Semantic Rendering of Data Tables: Multivalued Information Systems Revisited

Marcin Wolski[1] and Anna Gomolińska[2]

[1] Maria Curie-Skłodowska University, Department of Logic and Cognitive Science,
Pl. Marii Curie-Skłodowskiej 4, 20-031 Lublin, Poland
`marcin.wolski@umcs.lublin.pl`
[2] University of Białystok, Faculty of Mathematics and Informatics,
Konstantego Ciołkowskiego 1M, 15-245 Białystok, Poland
`anna.gom@math.uwb.edu.pl`

Abstract. Data tables provide a convenient means of representation of descriptive information about objects. They serve also as standard input for data analysis tools or theories. In this paper we focus our attention upon the special class of data tables, namely multivalued information systems introduced by Z. Pawlak and E. Orłowska in the early 80s. The main idea presented in the paper is to interpret multivalued information systems as semantically processed single valued data tables. This interpretation allows us to describe classical rough set theory, dominance-based rough set theory, and formal concept analysis within the framework of multivalued information systems.

Keywords: information system, rough set, dominance relation, formal concept

1 Introduction

Data tables provide a simple and effective means of representation of collected pieces of information about a given set of objects. They serve also as standard input for data analysis tools or theories, output of which is often referred to as *knowledge*. In the present paper we shall focus our attention upon the special class of data tables, namely multivalued/approximate/nondeterministic information systems, introduced by Z. Pawlak and E. Orłowska in the early 80s [3, 4]. These systems are generalisations of the standard data tables/information systems to the case in which for an object x and an attribute A we are given (as the entry in the table) a set V_A of attribute values instead of a single value. The formal definitions of both multivalued and approximate information systems are actually the same (see [4]); it is the interpretation/semantics of the entries of these tables which makes the difference: the first interpretation is *the object x has all values from V_A for the attribute A* (multivalued systems), whereas the second reads as *the object x has a single value from the set V_A for the attribute A* (approximate systems). These systems equipped with a generalised semantics (which reads *for the object x and the attribute A the set V_A provides some possible values*) are called by Z. Pawlak and E. Orłowska nondeterministic information systems [3]. Of course, the semantics of the entries in the table determines how information is further processed; in other words, which relations between objects are used to construct information granules

being the building blocks of knowledge. The main concern of [4] is the informational indiscerniblity between objects (an equivalence relation), whereas the main focus of [3] are the information inclusion and the informational connection (a preorder and a tolerance relation, respectively).

The main novelty of the present paper consists in taking multivalued information systems as semantically enriched single valued data tables; this idea can be summarised by the following equation:

$$\text{data table} + \text{semantics} = \text{multivalued information system.}$$

Thus multivalued information systems represent semantically processed data. We start our study with some standard data tables used in the leading theories of data analysis: single valued information systems from rough set theory (RST) [4–6], single valued information systems enriched by dominance relations taken from dominance based rough set theory (DRS) [7, 8], and formal contexts from formal concept analysis (FCA) [1, 9]. Then we provide these structures with some specific interpretation/semantics. To this end we use *scales* from FCA, which are tools to convert a multivalued formal context (which is actually a standard information system) into a (single valued) formal context. We are going to employ these scales in order to obtain multivalued information systems. Then we focus upon informational relations of indiscerniblity and inclusion. These steps allow us to consider RST, DRS, and FCA within a single conceptual framework of multivalued information systems and to emphasise how these theories differ semantically. Finally, we shall discuss different interpretations of RST, DRS, and FCA based operators in the context of John Stuart Mill inductive reasoning [2].

2   Data Tables and Semantics

In the present section we discuss different forms of data tables considered in the leading theories of data analysis: rough set theory (RST) [4–6], dominance based rough sets (DRS) [7, 8], and formal concept analysis (FCA) [1, 9]. Of course, apart from data tables (input), each theory provides special tools to process the tables and produce some meaningful output. However, in the present section we shall discuss only tables, whereas the ways they may be further processed will be presented in the next section.

Definition 1 (Formal Context [1, 9]). *A formal context is a triple (U, Att, R), where U is a set of objects, Att a set of binary attributes and $R \subseteq U \times Att$.*

If an object x stands in the relation R to A, then we mark it in the data table by ✗. Table depicted by Fig. 1 presents a very simple context; Bob is a good mathematician whereas Agnes is not; on the bright side, she is rich.

Definition 2 (Information System [3, 4]). *A quadruple $\mathcal{I} = (U, Att, Val, f)$ is called an* information system*, where:*

– *U is a nonempty finite set of objects,*
– *A is a nonempty finite set of attributes,*

	good mathematician	rich
Steven	✘	
Bob	✘	✘
Agnes		✘

Fig. 1. A simple formal context

- $V = \bigcup_{A \in Att} Val_A$, where Val_A is the value–domain of the attribute A, and $Val_A \cap Val_B = \emptyset$, for all $A, B \in Att$ (*the last condition is not necessary, yet it is mathematically convenient*),
- $f : U \times Att \to Val$ is an information function, such that for all $A \in Att$ and $x \in U$ it holds that $f(x, A) \in Val_A$.

If f is a partial function then the information system $\mathcal{I}$ is called incomplete. *If the codomain of f is the powerset of Val, then the system is called* multivalued.

In what follows we shall confine our attention to the complete information (a simple example is depicted by Fig. 2) systems and their multivalued version being the result of scaling.

	mathematician (1-6)	rich
Steven	5	$ 0.1 million
Bob	4	$ 0.8 million
Agnes	2	$ 1 million

Fig. 2. An information system

Definition 3 (Dominance-Based Data Table [7, 8]). *A dominance-based data table is a system $\mathcal{I} = (U, Att, Val, f, \leq_D)$, where (U, Att, Val, f) is an information system such that all attribute values are real numbers. Let us define: $x \leq_A y$ iff $f(A, x) \leq f(A, y)$, for $x, y \in U$. Then $x \leq_D y$ iff $x \leq_A y$ for all $A \in Att$.*

Thus, information systems allow one to be more specific about the meaning of attributes. In the case of dominance-based data tables, we additionally assume that the attribute values are comparable with respect to some complete order. If $f(A, x) \leq f(A, y)$, then we say that *y is at least as good as x with respect to A*. If y is at least as good as x with respect to all attributes then we say that y *dominates* x.

Of course, any information system may be converted into a formal context. The easiest way of doing this is called in FCA a *nominal scaling*. Formally, a scale for A is a formal context (Val_A, Val_A, R_A) having Val_A as both the set of objects and the set of attributes. We also assume that the identity id_{Val_A} is included in R_A. After scaling each pair attribute-value (A, v) is regarded as a separate attribute of the new context $C_I = (U, \{(A, v)\}_{A \in Att\ v \in Val_A}, R)$. For the fundamental relation in rough set

theory is the indistinguishability, the values in (Val_A, Val_A, R_A) are compared by the equality relation: R_A is $=$, for every A; and in consequence 1 is interpreted as 1, 2 as 2, and so on. That is why in the scale depicted in Fig. 3 only the diagonal is marked by ✘. This type of scaling is called *nominal*. However, since all values in our exemplary data

	1	2	3	4	5	6
1	✘					
2		✘				
3			✘			
4				✘		
5					✘	
6						✘

	$ 0.1	$ 0.8	$ 1
$ 0.1	✘		
$ 0.8		✘	
$ 1			✘

	mathematician						rich		
	1	2	3	4	5	6	$ 0.1	$ 0.8	$ 1
Steven				✘			✘		
Bob			✘					✘	
Agnes		✘							✘

Fig. 3. Nominal scaling: the common ground of FCA and RST

table are real numbers, we can compare them with respect to $\leq$ instead of $=$. Actually, when we order these values according to $\leq$ we do change the scaling, or better still, the semantics (interpretation) of attribute values. Now, e.g., Bob's score in physics 5 and Agnes's score 2 means that Bob is a better mathematician than Agnes. In other words, whatever Agnes can solve, Bob can as well. Following DRS, we may say that Bob is at least as good as Agnes with respect to the attribute *mathematician*. Formally, the higher value (e.g., 5) with respect to $\leq$ implies the lower value (e.g., 2). Such scales are called *ordinal scales*. Thus this time the mark ✘ on 5 may mean that Bob has solved at least 5 problems. Under this reading Bob has solved at least 4 problems too. Therefore the scaling representing this interpretation (semantics) may be defined as depicted in Fig. 4. Please note that the values $ 0.1 or $ 0.8 are interpreted in the same fashion. Thus ✘ on $ 0.8 means that Bob has at least $ 0.8 on his bank account, and in consequence he has at least $ 0.1 too. Following DRS, we may say that Bob dominates Steven. In consequence, both RST and DRS start with the same date table but use different scales (interpretations). Of course, there are also (many) other possibilities. We conclude this

	1	2	3	4	5	6
1	✘					
2	✘	✘				
3	✘	✘	✘			
4	✘	✘	✘	✘		
5	✘	✘	✘	✘	✘	
6	✘	✘	✘	✘	✘	✘

	$ 0.1	$ 0.8	$ 1
$ 0.1	✘		
$ 0.8	✘	✘	
$ 1	✘	✘	✘

	mathematician						rich		
	1	2	3	4	5	6	$ 0.1	$ 0.8	$ 1
Steven	✘	✘	✘	✘	✘		✘		
Bob	✘	✘	✘	✘			✘	✘	
Agnes	✘	✘					✘	✘	✘

Fig. 4. DRS in terms of formal contexts

part with a nontrivial interpretation offered by B. Ganter during his seminar at Warsaw

University (a few years ago). He suggested to read the exam scores in a natural language and take the "natural" scale. Under this reading, someone who has done a very good

	bad
1 may be interpreted as	*bad*
2 may be interpreted as	*unsatisfactory*
3 may be interpreted as	*satisfactory*
4 may be interpreted as	*good*
5 may be interpreted as	*very good*
6 may be interpreted as	*excellent*

	1	2	3	4	5	6
1	✗	✗				
2		✗				
3			✗			
4			✗	✗		
5			✗	✗	✗	
6			✗	✗	✗	✗

Fig. 5. B. Ganter's semantics of exam's scores and scaling

job has done also a good job. The job of course is also satisfactory. However, someone who has done a bad job has also done an unsatisfactory job, but not *vice versa*. In consequence we obtain yet another scale, as depicted by Fig. 5. In this scale we use two orderings, and this type of scaling is therefore called *bi-ordinal*. Thus, starting from the information system depicted by Fig. 2 we can obtain a number of different multivalued information systems, depending on which scale is applied to the original system (see Fig. 6).

Nominal scaling $\Rightarrow$

	mathematician (1-6)	rich
Steven	{5}	{$0.1 million }
Bob	{4}	{$0.8 million }
Agnes	{2}	{$1 million }

Ordinal scaling $\Rightarrow$

	mathematician (1-6)	rich
Steven	{1, 2, 3, 4, 5}	{$0.1 million }
Bob	{1, 2, 3, 4}	{$0.1 million , $0.8 million }
Agnes	{1, 2}	{$0.1 million , $0.8 million $1 million }

B. Ganter's scaling $\Rightarrow$

	mathematician (1-6)	rich
Steven	{3, 4, 5}	{$0.1 million }
Bob	{3, 4}	{$0.1 million , $0.8 million }
Agnes	{2}	{$0.1 million , $0.8 million $1 million }

Fig. 6. Multivalued information systems as the results of scaling, i.e. providing data tables with interpretation/semantics

More formally, different interpretations/semantics of attribute values lead to different formal contexts. When one starts with a multivalued context (an information system)

$\mathcal{I} = (U, Att, Val, f)$ and a set of scales $S = \{(Val_A, Val_A, R_A) : A \in Att\}$ (as discussed above), then every pair (A, v), where $A \in Att$ and $v \in Val_A$, is regarded as the attribute in the induced formal context $C_I = (U, \{(A, v)\}_{A \in Att\ v \in Val_A}, R)$, where R is defined by

$$R = \{(x, (A, v)) : v \in f_s(x, A)\},$$

$$f_s(x, A) = \{v_i \in Val_A : f(x, A) = v\ \&\ (v, v_i) \in R_A\}.$$

Of course every (multivalued) context $\mathcal{I} = (U, Att, Val, f)$ may also be converted into a multivalued information system $\mathcal{I}_S = (U, Att, Val, f_s)$. The whole process of providing $\mathcal{I}$ with semantics given by S is depicted by Fig.6.

3 Information Processing: Approximation Operators

In the previous section we have discussed information systems (collected pieces of data about objects) which must be further processed to give some meaningful output (knowledge). The information processing in rough set theory [4–6] and dominance based rough set theory (DRS) [7, 8] is done by means of binary relations between objects. However, formal concept analysis (FCA) [1, 9] is based upon a relation between objects and attributes; in some special cases this relation may be reduced to the relation between objects, but it is not always possible. In multivalued information systems, in contrast to the classical information systems where there is considered a single relation, there are three important relations between objects [3, 4].

Definition 4. *Let (U, Att, Val, f) be a multivalued information system; then one can define:*

- *Informational Indiscerniblity: $x\ Ind\ y$ iff $f(x, A) = f(y, A)$,*
- *Informational Connectivity (Similarity): $x\ Sim\ y$ iff $f(x, A) \cap f(y, A) \neq \emptyset$,*
- *Informational Inclusion: $x\ Incl\ y$ iff $f(x, A) \subseteq f(y, A)$,*

for all $A \in Att$ and $x, y \in U$.

Usually, the output of data analysis is related to some aspect of reality represented by a *decision attribute*. For example, we could take the *subjective quality of life* as the decision attribute. Then our aim would be to "express" the subjective quality of life in terms of the attributes *mathematician* and *rich*, which are (in this case) called *conditional attributes* (we would like to obtain knowledge how the subjective quality of life "depends" on wealth and education in mathematics). Information systems having a single attribute distinguished as a decision attribute are called *decision tables*. In such a case, all informational relations (indiscerniblity, connectivity, and inclusion) are defined with respect to (only) conditional attributes.

As earlier, we shall start discussion in this section with FCA (and its operators).

Definition 5 (Derivation Operators). *For a formal context $\mathcal{C} = (U, Att, R)$, define:*

$$R'(X) = \{A \in Att : \forall x \in X\ ((x, A) \in R)\},$$

$$R'(\mathcal{A}) = \{x \in U : \forall A \in \mathcal{A}\ ((x, A) \in R)\},$$

for all $X \subseteq U$ and $\mathcal{A} \subseteq Att$.

	mathematician (1-6)	rich	sub. quality of life (1-3)
Steven	5	$ 0.1 million	3
Bob	4	$ 0.8 million	3
Agnes	2	$ 1 million	1

Fig. 7. A decision table

Definition 6 (Formal Concept). *A formal concept is a pair $(X, \mathcal{A})$ such that $X = R'(\mathcal{A})$ and $\mathcal{A} = R'(X)$. X is called an extension and $\mathcal{A}$ is called an intention of this concept.*

Let us start with a complete information system $\mathcal{I} = (U, Att, Val, f)$ and a semantics S, that is, a family of scales R_A, for all $A \in Att$. As one can easily observe, for every object x in $C_I = (U, \{(A, v)\}_{A \in Att \ v \in Val_A}, R)$, a pair $(R'(R'(\{x\})), R'(\{x\}))$ is a concept, and $y \in R'(R'(\{x\}))$ provided that $x \ Incl \ y$ in the corresponding multivalued information system $\mathcal{I}_S = (U, Att, Val, f_s)$:

$$R'(R'(\{x\})) = \{y \in U : x \ Incl \ y\}.$$

However, it usually happens that

$$R'(R'(X)) \neq \{y \in U : \exists x \ (x \ Incl \ y \ \& \ x \in X)\}.$$

But it always holds that

$$X \subseteq \{y \in U : \exists x \ (x \ Incl \ y \ \& \ x \in X)\} \subseteq R'(R'(X)).$$

However, after conversion of an information system $\mathcal{I} = (U, Att, Val, f)$ to the formal context $C_I = (U, \{(A, v)\}_{A \in Att \ v \in Val_A}, R)$ we lose the contact with original attributes (we shall discuss this issue in detail later in the paper).

Definition 7 (Lower and Upper Approximations). *A pair (U, E), where E is an equivalence relation, is called an approximation space. Define after Z. Pawlak:*

$$Low_E(X) = \{x \in U : [x]_E \subseteq X\},$$

$$Upp_E(X) = \{x \in U : [x]_E \cap X \neq \emptyset\}.$$

$Low_E(X)$ is called the lower approximation of X, whereas $Upp_E(X)$ is called the upper approximation of X.

Coming back to information systems, every set of attributes $\mathcal{A} \subseteq Att$ of an information system $\mathcal{I} = (U, Att, Val, f)$ induces an approximation space $(U, E_\mathcal{A})$, where $E_\mathcal{A} = \{(x, y) : f(x, A) = f(y, A) \ \text{for all} \ A \in \mathcal{A}\}$. In order to simplify the notation, we shall write $Low_\mathcal{A}(X)$ and $Upp_\mathcal{A}(X)$ for $Low_{E_\mathcal{A}}(X)$ and $Upp_{E_\mathcal{A}}(X)$, respectively. In the case when $\mathcal{A} = Att$, we shall leave E without any subscript.

Every information system $\mathcal{I} = (U, Att, Val, f)$ (together with a family of scales S) induces also a multivalued information system $\mathcal{I}_S = (U, Att, Val, f_s)$ and another approximation space (U, Ind). Due to scaling, Ind and E may be two different relations.

For any $\mathcal{A} \subseteq Att$ of (U, Att, Val, f_s) the corresponding indiscernibility relation will be denoted by $Ind_\mathcal{A}$. This notational convention will also be used for other relations.

As usual, we can generalise E to any reflexive relation P (e.g. Sim or $Incl$) and obtain generalised approximation operators. Let $[x]_P = \{y \in U : (x,y) \in P\}$ and define:

$$Low_P(X) = \{x \in U : [x]_P \subseteq X\},$$
$$Upp_P(X) = \{x \in U : [x]_P \cap X \neq \emptyset\}.$$

As one can note, it holds that

$$Upp_{Incl}(x) = R'(R'(\{x\})) = \{y \in U : x \; Incl \; y\},$$

and

$$Upp_{Incl}(X) = \{y \in U : \exists x \; (x \; Incl \; y \; \& \; x \in X)\}.$$

However, $R'(R'(X))$ is much more complex in the settings of information systems than it might seem at the first sight. It is worth noting that $R'(X)$, for $X \subseteq U$, is a set consisting of pairs (A, v). So, in order to go back to the level of information systems we need a method of retrieving the original attributes from this set, so as it would act as $\mathcal{A} \subseteq Att$. Let $Atex$ (attribute extraction) be defined by $Atex(H) = \{A : (A, v) \in H\}$ for $H \subseteq Att \times Val$. Obviously, this a projection operation on the first coordinate and it makes sense only for a family of *regular* scales. Consider the following example. Let $Val = \{1, 2, 3, 4, 5, 6\}$ be a set of values of some attribute A. Assume that a scaling converts 1 to $\{1, 2\}$, 2 to $\{2, 3\}$, and all other values to $\{3, 4, 5, 6\}$. So after scaling A has three value sets: $\{1, 2\}$, $\{2, 3\}$, $\{3, 4, 5, 6\}$. Let $f(x, A) = \{1, 2\}$ and $f(y, A) = \{2, 3\}$. Now, let us start with $(U, \{A\}, Val, f)$, then go to the corresponding C_I, and compute $R'(\{x, y\})$, which is $\{(A, 2)\}$ – but this item does not make sense in our semantics S: $\{2\}$ is the meaning of neither element of Val. Thus, using set intersection $\cap$ we may produce a new non-empty value set, which is not present in $\mathcal{I}_S = (U, Att, Val, f_s)$. A scale is *regular* if that is not possible. Nominal and ordinal scales are regular.

Only for regular scales we are able to define the concepts of the formal context on the level of attributes of information systems. Let $\mathcal{I}_S = (U, Att, Val, f_s)$ be a multi-valued information system obtained from an information system $I = (U, Att, Val, f)$ by means of regular scales S; then

$$R'(R'(X)) = \{y \in U : \forall A \in Atex(R'(X)) \; \exists x \in X \; (x \; Incl_A \; y)\}.$$

If the scale is not regular, then the following inclusion holds only:

$$R''(X) = \{y \in U : \forall A \in Atex(R'(X)) \; \exists x \in X \; (x \; Incl_A \; y)\} \subseteq R'(R'(X)).$$

Therefore, in such a case we need a new name R'' for this operator.

Let us now consider a decision table $\mathcal{I}_G = (U, Att, Val, G, f)$, that is, an information system $\mathcal{I} = (U, Att, Val, f)$ equipped with a decision (goal) attribute $G \notin Att$ and f being defined on $Att \cup \{G\}$. The semantics S for $\mathcal{I}_G$ needs now to include a scale $R_G = (Val_G, Val_G, R_G)$ for the decision attribute. The main goal is to approximate a given pair $(G, [v]_{R_G})$, where $v \in Val_G$ is a specific distinguished value. More precisely, we want to approximate the set $X = \{x \in U : (v, f(x, G)) \in R_G\}$.

Let us take two scaling methods for the decision attribute *subjective quality of life*:

nominal scale Nom: $R_G = \{(i,j) : i,j \in \{1,2,3\} \ \& \ i = j\}$;
ordinal scale Ord: $R_G = \{(i,j) : i,j \in \{1,2,3\} \ \& \ j \leq i\}$.

The nominal scale Nom_G interprets 1 as *low quality of life*, 2 as *average quality of life*, and 3 as *high quality of life*. As expected, the ordinal scale Ord_G interprets $1, 2$, and 3 as: *at least low quality of life*, *at least average quality of life*, and *at least high quality of life*, respectively. Now, let us take the value 3 as the distinguished value of the decision attribute. Then under Nom_G we are going to approximate the set $\{Bob, Steven\}$, however, under Ord_G the set to be approximated is $\{Agnes, Bob, Steven\}$.

The dominance-based rough set approach (DRS) [7, 8] is actually a kind of rough set theory rendered according to the above ideas and the ordinal scaling method. An information system is equipped with a dominance relation $\leq_D$, that is, we consider the system $\mathcal{I} = (U, Att, Val, f, \leq_D)$ (see Section 2). This system induces a multivalued information system $\mathcal{I}_S = (U, Att, Val, f_s)$, where S consists of ordinal scales (Val_A, Val_A, R_A) for every $A \in Att$. Before we recall the definitions of the approximation operators, we need a few auxiliary concepts:

- $D^+(x) = \{y \in U : x \leq_D y\}$ (a set of objects dominating x, or *better than x*);
- $D^-(x) = \{y \in U : y \leq_D x\}$ (a set of objects dominated by x, or *worse than x*);
- Decision attribute G, $V_G = T$, $Cl_t = \{x \in U : f(x, G) = t\}$,

$$Cl_t^\leq = \bigcup_{s \leq_G t} Cl_s \ \text{ and } \ Cl_t^\geq = \bigcup_{t \leq_G s} Cl_s,$$

where $t, s \in T$. It is additionally assumed that

$$Cl_s \cap Cl_t = \emptyset \ \text{ for } s \neq t \text{ and } \bigcup_{s \in T} Cl_s = U.$$

Classification patterns to be discovered are functions representing granules $Cl^\leq{}_t$ and $Cl^\geq{}_t$ by means of granules $D^+(x)$ and $D^-(x)$. It is worth emphasising that due to the preference order, the sets to be approximated are not the particular Cl_t (for some $t \in V_G$), but the upward and downward unions.

As said in the previous section, DRS may be represented in terms of ordinal scaling. In what follows we would like to make this scaling explicit in DRS and use relations Ind and $Incl$ (from multivalued information systems) rather than the dominance relation $\leq_D$. Let us consider a multivalued information system $\mathcal{I}_S = (U, Att, Val, f_s)$ obtained from $\mathcal{I} = (U, Att, Val, f, \leq_D)$ by means of a scaling set S. Please note that due to the ordinal scaling of all attributes of $\mathcal{I}$, it holds that:

$$D^+(x) = \{y \in U : x \ Incl \ y\} = R'(R'(\{x\})),$$

$$D^-(x) = \{y \in U : x \ Incl^{-1} \ y\} = \{y \in U : y \ Incl \ x\},$$

$$Cl_t^\geq = \{y \in U : \exists x \ (x \ Incl_{\{G\}} y \ \& \ x \in Cl_t)\},$$

$$Cl_t^\leq = \{y \in U : \exists x \ (x \ Incl_{\{G\}}^{-1} y \ \& \ x \in Cl_t)\}.$$

The specific interpretation of ordinal scaling in DRS makes a new type of inconsistency in data tables possible:

(a) an object x belongs to $Cl^{\geq}_t$ (that is, it belongs to Cl_t or a class better than Cl_t), but it is dominated by some objects $y \notin Cl^{\geq}_t$ (it is dominated by some object from a worse class),

(b) an object x belongs to the class $Cl^{\leq}_t$ (that is, it belongs Cl_t or a class worse than Cl_t), but it dominates some object $y \notin Cl^{\leq}_t$ (it dominates some object from a better class).

These objects are regarded as borderline cases: they might or might not belong to a given class. In consequence, in DRS we consider the following approximations:

$$\underline{Cl^{\leq}_t} = \{x \in U : D^-(x) \subseteq Cl^{\leq}_t\},$$

$$\overline{Cl^{\leq}_t} = \{x \in U : D^+(x) \cap Cl^{\leq}_t \neq \emptyset\},$$

$$\underline{Cl^{\geq}_t} = \{x \in U : D^+(x) \subseteq Cl^{\geq}_t\},$$

$$\overline{Cl^{\geq}_t} = \{x \in U : D^-(x) \cap Cl^{\geq}_t \neq \emptyset\}.$$

As earlier, our aim is to express these approximation operators by means of $Incl$. Thus, let be given an information system $\mathcal{I} = (U, Att, Val, f, \leq_D)$ and its corresponding multivalued information system (U, Att, Val, f_s), obtained by means of the ordinal scaling method Ord. Then

$$\underline{Cl^{\leq}_t} = \bigcup_{D^-(x) \subseteq Cl^{\leq}_t} D^-(x) = \{y \in U : \forall x \in U \ (x \ Incl \ y \Rightarrow x \in Cl^{\leq}_t)\},$$

$$\overline{Cl^{\leq}_t} = \bigcup_{x \in Cl^{\leq}_t} D^-(x) = \{y \in U : \exists x \in U \ (x \ Incl^{-1} \ y \ \& \ x \in Cl^{\leq}_t)\},$$

$$\underline{Cl^{\geq}_t} = \bigcup_{D^+(x) \subseteq Cl^{\geq}_t} D^+(x) = \{y \in U : \forall x \in U \ (x \ Incl^{-1} \ y \Rightarrow x \in Cl^{\geq}_t)\},$$

$$\overline{Cl^{\geq}_t} = \bigcup_{x \in Cl^{\geq}_t} D^+(x) = \{y \in U : \exists x \in U \ (x \ Incl \ y \ \& \ x \in Cl^{\geq}_t)\}.$$

So, we are able to transfer FCA, RST, and DRS, along with explicitly given semantics S (a family of scales R_A for every attribute $A \in Att$) into the framework of multivalued information systems. The connections between the operators discussed above are as follows:

$$\underline{Cl^{\leq}_t} = Low_{Incl^{-1}}(Cl^{\leq}_t),$$

$$\overline{Cl^{\leq}_t} = Upp_{Incl}(Cl^{\leq}_t),$$

$$\underline{Cl^{\geq}_t} = Low_{Incl}(Cl^{\geq}_t),$$

$$\overline{Cl^{\geq}_t} = Upp_{Incl^{-1}}(Cl^{\geq}_t) \subseteq R'(R'(Cl^{\geq}_t)).$$

Two important comments are needed. As long as we regard all above operators as *approximation* operators, RST and DRS based results are better than that coming from

FCA. However, when we change the context, then the FCA operator may be more preferable. Secondly, in DRS, $x\ Incl\ y$ is read as y *is better than* x. However, there are other readings possible, e.g., *we have more pieces of information about y than about x*.

Let us consider an example which brings new meanings for relations and operators we have discussed so far. Let **w** be a serious disease which we have an antibiotic working against. Let us assume that we have a test checking whether someone is ill, but the antibiotic should be given to a patient before the disease develops. So, our aim is to select people who will be given the medicine. One very expensive solution is to give the medicine to all people (that is, all elements of the universe U). On the other extreme, we could give the medicine only to people who test positive for **w**. Of course, both solutions are not good and we need to find another method of selection. We are going to employ the John Stuart Mill inductive reasoning [2] here, which was designed to solve problems of this type, but under complete knowledge. We shall focus our attention on the very basic cannon, namely the direct method of agreement (Fig. 8). We are not going to use the pure form of this cannon, but rather its rough set based rendering.

A B C D	**occur together with**		**w, v**
A E F G	**occur together with**		**w, z**
	Therefore A is the cause of		**w**

Fig. 8. Direct method of agreement: *If two or more instances of the phenomenon under investigation have only one circumstance in common, the circumstance in which alone all the instances agree, is the cause (or effect) of the given phenomenon* [2].

Let be given an information system $\mathcal{I} = (U, Att, Val, f)$ representing our (medical) knowledge about people. At first, we employ a semantics S consisting only of ordinal scales such that for each attribute A of $\mathcal{I}_S = (U, Att, Val, f_S)$, the relation $x\ Incl_A\ y$ means that y *is in a worse medical condition than* x with respect to A and **w**. In our settings, Mill's inductive reasoning is modelled in the following way. The concept **w** is actually a set $\{x \in U : f(x, \mathbf{w}) = \top\}$, where $Val_{\mathbf{w}} = \{\top, \bot\}$ (truth and false, respectively). The scale $R_{\mathbf{w}}$ is given by $\{(\top, \top), (\top, \bot), (\bot, \bot)\}$, so if $f(x, \mathbf{w}) = \top$, then $f_S(x, \mathbf{w}) = \{\top, \bot\}$, and if $f(x, \mathbf{w}) = \bot$, then $f_S(x, \mathbf{w}) = \{\bot\}$. Thus, as required, if $x\ Incl_{\mathbf{w}}\ y$, then y is in the same or worse medical condition than x, so if x is ill, then y must be ill as well. Let $Cl_\top$ be a set of positive examples of **w** (direct method of agreement). The term "in common" (Fig. 8) is beyond the expressive power of pure RST and DRS. However, in our case we can retrieve common attributes by means of $Atex(R'(Cl_\top))$. Now we can compute possible solutions to our problem:

$$Cl_\top \subseteq Upp_{Ind}(Cl_\top) \subseteq Upp_{Incl}^{-1}(Cl_\top) = \overline{Cl_\top^{\geqq}} \subseteq R'(R'(Cl_\top^{\geq})).$$

As said $Cl_\top$ is an extreme solution, another one may be given by $Upp_{Ind}(Cl_\top)$ (that is, we give the antibiotic to all people with exactly the same medical description in terms of conditional attributes as some patients having positive test for **w**). It seems reasonable, however the next solution $Upp_{Incl}^{-1}(Cl_\top) = \overline{Cl_\top^{\geqq}}$ is much better: we give the medicine to all people with the same or worse medical condition than some patients who have

positively tested for **w**. Better still, we may apply the direct method of agreement and give medicine to all people in $R'(R'(Cl_{\top}^{\geq}))$ having medically worse results only for attributes which seem to be relevant to **w**. It is worth emphasising that regarded as an approximation of $Cl_{\top}$, the set $R'(R'(Cl_{\top}^{\geq}))$ is the worst candidate, but in this very settings it is the best solution for the problem at issue. Consider a non-regular scale now and assume that all people who positively tested on **w** have problems with blood pressure A. So scaling of A shows how unstable is the pressure. Some patients may have value $\{normal, high\}$, some $\{low, high\}$, and some $\{low, normal, high\}$, but none of them has $\{normal\}$. This time $R'(R'(Cl_{\top}^{\geq}))$ is a very bad solution, because it may include people with a normal blood pressure. However, we can still use the direct method of agreement in a modified version, and take $R''(Cl_{\top}^{\geq})$ as the solution.

4 Conclusions

In the paper we have investigated the implicit semantics used in some leading theories of data analysis: rough set theory (RST) [4–6], dominance based rough set theory (DRS) [7, 8], and formal contexts from formal concept analysis (FCA) [1, 9]. We have presented all theories within the unifying framework of multivalued information systems [3, 4], enriched with the scaling methods from FCA. We have also discussed the relations between the operators coming from these theories, and presented their different interpretations in the context of John Stuart Mill inductive reasoning [2].

References

1. Ganter, B., Wille, R.: Formal Concept Analysis. Mathematical Foundations, Springer Verlag (1999)
2. Mill, J.S.: A System of Logic. Vol. 1. London (1843)
3. Orłowska, E., Pawlak, Z.: Representation of nondeterministic information. Theoretical Computer Science 29, pp. 27–39 (1984)
4. Pawlak, Z.: Systemy informacyjne. Podstawy teoretyczne. Warszawa, WNT (1983)
5. Pawlak, Z.: Rough Sets: Theoretical Aspects of Reasoning about Data. Kluwer Academic Publisher (1991)
6. Pawlak, Z.: Wiedza z perspektywy zbiorów przybliżonych. Institute of Computer Science Report 23, (1992)
7. Słowinski, R., Greco, S., Matarazzo, B.: Dominance-based rough set approach to reasoning about ordinal data. Lecture Notes in Artificial Intelligence 4585, 5–11, (2007)
8. Słowiński, R., Greco, S., Matarazzo, B.: Rough sets in decision making. In: R.A.Meyer y (Ed.), Encyclopedia of Complexity and Systems Science, Springer, NY, 7753–7786, (2009)
9. Wille, R.: Concept lattices and conceptual knowledge systems. Computers & Mathematics with Applications 23, 493–515, (1992)