying below the curve for which $c=0$
the premise only disconfirms the conclusion

Computational experiment: general info about the dataset

- Dataset ***adult***, created in ’96 by B. Becker & R. Kohavi from census database
- **32 561 instances**
- 9 nominal attributes
  - workclass: Private, Local-gov, etc.;
  - education: Bachelors, Some-college, etc.;
  - marital-status: Married, Divorced, Never-married, et.;
  - occupation: Tech-support, Craft-repair, etc.;
  - relationship: Wife, Own-child, Husband, etc.;
  - race: White, Asian-Pac-Islander, etc.;
  - sex: Female, Male;
  - native-country: United-States, Cambodia, England, etc.;
  - salary: >50K, <=50K
- throughout the experiment, $sup(\phi \rightarrow \psi)$ is denotes relative rule support [0,1]

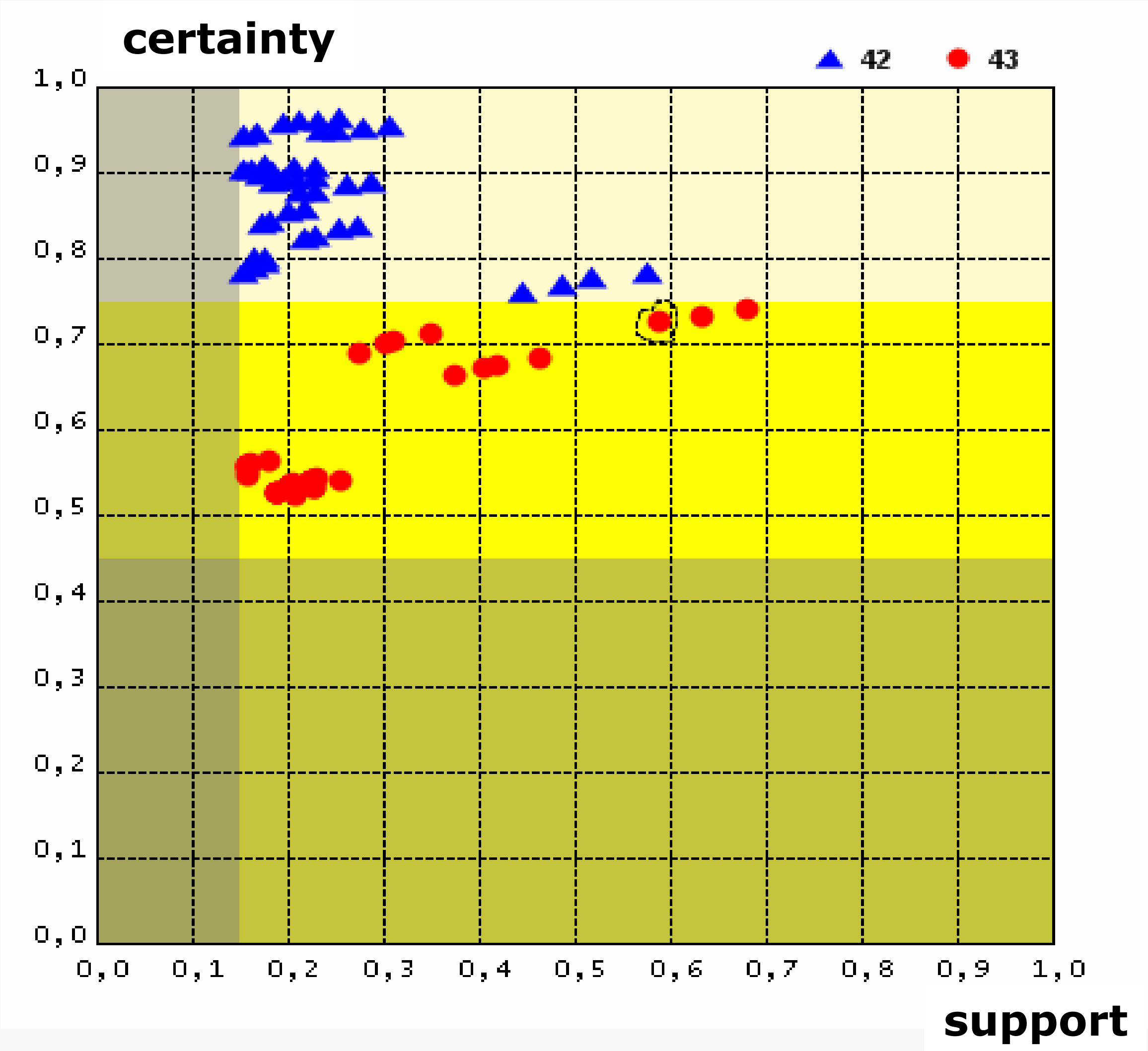
Support-certainty vs. support-confirmation Pareto border

- Example of „CENSUS” dataset:
  - 9 attributes
  - 32.561 instances (objects)

Association rules

premise	conclusion	support	certainty	confirmation s	confirmation f
race is White	native-country is United-States	0,80	0,93	0,16	0,15
native-country is United-States	race is White	0,80	0,88	0,24	0,09
class is <=50K	native-country is United-States	0,68	0,91	-0,03	-0,04
native-country is United-States	class is <=50K	0,68	0,75	-0,06	-0,01
native-country is United-States	workclass is Private	0,67	0,73	-0,08	-0,02
workclass is Private	native-country is United-States	0,67	0,90	-0,03	-0,05
race is White	workclass is Private	0,63	0,74	-0,01	0,00
workclass is Private	race is White	0,63	0,86	0,00	0,00
race is White	class is <=50K	0,63	0,74	-0,11	-0,04
class is <=50K	race is White	0,63	0,84	-0,07	-0,07
native-country is United-States	sex is Male	0,62	0,68	0,00	0,00
sex is Male	native-country is United-States	0,62	0,91	0,00	0,00
race is White	sex is Male	0,60	0,70	0,14	0,05
sex is Male	race is White	0,60	0,89	0,08	0,11
workclass is Private	native-country is United-States and race is White	0,59	0,80	-0,03	-0,02
native-country is United-States and workclass is Private	race is White	0,59	0,88	0,06	0,09
race is White and workclass is Private	native-country is United-States	0,59	0,93	0,04	0,10

Support-certainty vs. support-confirmation Pareto border



„CENSUS” dataset

association rules

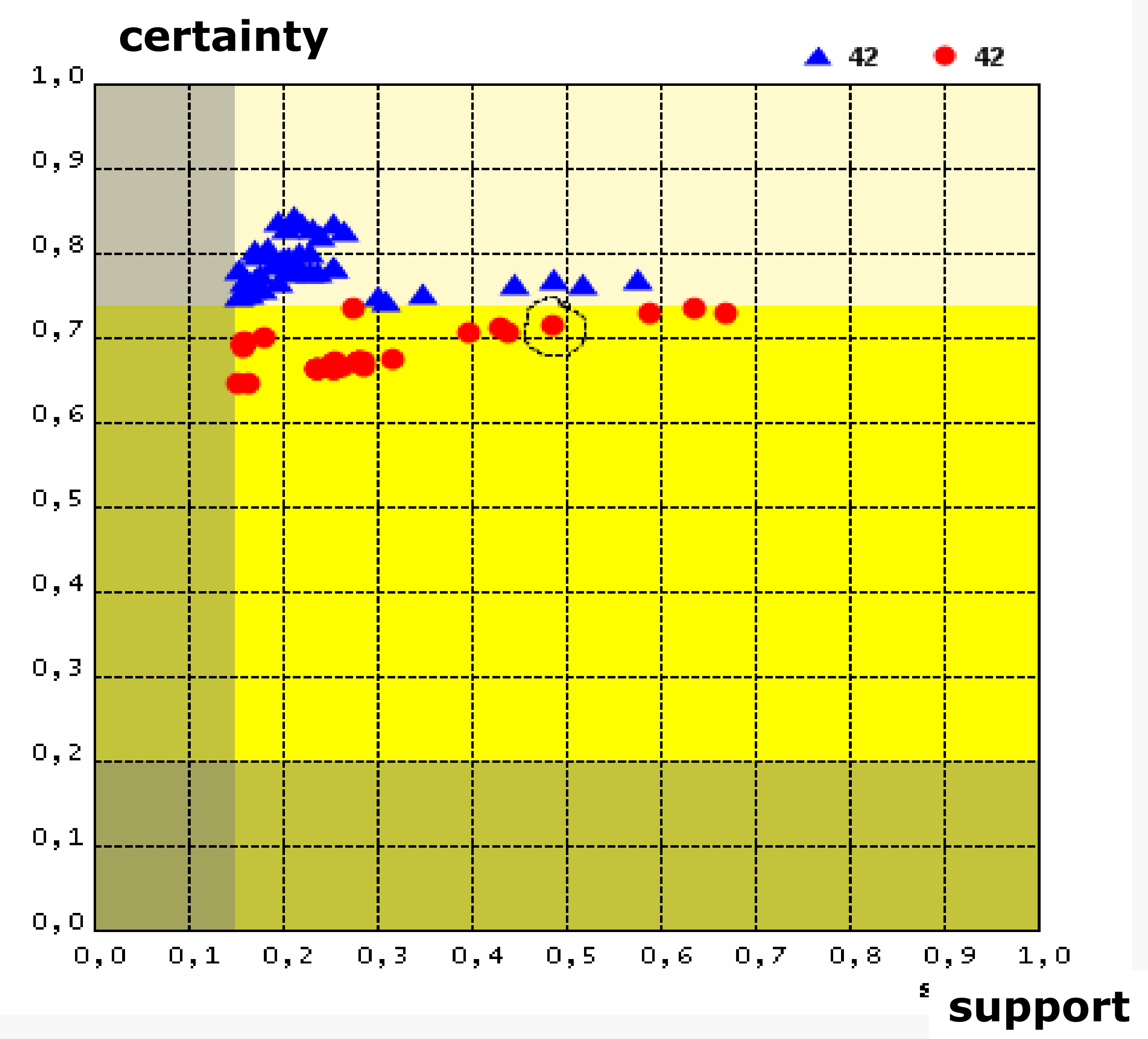
supp ≥ 15%

cer ≥ 45%

● **confirmation<=0**

premise	conclusion	supp	conf	s	f
native-country is United-States and race is White	class is <=50K	0,59	0,73	-0,11	-0,05

Support-certainty vs. support-confirmation Pareto border



„CENSUS" dataset

association rules

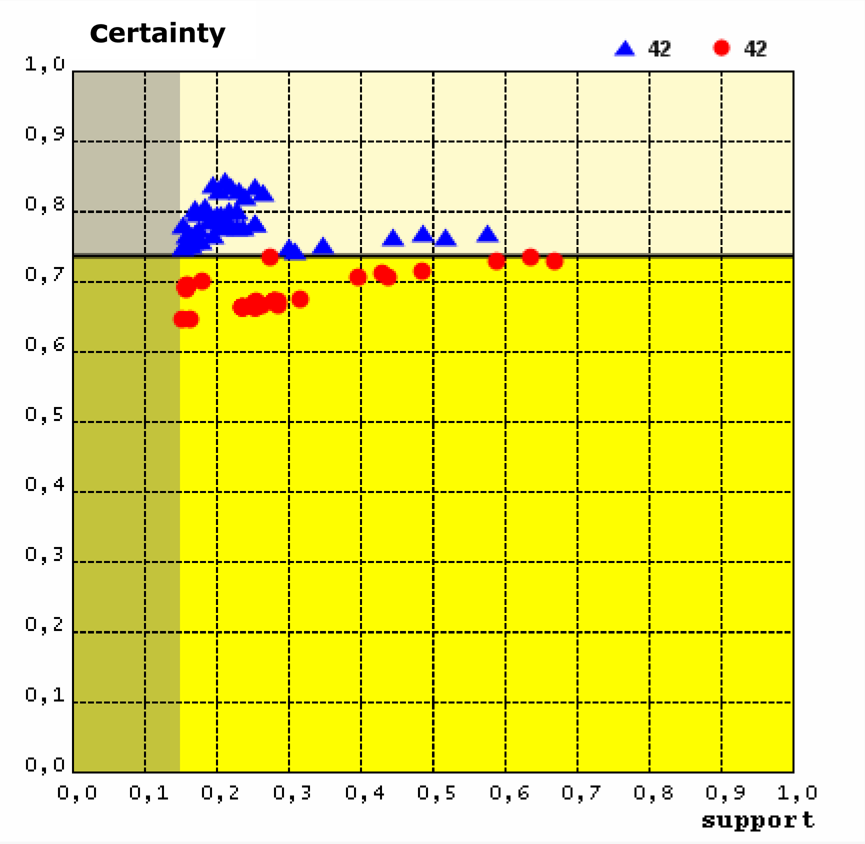
supp ≥ 15%

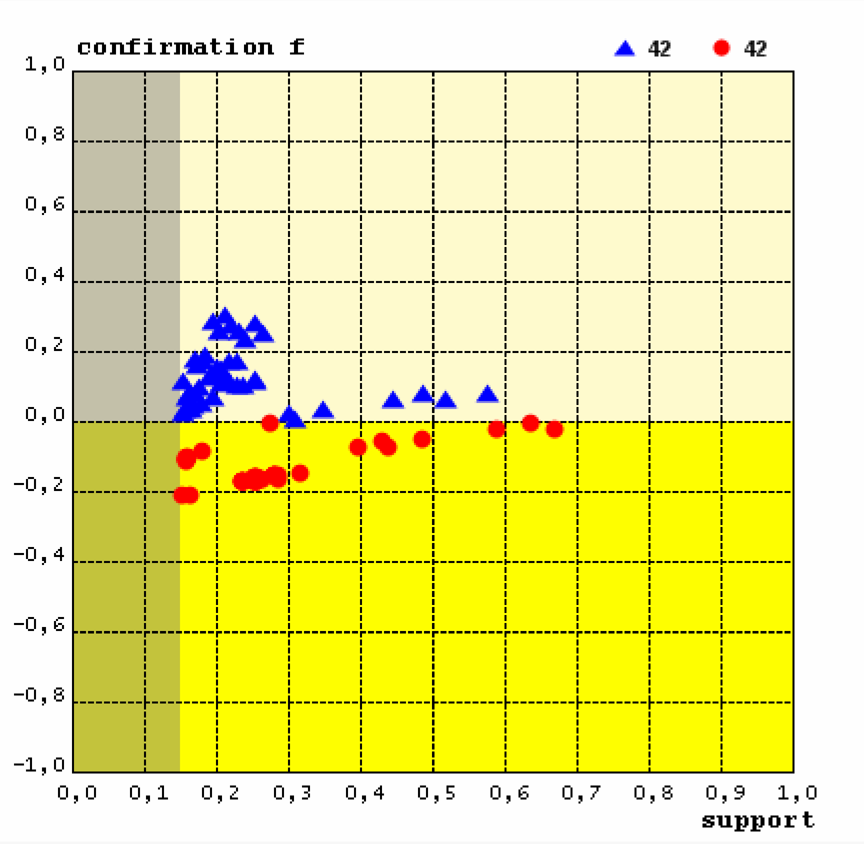
cer ≥ 20%

- **confirmation<=0**

premise	conclusion	supp	conf	s	f
sex is Male	workclass is Private	0,49	0,72	-0,06	-0,05

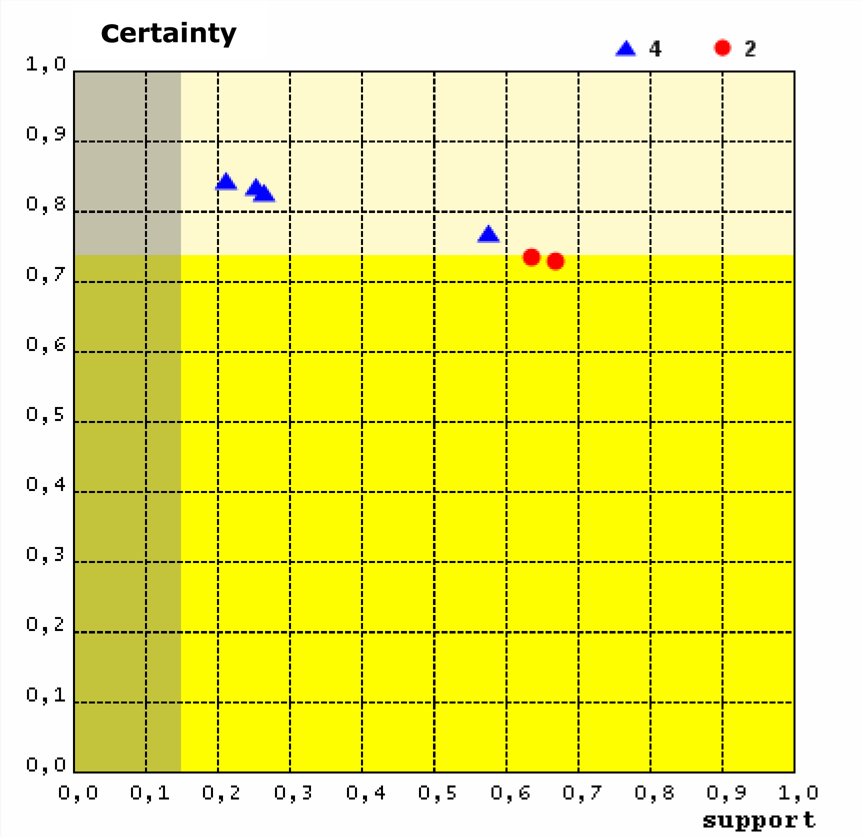
Support-certainty vs. support-confirmation Pareto border

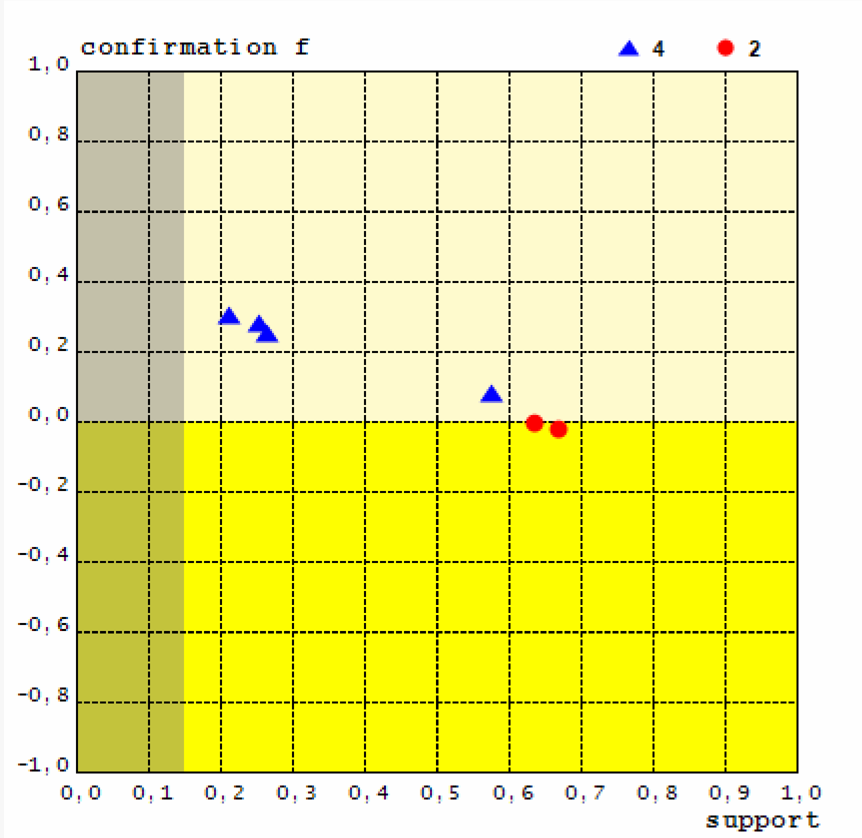




- ● indicates rules with negative confirmation
- the decision class constitutes over 70% of the whole dataset
- rules with high certainty can be disconfirming
- even some rules from the Pareto border need to be discarded

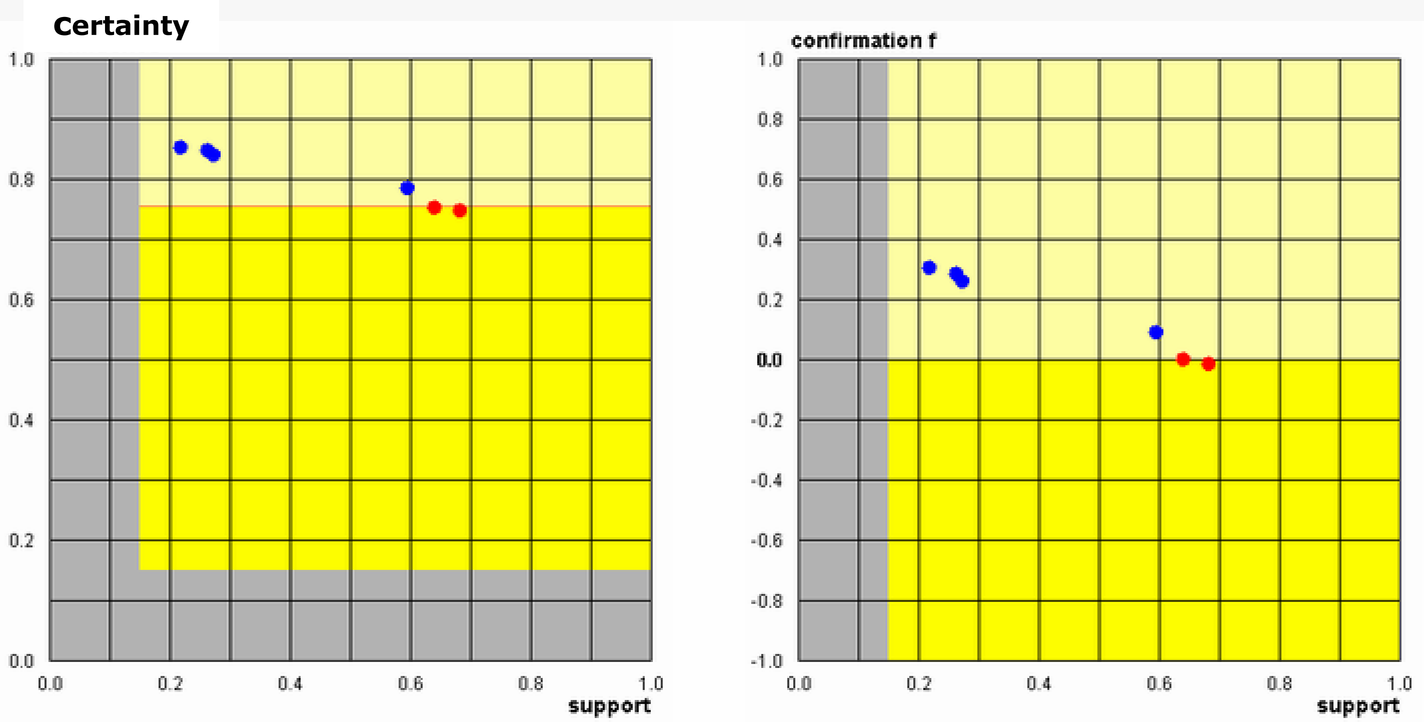
Support-certainty vs. support-confirmation Pareto border





- ● indicates rules with negative confirmation
- both Pareto borders contain the same rules

Support-certainty vs. support-confirmation Pareto border



premise	conclusion	supp	conf	s	f	a-supp
marital-status is Never-married **and race is** White **and class is** =50K	**workclass is** Private	0.22	0.85	0.13	0.30	0.04
marital-status is Never-married **and class is** =50K	**workclass is** Private	0.26	0.85	0.13	0.28	0.05
marital-status is Never-married	**workclass is** Private	0.27	0.84	0.13	0.26	0.05
race is White	**workclass is** Private	0.64	0.75	-0.01	-0.00	0.21
native-country is United-States	**workclass is** Private	0.68	0.75	-0.07	-0.02	0.23
class is =50K	**workclass is** Private	0.60	0.78	0.13	0.09	0.16

Measures with the property M in support-confidence space

- <u>Theorem:</u>
  When the value of support is held fixed, then $F(a, b, c, d)$ is monotone in confidence.

- <u>Theorem:</u>
  When the value of confidence is held fixed, then $F(a, b, c, d)$ admitting derivative with respect to all its variables a, b, c and d, is monotone in support if:

$$\frac{\partial F}{\partial c} = \frac{\partial F}{\partial d} = 0 \quad or \quad \frac{\frac{\partial F}{\partial a} - \frac{\partial F}{\partial b}}{\frac{\partial F}{\partial d} - \frac{\partial F}{\partial c}} \geq \frac{1}{conf(\phi \to \psi)} - 1$$

Measures with the property M in support-confidence space

- Conclusions:
  - *F*or a set of rules with the same conclusion, any interestingness measure with property M is always non-decreasing with respect to confidence when the value of support is kept fixed
  - All those interestingness measures that are independent of $c = sup(\phi \rightarrow \neg\psi)$ and $d = sup(\neg\phi \rightarrow \neg\psi)$ are always monotone in support when the value of confidence remains unchanged
  - There are some measures with property M whose optimal rules will not be on the support-confidence Pareto border.

Support-anti-support Pareto border

- How to find rules optimal according to any confirmation measure with the property of monotonicity (M) ?

- Theorem (Greco, Brzezińska, Słowiński 2006):
  When the value of support is held fixed, then $F(a, b, c, d)$ with property (M) is anti-monotone (non-increasing) in anti-support

- Theorem (Greco, Brzezińska, Słowiński 2006):
  When the value of anti-support is held fixed, then $F(a, b, c, d)$ with property (M) is monotone (non-decreasing) in support

- Anti-support is the number of examples which satisfy the premise of the rule but not its conclusion: $supp(\phi \rightarrow \neg\psi)$

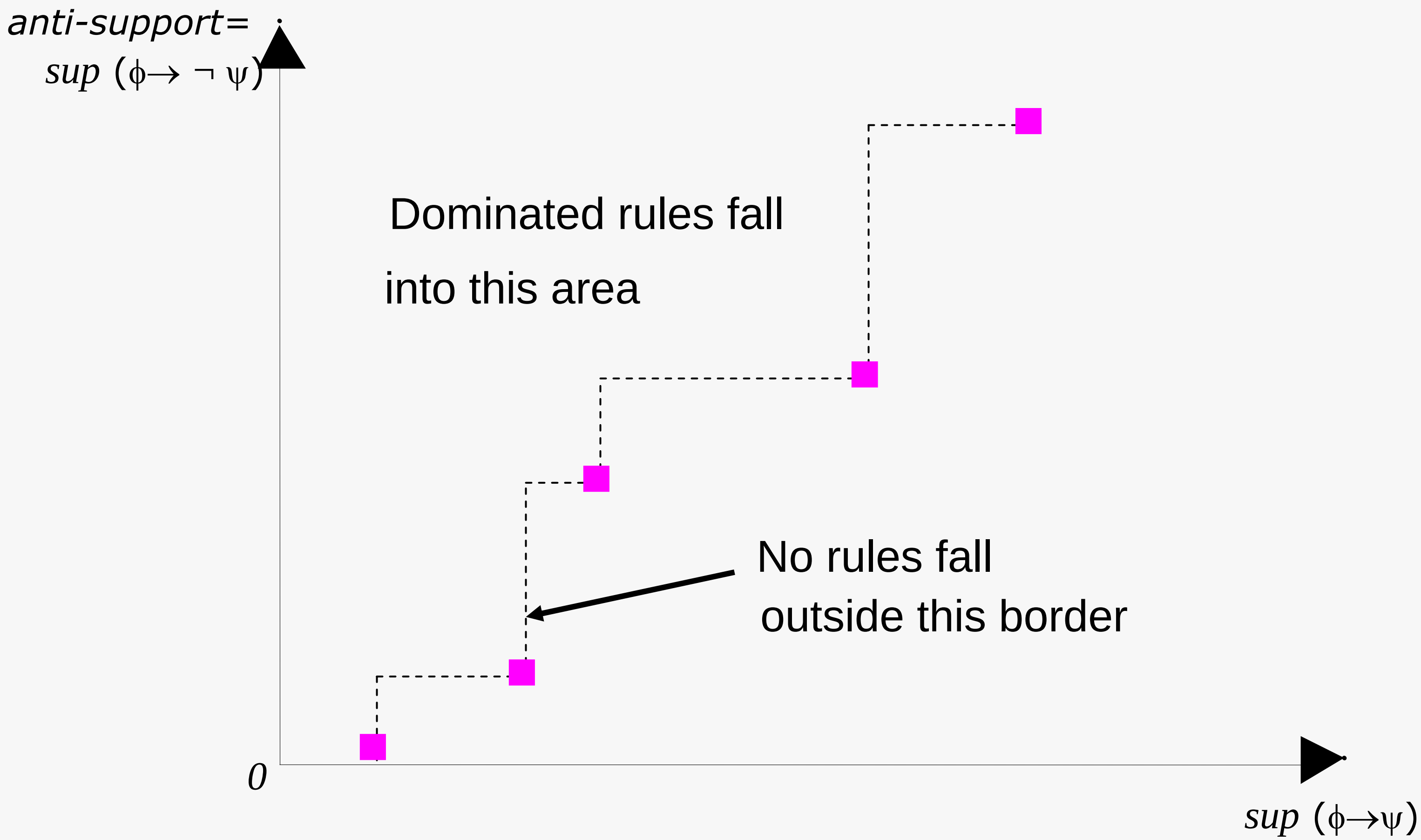
Support - anti-support Pareto border

- Theorem:
  For rules with the same conclusion,
  the best rules according to any measure with the property M
  must reside on the support-anti-support Pareto border

- The support-anti-support Pareto border is the set of rules such that there is no other rule having greater support and smaller anti-support

- Theorem:
  The support - anti-support Pareto border is, in general, not smaller than the support-confidence Pareto border

Support - anti-support Pareto border



The best rules according to any measure with the property M must reside on the support - anti-support Pareto border

Confirmation perspective on support - anti-support border

- Is there a curve separating rules with negative value of any confirmation measure in the support-anti-support space?

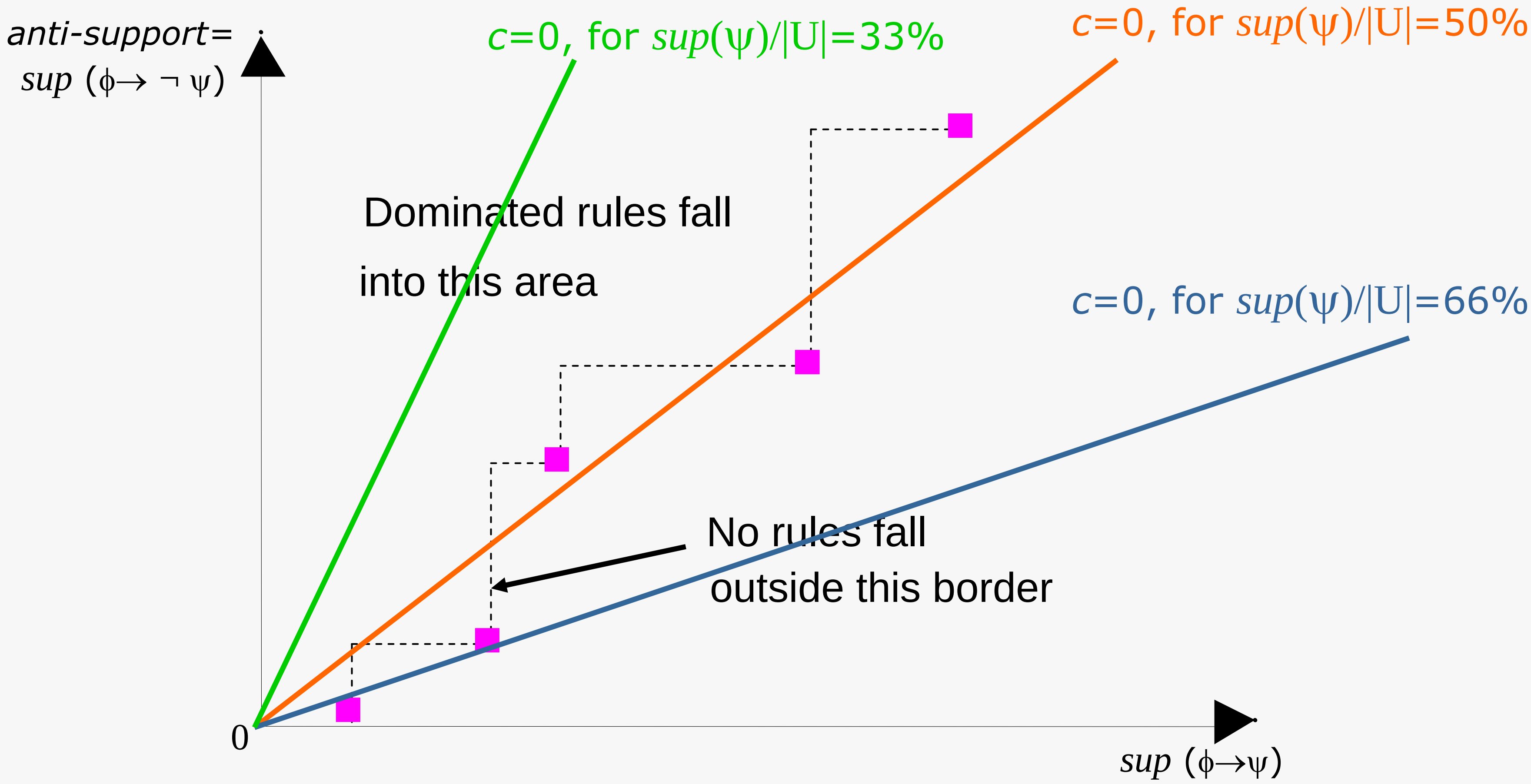
- *Theorem:*
  Rules lying above a linear function:

$$sup(\phi \rightarrow \psi)[|U|/sup(\psi)-1]$$

  have a negative value of any confirmation measure.

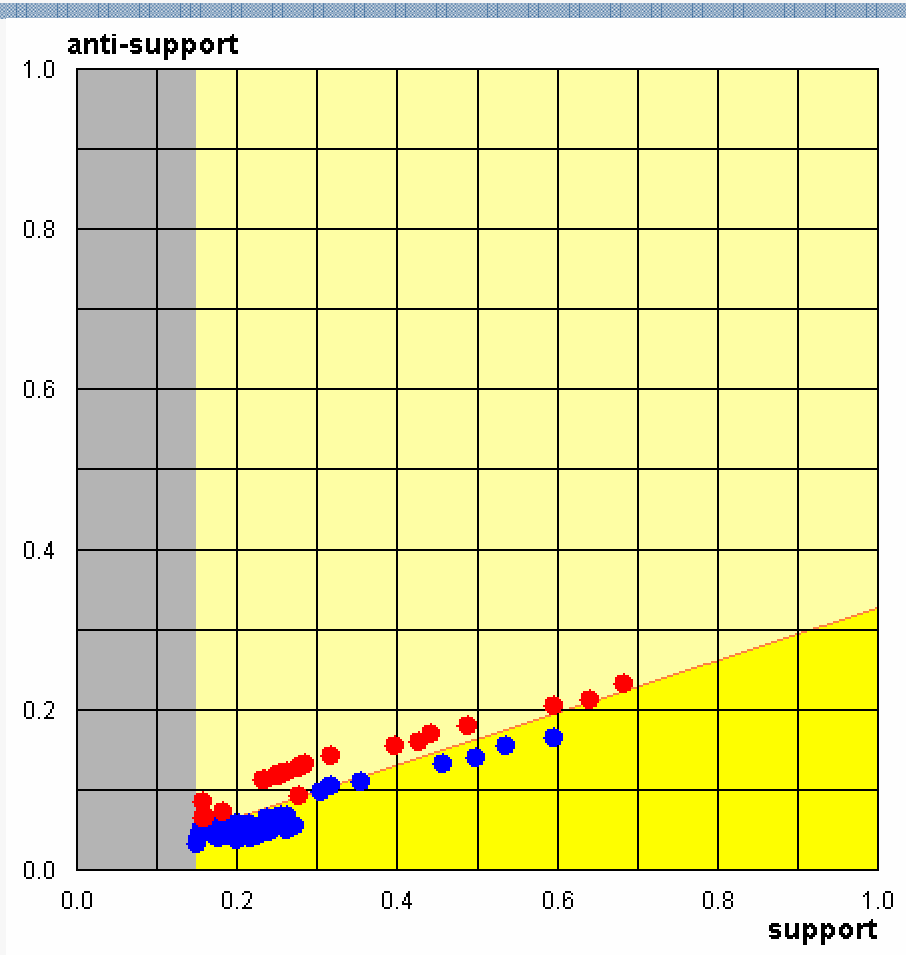
  For those rules, the premise only disconfirms the conclusion!

Confirmation perspective on support - anti-support border



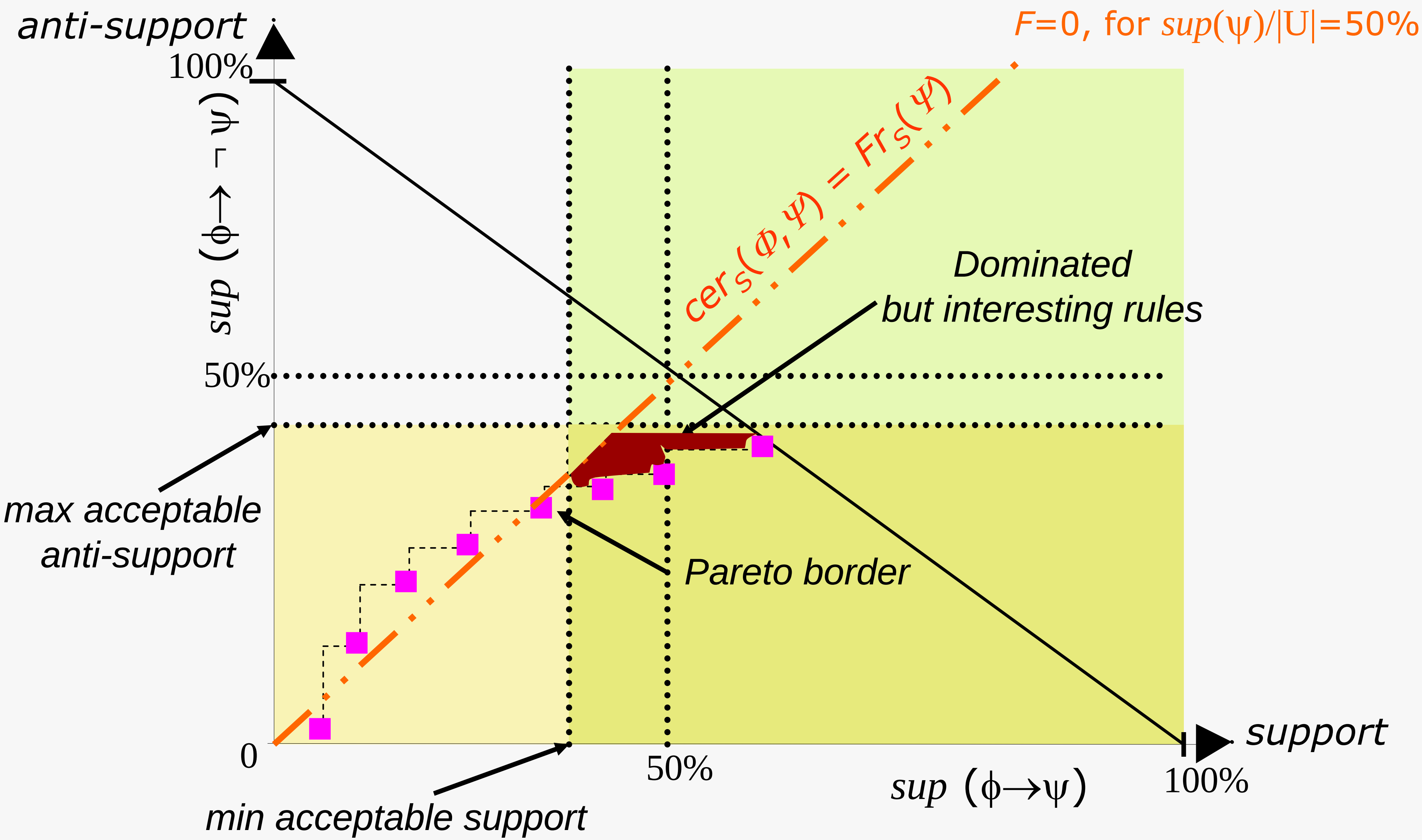
For rules lying above the curve for which c=0
the premise only disconfirms the conclusion

Support - anti-support (workclass=Private)



- ● indicates rules with negative confirmation
- even some rules from the Pareto border need to be discarded

Confirmation perspective on support-anti-support border



Decision rules – efficiency of intervention

- Intervention is a three-stage process:
  (Greco, Matarazzo, Pappalardo, Słowiński 2005)
  - mining rules in universe *U*
  - modification (manipulation) of universe *U*′, based on a rule mined in *U*, with the aim of getting a desired result
  - transition from universe *U*′ to universe *U*′′ due to the modification

S.Greco, B.Matarazzo, N.Pappalardo, R.Słowiński: Measuring expected effects of interventions based on decision rules. *J. of Experimental and Applied Artificial Intelligence*, 17 (2005) no. 1-2, 103-118

Efficiency of intervention – the playground of three universes

- For example, suppose the following rule mined from U:

  $r \equiv$ *if absence of symptom Φ, then no disease Ψ* with 90% certainty

  (i.e. in 90% of cases where symptom Φ is absent there is no disease Ψ)

- On the basis of r, intervention T in U' can be taken:

  $T \equiv$ *eliminate symptom Φ to get out from disease Ψ in U''*

- T is based on a hypothesis of homogeneity of universes U and U'

- Homogeneity means that r is also valid in U':

  one can expect that 90% of sick patients with symptom Φ

  will get out from the disease due to the intervention T

- $S = (U, A)$, $S' = (U', A)$: two data tables referring to universes U, U'

Decision rules – efficiency of intervention

- If we modify property $\neg\Phi$ to property Φ in set $\|\neg\Phi\wedge\neg\Psi\|_{S'}$ we may reasonably expect that:

  $$cer_S(\Phi,\Psi) \times supp_{S'}(\neg\Phi,\neg\Psi)$$

  objects from set $\|\neg\Phi\wedge\neg\Psi\|_{S'}$ in U' will enter decision class Ψ in U''

- Expected *relative increment* of objects from U' entering decision class Ψ in universe U'':

  $$incr_{SS'}(\Psi) = cer_S(\Phi,\Psi) \times cer_{S'}(\neg\Psi,\neg\Phi) \times \frac{card(\|\neg\Psi\|_{S'})}{card(U')}$$

  where $cer_{S'}(\neg\Psi,\neg\Phi)$ is a certainty factor of the *contrapositive rule* $r^{cp} \equiv \neg\Psi\rightarrow\neg\Phi$ in U'

- *Efficiency of the intervention*:

  $$eff_{SS'}(\Phi,\Psi) = cer_S(\Phi,\Psi) \times cer_{S'}(\neg\Psi,\neg\Phi)$$

Efficiency of intervention – multi-attribute intervention

- If condition formula Φ is composed of *n elementary conditions*, we consider rule $r \equiv \Phi_1 \wedge \Phi_2 \wedge \ldots \wedge \Phi_n \rightarrow \Psi$, with $cer_S(\Phi, \Psi)$

- *Relative increment*, for $P \subseteq N = \{1,\ldots,n\}$:

$$incr_{SS'}(\Psi) = \sum_{\emptyset \subset P \subseteq N} \left[cer_S(\Phi, \Psi) \times cer_{S'}\left(\neg\Psi, \bigwedge_{i \in P} \neg\Phi_i \wedge \bigwedge_{j \notin P} \Phi_j \right) \right] \times \frac{card\left(\left\| \neg\Psi \right\|_{S'}\right)}{card(U')}$$

where $cer_{S'}\left(\neg\Psi, \bigwedge_{i \in P} \neg\Phi_i \wedge \bigwedge_{j \notin P} \Phi_j \right)$ is a certainty factor

of the *contrapositive rule* $r_P^{cp} \equiv \neg\Psi \rightarrow \bigwedge_{i \in P} \neg\Phi_i \wedge \bigwedge_{j \notin P} \Phi_j$ in U'

- *Efficiency of the multi-attribute intervention*:

$$eff_{SS'}(\Phi, \Psi) = cer_S(\Phi, \Psi) \times \sum_{\emptyset \subset P \subseteq N} cer_{S'}\left(\neg\Psi, \bigwedge_{i \in P} \neg\Phi_i \wedge \bigwedge_{j \notin P} \Phi_j \right)$$

Intervention based on „at least” and „at most” rules

- Interpretation of the intervention based on „at least” and „at most” rules obtained from the Dominance-based Rough Set Approach

  - „at least” rules

    if $x_{q1} \succeq_{q1} r_{q1}$ *and* $x_{q2} \succeq_{q2} r_{q2}$ *and* … $x_{qp} \succeq_{qp} r_{qp}$, *then* $x \in Class_t^{\geq}$

    where for $w_q, z_q \in X_q$, „$w_q \succeq_q z_q$” means „w_q is at least as good as z_q” and $x \in Class_t^{\geq}$ means „x belongs to class $Class_t$ or better”

  - „at most” rules

    if $x_{q1} \preceq_{q1} r_{q1}$ *and* $x_{q2} \preceq_{q2} r_{q2}$ *and* … $x_{qp} \preceq_{qp} r_{qp}$, *then* $x \in Class_t^{\leq}$

    where for $w_q, z_q \in X_q$, „$w_q \preceq_q z_q$” means „w_q is at most as good as z_q” and $x \in Class_t^{\leq}$ means „x belongs to class $Class_t$ or worse”

Intervention based on „at least” and „at most” rules

- The „at least” rules indicate **opportunities for improving** the assignment of object x to $Class_t$ or better, if it was not assigned as high, and its score on $q_1,\ldots,q_p$ would grow to $r_{q1},\ldots,r_{qp}$

- The „at most” rules indicate **threats for deteriorating** the assignment of object x to $Class_t$ or worse, if it was not assigned as low, and its score on $q_1,\ldots,q_p$ would drop to $r_{q1},\ldots,r_{qp}$

Intervention based on „at least” and „at most” rules - example

- Example: customer satisfaction analysis by a *Company*
- 19 questions and 3 classes of overall satisfaction: *High*, *Medium*, *Low*

Threats of deterioration of satisfaction	Opportunities for improvement of satisfaction
Deterioration from ***High*** or ***Medium*** to ***Low*** satisfaction	Improvement from ***Low*** to ***Medium*** or ***High*** satisfaction
Deterioration from ***High*** to ***Medium*** or ***Low*** satisfaction	Improvement from ***Low*** or ***Medium*** to ***High*** satisfaction

Intervention based on monotonic rules - example

At least rule:

If (A3 ≥ 4) & (C3 ≥ 3), then Satisfaction $\succeq$ *Medium*

$$incr_{SS'}(Medium) = 77\%$$

Opportunity: if

- invoicing is at least *mostly accurate and errors are rare*, and
- Company is involved in at least *some advertising / promotions*,

then satisfaction of 77% of customers with Satisfaction = *Low* will improve to *Medium* or *High*

Intervention based on monotonic rules - example

At most rule:

If (A2 ≤ 3) & (E4 ≤ 4) , then Satisfaction $\preceq$ *Low*

$$incr_{SS'}(Low) = 89\%$$

Threat: if

- products are *not in good condition*, and
- Company is *not always the first to come out with technologically advanced products and better solutions*,

then satisfaction of 89% of customers with Satisfaction = *High* or *Medium* will deteriorate to *Low*

Intervention based on monotonic rules

- In practice, the choice of rules used for intervention can be supported by additional measures, like:
  - length of the rule – the shorter the better,
  - cost of intervention on attributes present in the rule,
  - priority of intervention on some types of attributes, like: short-term before long-term actions

Examples of Application

DRSA – example of technical diagnostics

- 176 vehicles (objects)
- 8 symptoms
- decision = technical state:
  - **3** – good state (in use)
  - **2** – minor repair
  - **1** – major repair (out of use)
- there is a monotonic relationship between each symptom and the decision
- inconsistent objects:

  11, 12, 39

Examples:

	MaxSpeed	ComprPressure	Blacking	Torque	SummerCons	WinterCons	OilCons	HorsePower	State
1.	90	2	38	481	21	26	0	145	3
2.	76	2	70	420	22	25	2	110	1
3.	63	1	82	400	22	24	3	101	1
4.	90	2	49	477	21	25	1	138	3
5.	85	2	52	460	21	25	1	130	2
6.	72	2	73	425	23	27	2	112	1
7.	88	2	50	480	21	24	1	140	3
8.	87	2	56	465	22	27	1	135	3
9.	90	2	16	486	26	27	0	150	3
10.	60	1	95	400	23	24	4	96	1
11.	80	2	60	451	21	26	1	125	1
12.	78	2	63	448	21	26	1	120	2
13.	90	2	26	482	22	24	0	148	3
14.	62	1	93	400	22	28	3	100	1
15.	82	2	54	461	22	26	1	132	2
16.	65	2	67	402	22	23	2	103	1
17.	90	2	51	468	22	26	1	138	3
18.	90	2	15	488	20	23	0	150	3
19.	76	2	65	428	27	33	2	116	1
20.	85	2	50	454	21	26	1	129	2
21.	85	2	58	450	22	25	1	126	2
22.	88	2	48	458	22	25	1	130	3
23.	60	1	90	400	24	28	4	95	1
24.	64	2	71	420	23	25	2	105	1
25.	75	2	64	432	22	25	1	114	2
26.	74	2	64	420	21	25	1	110	2
27.	68	2	70	400	22	26	2	100	1

Attributes: 9 of 10 | Examples: 76 | Decision: State | Missing Values: No

Unions of Classes

Quality of Approximation: 0,960000

Unions of Classes:

Union Name	Examples	Accur...	Card. ...	Lower...	Upper...
At most 1		0,8900	26	25	28
Lower:	2, 3, 6, 10, 14, 16, 19, 23, 24, 27, 28, 38, 40,...				
Upper:	2, 3, 6, 10, 11, 12, 14, 16, 19, 23, 24, 27, 28,...				
Boundary:	11, 12, 39				
At most 2		1,0000	42	42	42
Lower:	2, 3, 5, 6, 10, 11, 12, 14, 15, 16, 19, 20, 21, ...				
Upper:	2, 3, 5, 6, 10, 11, 12, 14, 15, 16, 19, 20, 21, ...				
Boundary:					
At least 2		0,9400	50	48	51
Lower:	1, 4, 5, 7, 8, 9, 13, 15, 17, 18, 20, 21, 22, 25,...				
Upper:	1, 4, 5, 7, 8, 9, 11, 12, 13, 15, 17, 18, 20, 21,...				
Boundary:	11, 12, 39				
At least 3		1,0000	34	34	34
Lower:	1, 4, 7, 8, 9, 13, 17, 18, 22, 29, 32, 33, 35, 3...				
Upper:	1, 4, 7, 8, 9, 13, 17, 18, 22, 29, 32, 33, 35, 3...				
Boundary:					

Reducts:

	Cardinality	Attributes Set
Core:	2	MaxSpeed, WinterCons
Reducts:		
1.	4	MaxSpeed, Blacking, Torque, WinterCons
2.	4	MaxSpeed, Blacking, SummerCons, WinterCons
3.	4	MaxSpeed, Blacking, WinterCons, HorsePower
4.	4	MaxSpeed, Torque, WinterCons, OilCons
5.	4	MaxSpeed, WinterCons, OilCons, HorsePower

Generated Rules: 10 **Displayed Rules: 10**

Number	Condition	Decision	Support	Relative Strength [%]
1.	(OilCons >= 2) & (HorsePower <= 119)	State at most 1	25	100,00
2.	(HorsePower <= 122)	State at most 2	35	83,33
3.	(MaxSpeed <= 85) & (WinterCons >= 25)	State at most 2	38	90,48
4.	(MaxSpeed >= 86) & (HorsePower >= 125)	State at least 3	33	97,06
5.	(WinterCons <= 24) & (HorsePower >= 123)	State at least 3	14	41,18
6.	(Blacking <= 54)	State at least 2	32	66,67
7.	(OilCons <= 1) & (WinterCons <= 25)	State at least 2	37	77,08
8.	(MaxSpeed >= 83) & (HorsePower >= 120)	State at least 2	44	91,67
9.	(WinterCons >= 26) & (SummerCons <= 21) & (MaxSpeed <= 80)	State = 1 OR 2	2	66,67
10.	(MaxSpeed <= 75) & (HorsePower >= 120)	State = 1 OR 2	1	33,33

Supporting Examples:

	MaxSpeed	ComprPressure	Blacking	Torque	SummerCons	WinterCons	OilCons	HorsePower	State
7.	88	2	50	480	21	24 *	1	140	3
13.	90	2	26	482	22	24 *	0	148	3
18.	90	2	15	488	20	23	0	150	3
29.	90	2	18	480	20	23	0	146	3
42.	86	2	52	462	22	24 *	1	129	3
44.	88	2	48	475	22	24 *	1	140	3
49.	90	2	38	482	20	24 *	0	146	3
55.	90	2	47	481	22	24 *	1	145	3
56. *	85	2	60	446	21	24 *	1	123 *	3
57.	88	2	50	465	21	24 *	1	137	3

Classification results for file: Buses.isf

Examples:

	Possible Decision	No. of matching rules	MaxSpeed	Compr	Blacking	Torque	SummerC	WinterC	OilCons	HorsePower
14.	1	3	62	1	93	400	22	28	3	100
15.	2	2	82	2	54	461	22	26	1	132
16.	1	2	65	2	67	402	22	23	2	103
17.	3	3	90	2	51	468	22	26	1	138
18.	3	5	90	2	15	488	20	23	0	150
19.	1	3	76	2	65	428	27	33	2	116
20.	2	3	85	2	50	454	21	26	1	129
21.	2	3	85	2	58	450	22	25	1	126
22.	3	4	88	2	48	458	22	25	1	130
23.	1	3	60	1	90	400	24	28	4	95
24.	1	3	64	2	71	420	23	25	2	105
25.	2	3	75	2	64	432	22	25	1	114

Used Rules:

Number	Condition	Decision	Support	Relative Strength [%]
3.	(MaxSpeed <= 85) & (WinterCons >= 25)	State at most 2	38	90,48
6.	(Blacking <= 54)	State at least 2	32	66,67

*M*obile *E*mergency *T*riage System - *MET* System

- MET – Mobile Emergency Triage
  - Facilitates triage disposition for presentations of acute pain (abdominal and scrotal pain, hip pain)
  - Supports triage decision with or without complete clinical information
  - Provides mobile support through handheld devices
  - http://www.mobiledss.uottawa.ca

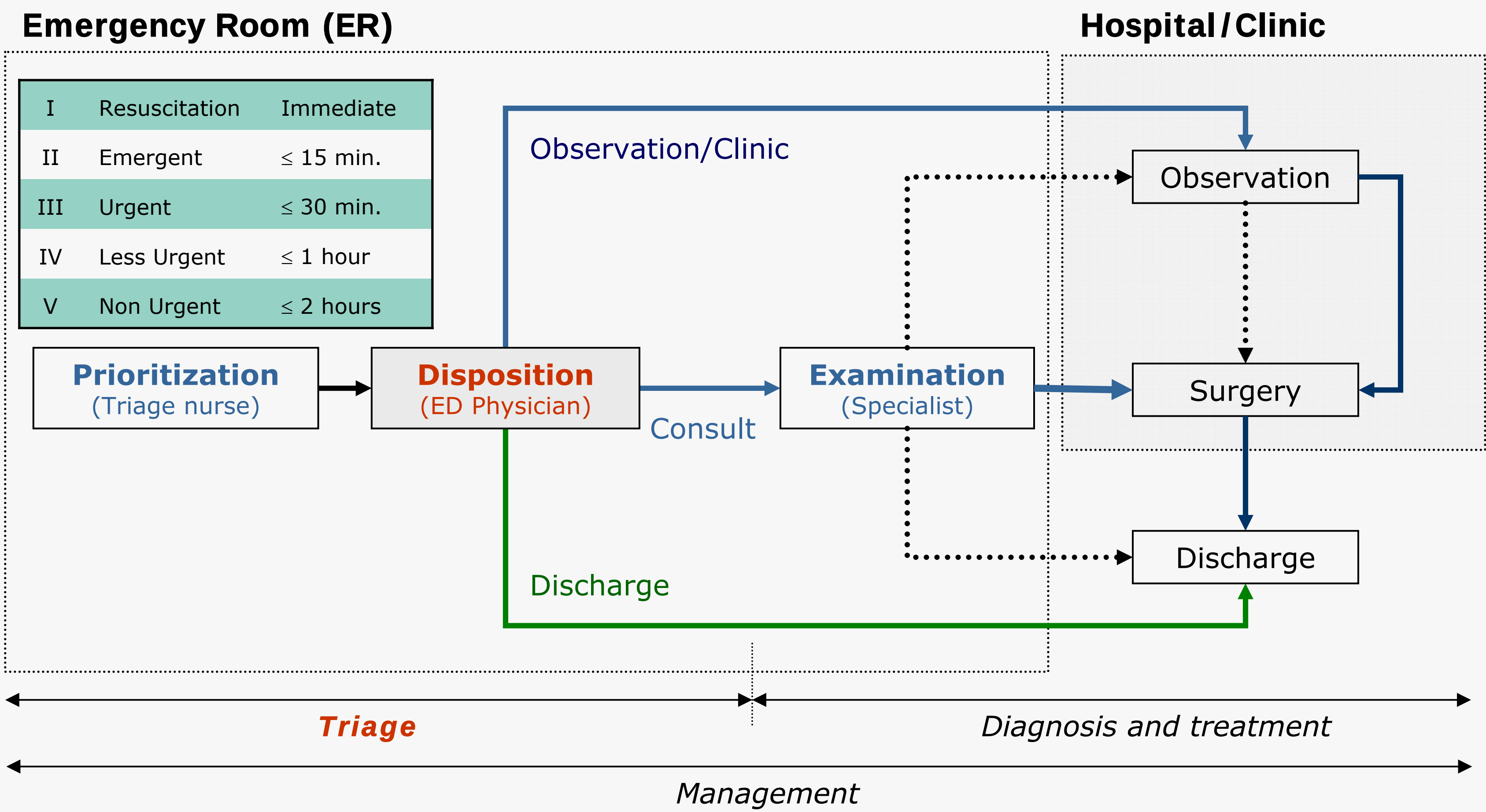
W. Michalowski

University of Ottawa

R. Słowiński, Sz. Wilk

Poznań University of Technology

Triage Process

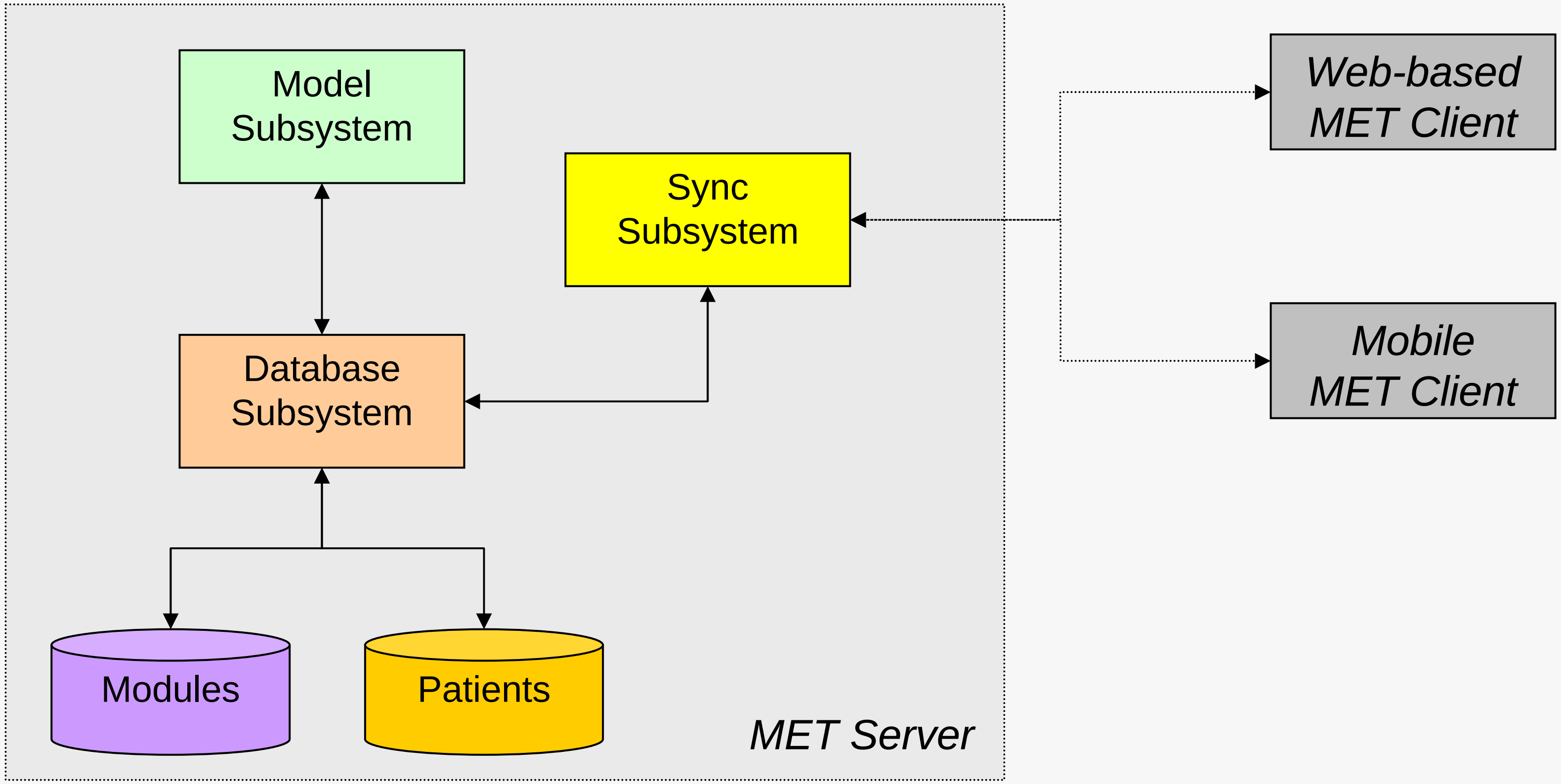


I	Resuscitation	Immediate
II	Emergent	≤ 15 min.
III	Urgent	≤ 30 min.
IV	Less Urgent	≤ 1 hour
V	Non Urgent	≤ 2 hours

Clinical Attributes

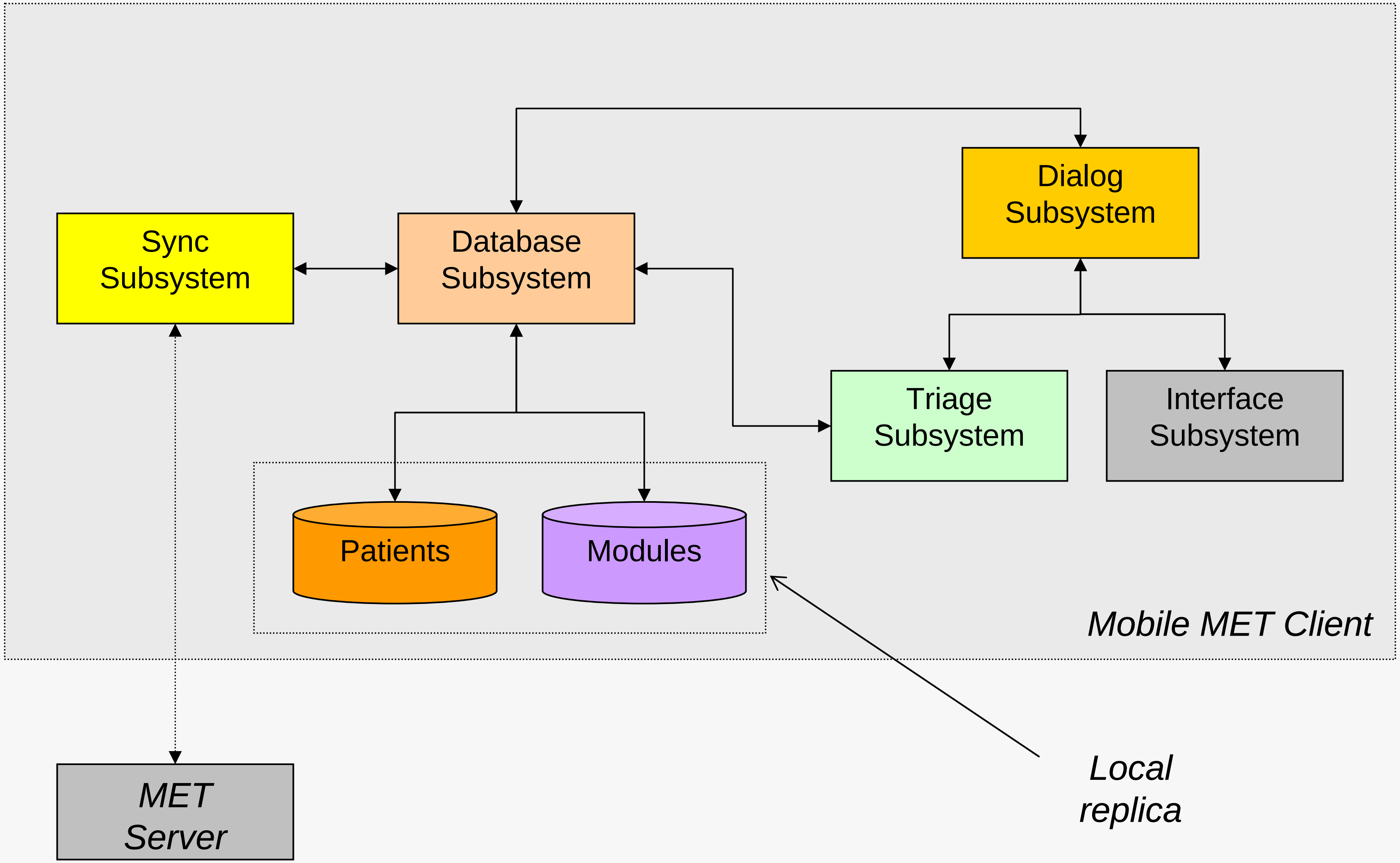
Age	Age	numeric – discretized: ≤ 5 years; > 5 years
Guard	Muscle guarding	yes, no
PainDur	Duration of pain	numeric – discretized: ≤ 24h, > 24h and ≤7 days, > 7 days
PainShift	Shifting of pain	yes, no
PainSite	Site of pain	RLQ, lower_abdomenomen, other
PainType	Type of pain	constant, intermittent
PrevVisit	Previous visit to ER	yes, no
RebTend	Rebound tenderness	yes, no
Sex	Sex	male, female
Tempr	Temperature	numeric – discretized: < 37°C, ≥ 37°C and ≤39°C, > 39°C
TendSite	Site of tenderness	RLQ, lower_abdomenomen, other
Vomiting	Vomiting	yes, no
WBC	White blood cells	numeric – discretized: ≤ 4, > 4 and < 12, ≥ 12

MET Server



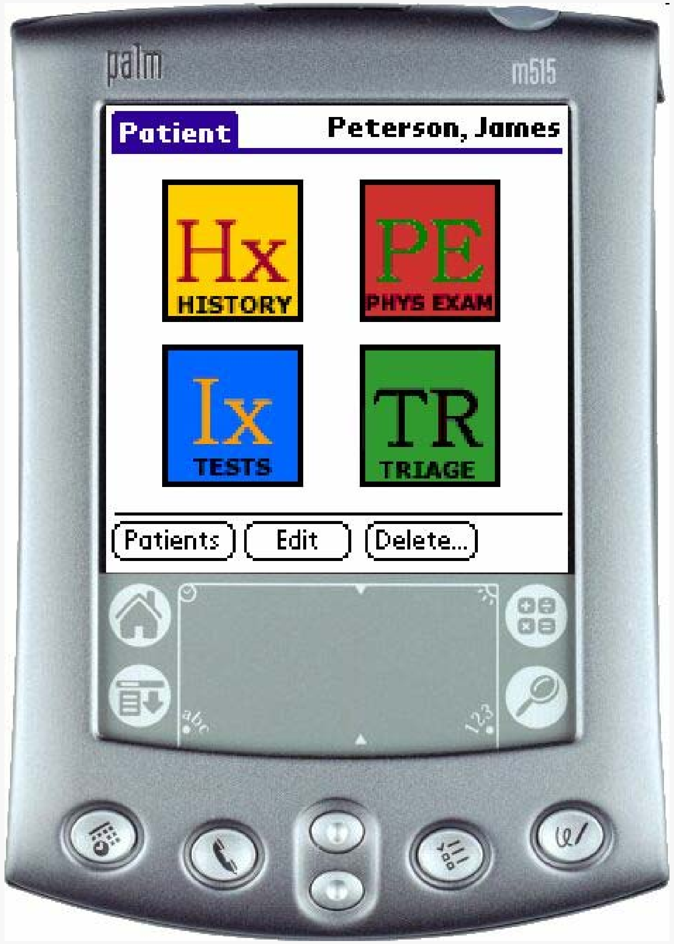
Model Subsystem
Sync Subsystem
Database Subsystem
Modules
Patients
MET Server
Web-based MET Client
Mobile MET Client

Mobile *MET* Client



MET interactions

Navigation between screens/activities

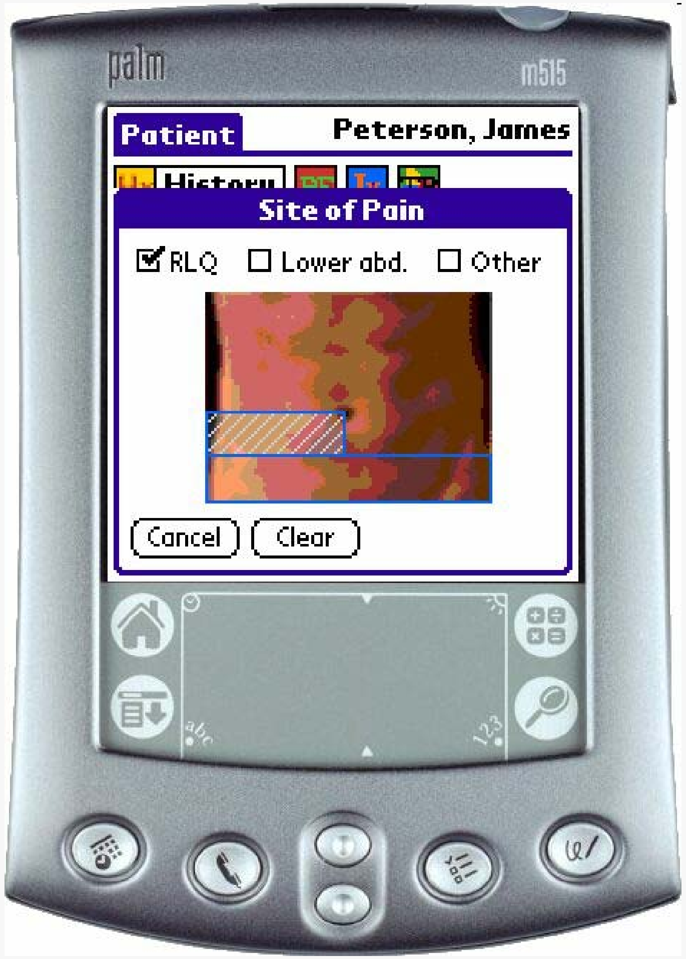


Using icon-based models

MET interactions

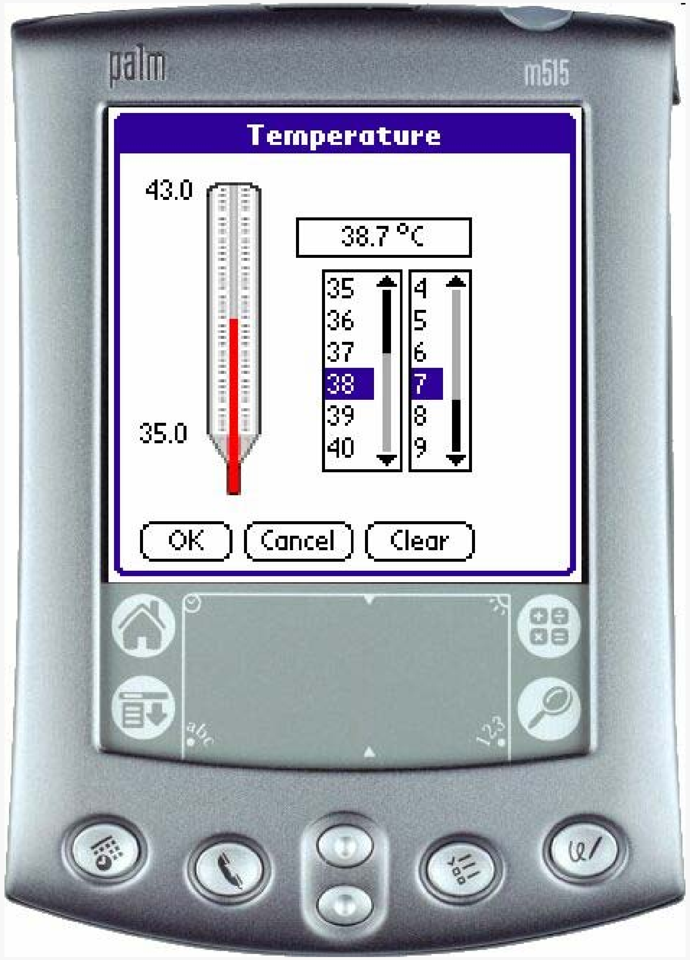
Inputting data

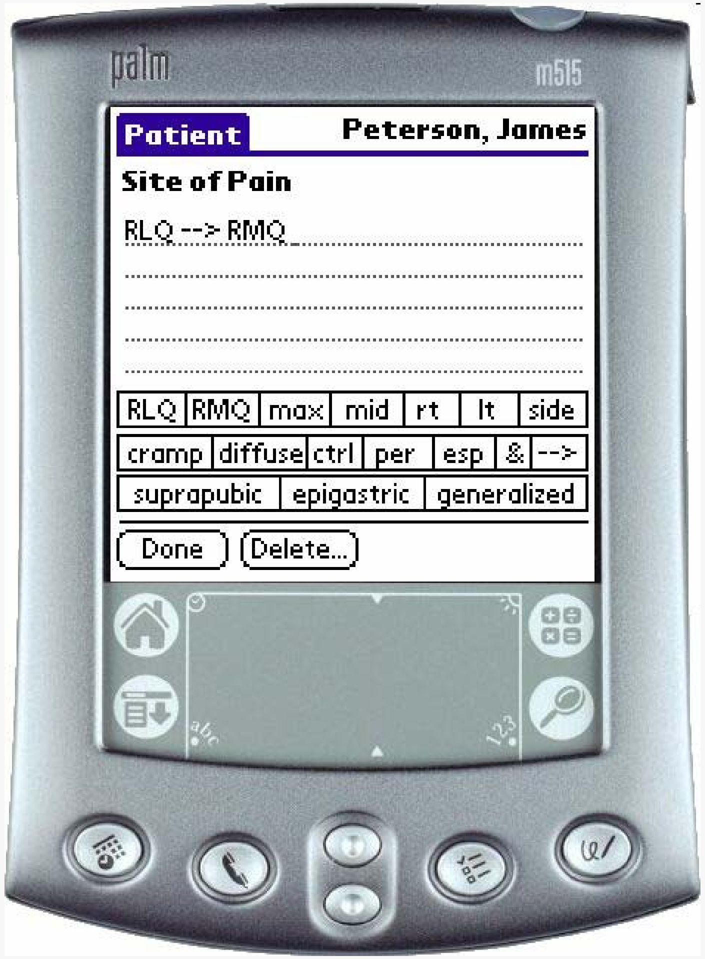
Using checkboxes



Using pictograms

MET interactions





Entering numerical values

Writing comments

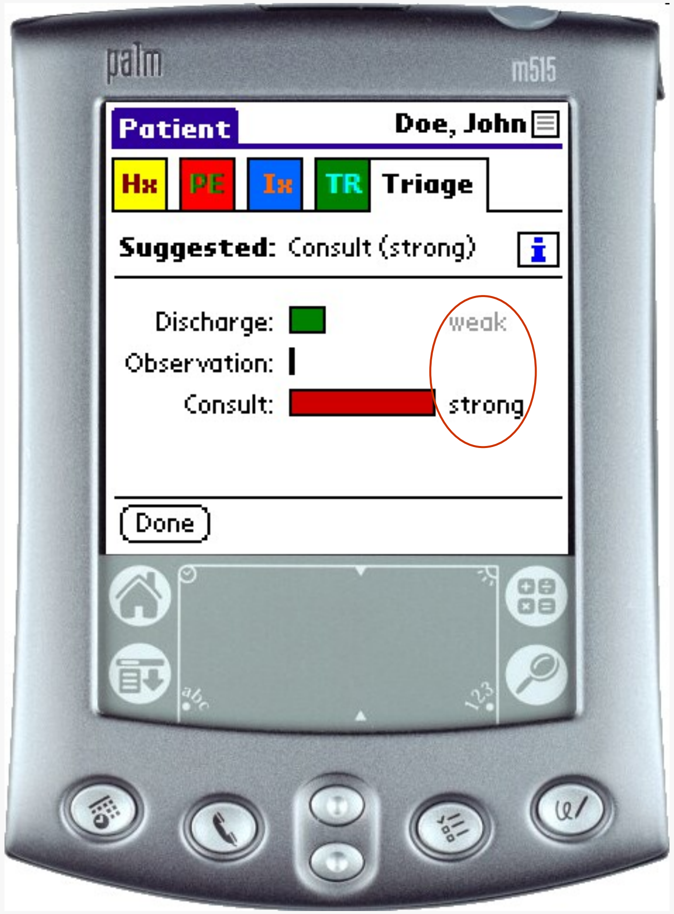
Decision Rules

- **if** (Age < 5 years) **and** (PainSite = lower_abdomen) **and** (RebTend = yes) **and** (4 < WBC < 12) **then** (Triage = discharge)

- **if** (PainDur > 7 days) **and** (PainSite = lower_abdomen) **and** (37 ≤ Tempr ≤ 39) **and** (TendSite = lower_abdomen) **then** (Triage = observation)

- **if** (Sex = male) **and** (PainSite = lower_abdomen) **and** (PainType = constant) **and** (RebTend = yes) **and** (WBCC ≥ 12) **then** (Triage = consult)

MET System - suggesting triage disposition



- Strength factors are presented instead of a definite and univocal answer (debiaser, not oracle)
- Strength factors are established with decision rules

Trial Location

CHEO Children's Hospital of Eastern Ontario
Centre hospitalier pour enfants de l'est de l'Ontario

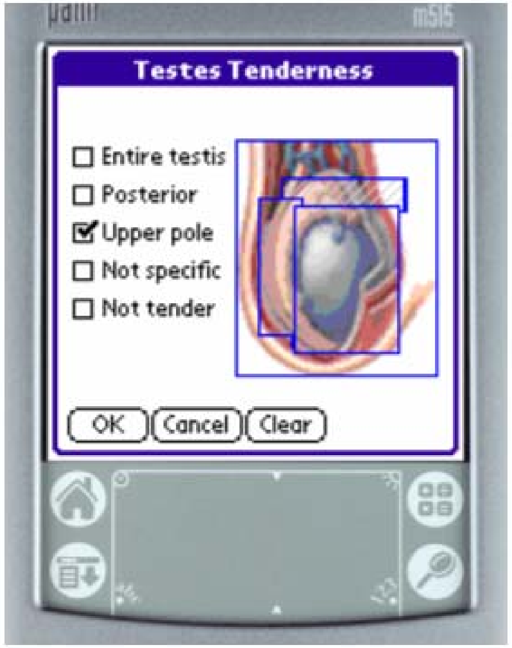
- Total pediatric population >400,000
- 55,000 patient visits in the ER per year
- 3 pediatric general surgeons (supported by emergency physicians and residents)

Trial Results

Accuracy of disposition for ED physicians and MET

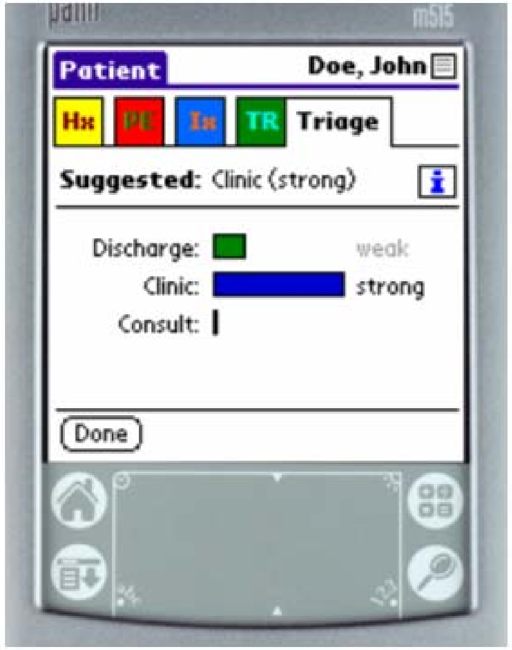
	Overall	Discharge	Observation	Consult
Physicians	64.6%	64.8%	63.0%	65.2%
MET	66.3%	75.8%	18.5%	69.6%

MET System – scrotal pain triage



List with a pictogram

Numeric keypad



Triage recommendations

DRSA to Multicriteria Choice and Ranking

DRSA to multicriteria choice & ranking

- Preference information is given by the DM as a set $B \subseteq A^R \times A^R$ of pairwise comparisons of reference actions
- The preference model is a set of decision rules induced from rough approximations of the holistic preference relation, e.g. S and S^c

$B \subseteq A^R \times A^R$

Pairwise Comparison Table (PCT)

Pairs of ref. actions	Evaluation on criteria				Preference relation
	q_1	q_2	...	q_m	
(x,y)	$q_1(x),q_1(y)$	$q_2(x),q_2(y)$	...	$q_m(x),q_m(y)$	xSy
(y,x)	$q_1(y),q_1(x)$	$q_2(y),q_2(x)$	...	$q_m(y),q_m(x)$	yS^cx
(y,u)	$q_1(y),q_1(u)$	$q_2(y),q_2(u)$	...	$q_m(y),q_m(u)$	ySu
...	...	...	...	...	...
(v,z)	$q_1(v),q_1(z)$	$q_2(v),q_2(z)$	...	$q_m(v),q_m(z)$	vS^cz

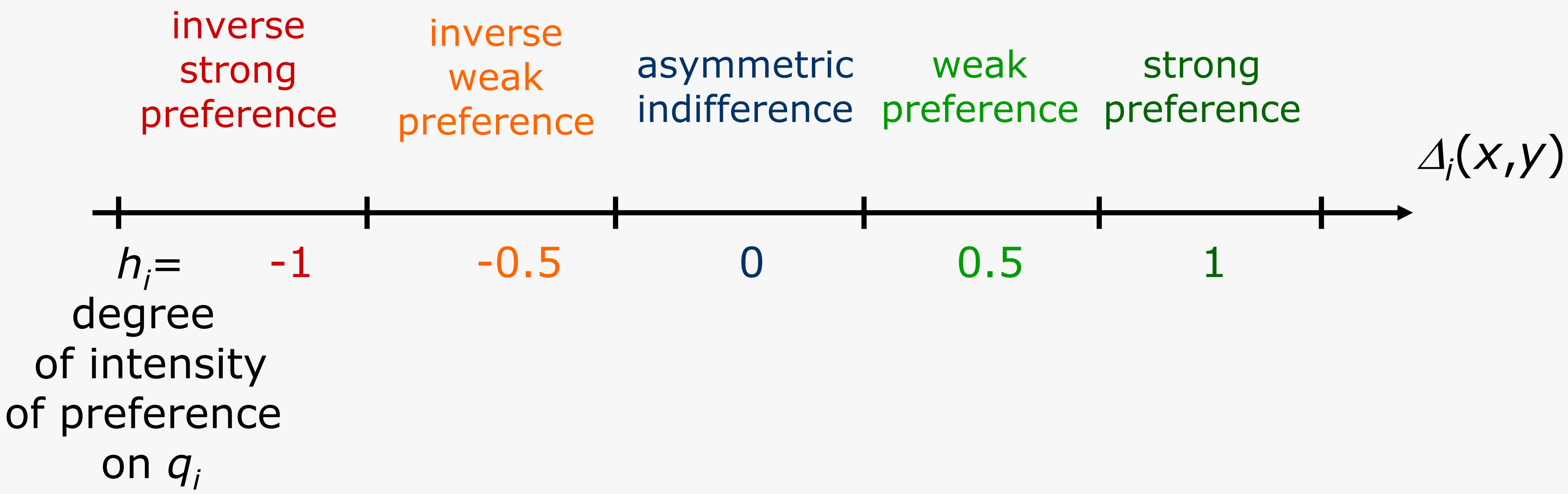
S – ***outranking***

S^c – ***non-outranking***

$F = \{q_1, q_2, \ldots, q_m\}$

Pairwise Comparison Table (PCT)

- If q_i is a cardinal criterion, the pair of evaluations $[q_i(x); q_i(y)]$, is replaced by the difference $\Delta_i(x,y)=q_i(x)-q_i(y)$, which may be translated to a degree of intensity of preference of x over y, e.g.:



- If q_i is an ordinal criterion, one keeps in PCT the pair of evaluations: $[q_i(x); q_i(y)]$, e.g. [*Medium*; *Basic*]

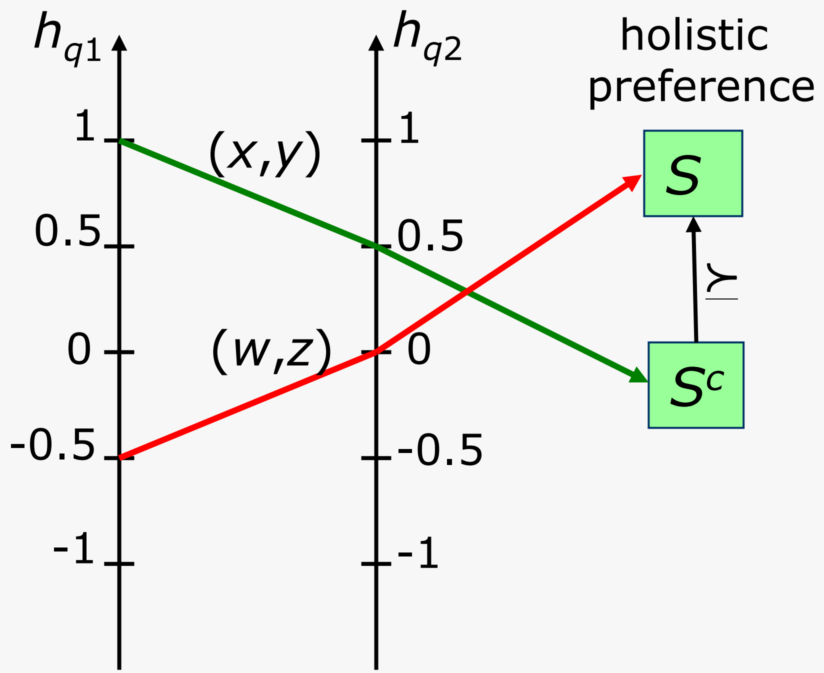
DRSA to multicriteria choice & ranking

- Problem → inconsistencies in the preference information, due to:
  - uncertainty of information – hesitation, unstable preferences,
  - incompleteness of the family of criteria,
  - granularity of information
- Inconsistency w.r.t. dominance principle:

$x\ S^c\ y$

while

$w\ S\ z$



DRSA to multicriteria choice & ranking

- Dominance relation for pairs of actions $(x,y),(w,z) \in A \times A$

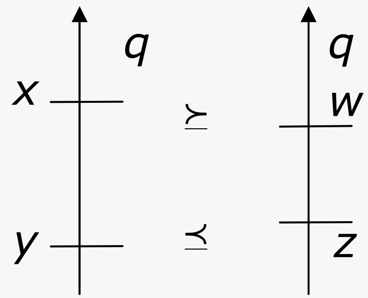
For cardinal criterion $q \in C$:

$\boxed{(x,y)D_q(w,z)}$

if $xP_q^{h_q}y$ and $wP_q^{k_q}z$

where $h_q \geq k_q$

For ordinal criterion $q \in C$:



$\boxed{(x,y)D_q(w,z)}$

- For subset $P \subseteq C$ of criteria: P-dominance relation on pairs of actions:

$(x,y)D_P(w,z)$ if $(x,y)D_q(w,z)$ for all $q \in P$, i.e.,

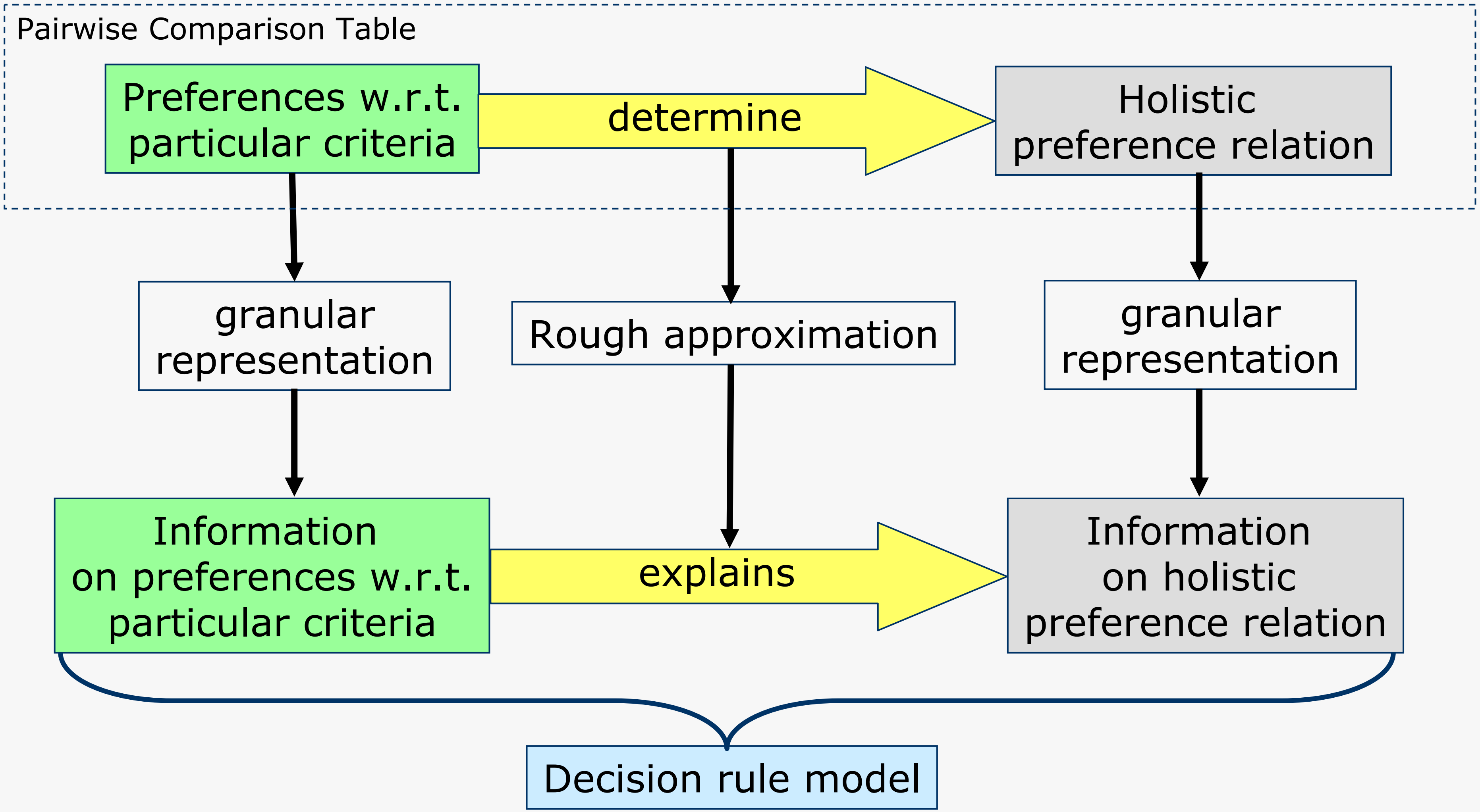
if x is preferred to y at least as much as w is preferred to z for all $q \in P$

- D_q is reflexive, transitive, but not necessarily complete (partial preorder)

$D_P = \bigcap_{q \in P} D_q$ is a partial preorder on $A \times A$

DRSA to multicriteria choice & ranking

- Basic idea of rough approximation applied to MCDA:



DRSA to multicriteria choice & ranking

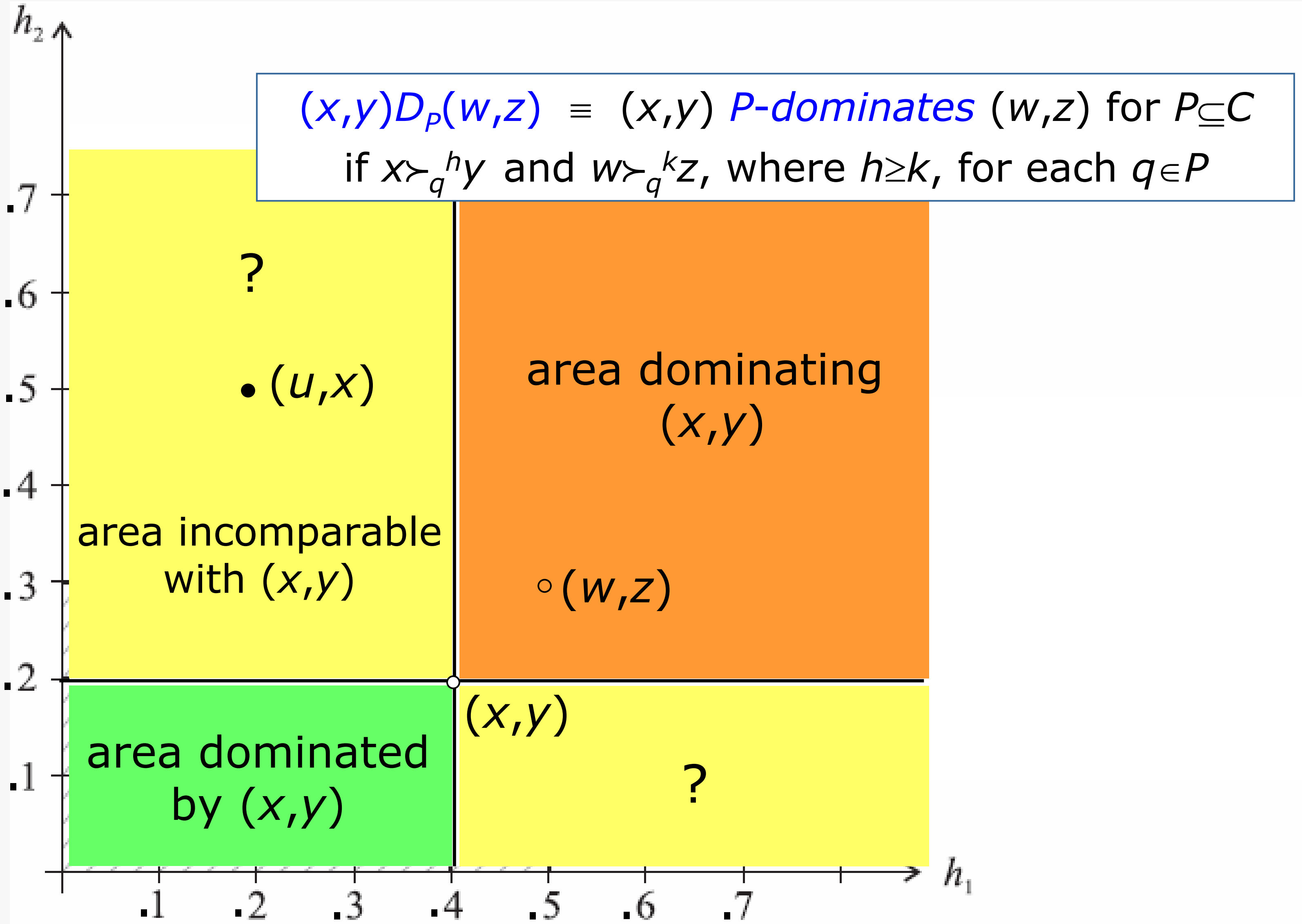
- Let $B \subseteq A^R \times A^R$ be a set of pairs of reference objects in a given PCT
- Granules of knowledge relative to preferences on particular criteria
  - *positive dominance cone*

$$D_P^+(x,y) = \{(w,z) \in B : (w,z) D_P (x,y)\}$$

  - *negative dominance cone*

$$D_P^-(x,y) = \{(w,z) \in B : (x,y) D_P (w,z)\}$$

DRSA – positive and negative dominance cones w.r.t. (x,y)



DRSA to multicriteria choice & ranking– formal definitions

- *P*-lower and *P*-upper approximations of outranking relations *S*:

$$\underline{P}(S)=\left\{(x,y)\in B : D_P^+(x,y)\subseteq S\right\}$$
$$\overline{P}(S)=\bigcup_{(x,y)\in S} D_P^+(x,y)$$

- *P*-lower and *P*-upper approximations of non-outranking relation S^c:

$$\underline{P}\left(S^c\right)=\left\{(x,y)\in B : D_P^-(x,y)\subseteq S^c\right\}$$
$$\overline{P}\left(S^c\right)=\bigcup_{(x,y)\in S^c} D_P^-(x,y)$$

- *P*-boundaries of *S* and S^c:

$$Bn_P(S)=\overline{P}(S)-\underline{P}(S),\qquad Bn_P\left(S^c\right)=\overline{P}\left(S^c\right)-\underline{P}\left(S^c\right)$$
$$Bn_P(S)=Bn_P\left(S^c\right)$$

$P\subseteq C$

DRSA for multiple-criteria choice and ranking – formal definitions

- Basic properties: $$\underline{P}(S) \subseteq S \subseteq \overline{P}(S), \quad \underline{P}(S^c) \subseteq S^c \subseteq \overline{P}(S^c)$$ $$\underline{P}(S) = B - \overline{P}(S^c), \quad \overline{P}(S) = B - \underline{P}(S^c)$$ $$\underline{P}(S^c) = B - \overline{P}(S), \quad \overline{P}(S^c) = B - \underline{P}(S)$$
- Quality of approximation of S and S^c: $\gamma_P = \dfrac{card\left(\underline{P}(S) \cup \underline{P}(S^c)\right)}{card(B)}$
- (S,S^c)-reduct and (S,S^c)-core
- Variable-consistency rough approximations of S and S^c ($l \in (0,1]$) :

$$\underline{P}^l(S) = \left\{(x,y) \in B : \frac{card\left(D_P^+(x,y) \cap S\right)}{card\left(D_P^+(x,y)\right)} \geq l\right\}$$

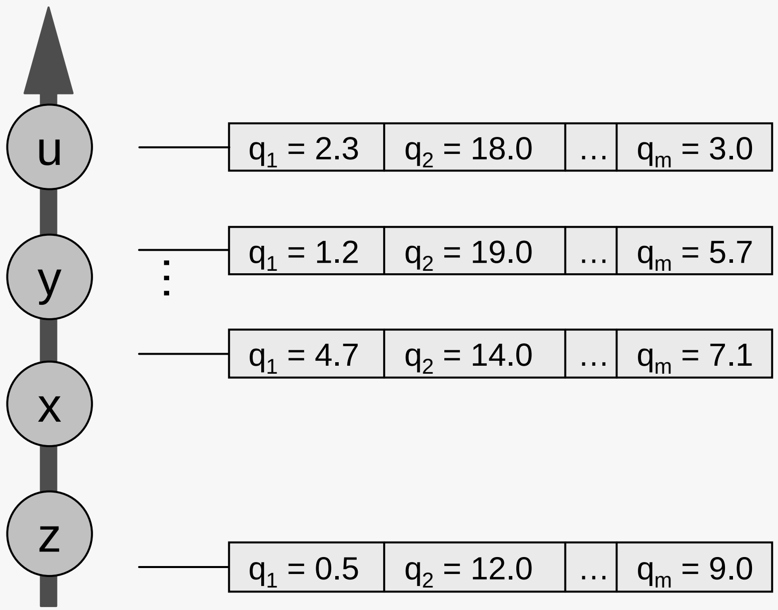
$$\overline{P}^l(S) = B - \underline{P}^l(S)$$

$$\underline{P}^l(S^c) = \left\{(x,y) \in B : \frac{card\left(D_P^-(x,y) \cap S^c\right)}{card\left(D_P^-(x,y)\right)} \geq l\right\}$$

$$\overline{P}^l(S^c) = B - \underline{P}^l(S^c)$$

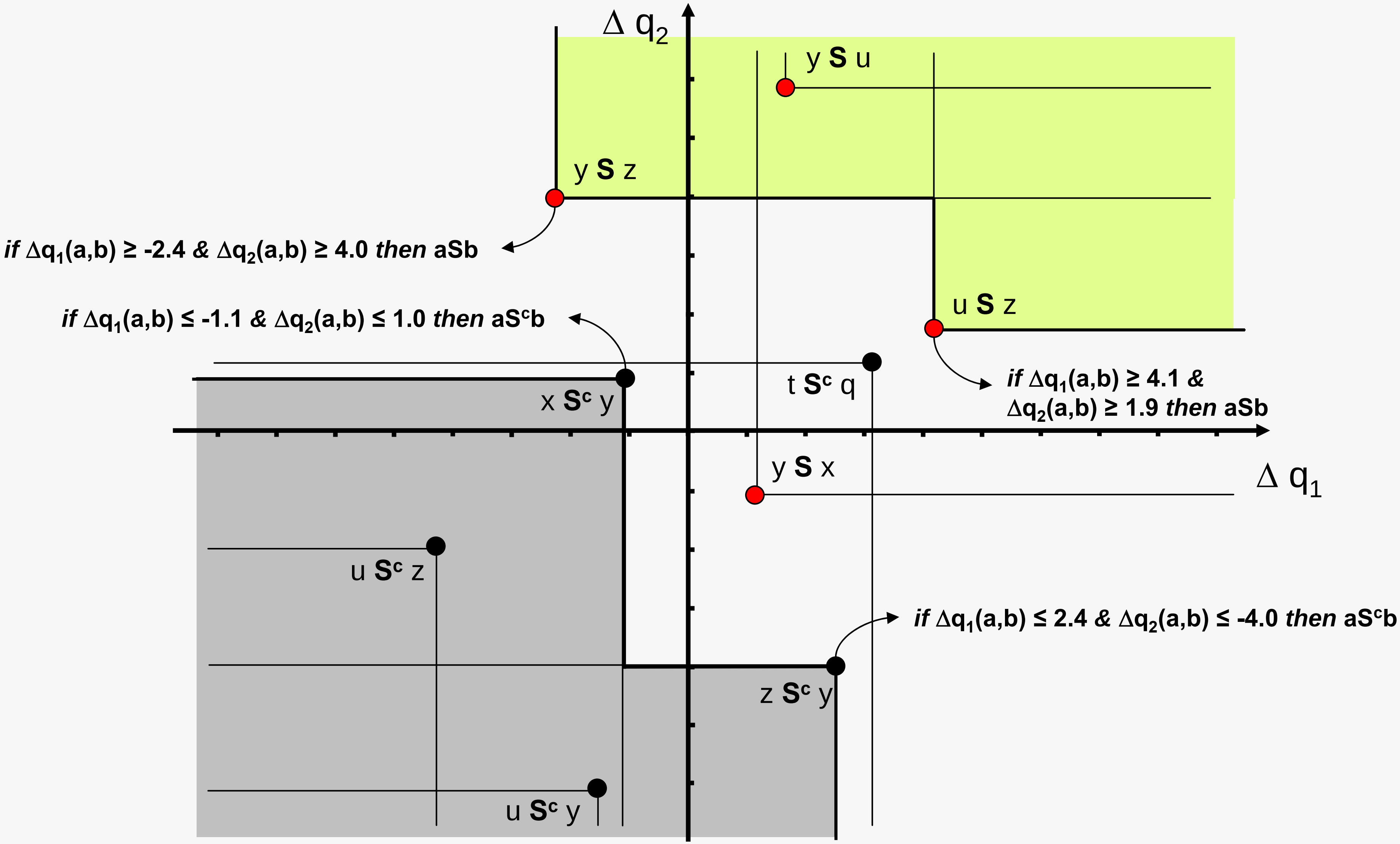
Example of application of DRSA

- Acquiring reference objects
- Preference information on reference objects:
  - making a ranking
  - pairwise comparison of the objects ($x\ S\ y$) or ($x\ S^c\ y$)
- Building the Pairwise Comparison Table (PCT)
- Inducing rules from rough approximations of relations S and S^c



Pairs of objects	Difference of evaluations on each criterion				Preference relation
	$\Delta\ q_1$	$\Delta\ q_2$	…	$\Delta\ q_n$	
(x,y)	Δ_1(x,y) = -1.1	Δ_2(x,y) = 1.0	...	Δ_n(x,y) = 2.7	x S^c y
(y,z)	Δ_1(y,z) = -2.4	Δ_2(y,z) = 4.0	...	Δ_n(y,z) = -4.1	y **S** z
(y,u)	Δ_1(y,u) = 1.8	Δ_2(y,u) = 6.0	...	Δ_n(y,u) = -6.0	y **S** u
...	...	...	...	...	...
(u,z)	Δ_1(u,z) = -4.2	Δ_2(u,z) = -2.0	...	Δ_n(u,z) = 1.9	u S^c z

Induction of decision rules from rough approximations of S and S^c



DRSA to multicriteria choice & ranking – decision rules

- Decision rules (for criteria with cardinal scales)

  - Certain $D_{\geq}$-decision rules (induced from $\underline{P}(S)$)

if $(x \succ_{q1}^{\geq h(q1)} y)$ *and* $(x \succ_{q2}^{\geq h(q2)} y)$ *and* ... $(x \succ_{qp}^{\geq h(qp)} y)$, *then certainly* xSy

  - Possible $D_{\geq}$-decision rules (induced from $\overline{P}(S)$)

if $(x \succ_{q1}^{\geq h(q1)} y)$ *and* $(x \succ_{q2}^{\geq h(q2)} y)$ *and* ... $(x \succ_{qp}^{\geq h(qp)} y)$, *then possibly* xSy

where $\succ_{q}^{\geq h(q)}$ = preference in degree „at least" $h(q)$ on criterion q

e.g. if car x is at least weakly preferred to y w.r.t. maximum speed & strongly preferred w.r.t. price, then x is at least as good as y

DRSA to multicriteria choice & ranking – decision rules

- Decision rules (for criteria with cardinal scales)
  - Certain $D_{\leq}$-decision rules (induced from $\underline{P}(S^c)$)

$$\text{if } (x \succ_{q1}^{\leq h(q1)} y) \text{ and } (x \succ_{q2}^{\leq h(q2)} y) \text{ and } \ldots (x \succ_{qp}^{\leq h(qp)} y), \text{ then certainly } xS^cy$$

  - Possible $D_{\leq}$-decision rules (induced from $\overline{P}(S^c)$)

$$\text{if } (x \succ_{q1}^{\leq h(q1)} y) \text{ and } (x \succ_{q2}^{\leq h(q2)} y) \text{ and } \ldots (x \succ_{qp}^{\leq h(qp)} y), \text{ then possibly } xS^cy$$

  - Approximate $D_{\geq\leq}$-decision rules (induced from $Bn_P(S)=Bn_P(S^c)$)

$$\text{if } (x \succ_{q1}^{\geq h(q1)} y) \text{ and } (x \succ_{q2}^{\geq h(q2)} y) \text{ and } \ldots (x \succ_{qk}^{\geq h(qk)} y)$$

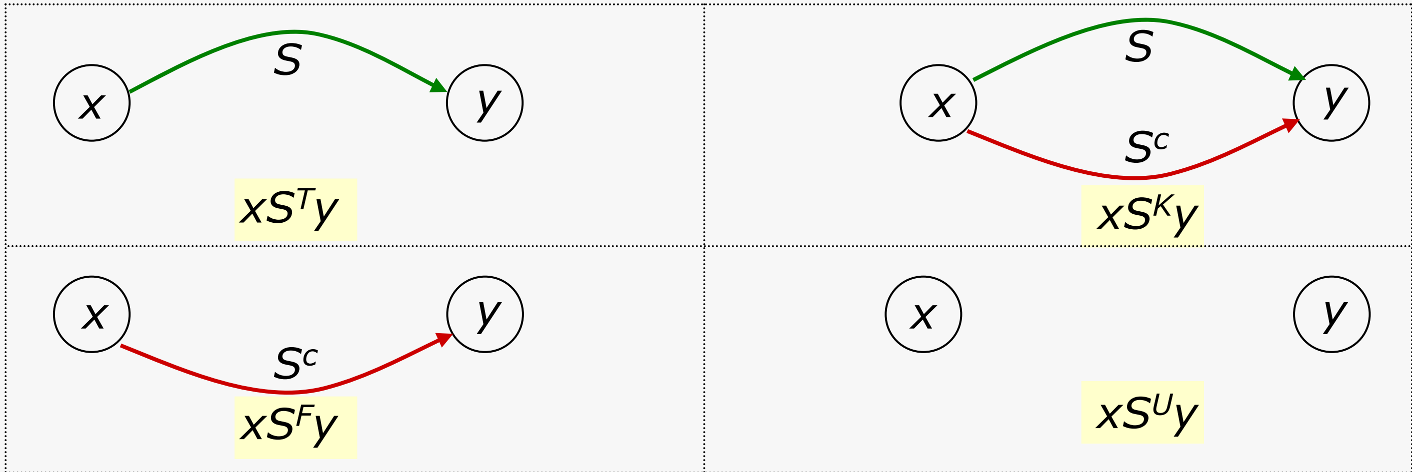
$$\text{and } (x \succ_{q(k+1)}^{\leq h(q(k+1))} y) \text{ and } \ldots (x \succ_{qp}^{\leq h(qp)} y), \text{ then } xSy \text{ or } xS^cy$$

where $\succ_q^{\geq h(q)}$ = preference in degree „at least” $h(q)$ on criterion q

$\succ_q^{\leq h(q)}$ = preference in degree „at most” $h(q)$ on criterion q

Application of decision rules for multicriteria choice & ranking

- Application of decision rules (preference model) on the whole set A induces a specific preference structure on A
- Any pair of objects $(x,y)\in A\times A$ can match the decision rules in one of four ways:
  - xSy and *not* xS^cy, that is *true* outranking (xS^Ty),
  - xS^cy and *not* xSy, that is *false* outranking (xS^Fy),
  - xSy and xS^cy, that is *contradictory* outranking (xS^Ky),
  - *not* xSy and *not* xS^cy, that is *unknown* outranking (xS^Uy).

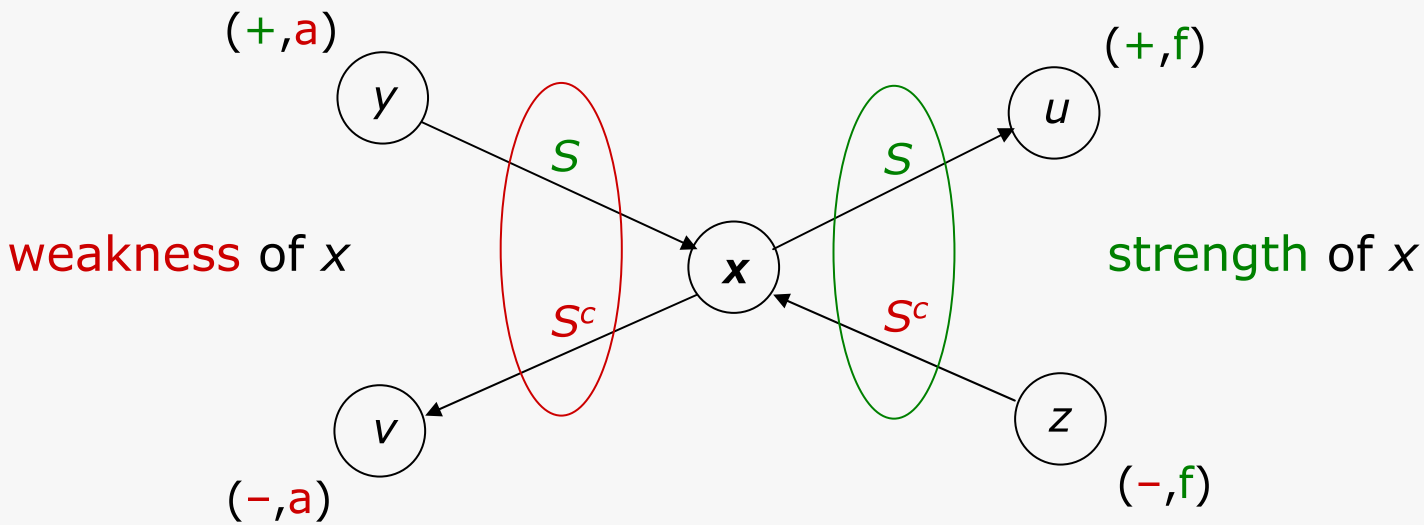


The 4-valued outranking underlines the presence and the absence of positive and negative reasons of outranking

DRSA for multicriteria choice & ranking – *Net Flow Score*

xSy – positive (+) argument in favor of x but against y

xS^cy – negative (–) argument against x *but* in favor of y



$$NFS(x) = \text{strength}(x) - \text{weakness}(x)$$

DRSA for multiple-criteria choice and ranking – final recommendation

- Exploitation of the preference structure by the *Net Flow Score* procedure for each action $x \in A$:

$$NFS(x) = \text{strength}(x) - \text{weakness}(x)$$

- Final recommendation:

ranking: complete preorder determined by $NSF(x)$ in A

best choice: action(s) $x^* \in A$ such that $NSF(x^*) = \max_{x \in A} \{NSF(x)\}$

DRSA for multiple-criteria choice and ranking – example

- Decision table with reference objects (warehouses)

Warehouse	A_1	A_2	A_3	d (ROE %)
1	good	medium	good	10.35
2	good	sufficient	good	4.58
3	medium	medium	good	5.15
4	sufficient	medium	medium	-5
5	sufficient	medium	medium	2.42
6	sufficient	sufficient	good	2.98
7	good	medium	good	15
8	good	sufficient	good	-1.55

A_1 - capacity of the sales staff, A_2 - perceived quality of goods
A_3 - high traffic location, *ROE - Return On Equity*

$x P_i^0 y$ (and $y P_i^0 x$): x is *indifferent* to y w.r.t. A_i

$x P_i^1 y$ (and $y P_i^{-1} x$): x is *preferred* to y w.r.t. A_i

$x P_i^2 y$ (and $y P_i^{-2} x$): x is *strongly preferred* to y w.r.t. A_i

xSy if $ROE(x) \geq ROE(y) - 2\%$

xS^cy if $ROE(x) < ROE(y) - 2\%$

DRSA for multicriteria choice and ranking – example

- Pairwise comparison table (PCT)

Pairs of warehouses	P_1^h on A_1	P_2^h on A_2	P_3^h on A_3	Comprehensive outranking
(1, 1)	P_1^0	P_2^0	P_3^0	S
(1, 2)	P_1^0	P_2^1	P_3^0	S
(1, 3)	P_1^1	P_2^0	P_3^0	S
(1, 4)	P_1^2	P_2^0	P_3^1	S
(1, 5)	P_1^2	P_2^0	P_3^1	S
(1, 6)	P_1^2	P_2^1	P_3^0	S
(1, 7)	P_1^0	P_2^0	P_3^0	S^c
(1, 8)	P_1^0	P_2^1	P_3^0	S
(2, 1)	P_1^0	P_2^{-1}	P_3^0	S^c
(2, 2)	P_1^0	P_2^0	P_3^0	S
…	…	…	…	…
(8, 7)	P_1^0	P_2^{-1}	P_3^0	S^c
(8, 8)	P_1^0	P_2^0	P_3^0	S

DRSA for multicriteria choice and ranking – example

- Quality of approximation of *S* and *Sc* by criteria from set *C* is 0.44
- RED_{PCT} = $CORE_{PCT}$ = $\{A_1,A_2,A_3\}$

$$\underline{C}(S) = \{(1,2),(1,4),(1,5),(1,6),(1,8),(3,2),(3,4),(3,5),(3,6),(3,8),(7,2),(7,4),(7,5),(7,6),(7,8)\}$$

$$\underline{C}(S^c) = \{(2,1),(2,7),(4,1),(4,3),(4,7),(5,1),(5,3),(5,7),(6,1),(6,3),(6,7),(8,1),(8,7)\}.$$

- $D_{\geq}$-decision rules and $D_{\leq}$-decision rules

$$\textit{if } x P_1^{\geq 1} y \textit{ and } x P_2^{\geq 1} y, \textit{ then } xSy; \qquad ((1,6),(3,6),(7,6))$$

$$\textit{if } x P_2^{\geq 1} y \textit{ and } x P_3^{\geq 0} y, \textit{ then } xSy; \qquad ((1,2),(1,6),(1,8),(3,2),(3,6),(3,8),(7,2),(7,6),(7,8))$$

$$\textit{if } x P_2^{\geq 0} y \textit{ and } x P_3^{\geq 1} y, \textit{ then } xSy; \qquad ((1,4),(1,5),(3,4),(3,5),(7,4),(7,5))$$

$$\textit{if } x P_1^{\leq -1} y \textit{ and } x P_2^{\leq -1} y, \textit{ then } xS^c y; \qquad ((6,1),(6,3),(6,7))$$

$$\textit{if } x P_2^{\leq 0} y \textit{ and } x P_3^{\leq -1} y, \textit{ then } xS^c y; \qquad ((4,1),(4,3),(4,7),(5,1),(5,3),(5,7))$$

$$\textit{if } x P_1^{\leq 0} y \textit{ and } x P_2^{\leq -1} y \textit{ and } x P_3^{\leq 0} y, \textit{ then } xS^c y; \qquad ((2,1),(2,7),(6,1),(6,3),(6,7),(8,1),(8,7))$$

DRSA for multicriteria choice and ranking – example

- $D_{\geq\leq}$-decision rules

$if\ x\,P_2^{\leq 0}\,y\ and\ x\,P_2^{\geq 0}\,y\ and\ x\,P_3^{\leq 0}\,y\ and\ x\,P_3^{\geq 0}\,y,\ then\ xSy\ or\ xS^c y;$

((1,1),(1,3),(1,7),(2,2),(2,6),(2,8),(3,1),(3,3),(3,7),(4,4),(4,5),
(5,4),(5,5),(6,2),(6,6),(6,8),(7,1),(7,3),(7,7),(8,2),(8,6),(8,8))

$if\ x\,P_2^{\leq -1}\,y\ and\ x\,P_3^{\geq 1}\,y,\ then\ xSy\ or\ xS^c y;$ ((2,4),(2,5),(6,4),(6,5),(8,4),(8,5))

$if\ x\,P_2^{\geq 1}\,y\ and\ x\,P_3^{\leq -1}\,y,\ then\ xSy\ or\ xS^c y;$ ((4,2),(4,6),(4,8),(5,2),(5,6),(5,8))

$if\ x\,P_1^{\geq 1}\,y\ and\ x\,P_2^{\leq 0}\,y\ and\ x\,P_3^{\leq 0}\,y,\ then\ xSy\ or\ xS^c y;$ ((1,3),(2,3),(2,6),(7,3),(8,3),(8,6))

$if\ x\,P_1^{\geq 1}\,y\ and\ x\,P_2^{\leq -1}\,y,\ then\ xSy\ or\ xS^c y;$ ((2,3),(2,4),(2,5),(8,3),(8,4),(8,5))

DRSA for multicriteria choice and ranking – example

- Ranking of warehouses for sale by decision rules and the *NFS* procedure

Warehouse for sale	A_1	A_2	A_3	Net Flow Score	Ranking
1'	good	sufficient	medium	1	5
2'	sufficient	good	good	11	1
3'	sufficient	medium	sufficient	-8	8
4'	sufficient	good	sufficient	0	6
5'	sufficient	sufficient	medium	-4	7
6'	sufficient	good	good	11	1
7'	medium	sufficient	sufficient	-11	9
8'	medium	medium	medium	7	3
9'	medium	good	sufficient	4	4
10'	medium	sufficient	sufficient	-11	9

- **Final ranking:** (2',6') → (8') → (9') → (1') → (4') → (5') → (3') → (7',10')
- **Best choice:** select warehouse 2' and 6' having maximum score (11)

DRSA for multiple-criteria choice and ranking - examples

- Input (preference) information:

A^R

Rerefence actions	Criteria				Rank in order
	q_1	q_2	…	q_m	
x	$f(x, q_1)$	$f(x, q_2)$	…	$f(x, q_m)$	1st
y	$f(y, q_1)$	$f(y, q_2)$	…	$f(y, q_m)$	2nd
u	$f(u, q_1)$	$f(u, q_2)$	…	$f(u, q_m)$	3rd
…	…	…	…	…	…
z	$f(z, q_1)$	$f(z, q_2)$	…	$f(z, q_m)$	k-th

$B \subseteq A^R \times A^R$

Pairwise Comparison Table (PCT)

Pairs of ref. actions	Difference on criteria				Preference relation
	q_1	q_2	…	q_m	
(x,y)	$\Delta_1(x,y)$	$\Delta_2(x,y)$	…	$\Delta_m(x,y)$	xSy
(y,x)	$\Delta_1(y,x)$	$\Delta_2(y,x)$	…	$\Delta_m(y,x)$	yS^cx
(y,u)	$\Delta_1(y,u)$	$\Delta_2(y,u)$	…	$\Delta_m(y,u)$	ySu
…	…	…	…	…	…
(v,z)	$\Delta_1(v,z)$	$\Delta_2(v,z)$	…	$\Delta_m(v,z)$	vS^cz

S – ***outranking***

S^c – ***non-outranking***

Δ_i – ***difference on q_i***

Association rules representing the Pareto optimal set

Association rules

- Relationships between attainable values of different objective functions (criteria) in the set of Pareto optimal (efficient) solutions

- Formal syntax (in case of maximization of objectives w.l.g.):
  - *If* $f_{i1}(x) \geq r_{i1}$ *and* $f_{i2}(x) \geq r_{i2}$ *and …* $f_{ip}(x) \geq r_{ip}$,

    then $f_{ip+1}(x) \leq r_{ip+1}$ *and* $f_{ip+2}(x) \leq r_{ip+2}$ *and …* $f_{iq}(x) \leq r_{iq}$

- Example:
  - *„if the maximum speed is at least 200 km/h and the time to reach 100 km/h is at most 7 seconds,*

    then the price is not less than 40,000$ and the fuel consumption is not less than 9 litres per 100 km"

Dominance-based Rough Set Approach to Interactive Multiple Objective Optimization (DRSA-IMO)

DRSA within Interactive Multiple Objective Optimization

1) Present to the DM a representative set of efficient solutions
2) Present association rules showing relationships between the attainable values of the objective functions in the Pareto optimal set
3) If the DM finds a satisfactory solution, then end; else go to the next step
4) The DM marks efficient solutions considered as (relatively) **good**
5) DRSA "*if...,then...*" decision rules are induced
6) The most interesting decision rules are presented to the DM
7) The DM selects one decision rule being the most adequate to his/her preferences
8) Constraints relative to this decision rule are adjoined
9) Go back to step 1

Greco, S., Matarazzo, B., Slowinski, R.: Dominance-Based Rough Set Approach to Interactive Multiobjective Optimization, Chapter 5 in J.Branke, K.Deb, K.Miettinen, R.Slowinski (eds.), *Multiobjective Optimization: Interactive and Evolutionary Approaches*. Springer-Verlag, to appear

Example of Product Mix Problem: Data

- Three products: A, B, C
- Produced quantity: x_A , x_B , x_C
- Price: $p_A=20$, $p_B=30$, $p_C=25$
- Time machine 1: $t_{1A}=5$, $t_{1B}=8$, $t_{1C}=10$
- Time machine 2: $t_{2A}=8$, $t_{2B}=6$, $t_{2C}=2$
- Raw material 1: $r_{1A}=1$, $r_{1B}=2$, $r_{1C}=0.75$; unit cost: 6
- Raw material 2: $r_{2A}=0.5$, $r_{2B}=1$, $r_{2C}=0.5$; unit cost: 8
- Market limit: $x^*_A=10$, $x^*_B=20$, $x^*_C=10$

Example of Product Mix Problem: Mathematical formulation

- Max Profit
- Min Total time (machine 1 + machine 2)
- Max Produced quantity of A
- Max Produced quantity of B
- Max Produced quantity of C
- Max Sales

Example of Product Mix Problem: Objectives and Constraints

- Max $20x_A + 30x_B + 25x_C - (1x_A + 2x_B + 0.75x_C)6 +$
  $- (0.5x_A + x_B + 0.5\, x_C)8$ [Profit]
- Min $5x_A + 8x_B + 10x_C + 8x_A + 6x_B + 2x_C$
  [Total time machine 1 + machine 2]
- Max x_A [Produced quantity of A]
- Max x_B [Produced quantity of B]
- Max x_C [Produced quantity of C]
- Max $20x_A + 30x_B + 25x_C$ [Sales]
- $x_A \leq 10$, $x_B \leq 20$, $x_C \leq 10$ [Market Limits]
- $x_A \geq 0$, $x_B \geq 0$, $x_C \geq 0$ [Non-negativity]

Set of representative efficient solutions

Solutions	Profit	Total time	x_A	x_B	x_C	Sales
S1	165	120	0	0	10	250
S2	172.6923	130	0.7692	0	10	265.3846
S3	180.3846	140	1.5384	0	10	280.7692
S4	141.125	140	3	3	4.916667	272.9167
S5	148.375	150	5	2	4.75	278.75
S6	139.125	150	5	3	3.583333	279.5833
S7	188.0769	150	2.3076	0	10	296.1538
S8	159	150	6	0	6	270
S9	140.5	150	6	2	3.666667	271.6667
S10	209.25	200	6	2	7.833333	375.8333
S11	189.375	200	5	5	5.416667	385.4167
S12	127.375	130	3	3	4.083333	252.0833
S13	113.625	120	3	3	3.25	231.25

The most interesting association rules

- If time≤140, then profit≤180.38 and sales≤280.77

  (s1,s2,s3,s4,s12,s13)

- If time≤150, then profit≤188.08 and sales≤296.15

  (s1,s2,s3,s4,s5,s6,s7,s8,s9,s12,s13)

- If $x_B \geq 2$, then profit≤209.25 and $x_A \leq 6$ and $x_C \leq 7.83$

  (s4,s5,s6,s9,s10,s11,s12,s13)

- If time≤150, then $x_B \leq 3$

  (s1,s2,s3,s4,s5,s6,s7,s8,s9,s12,s13)

- If profit≥148.38 and time≤150, then $x_B \leq 2$

  (s1,s2,s3,s5,s7,s8)

The most interesting association rules

- If $x_A \geq 5$, then time$\geq$150

  (s5,s6,s8,s9,s10,s11)

- If profit$\geq$127.38 and $x_A \geq 3$, then time$\geq$130

  (s4,s5,s6,s8,s9,s10,s11,s12)

- If time$\leq$150 and $x_B \geq 2$, then profit$\leq$148.38

  (s4,s5,s6,s9,s12,s13)

- If $x_A \geq 3$ and $x_C \geq 4.08$, then time$\geq$130

  (s4,s5,s8,s10,s11,s12)

- If sales$\geq$256.38, then time $\geq$130

  (s2,s3,s4,s5,s6,s7,s8,s9,s10,s11)

Sorting of representative efficient solutions

Solutions	Profit	Total time	x_A	x_B	x_C	Sales	Class
S1	165	120	0	0	10	250	*
S2	172.6923	130	0.7692	0	10	265.3846	*
S3	180.3846	140	1.5384	0	10	280.7692	**Good**
S4	141.125	140	3	3	4.916667	272.9167	**Good**
S5	148.375	150	5	2	4.75	278.75	**Good**
S6	139.125	150	5	3	3.583333	279.5833	*
S7	188.0769	150	2.3076	0	10	296.1538	*
S8	159	150	6	0	6	270	*
S9	140.5	150	6	2	3.666667	271.6667	**Good**
S10	209.25	200	6	2	7.833333	375.8333	*
S11	189.375	200	5	5	5.416667	385.4167	*
S12	127.375	130	3	3	4.083333	252.0833	*
S13	113.625	120	3	3	3.25	231.25	*

DRSA decision rule induction

- 12 rules were induced with the following frequency of the presence of objectives in the premise:
- Profit: 4/12
- Total time: 12/12
- Produced quantity A: 7/12
- Produced quantity B: 4/12
- Produced quantity C: 5/12
- Sales: 5/12

The most interesting DRSA decision rules

- If profit$\geq$140.5 and time$\leq$150 and $x_B\geq 2$,
  then product mix is good (s4,s5,s9)
- If time$\leq$140 and $x_A\geq 1.538462$ and $x_C\geq 10$,
  then product mix is good (s3)
- If time$\leq$150 and $x_B\geq 2$ and $x_C\geq 4.75$,
  then product mix is good (s4,s5)
- If time$\leq$140 and sales$\geq$272.9167,
  then product mix is good (s3,s4)
- If time$\leq$150 and $x_B\geq 2$ and $x_C\geq 3.666667$ and sales$\geq$271.6667,
  then product mix is good (s4,s5,s9)

Selected decision rule and relative added constraints

- The DM selected the following rule as the most adequate to his/her preferences:

  If profit $\geq$ 140.5 and time $\leq$ 150 and $x_B \geq 2$,
  then product mix is good (s4,s5,s9)

- Added constraints to the production mix problem:
- $20x_A + 30x_B + 25x_C - (1x_A + 2x_B + 0.75x_C)6 +$
  $- (0.5x_A + x_B + 0.5\, x_C)8 \geq 140.5$ [Profit $\geq$ 140.5]
- $5x_A + 8x_B + 10x_C + 8x_A + 6x_B + 2x_C \leq 150$ [time $\leq$ 150]
- $x_B \geq 2$ [Produced quantity of B $\geq$ 2]

Set of representative efficient solutions (second iteration)

Solutions	Profit	Total time	x_A	x_B	x_C	Sales
S1′	186.53	150	0.15	2	10	313.07
S2′	154.87	150	3	3	5.75	293.75
S3′	172	150	2	2	8	300
S4′	162.75	150	2	3	6.83	300.83
S5′	174	140	0	2	9.33	293.33
S6′	158.25	140	2	2	7.16	279.16
S7′	149	140	2	3	6	280
S8′	160.25	130	0	2	8.5	272.5
S9′	144.5	130	2	2	6.33	258.33
S10′	153.375	125	0	2	8.08	262.08
S11′	145.5	125	1	2	7	255
S12′	141.5625	125	1.5	2	6.45	251.45

The most interesting association rules

- If time$\leq$140, then profit$\leq$174 and $x_C\leq$9.33 and sales$\leq$293.33

  (s5′,s6′,s7′,s8′,s9′,s10′,s11′,s12′)

- If $x_A\geq$2, then $x_B\leq$3 and sales$\leq$300.83

  (s2′,s3′,s4′,s6′,s7′,s9′)

- If $x_A\geq$2, then profit$\leq$172 and $x_C\leq$8

  (s2′,s3′,s4′,s6′,s7′,s9′)

- If time$\leq$140, then $x_A\leq$2 and $x_B\leq$3

  (s5′,s6′,s7′,s8′,s9′,s10′,s11′,s12′)

- If profit$\geq$158.25, then $x_A\leq$2

  (s1′,s3′,s4′,s5′,s6′,s8′)

- If $x_A\geq$2, then time$\geq$130

  (s2′,s3′,s4′,s6′,s7′,s9′)

The most interesting association rules

- If $x_C \geq 7.17$, then $x_A \leq 2$ and $x_B \leq 2$

  (s1′,s3′,s5′,s6′,s8′,s10′)

- If $x_C \geq 6$, then $x_A \leq 2$ and $x_B \leq 3$

  (s1′,s3′,s4′,s5′,s6′,s7′,s8′,s9′,s10′,s11′,s12′)

- If $x_C \geq 7$, then time≥125 and $x_B \leq 2$

  (s1′,s3′,s5′,s6′,s8′,s10′,s11′)

- If sales≥280, then time≥140 and $x_B \leq 3$

  (s1′,s2′,s3′,s4′,s5′,s7′)

- If sales≥279.17, then time≥140

  (s1′,s2′,s3′,s4′,s5′,s6′,s7′)

- If sales≥272, then time≥130

  (s1′,s2′,s3′,s4′,s5′,s6′,s7′,s8′)

Sorting of representative efficient solutions (second iteration)

Solutions	Profit	Total time	x_A	x_B	x_C	Sales	Class
S1′	186.53	150	0.15	2	10	313.07	*
S2′	154.87	150	3	3	5.75	293.75	*
S3′	172	150	2	2	8	300	**Good**
S4′	162.75	150	2	3	6.83	300.83	**Good**
S5′	174	140	0	2	9.33	293.33	*
S6′	158.25	140	2	2	7.16	279.16	**Good**
S7′	149	140	2	3	6	280	*
S8′	160.25	130	0	2	8.5	272.5	*
S9′	144.5	130	2	2	6.33	258.33	*
S10′	153.375	125	0	2	8.08	262.08	*
S11′	145.5	125	1	2	7	255	**Good**
S12′	141.5625	125	1.5	2	6.45	251.45	**Good**

DRSA decision rule induction

- 8 rules were induced with the following frequency of the presence of objectives in the premise:
- Profit: 2/8
- Total time: 1/8
- Produced quantity A: 5/8
- Produced quantity B: 3/8
- Produced quantity C: 3/8
- Sales: 2/8

The most interesting DRSA decision rules

- If profit≥158.25 and $x_A \geq 2$,
  then product mix is good (s3′,s4′,s6′)
- If time≤125 and $x_A \geq 1$,
  then product mix is good (s11′,s12′)
- If $x_A \geq 1$ and $x_C \geq 7$,
  then product mix is good (s3′,s6′,s11′)
- If $x_A \geq 1.5$ and $x_C \geq 6.46$,
  then product mix is good (s3′,s4′,s6′,s12′)
- If $x_A \geq 2$ and sales≥300,
  then product mix is good (s3′,s4′)

Selected decision rule and relative added constraints

- The DM selected the following rule as the most adequate to his/her preferences:

  If profit $\geq$ 158.25 and $x_A \geq 2$,
  then product mix is good (s3′,s4′,s6′)

- Added constraints to the production mix problem:
- $20x_A + 30x_B + 25x_C - (1x_A + 2x_B + 0.75x_C)6 + - (0.5x_A + x_B + 0.5x_C)8 \geq 158.25$ [Profit $\geq$ 158.25]
- $x_A \geq 2$ [Produced quantity of A $\geq$ 2]

Set of representative efficient solutions (third iteration)

Solutions	Profit	Total time	x_A	x_B	x_C	Sales
S1″	165.125	145	2	2	7.58	289.58
S2″	158.25	150	2	3.48	6.26	301.23
S3″	158.25	145	2	2.74	6.71	290.20
S4″	158.25	140	2	2	7.16	279.16
S5″	164.125	150	3	2	6.91	292.91
S6″	158.25	145.72	3	2	6.56	284.01

The most interesting association rules

- If time≤145, then x_A≤2 and x_B≤2.74 and sales≤290.2
  (s2″,s3″,s4″)
- If x_C≥6.92, then x_A≤3 and x_B≤2 and sales≤292.92
  (s3″,s4″,s5″)
- If time≤145, then profit≤165.13 and x_A≤2 and x_C≤7.58
  (s2″,s3″,s4″)
- If x_C≥6.72, then x_B≤2.74
  (s2″,s3″,s4″,s5″)
- If sales≥289.58, then profit≤165.13 and time≥145 and x_C≤7.58
  (s1″,s2″,s3″,s5″)

Set of representative efficient solutions (third iteration) and the selected solution

Solutions	Profit	Total time	x_A	x_B	x_C	Sales	Class
S1″	165.125	145	2	2	7.58	289.58	**Selected**
S2″	158.25	150	2	3.48	6.26	301.23	*
S3″	158.25	145	2	2.74	6.71	290.20	*
S4″	158.25	140	2	2	7.16	279.16	*
S5″	164.125	150	3	2	6.91	292.91	*
S6″	158.25	145.72	3	2	6.56	284.01	*

Conclusions 1

Main features of the interactive method:

- The method is based on ordinal properties of values of objective functions only
- At each step, the method does not aggregate the objective functions into a single value (no scalarization is involved)
- DM gives preference information by answering easy questions in terms of holistic sorting, without use of any technical parameters, such as weights, tradeoffs, thresholds,...

Conclusions 2

Main advantages of DRSA involving rules:

- Association rules
  - They represent relationships between attainable values of objective functions
  - DM learns from them about the shape of the Pareto optimal set
- Both association and decision rules are important in a learning oriented perspective:
  - They are easily understandable and intelligible for the DM ("glass box")
  - They permit the DM to identify Pareto optimal solutions supporting each rule
  - They enable argumentation, explanation and justification of the final decision
    (as a conclusion of a decision process,
    not just as a mechanical application of a technical approach)

Financial Portfolio Decision Analysis using Dominance-based Decision Rules

Portfolio analysis: basic data

- Three securities: S_1, S_2, S_3
- Expected returns of the securities: r_1=12%, r_2=14%, r_3=16%.
- Variance-Covariance matrix

	S_1	S_2	S_3
S_1	100	50	-20
S_2	50	200	10
S_3	-20	10	300

- Weights of three securities in a portfolio P: w_1, w_2, w_3;
  - $w_1 \geq 0$, $w_2 \geq 0$, $w_3 \geq 0$
  - $w_1+w_2+w_3 = 1$

Portfolio Risk and Return

- Expected return on a portfolio [$E(R_P)$] is a linear combination of expected returns [$E(R_i)$] of N component securities using weights (w_i):

$$E(R_p) = \sum_{i=1}^{N} w_i E(R_i)$$

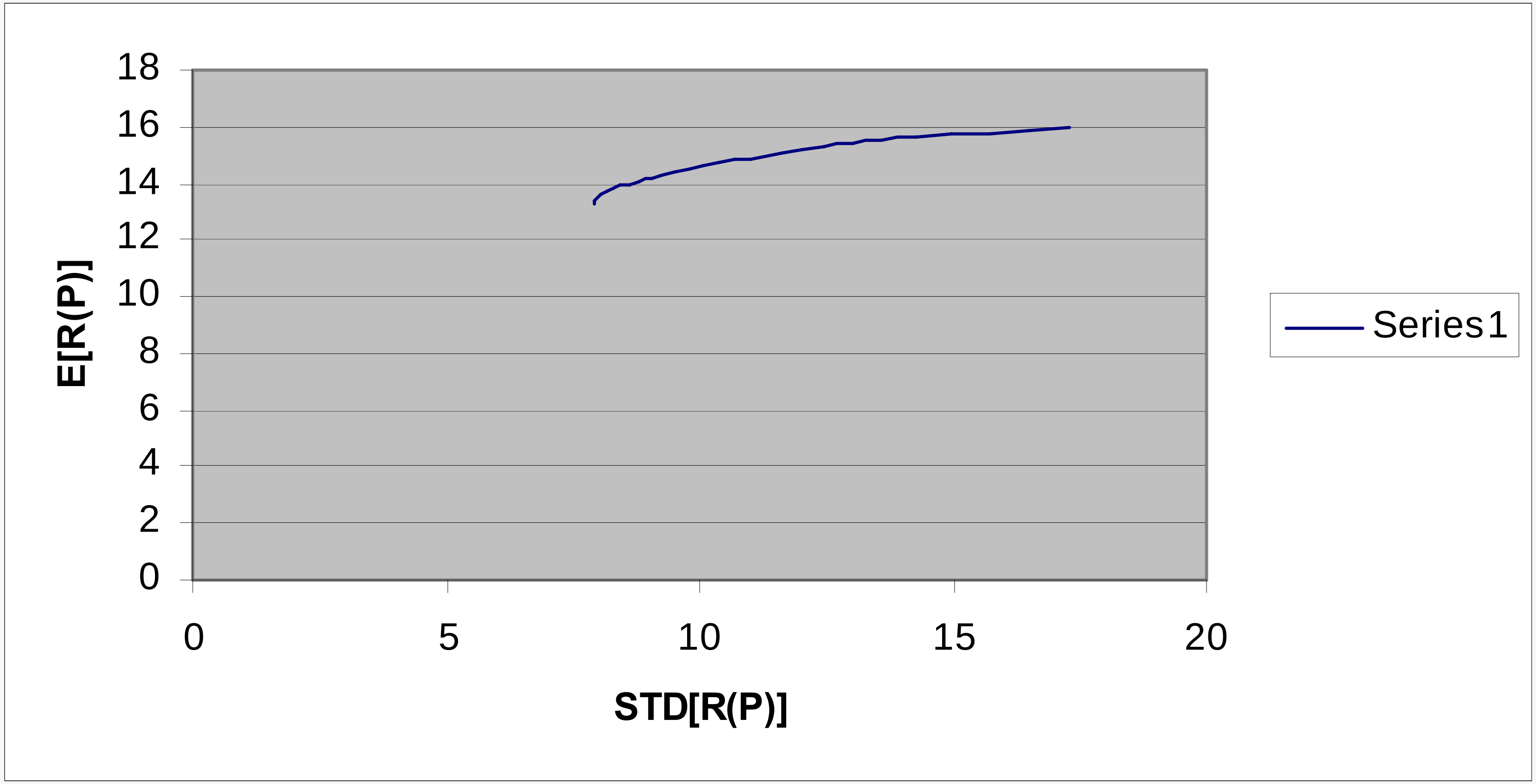
- Variance of a portfolio

$$\sigma_p^2 = \sum_{i=1}^{N} w_i^2 \sigma_i^2 + \sum_{i=1}^{N} \sum_{\substack{j=1 \\ i \neq j}}^{N} w_i w_j \sigma_{ij}$$

- Standard deviation of a portfolio

$$STD[R(P)] = \sqrt{\sigma_P^2}$$

Efficient frontier



Portfolio selection: mathematical formulation

- $r_{1\%}(P)$ =Max return at 1% (E[R(P)] + 2.33×STD[R(P)])
- $r_{25\%}(P)$=Max return at 25% (E[R(P)] + 0.67×STD[R(P)])
- $r_{50\%}(P)$=Max return at 50% (E[R(P)])
- $r_{75\%}(P)$=Max return at 75% (E[R(P)] - 0.67×STD[R(P)])
- $r_{99\%}(P)$=Max return at 99% (E[R(P)] - 2.33×STD[R(P)])

Set of representative solutions (first iteration)

	w_1	w_2	w_3	r	σ	$r_{1\%}(P)$	$r_{25\%}(P)$	$r_{50\%}(P)$	$r_{75\%}(P)$	$r_{99\%}(P)$	Class
P1	0.39	0.29	0.32	13.86	8.43	33.50	19.51	13.86	8.21	-5.78	*
P2	0.21	0.22	0.57	14.71	10.64	39.49	21.84	14.71	7.58	-10.07	*
P3	0.01	0.48	0.51	15.01	11.39	41.55	22.64	15.01	7.37	-11.54	*
P4	0.61	0.04	0.35	13.50	8.30	32.82	19.05	13.50	7.94	-5.83	*
P5	0.43	0.39	0.18	13.52	8.58	33.51	19.27	13.52	7.77	-6.48	**Good**
P6	0.51	0.46	0.03	13.04	9.58	35.37	19.46	13.04	6.62	-9.29	*
P7	0.52	0.20	0.29	13.54	8.03	32.24	18.92	13.54	8.16	-5.16	**Good**
P8	0.54	0.04	0.42	13.75	8.70	34.03	19.58	13.75	7.92	-6.53	**Good**
P9	0.34	0.21	0.45	14.22	9.16	35.57	20.36	14.22	8.08	-7.13	*
P10	0.54	0.22	0.23	13.38	7.99	32.01	18.74	13.38	8.03	-5.24	**Good**
P11	0.60	0.15	0.25	13.28	7.94	31.78	18.60	13.28	7.97	-5.21	*
P12	0.53	0.19	0.28	13.5	8.00	32.14	18.86	13.5	8.14	-5.14	**Good**
P13	0.37	0.26	0.37	14	8.62	34.09	19.78	14	8.224	-6.09	**Good**
P14	0.21	0.34	0.46	14.5	9.79	37.30	21.06	14.5	7.94	-8.30	**Good**
P15	0.04	0.41	0.54	15	11.33	41.39	22.59	15	7.41	-11.39	*
P16	0	0.25	0.75	15.5	13.60	47.19	24.61	15.5	6.39	-16.19	*
P17	0	0	1	16	17.32	56.36	27.60	16	4.40	-24.36	**Good**

DRSA decision rule induction

- 19 rules were induced with the following frequency of the presence of objectives in the premise:
- $r_{1\%}(P)$: 6/19
- $r_{25\%}(P)$: 5/19
- $r_{50\%}(P)$: 5/19
- $r_{75\%}(P)$: 5/19
- $r_{99\%}(P)$: 12/19

The most interesting DRSA decision rules (1)

- If $r_{1\%}(P) \geq 32.01\%$ and $r_{99\%}(P) \geq -5.24\%$,
  then portfolio is good (P7, P10, P12)
- If $r_{25\%}(P) \geq 18.74\%$ and $r_{99\%}(P) \geq -5.24\%$,
  then portfolio is good (P7, P10, P12)
- If $r_{50\%}(P) \geq 13.38\%$ and $r_{99\%}(P) \geq -5.24\%$,
  then portfolio is good (P7, P10, P12)
- If $r_{75\%}(P) \geq 8.03\%$ and $r_{99\%}(P) \geq -5.24\%$,
  then portfolio is good (P7, P10, P12)

The most interesting DRSA decision rules (2)

- If $r_{1\%}(P) \geq 33.51\%$ and $r_{99\%}(P) \geq -6.48\%$, then portfolio is good (P5, P13)
- If $r_{1\%}(P) \geq 34.03\%$ and $r_{99\%}(P) \geq -6.53\%$, then portfolio is good (P8, P13)
- If $r_{50\%}(P) \geq 16\%$, then portfolio is good (P17)
- If $r_{50\%}(P) \geq 14.5\%$ and $r_{99\%}(P) \geq -8.3\%$, then portfolio is good (P14)

Selected decision rule and relative added constraints

- The DM selected the following rule as the most adequate to his preferences:

  If $r_{75\%}(P) \geq 8.03\%$ and $r_{99\%}(P) \geq -5.24\%$,
  then portfolio is good (P7, P10, P12)

- Added constraints to the portfolio selection problem:
  - $r_{75\%}(P)=E[R(P)] - 0.67 \times STD[R(P)] \geq 8.03\%$,
  - $r_{99\%}(P)=E[R(P)] - 2.33 \times STD[R(P)] \geq -5.24\%$.

Set of representative solutions (second iteration)

	w_1	w_2	w_3	r	σ	$r_{1\%}(P)$	$r_{25\%}(P)$	$r_{50\%}(P)$	$r_{75\%}(P)$	$r_{99\%}(P)$	Class
P1’	0.52	0.20	0.29	13.86	8.03	32.24	18.92	13.54	8.16	-5.16	*
P2’	0.54	0.19	0.27	14.71	7.98	32.04	18.80	13.45	8.11	-5.13	**Good**
P3’	0.54	0.20	0.26	15.01	7.98	32.05	18.80	13.45	8.10	-5.15	*
P4’	0.50	0.23	0.27	13.50	8.05	32.29	18.93	13.53	8.14	-5.22	**Good**
P5’	0.53	0.18	0.29	13.52	8.02	32.20	18.89	13.52	8.15	-5.16	**Good**
P6’	0.57	0.16	0.27	13.04	7.96	31.93	18.72	13.39	8.06	-5.14	**Good**
P7’	0.54	0.16	0.30	13.54	8.02	32.20	18.89	13.51	8.14	-5.18	*
P8’	0.52	0.21	0.27	13.75	8.01	32.14	18.85	13.49	8.12	-5.17	*
P9’	0.59	0.12	0.29	14.22	7.99	32.00	18.74	13.39	8.04	-5.22	*
P10’	0.59	0.12	0.30	13.38	8.00	32.06	18.78	13.42	8.05	-5.23	*
P11’	0.58	0.16	0.26	13.35	7.94	31.86	18.67	13.35	8.03	-5.16	*
P12’	0.49	0.20	0.30	13.62	8.10	32.49	19.05	13.62	8.20	-5.24	**Good**
P13’	0.57	0.17	0.27	13.40	7.96	31.94	18.73	13.4	8.07	-5.14	*
P14’	0.55	0.18	0.27	13.45	7.97	32.03	18.79	13.45	8.11	-5.13	**Good**
P15’	0.53	0.18	0.28	13.50	8.00	32.14	18.86	13.5	8.14	-5.14	*
P16’	0.50	0.20	0.30	13.60	8.07	32.41	19.01	13.6	8.19	-5.21	**Good**

DRSA decision rule induction

- 5 rules were induced with the following frequency of the presence of objectives in the premise:
- $r_{1\%}(P)$: 1/5
- $r_{25\%}(P)$: 1/5
- $r_{50\%}(P)$: 1/5
- $r_{75\%}(P)$: 1/5
- $r_{99\%}(P)$: 1/5

The most interesting DRSA decision rules

- If $r_{1\%}(P) \geq 32.29\%$,
  then portfolio is good (P4', P12', P16')
- If $r_{25\%}(P) \geq 18.93\%$,
  then portfolio is good (P4', P12', P16')
- If $r_{50\%}(P) \geq 13.6\%$,
  then portfolio is good (P12', P16')
- If $r_{75\%}(P) \geq 8.19\%$,
  then portfolio is good (P12', P16')
- If $r_{99\%}(P) \geq -5.13\%$,
  then portfolio is good (P2', P14')

Selected decision rule and relative added constraints

- The DM selected the following rule as the most adequate to his preferences:

  If $r_{25\%}(P) \geq 18.93\%$,
  then portfolio is good (P4', P12', P16')

- Added constraint to the portfolio selection problem:

  $r_{25\%}(P) = E[R(P)] - 0.67 \times STD[R(P)] \geq 18.93\%$.

Set of representative solutions (third iteration) and the selected solution

	w_1	w_2	w_3	r	σ	$r_{1\%}(P)$	$r_{25\%}(P)$	$r_{50\%}(P)$	$r_{75\%}(P)$	$r_{99\%}(P)$	Class
P1''	0.50	0.20	0.30	13.59	8.07	32.38	18.99	13.59	8.18	-5.20	*
P2''	0.49	0.20	0.30	13.62	8.09	32.48	19.04	13.62	8.20	-5.24	*
P3''	0.50	0.19	0.31	13.62	8.09	32.47	19.04	13.62	8.20	-5.23	*
P4''	0.51	0.20	0.29	13.55	8.03	32.27	18.93	13.55	8.17	-5.17	*
P5''	0.50	0.22	0.28	13.55	8.05	32.31	18.95	13.55	8.16	-5.20	*
P6''	0.50	0.21	0.28	13.55	8.04	32.29	18.94	13.55	8.16	-5.19	*
P7''	0.52	0.17	0.30	13.56	8.04	32.30	18.95	13.56	8.17	-5.19	*
P8''	0.50	0.21	0.29	13.59	8.07	32.38	18.99	13.59	8.18	-5.21	*
P9''	0.49	0.23	0.28	13.58	8.07	32.39	18.99	13.58	8.17	-5.23	*
P10'	0.50	0.20	0.30	13.56	8.05	32.33	18.96	13.56	8.16	-5.21	*
P11''	0.52	0.19	0.29	13.55	8.03	32.26	18.93	13.55	8.17	-5.17	*
P12''	0.49	0.20	0.30	13.62	8.10	32.49	19.05	13.62	8.20	-5.24	**Selected**
P13''	0.51	0.20	0.29	13.57	8.05	32.33	18.96	13.57	8.18	-5.19	*
P14''	0.5	0.2	0.3	13.60	8.07	32.41	19.01	13.60	8.19	-5.21	*

DRSA to Decision under Risk and Uncertainty

DRSA to decision under risk and uncertainty

- $A=\{A_1, A_2, A_3, A_4, A_5, A_6, ...\}$ – set of acts
- $ST=\{st_1, st_2, st_3, ...\}$ – set of elementary states of the world
- Pr – a priori probability distribution over ST
  e.g.: $pr_1=0.25$, $pr_2=0.35$, $pr_3=0.40$, ...
- $X=\{0, 10, 15, 20, 30, ...\}$ – set of possible outcomes (gains)
- $\boldsymbol{Cl}=\{Cl_1, Cl_2, Cl_3, ...\}$ – set of quality classes of the acts,
  e.g.: Cl_1=bad acts, Cl_2=medium acts, Cl_3=good acts
- $\rho(A_i,\pi)=x$ means that by act A_i one can gain *at least x* with probability $\pi=Pr(W)$, where $W\subseteq ST$ is an event
- There is a partial preorder on probabilities π of events
- Act A_i ***stochastically dominates*** A_j iff $\rho(A_i,\pi) \geq \rho(A_j,\pi)$ for each probability $\pi\in\Pi$

DRSA to decision under risk and uncertainty

- Preference information given by a Decision Maker:

  assignment to acts to quality classes

- Example:

π/*Act*	A_1	A_2	A_3	A_4	A_5	A_6
.25	30	20	20	20	20	20
.35	10	20	20	20	20	20
.40	10	20	20	20	20	20
.60	10	20	15	15	20	20
.65	10	20	15	15	20	20
.75	10	20	0	15	10	20
1	10	0	0	0	10	10
Class	good	medium	medium	bad	medium	good

DRSA to decision under risk and uncertainty

- Decision rules induced from rough approximations of quality classes

$$\text{if } \rho(A_i, 0.75) \geq 20 \text{ and } \rho(A_i, 1) \geq 10, \text{ then } A_i \in Cl_3^{\geq}$$

"if the probability of gaining at least 20 is 0.75 and the probability of gaining at least 10 is 1, then act A_i is at least good"

$$\text{if } \rho(A_i, 0.25) \leq 20 \text{ and } \rho(A_i, 0.75) \leq 15, \text{ then } A_i \in Cl_2^{\leq}$$

"if the probability of gaining at most 20 is 1 and the probability of gaining at most 15 is 0.75, then act A_i is at most medium"

- Generalization:
  DRSA for decision under risk with outcomes distributed over time

Greco S., Matarazzo B., Slowinski R., Rough set approach to decisions under risk. [In]: W.Ziarko, Y.Yao (eds.): *Rough Sets and Current Trends in Computing*, LNAI 2005, Springer-Verlag, Berlin, 2001, pp. 160-169

DRSA to Case-Based Reasoning

DRSA to Fuzzy Case-Based Reasoning (CBR)

- Case-Based Reasoning regards the inference of some proper conclusions related to a new situation by the analysis of similar cases from a memory of previous cases

- It is based on three principles

  a) similar problems have similar solutions

  b) types of encountered problems tend to recur

  c) the more similar are the causes,
  the more similar the effects one can expect (DRSA!)

Fuzzy set approach to Case-Based Reasoning:

Dubois, D., Prade, H., Esteva, F., Garcia, P., Godo, L., Lopez de Mantara, R., Fuzzy Set Modelling in Case-based Reasoning, *Int. J. of Intelligent Systems*, 13 (1998) 345-373

DRSA to Fuzzy Case-Based Reasoning

- Measuring similarity is the essential point of all case-based reasoning and, particularly, of fuzzy set approach to case-based reasoning

- Problems of modelling similarity are relative to two levels:
  - at level of similarity with respect to single features: how to define a meaningful similarity measure with respect to a single feature ?
  - at level of similarity with respect to all features: how to propely aggregate the similarity measure with respect to single features in order to obtain a comprehensive similarity measure ?

S.Greco, B.Matarazzo, R.Słowiński: Dominance-based Rough Set Approach to Case-Based Reasoning. [In]: V. Torra, Y. Narukawa, A. Valls, J. Domingo-Ferrer (eds.), *Modelling Decisions for Artificial Intelligence*. LNAI 3885, Springer-Verlag, Berlin, 2006, pp. 7-18

DRSA to Fuzzy Case-Based Reasoning

- DRSA tends to be as „neutral" and „objective" as possible with respect to similarity relation
- **At level of similarity concerning single features**: only ordinal properties of similarity are exploited
- **At level of aggregation of similarity relative to single features**:
  - no specific functional aggregation (like weighted Lp norms, min, etc.) is used
  - a set of decision rules based on very general monotonicity relation between comprehensive similarity and similarity on single features
- Such an approach to Case-Based Reasoning is very little „invasive"

DRSA to Fuzzy Case-Based Reasoning

- Monotonicity: "The more similar are the descriptions, the more similar are the outcomes"
- Similarity is a concept concerning **pairs** of objects
- **Pairwise fuzzy information base**: $\boldsymbol{B}=<U, F, \sigma>$, where

  U – finite set of objects (universe)

  $F=\{f_1,f_2,\ldots,f_m\}$ – finite set of features

  $\sigma : U\times U\times F \rightarrow [0,1]$ – function expressing the credibility $\sigma(x,y,f_h)\in[0,1]$ that object x is similar to object y w.r.t. feature f_h $[\sigma(x,x,f_h)=1]$
- Each pair $(x,y)\in U\times U$ is described by: $Des_F(x,y)=[\sigma(x,y,f_1),\ldots,\sigma(x,y,f_m)]$
- For each subset of properties $E\subseteq F$: $Des_E(x,y)=[\sigma(x,y,f_h),\ f_h\in E]$

DRSA to Fuzzy Case-Based Reasoning

- Dominance relation on $U \times U$, concerning similarity between pairs of objects: for all $x,y,w,z \in U$, $E \subseteq F$

$(x,y)D_E(w,z)$: „x is similar to y at least as much as w is similar to z w.r.t. all the considered attributes from E"

- Dominance principle with respect to similarity

If x belongs to X and $(y,x)D_E(z,x)$, then y should belong to X with at least the same credibility as z belongs to X.

DRSA to Fuzzy Case-Based Reasoning

- For each $\emptyset \subset E \subseteq F$ and $x,y,w,z \in U$

$$(x,y)D_E(w,z) \iff \sigma(x,y,f_i) \geq \sigma(w,z,f_i) \text{ for all } f_i \in E$$

- For each $\emptyset \subset E \subseteq F$ and $x \in U$

positive cone: $D_E^+(y,x)=\{w \in U: (w,x)D_E(y,x)\}$

Interpretation:

set of objects being similar to x not less than y is similar to x

negative cone: $D_E^-(y,x)=\{w \in U: (y,x)D_E(w,x)\}$

Interpretation:

set of objects being similar to x not more than y is similar to x

- In the pair (y,x), x is a *reference object*, and y is a *limit object*, for y is conditioning the membership of w in $D_E^+(y,x)$ and $D_E^-(y,x)$

Fuzzy set of „similar objects”

- Fuzzy set X on U incl. objects with decision similar to reference object x

  Membership function of fuzzy set X (degree of similarity):

$$\mu_X: U \to [0,1]$$

- For each cutting level (limit degree of similarity) $\alpha \in [0,1]$:
  - upside cutting

    $X^{\geq\alpha} = \{y \in U: \mu_X(y) \geq \alpha\}$ $\qquad$ $X^{>\alpha} = \{y \in U: \mu_X(y) > \alpha\}$

  - downside cutting

    $X^{\leq\alpha} = \{y \in U: \mu_X(y) \leq \alpha\}$ $\qquad$ $X^{<\alpha} = \{y \in U: \mu_X(y) < \alpha\}$

- Complementarity :

  $U - X^{\geq\alpha} = X^{<\alpha}$, $\quad U - X^{\leq\alpha} = X^{>\alpha}$, $\quad U - X^{>\alpha} = X^{\leq\alpha}$, $\quad U - X^{<\alpha} = X^{\geq\alpha}$

Case-Based Rough approximations of a fuzzy set of similar objects

- For each reference object $x \in U$, cutting level $\alpha \in [0,1]$ and similarity function σ, we can define lower & upper approximations of $X^{\geq\alpha}$ with respect to features $E \subseteq F$:

  Upside lower approximation :

  $$\underline{E}(x)_\sigma\left(X^{\geq\alpha}\right) = \left\{y \in U : D_E^+(y,x) \subseteq X^{\geq\alpha}\right\}$$

  it contains all objects $y \in U$ such that any object w being similar to x at least as much as y is similar to x w.r.t. features from E, also belongs to $X^{\geq\alpha}$

  Upside upper approximation :

  $$\overline{E}(x)_\sigma\left(X^{\geq\alpha}\right) = \left\{y \in U : D_E^-(y,x) \cap X^{\geq\alpha} \neq \emptyset\right\}$$

  it contains all objects $y \in U$ such that there is at least one object w being similar to x at most as much as y is similar to x w.r.t. features from E, which belongs to $X^{\geq\alpha}$

Case-Based Rough approximations of a fuzzy set of similar objects

- For each reference object $x \in U$, cutting level $\alpha \in [0,1]$ and similarity function σ, we can define lower & upper approximations of $X^{\leq\alpha}$ with respect to features $E \subseteq F$:

  Downside lower approximation :

  $$\underline{E}(x)_{\sigma}\left(X^{\leq\alpha}\right) = \left\{ y \in U : D_E^{-}(y, x) \subseteq X^{\leq\alpha} \right\}$$

  Downside upper approximation :

  $$\overline{E}(x)_{\sigma}\left(X^{\leq\alpha}\right) = \left\{ y \in U : D_E^{+}(y, x) \cap X^{\leq\alpha} \neq \varnothing \right\}$$

Case-Based Rough approximations of a fuzzy set of similar objects

- Rough approximations can be rewritten in logical terms

Upside lower approximation :

$$\underline{E}(x)_\sigma\left(X^{\geq\alpha}\right)=\left\{y\in U:\forall w\in U \text{ such that } (w,x)D_E(y,x)\Rightarrow w\in X^{\geq\alpha}\right\}$$

Upside upper approximation :

$$\overline{E}(x)_\sigma\left(X^{\geq\alpha}\right)=\left\{y\in U:\exists w\in U \text{ such that } (y,x)D_E(w,x) \text{ and } w\in X^{\geq\alpha}\right\}$$

Downside lower approximation :

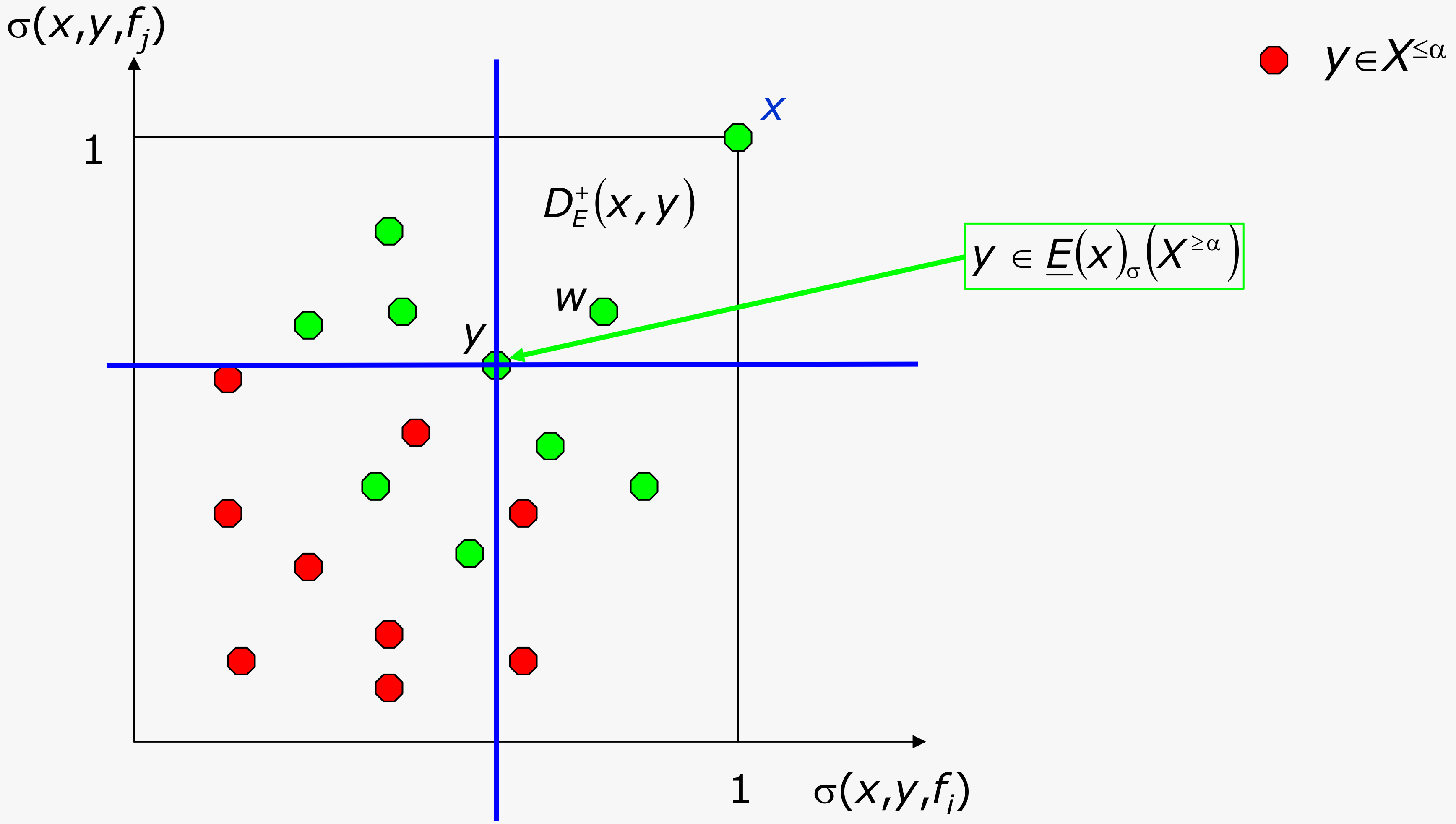
$$\underline{E}(x)_\sigma\left(X^{\leq\alpha}\right)=\left\{y\in U:\forall w\in U \text{ such that } (y,x)D_E(w,x)\Rightarrow w\in X^{\leq\alpha}\right\}$$

Downside upper approximation :

$$\overline{E}(x)_\sigma\left(X^{\leq\alpha}\right)=\left\{y\in U:\exists w\in U \text{ such that } (w,x)D_E(y,x) \text{ and } w\in X^{\leq\alpha}\right\}$$

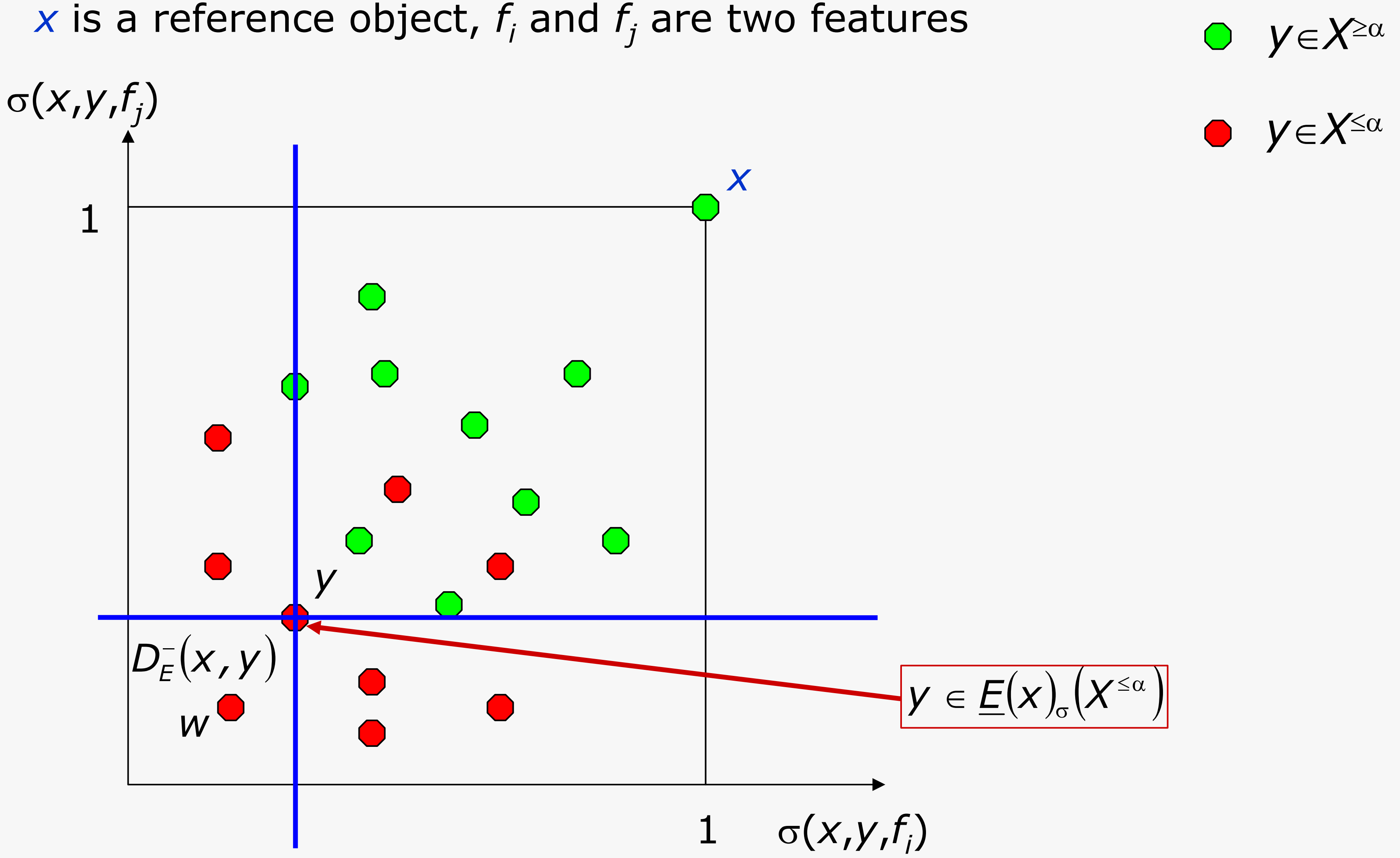
Similarity space in CBR

x is a reference object, f_i and f_j are two features



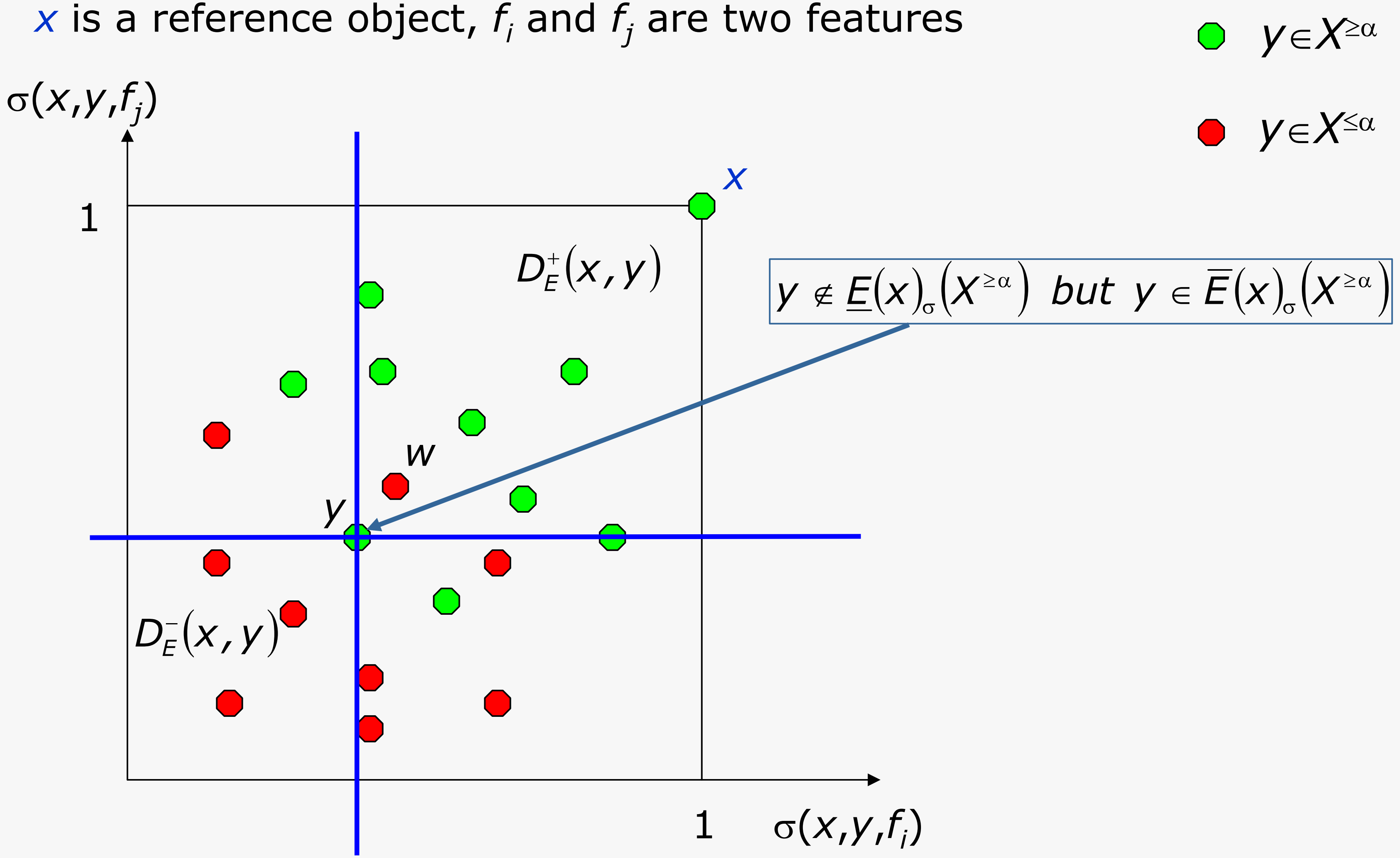
Similarity space in CBR

x is a reference object, f_i and f_j are two features



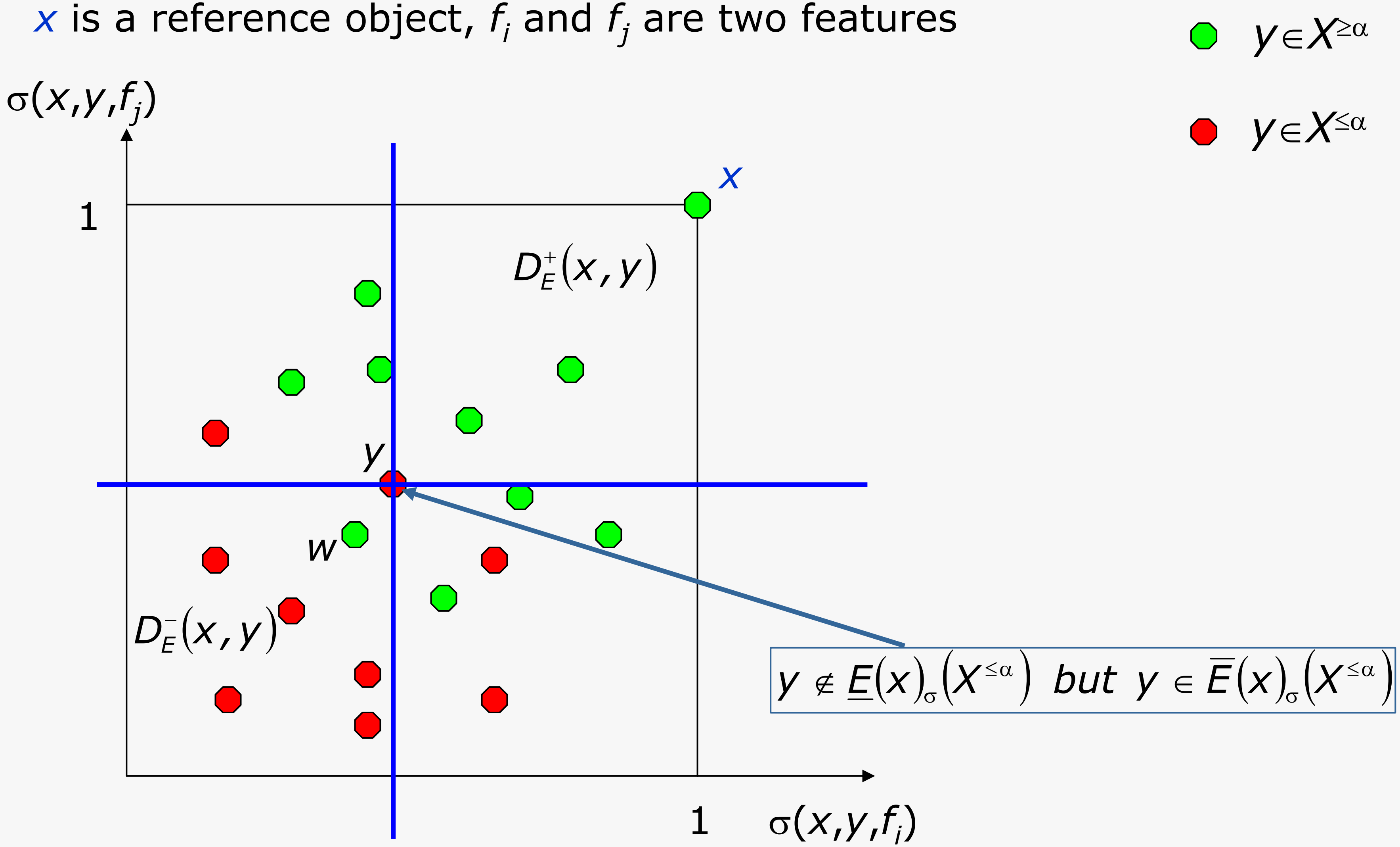
Similarity space in CBR

x is a reference object, f_i and f_j are two features



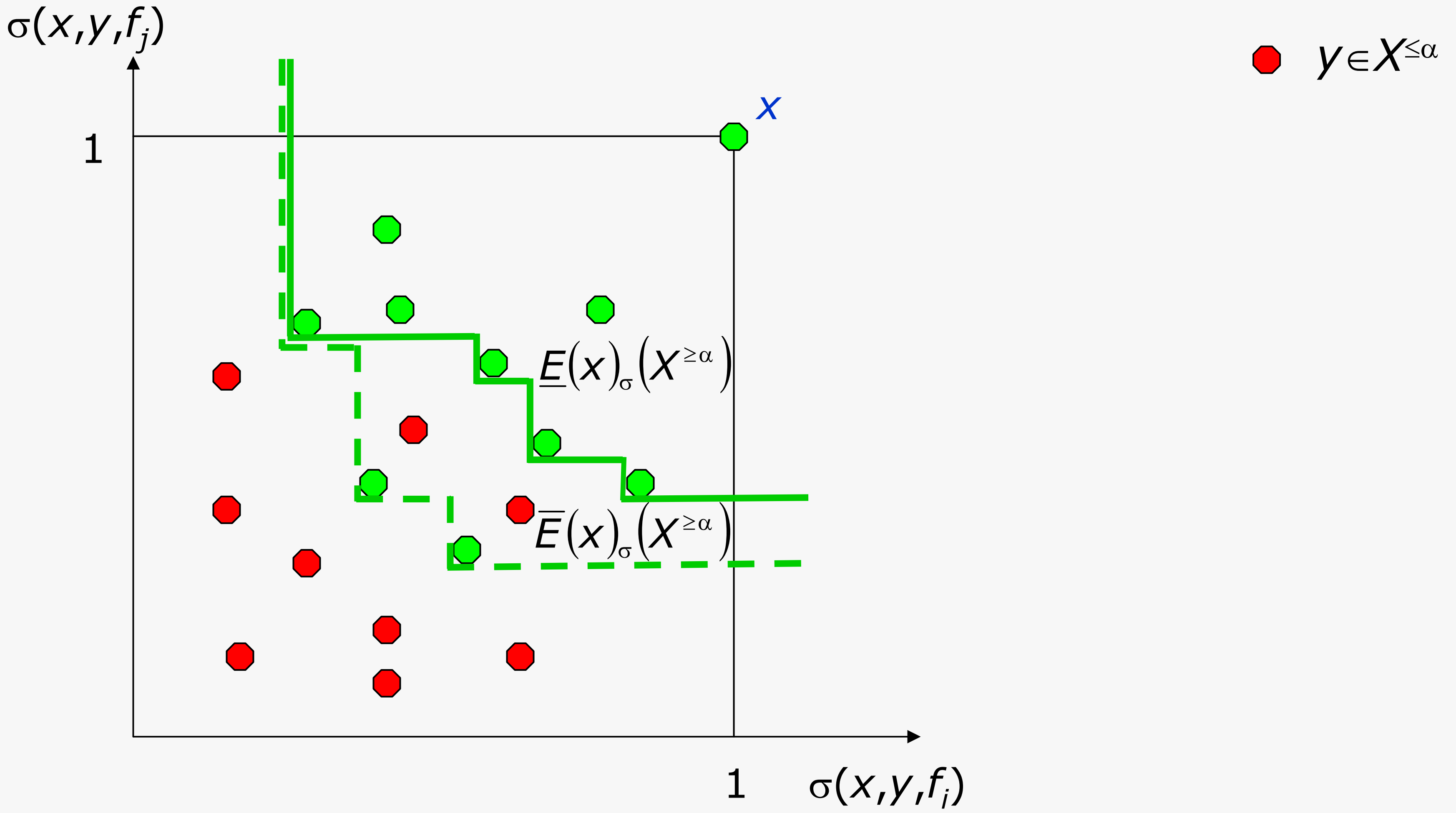
Similarity space in CBR

x is a reference object, f_i and f_j are two features



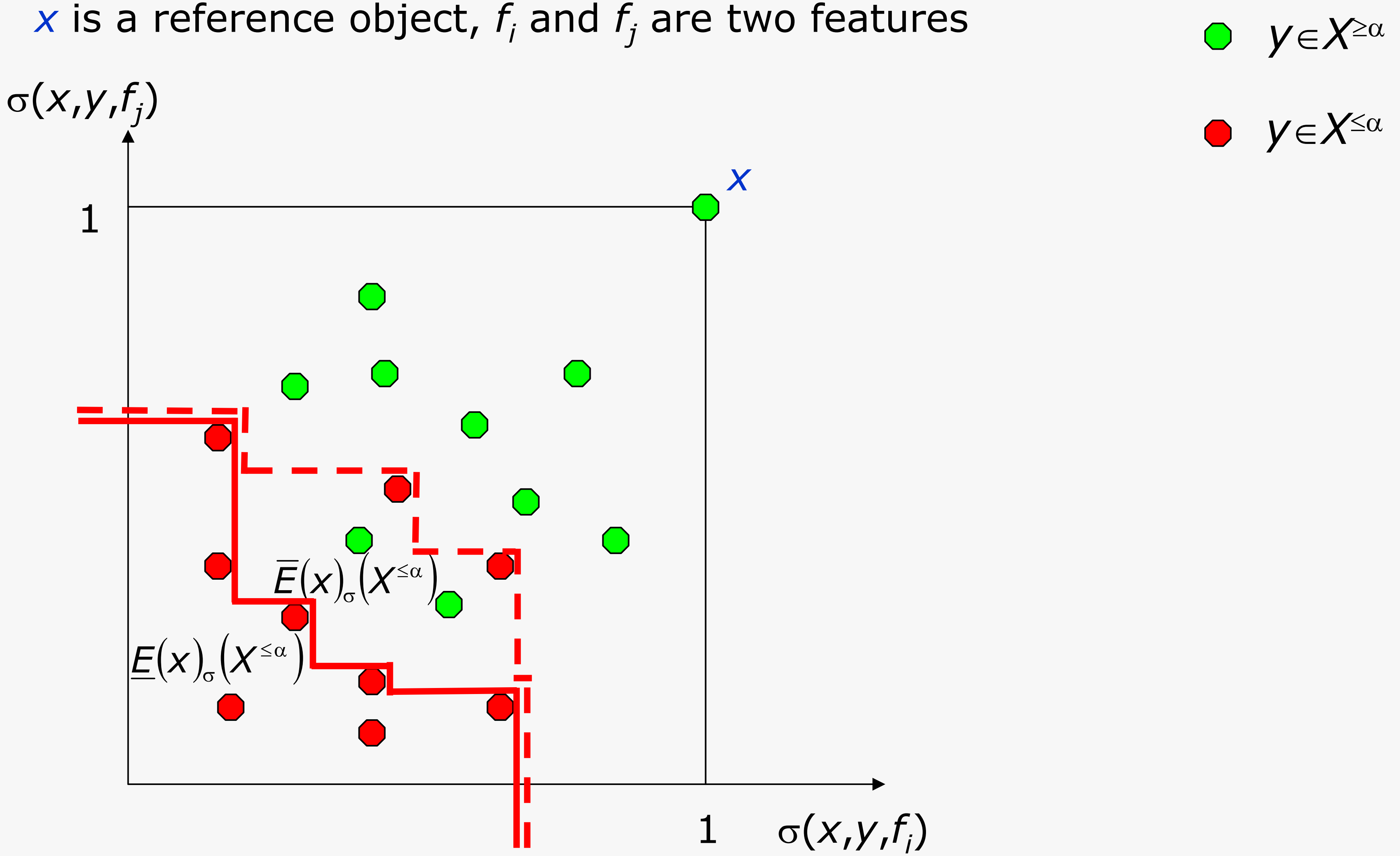
Similarity space in CBR

x is a reference object, f_i and f_j are two features



Similarity space in CBR

x is a reference object, f_i and f_j are two features



CBR-DRSA decision rules

- Decision rules induced by DRSA from pairwise fuzzy information base:

  $\underline{E}(x)_\sigma\left(X^{\geq\alpha}\right)$

  „if object w is similar to object x w.r.t. feature f_{i1} to degree at least h_{i1},
  and … and w.r.t. feature f_{im} to degree at least h_{im},
  then object w certainly belongs to set X to degree at least α"

  $\overline{E}(x)_\sigma\left(X^{\geq\alpha}\right)$

  „if object w is similar to object x w.r.t. feature f_{i1} to degree at least h_{i1},
  and … and w.r.t. feature f_{im} to degree at least h_{im},
  then object w possibly belongs to set X to degree at least α"

  where $\{f_{i1},\ldots,f_{im}\}=E$ and $h_{i1},\ldots,h_{im}\in[0, 1]$

CBR-DRSA decision rules

- Decision rules induced by DRSA from pairwise fuzzy information base:

$\underline{E}(x)_\sigma\left(X^{\leq\alpha}\right)$

„if object w is similar to object x w.r.t. feature f_{i1} to degree at most h_{i1}, and … and w.r.t. feature f_{im} to degree at most h_{im}, then object w certainly belongs to set X to degree at most α"

$\overline{E}(x)_\sigma\left(X^{\leq\alpha}\right)$

„if object w is similar to object x w.r.t. feature f_{i1} to degree at most h_{i1}, and … and w.r.t. feature f_{im} to degree at most h_{im}, then object w possibly belongs to set X to degree at most α"

where $\{f_{i1},\ldots,f_{im}\}=E$ and $h_{i1},\ldots,h_{im}\in[0, 1]$

Comparison of CBR-DRSA decision rules and CBR-gradual rules

- CBR-gradual rules: $s(z,x)\geq\alpha \Rightarrow t(z,x)\geq\alpha$

  where s and t measure the credibility of similarity with respect to condition attribute and decision attribute, respectively

- Advantages of CBR-DRSA decision rules:
  - The CBR-DRSA decision rules do not need the aggregation (always subjective and arbitrary to some extent) of the similarity w.r.t. different features in one comprehensive similarity function
  - The CBR-DRSA decision rules permit to consider different thresholds for degrees of credibility in the premise and in the conclusion

Other extensions of DRSA

- DRSA as a way of handling fuzzy-rough hybridization
- DRSA for choice and ranking with graded preference relations
- DRSA for choice and ranking with Lorenz dominance relation
- DRSA for decision with multiple decision makers
- DRSA with missing values of attributes and criteria
- DRSA for hierarchical decision making
- Discovering association rules in preference-ordered data sets

DRSA as a Way of Handling Fuzzy-Rough Hybridization

DRSA as a proper way of handling graduality in Rough Set Theory

- Rough set concept refers to some ideas of Leibniz (indiscernibility), Frege (vague concepts), Boole (reasoning methods) and Bayes (inductive reasoning)

- Gottfried Leibniz (Leibniz's law)
  „identity of indiscernibles":

  > *if **x** and **y** are indiscernible* (i.e. ***x*** *and* ***y*** *have the same properties*), *then* ***x*=*y***

  „indiscernibility of identicals":

  > *if* ***x*=*y***, *then* ***x*** *and* ***y*** *are indiscernible* (i.e. ***x*** *and* ***y*** *have the same properties*)

- Rough set theory by Zdzisław Pawlak uses Leibniz's law to classify objects falling under the same concept – reformulation of the „identity of indiscernibles":

  > *if **x** and **y** are indiscernible, then **x** and **y** **belong** to the same class*

  „Indiscernibility of identicals" cannot be reformulated analogously, because it is not true that *if **x** and **y** belong to the same class, then **x** and **y** are indiscernible*

- Rough set theory needs a still weaker form of „identity of indiscernibles"

DRSA as a proper way of handling graduality in Rough Set Theory

- According to Gottlob Frege:
  „A concept must have a sharp boundary.
  To the (vague) concept without a sharp boundary there would correspond an area that had not a sharp boundary-line all around"

- Following this intuition, one can further reformulate the „identity of indiscernibles":

  *if **x** and **y** are indiscernible, then **x** and **y** **should belong** to the same class*

  This formulation implies that there is an inconsistency if ***x*** and ***y*** are indiscernible and ***x*** and ***y*** belong to different classes

- The contribution of the ideas of Leibniz and Frege to the Pawlak's rough set should be completed by the idea of Georg Boole concerning presence (truth) or absence (falsity) of a property for an object

- It is natural, moreover, to weaken this principle by considering that a property can be present (true) **to some degree** (graduality)

DRSA as a proper way of handling graduality in Rough Set Theory

- The graduality of truth was considered by Jan Łukasiewicz in multi-valued logic, and then by Lotfi Zadeh within fuzzy set theory, where graduality concerns membership to a set

- Any proposal of **putting rough sets and fuzzy sets together** can be seen as a reconstruction of the rough set concept, where the Boole's binary logic is substituted by Łukasiewicz's multi-valued logic, such that the Leibniz's identity of indiscernibles and the Frege's intuition about vagueness are combined through the idea that a property is true to some degree:

*if the degree of each property for **x** is greater than or equal to the degree for **y**,*
*then **x** **should belong** to the considered class **in degree at least as high as y***

- This formulation is perfectly concordant with our **Dominance-based Rough Set Approach** – it handles the monotonic relationship in exacly the same way

Remarks on fuzzy extensions of rough sets

- Cattaneo 1998; Dubois & Prade 1992; Lin 1992; Greco, Matarazzo & Słowiński 1999, 2000; Inuiguchi & Tanino 2002; Morsi & Yakout 1998; Nakamura & Gao 1991; Polkowski 2002, Słowiński 1995; Słowiński & Stefanowski 1996; Yao 1997; Radzikowska & Kerre 2003; Thiele 2000; Wu, Mi & Zhang 2003; ...

- The fuzzy extensions of Pawlak's definition of lower and upper approximations use fuzzy connectives (t-norm, t-conorm, fuzzy implication)

- There is no "right" connective

- In general, fuzzy connectives depend on cardinal properties of membership degrees, i.e. the result is sensitive to order preserving transformation of membership degrees

Remarks on fuzzy extensions of rough sets

- A natural question arises: is it reasonable to expect from membership degree a cardinal content instead of ordinal only?
- In other words, is it realistic to think that human is able to express in a meaningful way not only that

  "object x belongs to fuzzy set X more likely than object y"

  but even something like

  "object x belongs to fuzzy set X two times more likely than object y"?

S.Greco, M.Inuiguchi, R.Słowiński: Fuzzy rough sets and multiple-premise gradual decision rules. *International Journal of Approximate Reasoning*, 41 (2005) 179-211

Remarks on fuzzy extensions of rough sets

- The dominance based rough approximation of a fuzzy set avoids arbitrary choice of fuzzy connectives and not meaningful operations on membership degrees
- Approximation of knowledge about *Y* using knowledge about *X* is based on positive or negative relationships between premises and conclusions, called *gradual rules*, i.e.:

  i) „the more *x* is *X*, the more it is *Y*" (positive relationship)

  ii) „the more *x* is *X*, the less it is *Y*" (negative relationship)
- Example:

  „the larger the market share of a company, the larger its profit"

  „the larger the debt of a company, the smaller its profit"

DRSA as an approach to computing with words

- Classical fuzzy set approach to **computing with words**:

  i) qualitative inputs, such as „very bad", „bad", „medium", „good", „very good"

  ii) numerical codification of the inputs (fuzzification): e.g.

  „very bad"=0, „bad"=0.25, „medium"=0.5, „good"=0.75, „very good"=1

  iii) algebraic operations on numerical codes : e.g.
  „comprehensive evalaution of a student good in mathematics and medium in physics"=(0.75+0.5)/2=0.625

  iv) recodification in qualitiative terms of the calculation result (defuzzification): e.g., 0.625=between medium and good

- Dominance-based Rough Set Approach does not need fuzzification and defuzzification: e.g.

  „*if* the student is at least medium in Mathematics *and* at least medium in Literature, *then* the student is at least medium"

Classical rough set as a particular case of dominance-based rough approximation of a fuzzy set

Classical rough set as a particular case of dominance-based rough approximation of a fuzzy set

- Given $E \subseteq U$, for each set $X \subseteq U$, we can define its upward lower approximation and its upward upper approximation :

$$\underline{E}^{(>)}(X) = \{x \in U : D_E^{\pm}(x) \subseteq X\} = \bigcup_{x \in U} \{D_E^{\pm}(x) : D_E^{\pm}(x) \subseteq X\}$$

$$\overline{E}^{(>)}(X) = \{x \in U : D_E^{\mp}(x) \cap X \neq \varnothing\} = \bigcup_{x \in U} \{D_E^{\pm}(x) : D_E^{\mp}(x) \cap X \neq \varnothing\}$$

- Analogously, we can define downward lower approximation and downward upper approximation of set $X \subseteq U$:

$$\underline{E}^{(<)}(X) = \{x \in U : D_E^{\mp}(x) \subseteq X\} = \bigcup_{x \in U} \{D_E^{\mp}(x) : D_E^{\mp}(x) \subseteq X\}$$

$$\overline{E}^{(<)}(X) = \{x \in U : D_E^{\pm}(x) \cap X \neq \varnothing\} = \bigcup_{x \in U} \{D_E^{\mp}(x) : D_E^{\pm}(x) \cap X \neq \varnothing\}$$

- The above approximations can be used to analyse data relative to gradual membership of objects to some concepts representing properties on one hand, and decision classes, on the other hand

Classical rough set as a particular case of dominance-based rough approximation of a fuzzy set

- Classical rough sets based on indiscernibility relation :

$$I_P = \{(x,y) \in U \times U : f(x,q)=f(y,q), \text{ for each } q \in P\}$$

$$I_P(x) = \{y \in U : f(x,q)=f(y,q), \text{ for each } q \in P\}$$

- For information table $\boldsymbol{S}$=<U, Q, V, f>, for set $X \subseteq U$ and for subset $P \subseteq Q$, the P-lower and the P-upper approximations of X are defined as follows :

$$\underline{P}(X) = \{x \in U : I_P(x) \subseteq X\}$$
$$\overline{P}(X) = \{x \in U : I_P(x) \cap X \neq \emptyset\}$$

- Let $\boldsymbol{B}$=<U, F, φ> be a Boolean information base, where $\varphi : U \times F \to \{0,1\}$

- Partition $\boldsymbol{F}=\{F_1,\dots,F_r\}$ of the set of properties F is called canonical, if for each $x \in U$ and for each $F_k \subseteq F$, $k=1,\dots,r$, there exists only one $f_j \in F_k$ such that $\varphi(x,f_j)=1$, and for all the others, $\varphi(x,f_h)=0$ ($h \neq j$) (N.B. $card(F_k) \geq 2$, $k=1,\dots,r$)

Classical rough set as a particular case of dominance-based rough approximation of a fuzzy set

- Any information table $\boldsymbol{S}$=<U, Q, V, f>, can be interpreted as a Boolean information base $\boldsymbol{B}$=<U, F, φ>, such that to each $v \in V_q$ there corresponds one property $f_{qv} \in F$ for which $\varphi(x,f_{qv})=1$ if $f(x,q)=v$, and $\varphi(x,f_{qv})=0$ otherwise
- $\boldsymbol{F}=\{F_1,\ldots,F_r\}$, with $F_q=\{f_{qv},\ v \in V_q\}$, $q \in Q$, is a canonical partition of F
- **Theorem** (Greco, Matarazzo, Słowiński 2006):
  Let $P \subseteq Q$ and let E^P be the set of all properties f_{qv} corresponding to values $v \in V_q$ for each attribute $q \in P$; for each set $X \subseteq U$, we have

$$\underline{E}^{P(>)}(X) = \underline{E}^{P(<)}(X) = \underline{P}(X)$$

$$\overline{E}^{P(>)}(X) = \overline{E}^{P(<)}(X) = \overline{P}(X)$$

- In fact,

$$D^{+}_{E^P}(x) = I_P(x)$$

$$D^{-}_{E^P}(x) = I_P(x)$$

DRSA for Multiple Decision Makers (DRSA-MDM)

Multiple Criteria Classification by Multiple Decision Makers

- Classification of objects described by multiple criteria is done by Multiple Decision Makers (MDM)
- Previous studies concentrated on convergence toward a consensus decision minimizing dissimilarities w.r.t. decisions of MDM (e.g. Inuiguchi, Miyajima 2006; Jelassi, Kersten, Zionts 1990; Nurmi, Kacprzyk, Fedrizzi 1996)
- Instead of supporting negotiation between MDM, we want to define conditions for a given scenario of a consensus decision, expressed in terms of decision rules
- To this aim, we extend the Dominance-based Rough Set Approach by introducing concepts related to dominance w.r.t. minimal profiles of evaluations given by MDM

Multiple Criteria Classification by Multiple Decision Makers

- Example: students described by scores (1–20) in mathematics (M), physics (Ph) and literature (L) are classified by 3 professors (P1, P2, P3) to preference ordered classes: Bad, Medium, Good
- Decisions of P1, P2, P3 have to be aggregated so as to designate students which will be finally accepted for a graduate program
- The aggregate decision represents a consensus between professors
- Possible consenus:
  - 2 professors classify as „at least Medium” + 1 professor classifies as „Good” [Medium, Medium, Good], [Medium, Good, Medium], [Good, Medium, Medium]
- Resulting rules, e.g.:

if student x gained at least 15 in M, and at least 18 in L, then x is accepted

if student x gained at most 10 in M, and at most 13 in Ph, then x is not accepted

DRSA for Multiple Decision Makers – definitions

- Set of criteria: $C=\{1,\ldots,q,\ldots,m\}$
- Set of decision makers (DM): $H=\{1,\ldots,i,\ldots,h\}$ (h decision attributes)
- Set of preference ordered classes for each DM $i \in H$:

$$\boldsymbol{Cl}_i=\{Cl_{t,i}\,,\ t\in T_i\},\quad T_i=\{1,\ldots,n_i\}$$

$$\bigcup_{t=1}^{n_i} Cl_{t,i} = U,\quad Cl_{t,i} \cap Cl_{r,i} = \varnothing,\quad \text{for all } r,t \in T_i$$

if $x\in Cl_{r,i}$, $y\in Cl_{s,i}$ and $r>s$, then x is better than y for DM $i\in H$

- For a single DM $i\in H$, the sets to be approximated are the upward and the downward unions of decision classes ($t=1,\ldots,n_i$):

$$Cl^{\geq}_{t,i} = \bigcup_{s\geq t} Cl_{s,i} \quad \text{(at least class } Cl_{t,i}\text{)}$$

$$Cl^{\leq}_{t,i} = \bigcup_{s\leq t} Cl_{s,i} \quad \text{(at most class } Cl_{t,i}\text{)}$$

DRSA for Multiple Decision Makers – definitions

- Considering the set of DMs as a whole, we need new concepts concerning minimal or maximal evaluation profiles,
  i.e. vectors of names of decision classes used by particular DMs

- Upward multi-union with respect to one configuration [t(1),…,t(h)]:

$$Cl^{\geq}_{[t(1),\dots,t(h)]} = \bigcap_{i \in H} Cl^{\geq}_{t(i),i}$$

- Downward multi-union with respect to one configuration [t(1),…,t(h)]:

$$Cl^{\leq}_{[t(1),\dots,t(h)]} = \bigcap_{i \in H} Cl^{\leq}_{t(i),i}$$

- Configuration [t(1),…,t(h)] means evaluation profile by h DMs

- E.g. Upward multi-union w.r.t. [*Bad*, *Medium*, *Average*] includes objects qualified as at least *Bad* by the 1st DM, and at least *Medium* by the 2nd DM, and at least *Average* by the 3rd DM

DRSA for Multiple Decision Makers – definitions

- Upward mega-union with respect to k configurations, $k=1,\dots,\prod_{i=1}^{h} n_i$, $\{[t_1(1),\dots,t_1(h)],\dots,[t_k(1),\dots,t_k(h)]\}$:

$$Cl^{\geq}_{\{[t_1(1),\dots,t_1(h)],\dots,[t_k(1),\dots,t_k(h)]\}} = \bigcup_{r=1}^{k} Cl^{\geq}_{[t_r(1),\dots,t_r(h)]}$$

- Downward mega-union with respect to k configurations, $k=1,\dots,\prod_{i=1}^{h} n_i$, $\{[t_1(1),\dots,t_1(h)],\dots,[t_k(1),\dots,t_k(h)]\}$:

$$Cl^{\leq}_{\{[t_1(1),\dots,t_1(h)],\dots,[t_k(1),\dots,t_k(h)]\}} = \bigcup_{r=1}^{k} Cl^{\leq}_{[t_r(1),\dots,t_r(h)]}$$

- $\prod_{i=1}^{h} n_i$ is the maximum number of all possible configurations $[t(1),\dots,t(h)]$, i.e. combinations of class names by particular DMs

- E.g. for 2 configurations [*Bad*, *Medium*, *Average*] and [*Medium*, *Bad*, *Average*], the upward mega-union includes objects qualified as at least *Bad* by the 1st DM, and at least *Medium* by the 2nd DM, and at least *Average* by the 3rd DM, **PLUS** objects qualified as at least *Medium* by the 1st DM, and at least *Bad* by the 2nd DM, and at least *Average* by the 3rd DM

DRSA for Multiple Decision Makers – definitions

- Using the concept of a mega-union, one can model a collective decision of majority type,
  e.g. for 3 DMs and YES/NO voting decisions for the objects,
  a „majority" mega-union is composed of such objects that at least 2 DMs voted YES for them: $Cl^{\geq}_{\{[YES,YES,NO],\ [YES,NO,YES],\ [NO,YES,YES]\}}$

- Principle of consistent representation of multi-unions: for any $P \subseteq C$
  - $x \in U$ belongs to $Cl^{\geq}_{[t(1),\ldots,t(h)]}$ without inconsistency if $x \in Cl^{\geq}_{[t(1),\ldots,t(h)]}$ and, for all $y \in U$ dominating x on P, also y belongs to $Cl^{\geq}_{[t(1),\ldots,t(h)]}$, i.e.

$$D_P^+(x) \subseteq Cl^{\geq}_{[t(1),\ldots,t(h)]}$$

  - $x \in U$ could belong to $Cl^{\geq}_{[t(1),\ldots,t(h)]}$ if there existed at least one $y \in Cl^{\geq}_{[t(1),\ldots,t(h)]}$ such that x dominates y on P, i.e.

$$x \in D_P^+(y)$$

DRSA for Multiple Decision Makers – definitions

- *P*-lower approximation of upward multi-union $Cl^{\geq}_{[t(1),\ldots,t(h)]}$:

$$\underline{P}\left(Cl^{\geq}_{[t(1),\ldots,t(h)]}\right) = \left\{x \in U :\ D_P^+(x) \subseteq Cl^{\geq}_{[t(1),\ldots,t(h)]}\right\}$$

- *P*-upper approximation of upward multi-union $Cl^{\geq}_{[t(1),\ldots,t(h)]}$:

$$\overline{P}\left(Cl^{\geq}_{[t(1),\ldots,t(h)]}\right) = \bigcup_{x \in Cl^{\geq}_{[t(1),\ldots,t(h)]}} D_P^+(x)$$

- Analogously, for downward multi-union $Cl^{\leq}_{[t(1),\ldots,t(h)]}$:

$$\underline{P}\left(Cl^{\leq}_{[t(1),\ldots,t(h)]}\right) = \left\{x \in U :\ D_P^-(x) \subseteq Cl^{\leq}_{[t(1),\ldots,t(h)]}\right\}$$

$$\overline{P}\left(Cl^{\leq}_{[t(1),\ldots,t(h)]}\right) = \bigcup_{x \in Cl^{\leq}_{[t(1),\ldots,t(h)]}} D_P^-(x)$$

DRSA for Multiple Decision Makers – definitions

- **Theorem 1**. For all $P \subseteq C$ and for any configuration $[t(1),\dots,t(h)]$:

$$\underline{P}\left(Cl^{\geq}_{[t(1),\dots,t(h)]}\right) = \bigcap_{i=1}^{h} \underline{P}\left(Cl^{\geq}_{t(i),i}\right), \qquad \overline{P}\left(Cl^{\geq}_{[t(1),\dots,t(h)]}\right) = \bigcup_{i=1}^{h} \overline{P}\left(Cl^{\geq}_{t(i),i}\right)$$

$$\underline{P}\left(Cl^{\leq}_{[t(1),\dots,t(h)]}\right) = \bigcap_{i=1}^{h} \underline{P}\left(Cl^{\leq}_{t(i),i}\right), \qquad \overline{P}\left(Cl^{\leq}_{[t(1),\dots,t(h)]}\right) = \bigcup_{i=1}^{h} \overline{P}\left(Cl^{\leq}_{t(i),i}\right)$$

„and”

„or”

DRSA for Multiple Decision Makers – definitions

- P-lower approximation of upward mega-union $Cl^{\geq}_{\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}}$

$$\underline{P}\left(Cl^{\geq}_{\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}}\right)=\left\{x\in U:\ D_P^{+}(x)\subseteq Cl^{\geq}_{\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}}\right\}$$

- P-upper approximation of upward mega-union $Cl^{\geq}_{\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}}$

$$\overline{P}\left(Cl^{\geq}_{\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}}\right)=\bigcup_{x\in Cl^{\geq}_{\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}}} D_P^{+}(x)$$

- Analogously, for downward mega-union $Cl^{\leq}_{\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}}$

$$\underline{P}\left(Cl^{\leq}_{\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}}\right)=\left\{x\in U:\ D_P^{-}(x)\subseteq Cl^{\leq}_{\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}}\right\}$$

$$\overline{P}\left(Cl^{\leq}_{\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}}\right)=\bigcup_{x\in Cl^{\leq}_{\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}}} D_P^{-}(x)$$

DRSA for Multiple Decision Makers – definitions

- **Theorem 2**. For all $P \subseteq C$ and for any set of configurations

$\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}$:

$$\underline{P}\left(Cl^{\geq}_{\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}}\right) = \bigcup_{r=1}^{k} \underline{P}\left(Cl^{\geq}_{[t_r(1),\ldots,t_r(h)]}\right)$$

$$\overline{P}\left(Cl^{\geq}_{\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}}\right) = \bigcup_{r=1}^{k} \overline{P}\left(Cl^{\geq}_{[t_r(1),\ldots,t_r(h)]}\right)$$

$$\underline{P}\left(Cl^{\leq}_{\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}}\right) = \bigcup_{r=1}^{k} \underline{P}\left(Cl^{\leq}_{[t_r(1),\ldots,t_r(h)]}\right)$$

$$\overline{P}\left(Cl^{\leq}_{\{[t_1(1),\ldots,t_1(h)],\ldots,[t_k(1),\ldots,t_k(h)]\}}\right) = \bigcup_{r=1}^{k} \overline{P}\left(Cl^{\leq}_{[t_r(1),\ldots,t_r(h)]}\right)$$

„or”

DRSA for multiple DMs – properties

- Each upward union $Cl_{t,i}^{\geq}$ is a particular upward multi-union:

$$Cl_{t,i}^{\geq} = Cl_{[1,\ldots,t(i),\ldots,1]}^{\geq}$$

- Each upward multi-union $Cl_{[t(1),\ldots,t(h)]}^{\geq}$ is a particular upward mega-union:

$$Cl_{[t(1),\ldots,t(h)]}^{\geq} = Cl_{\{[t(1),\ldots,t(h)]\}}^{\geq}$$

- All properties of mega-unions also hold for multi-unions and for single DM

- We present properties for all kinds of upward unions – the properties for all downward unions are analogous

DRSA for multiple DMs – property of inclusion

- Property of inclusion and the associated order relation between upward and downward unions, multi-unions and mega-unions
- There is an isomorphism between inclusion relation $\subseteq$ on the set of all upward unions $\boldsymbol{Cl}^{\geq} = \{Cl_t^{\geq},\ t \in T\}$ and order relation $\geq$ on the set of class indices $T = \{1,\ldots,n\}$:

$$Cl_r^{\geq} \subseteq Cl_s^{\geq} \Leftrightarrow r \geq s$$

- Inclusion relation $\subseteq$ on $\boldsymbol{Cl}^{\geq}$ is a complete preorder (strongly complete & transitive)

DRSA for multiple DMs – property of inclusion

- There is an isomorphism between inclusion relation $\subseteq$ on the set of all upward multi-unions

$$\boldsymbol{Cl}^{\geq \Pi} = \left\{ Cl^{\geq}_{[t(1),\dots,t(h)]},\ [t(1),\dots,t(h)] \in \prod_{i=1}^{h} T_i \right\}$$

and order relation $\geqq$ on the Cartesian product of class indices $\prod_{i=1}^{h} T_i$ expressed as follows:

for any two configurations $\boldsymbol{t}^1 = [t^1(1),\dots,t^1(h)], \boldsymbol{t}^2 = [t^2(1),\dots,t^2(h)] \in \prod_{i=1}^{h} T_i$

$$Cl^{\geq}_{\boldsymbol{t}^1} \subseteq Cl^{\geq}_{\boldsymbol{t}^2} \Leftrightarrow \boldsymbol{t}^1 \geqq \boldsymbol{t}^2$$

- The order relation $\geqq$ is the dominance relation (partial preorder) in the set of all configurations

- Inclusion relation $\subseteq$ on $\boldsymbol{Cl}^{\geq \Pi}$ is a partial preorder (reflexive & transitive)

DRSA for multiple DMs – property of inclusion

- For all two configurations $\boldsymbol{t}^1, \boldsymbol{t}^2 \in \prod_{i=1}^{h} T_i$, $\boldsymbol{t}^1 \geqq \boldsymbol{t}^2 \Leftrightarrow \boldsymbol{t}^2 \leqq \boldsymbol{t}^1$, which implies:

$$Cl_{\boldsymbol{t}^1}^{\geq} \subseteq Cl_{\boldsymbol{t}^2}^{\geq} \Leftrightarrow Cl_{\boldsymbol{t}^2}^{\leq} \subseteq Cl_{\boldsymbol{t}^1}^{\leq}$$

DRSA for multiple DMs – property of inclusion

- For upward mega-unions, we consider order relation $\langle \geqq \rangle$ defined in the power set of h-dimensional real space $2^{\boldsymbol{R}^h}$
- For any two sets of k_1 and k_2 configurations

$$\langle \boldsymbol{x}^1 \rangle = \left\{ \left[x_1^{1,1}, \ldots, x_h^{1,1} \right], \ldots, \left[x_1^{1,k_1}, \ldots, x_h^{1,k_1} \right] \right\},$$

$$\langle \boldsymbol{x}^2 \rangle = \left\{ \left[x_1^{2,1}, \ldots, x_h^{2,1} \right], \ldots, \left[x_1^{2,k_2}, \ldots, x_h^{2,k_2} \right] \right\} \in 2^{\boldsymbol{R}^h}$$

the order relation $\langle \geqq \rangle$ holds:

$\langle \boldsymbol{x}^1 \rangle \langle \geqq \rangle \langle \boldsymbol{x}^2 \rangle \Leftrightarrow$ for each $\left[x_1^{1,i}, \ldots, x_h^{1,i} \right] \in \langle \boldsymbol{x}^1 \rangle$ there exists $\left[x_1^{2,j}, \ldots, x_h^{2,j} \right] \in \langle \boldsymbol{x}^2 \rangle$

such that $\left[x_1^{1,i}, \ldots, x_h^{1,i} \right] \geqq \left[x_1^{2,j}, \ldots, x_h^{2,j} \right]$, $i = 1, \ldots, k_1$, $j = 1, \ldots, k_2$

- Similarily,

$\langle \boldsymbol{x}^2 \rangle \langle \leqq \rangle \langle \boldsymbol{x}^1 \rangle \Leftrightarrow$ for each $\left[x_1^{2,j}, \ldots, x_h^{2,j} \right] \in \langle \boldsymbol{x}^2 \rangle$ there exists $\left[x_1^{1,i}, \ldots, x_h^{1,i} \right] \in \langle \boldsymbol{x}^1 \rangle$

such that $\left[x_1^{2,j}, \ldots, x_h^{2,j} \right] \leqq \left[x_1^{1,i}, \ldots, x_h^{1,i} \right]$, $i = 1, \ldots, k_1$, $j = 1, \ldots, k_2$

DRSA for multiple DMs – property of inclusion

- The order relations $\langle\geqq\rangle$ and $\langle\leqq\rangle$ on $2^{\boldsymbol{R}^h}$ are independent:

$$\langle\boldsymbol{x}^1\rangle\langle\geqq\rangle\langle\boldsymbol{x}^2\rangle \text{ is not equivalent to } \langle\boldsymbol{x}^2\rangle\langle\leqq\rangle\langle\boldsymbol{x}^1\rangle$$

- E.g. $h=2$, $\langle\boldsymbol{x}^1\rangle=\{[3,3]\}$ and $\langle\boldsymbol{x}^2\rangle=\{[1,2],[4,1]\}$

  Then $\langle\boldsymbol{x}^1\rangle\langle\geqq\rangle\langle\boldsymbol{x}^2\rangle$, because for $[3,3]\in\langle\boldsymbol{x}^1\rangle$ there exists $[1,2]\in\langle\boldsymbol{x}^2\rangle$

  such that $[3,3]\geqq[1,2]$,

  but $\langle\boldsymbol{x}^2\rangle\langle\leqq\rangle\langle\boldsymbol{x}^1\rangle$ does not hold because for $[4,1]\in\langle\boldsymbol{x}^2\rangle$

  there is no configuration $\boldsymbol{x}^{1,i}\in\langle\boldsymbol{x}^1\rangle$ such that $[4,1]\leqq\boldsymbol{x}^{1,i}$

  (in fact, $[4,1]\leqq[3,3]$ is not true)

DRSA for multiple DMs – property of inclusion

- There is an isomorphism between inclusion relation $\subseteq$ on the set of all upward mega-unions

$$\boldsymbol{Cl}^{\geq 2^{\Pi}} = \left\{ Cl^{\geq}_{\{[t_1(1),\dots,t_1(h)],\dots,[t_k(1),\dots,t_k(h)]\}},\ [t_1(1),\dots,t_1(h)],\dots,[t_k(1),\dots,t_k(h)] \in \prod_{i=1}^{h} T_i \right\}$$

and order relation $\langle\geqq\rangle$ on the power set of Cartesian product of class indices $2^{\prod_{i=1}^{h} T_i}$ expressed as follows:

for any two sets of k_1 and k_2 configurations

$$\langle \boldsymbol{t}^1 \rangle = \left\{ [t_1^1(1),\dots,t_1^1(h)],\dots,[t_{k_1}^1(1),\dots,t_{k_1}^1(h)] \right\},$$

$$\langle \boldsymbol{t}^2 \rangle = \left\{ [t_1^2(1),\dots,t_1^2(h)],\dots,[t_{k_2}^2(1),\dots,t_{k_2}^2(h)] \right\} \in 2^{\prod_{i=1}^{h} T_i}$$

$$Cl^{\geq}_{\langle \boldsymbol{t}^1 \rangle} \subseteq Cl^{\geq}_{\langle \boldsymbol{t}^2 \rangle} \Leftrightarrow \langle \boldsymbol{t}^1 \rangle \langle\geqq\rangle \langle \boldsymbol{t}^2 \rangle$$

- Inclusion relation $\subseteq$ on $\boldsymbol{Cl}^{\geq 2^{\Pi}}$ is a partial preorder, however,

$Cl^{\geq}_{\langle \boldsymbol{t}^1 \rangle} \subseteq Cl^{\geq}_{\langle \boldsymbol{t}^2 \rangle}$ is not equivalent to $Cl^{\leq}_{\langle \boldsymbol{t}^2 \rangle} \subseteq Cl^{\leq}_{\langle \boldsymbol{t}^1 \rangle}$

DRSA for multiple DMs – properties

- The upward mega-unions satisfy the basic properties of rough approximations:

  for all $P \subseteq R \subseteq C$, and for all $\langle \boldsymbol{t} \rangle \in 2^{\prod_{i=1}^{h} T_i}$,

  - Rough inclusion

  $$\underline{P}\left(Cl_{\langle \mathbf{t} \rangle}^{\geq}\right) \subseteq Cl_{\langle \mathbf{t} \rangle}^{\geq} \subseteq \overline{P}\left(Cl_{\langle \mathbf{t} \rangle}^{\geq}\right)$$

  - Complementarity

  $$\underline{P}\left(Cl_{\langle \mathbf{t} \rangle}^{\geq}\right) = U - \overline{P}\left(Cl_{\langle \mathbf{t'} \rangle}^{\leq}\right) \quad \text{where} \quad Cl_{\langle \mathbf{t'} \rangle}^{\leq} = U - Cl_{\langle \mathbf{t} \rangle}^{\geq}$$

  - Monotonicity

  $$\underline{P}\left(Cl_{\langle \mathbf{t} \rangle}^{\geq}\right) \subseteq \underline{R}\left(Cl_{\langle \mathbf{t} \rangle}^{\geq}\right) \quad \text{and} \quad \overline{P}\left(Cl_{\langle \mathbf{t} \rangle}^{\geq}\right) \supseteq \overline{R}\left(Cl_{\langle \mathbf{t} \rangle}^{\geq}\right)$$

Conclusions to DRSA for Multiple Decision Makers

- DRSA for Multiple Decision Makers is based on a new definition of dominance w.r.t. profiles of classification (configurations) made by DMs
- DRSA for MDM permits to characterize conditions for objects to reach a given consensus
- These conditions are expressed in terms of decision rules
- Premises are formulated in multiple-criteria evaluation space
- Conclusions are formulated in multiple-DMs classification space
- DRSA for MDM exploits ordinal information only, and decision rules do not convert ordinal information into numeric one
- DRSA for MDM does not search for concordant decision rules for multiple DMs considered as individuals but rather characterizes conditions for a consensus attainable for multiple DMs considered as a whole

Other extensions of DRSA

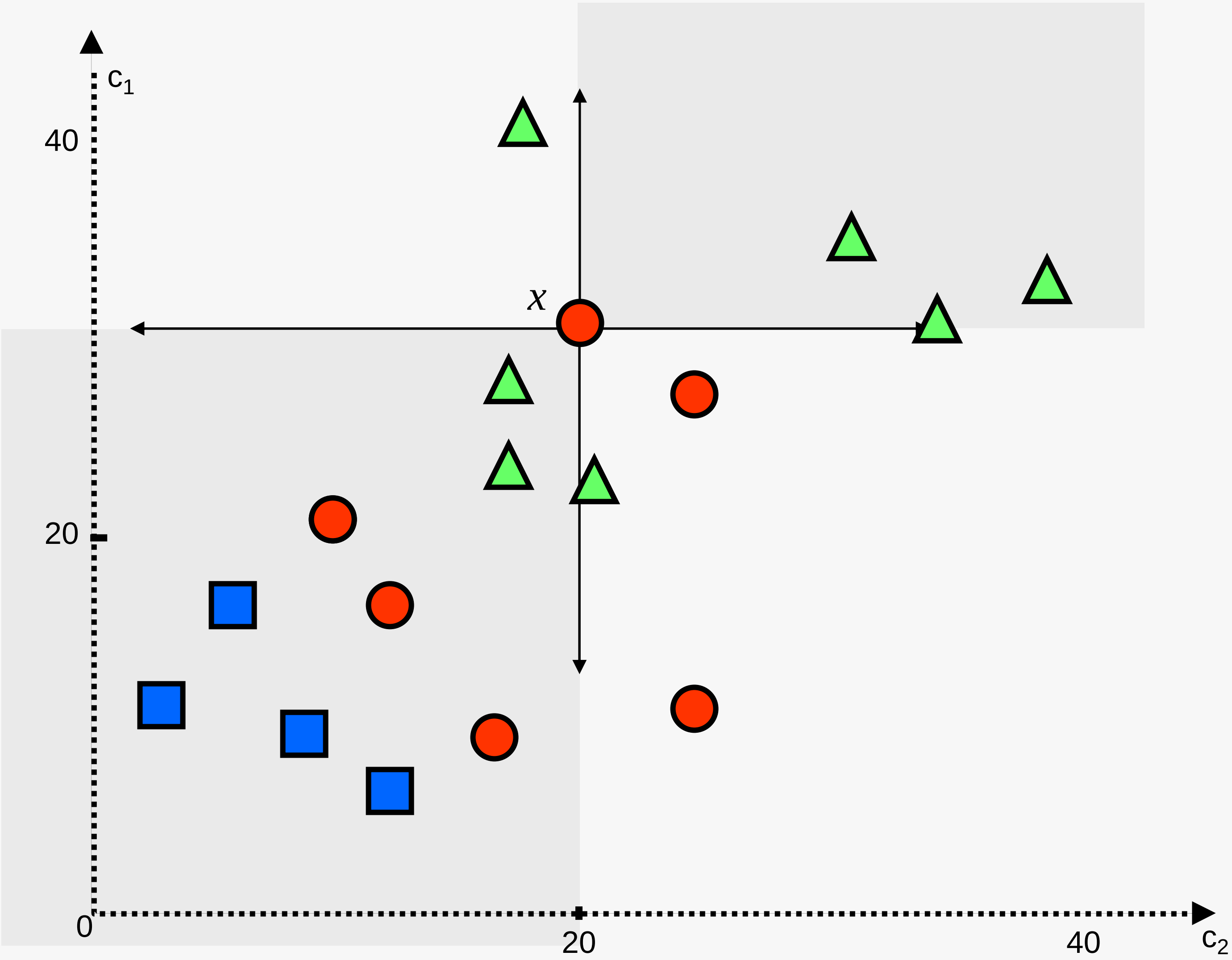
Extensions of DRSA dealing with preference-ordered data

- **Missing values of attributes and criteria**

Investments ↑	Sales ↑	Effectiveness ↑
40	17,8	High
35	30	High
32.5	39	High
31	35	High
27.5	17.5	High
24	17.5	High
22.5	20	High
*	19	Medium
27	25	Medium
21	9.5	Medium
18	12.5	Medium
10.5	25.5	Medium
9.75	17	Medium
17.5	5	Low
11	2	Low
10	9	Low
5	13	Low

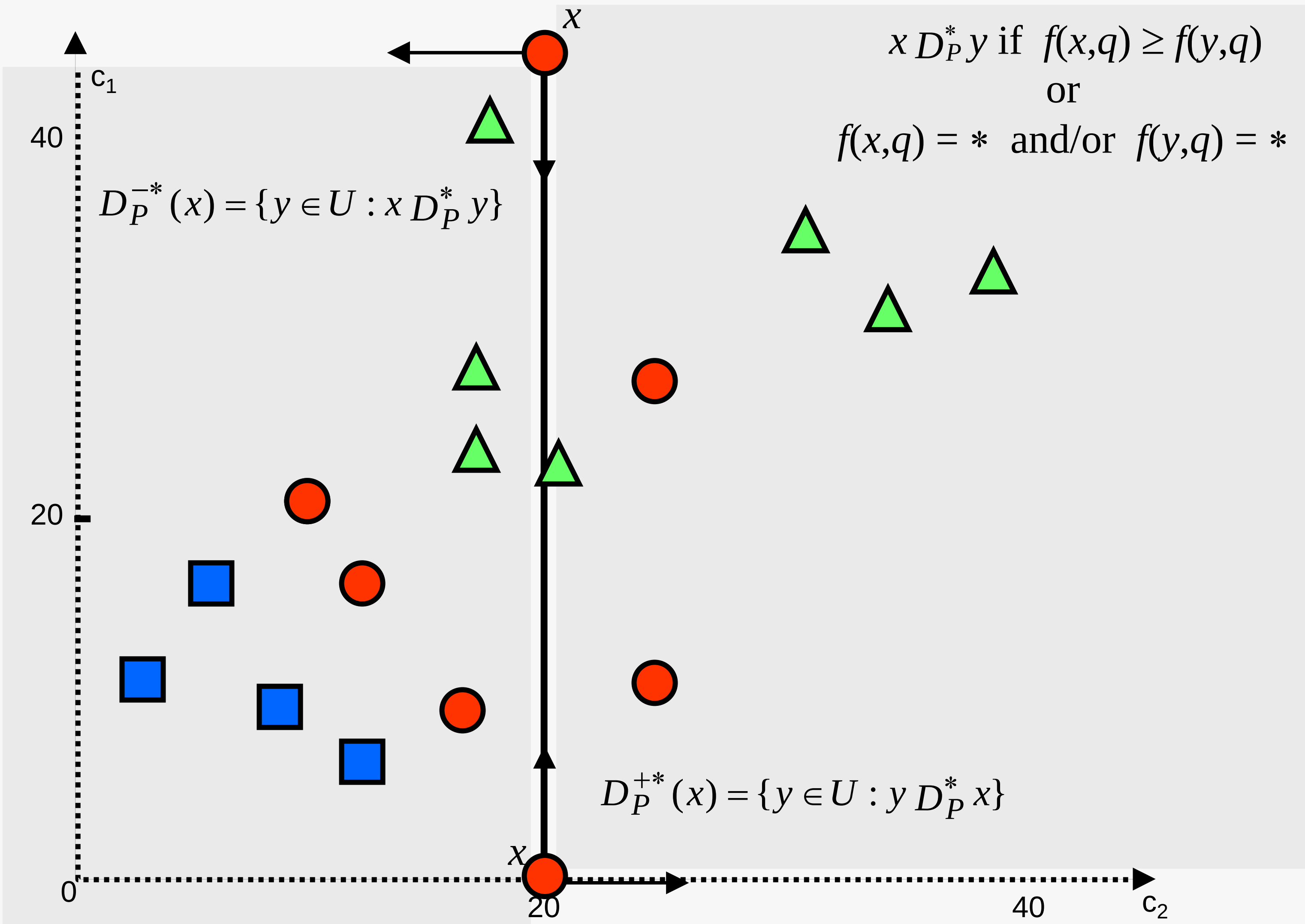
Missing values of attributes and criteria

Granules of knowledge



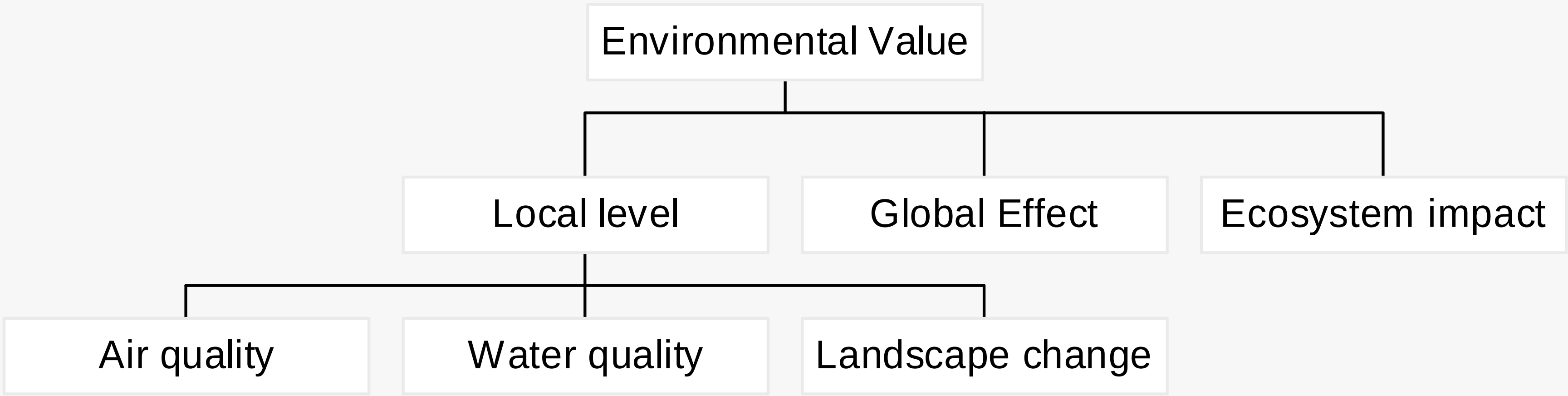
Missing values of attributes and criteria

Granules of knowledge



Hierarchical structure of attributes and criteria

- **Hierarchical structure of attributes and criteria**
  - Example



Hierarchical structure of attributes and criteria

Projects	Air quality	Water quality	Landscape change	Local level	Global effect	Ecosystem impact	Environmental value
P1	5	1	B	I	B	B	C
P2	2	3	M	I	M	B	C
P3	3	4	M	I	B	M	C
P4	1	6	G	II	B	M-G	B
P5	5	5	M	III	G	VG	A
P6	2	7	G	III	G	M-G	B
P7	6	6	G	III	G	G	A
P8	5	7	M	II	VG	G	B

$$VG \succ G \succ M \succ B$$

$$III \succ II \succ I$$

$$A \succ B \succ C$$

Hierarchical structure of attributes and criteria

Projects	Air quality	Water quality	Landscape change	Local level
P1	5	1	B	I
P2	2	3	M	I
P3	3	4	M	I
P4	1	6	G	II
P5	*5*	*5*	*M*	*III*
P6	2	7	G	III
P7	6	6	G	III
P8	*5*	*7*	*M*	*II*

P5 and P8 are inconsistenst

Projects	Local level	Global effect	Ecosystem impact	Environmental value
P1	I	B	B	C
P2	I	M	B	C
P3	I	B	M	C
P4	II	B	M-G	B
P5	II-III	G	VG	A
P6	*III*	*G*	*M-G*	*B*
P7	*III*	*G*	*G*	*A*
P8	*II-II*	*VG*	*G*	*B*

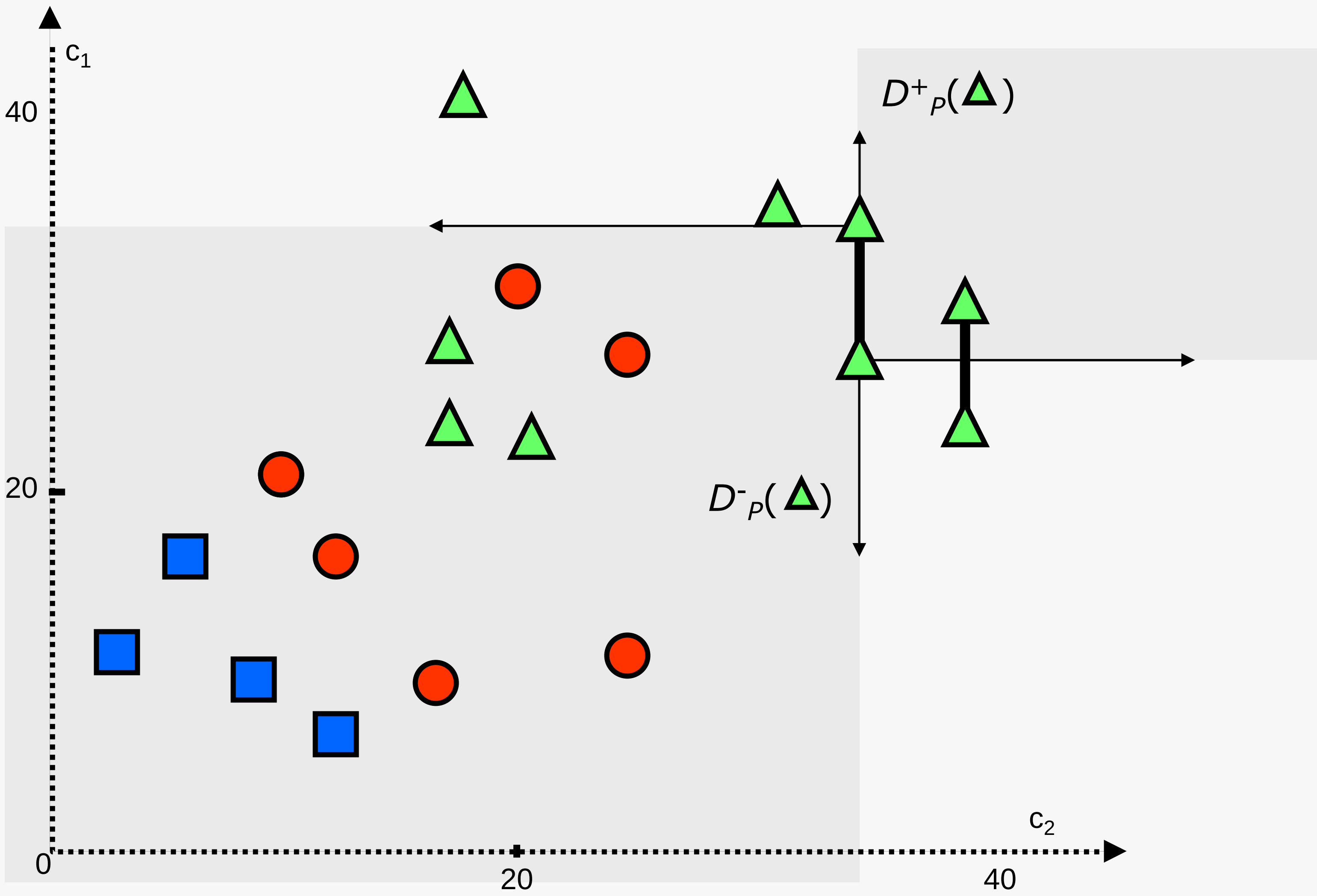
P6 and P7 are inconsistenst

P8 and P7 are inconsistenst

The second inconsistency does not appear in the original table – it is conditioned by the first level

Hierarchical structure of attributes and criteria

Interval order – dominance cones



Hierarchical structure of attributes and criteria

- Examples of decision rules with interval order

If u(x, Local) ≥ II, then x is at least B on EV (P4, P5, P6, P7, P8)

If l(x, Local) ≤ I, then x is at most C on EV (P1, P2, P3)

u(x)- upper bound of value x

l(x)- lower bound of value x

- **Discovering association rules in preference-ordered data sets**
  - Example

Client	Salary	Account status	Credit risk
A	9000	high	low
B	4000	medium	medium
C	5500	medium	high

  - ***Monotonic relationship*** between "salary" and "credit risk " : improvement of "salary" should not increase "credit risk"

    If so, B and C are ***inconsistent*** examples!

Discovering association rules in preference-ordered data sets

- **Technical diagnostics** – study of dependencies among values of symptoms

  Criteria = symptoms:

  $\uparrow a_1$ - maximum speed [km/h],

  $\uparrow a_2$ - compression pressure [Mpa],

  $\downarrow a_3$ - blacking components in exhaust gas [%],

  $\uparrow a_4$ - torque [Nm],

  $\downarrow a_5$ - summer fuel consumption [l/100 km],

  $\downarrow a_6$ - winter fuel consumption [l/100 km],

  $\downarrow a_7$ - oil consumption [l/1000 km],

  $\uparrow a_8$ - maximum horsepower of the engine [KM].

Discovering association rules in preference-ordered data sets

- Monotonic relationship (MR):

	Speed	Pressure	Blacking	Torque	FuelS	FuelW	Oil	HorsePower
Speed		x		x				x
Pressure	x		x	x	x	x	x	x
Blacking		x				x	x	x
Torque	x	x						x
FuelS		x				x		
FuelW		x	x		x		x	
Oil		x	x			x		
HorsePower	x	x	x	x				

- Looking for association rules with parameters:

minsupport = 50% (38 objects)

minconfidence = 75%

mincredibility = 75%

Discovering association rules in preference-ordered data sets

- **Without** considering MR among criteria: **40 association rules**
- Considering MR among criteria, **23** on **28** rules had to be removed because their credibility < 75% !
- Next 8 rules had to be deleted, because they are absorbed by others.
- Finally, **9 association rules** satisfied all requirements!
- An example of association rule:

„(pressure$\geq$2.4) $\rightarrow$ (torque$\geq$44.1)&(speed$\geq$74)”
with support 53.9%, confidence 97.6% and credibility 97.62%

- Ignoring the preference information may lead to wrong results – 78% of typical association rules are not valid!

Conclusions

- DRSA handles monotonic relationships between condition and decision attributes
- Classical rough set is a particular case of dominance-based rough approximation of a fuzzy set
- Preference model induced from rough approximations of unions of decision classes (or preference relations S and S^c) is expressed in a natural and comprehensible language of "*if*..., *then*..." decision rules
- Preference model built of decision rules is the most general, requires the weakest axioms, and can represent inconsistent preferences
- Heterogeneous information (attributes, criteria) and scales of preference (ordinal, cardinal) can be processed within DRSA
- DRSA exploits ordinal information only, and decision rules do not convert ordinal information into numeric one
- DRSA supplies useful elements of knowledge about decision situation:
  - certain and doubtful knowledge distinguished by lower and upper approximations
  - relevance of particular attributes or criteria and information about their interaction
  - reducts of attributes or criteria conveying important knowledge contained in data
  - the core of indispensable attributes and criteria
  - decision rules can be used for explanation of past decisions, for decision support and for strategic interventions

Free software available

ROSE

ROugh Set data Explorer

4eMka & JAMM & jMAF

New Decision Support Tools for Analysis and Solving
Multicriteria Classification Problems

http://idss.cs.put.poznan.pl/site/software.html