Brock University

Department of Computer Science

Modal-style Operators in Qualitative Data Analysis

Günther Gediga and Ivo Düntsch
Technical Report # CS-02-15
May 2002

Brock University
Department of Computer Science
St. Catharines, Ontario
Canada L2S 3A1
www.cosc.brocku.ca

Modal–style operators in qualitative data analysis*

Ivo Düntsch†
Department of Computer Science
Brock University
St. Catherines, Ontario, Canada, L2S 3A1
duentsch@cosc.brocku.ca

Günther Gediga†
Institut für Evaluation und Marktanalysen
Brinkstr. 19
49143 Jeggen, Germany
gediga@eval-institut.de

Summary

We explore the modal possibility operator and its dual necessity operator in qualitative data analysis, and show that it (quite literally) complements the derivation operator of formal concept analysis; we also propose a new generalisation of the rough set approximation operators. As an example for the applicability of the concepts we investigate the Morse data set which has been frequently studied in multidimensional scaling procedures.

1 Introduction

A frequently used operationalisation of data is an

$$\text{Object} \mapsto \text{Attribute}$$

relationship. Such operationalisation comes in various flavours: Examples include deterministic information systems a la Pawlak [19], the many–valued tables of Lipski [14] and Orłowska & Pawlak [18], in which each object is assigned a set of attribute values, the property systems of Vakarelov [24], or the relational attribute systems of Düntsch et al. [5] which incorporate semantical constraints. In its most general form, each object of the universe of discourse is related to one or more attribute values. Mathematically, one considers structures $\langle U, V, R\rangle$, where U and V are sets, and $R \subseteq U \times V$ is a binary relation between elements of U and elements of V. Based on the existential and universal quantifiers, one can define mappings $2^U \to 2^V$ in a natural way, namely,

$$\langle R\rangle(X) = \{y \in V : (\exists x \in X)xRy\} = \bigcup_{x\in X} R(x),$$
$$[[R]](X) = \{y \in V : (\forall x \in X)xRy\} = \bigcap_{x\in X} R(x),$$

where $R(x) = \{y \in V : xRy\}$. In a general mathematical setting, the mapping $[[R]]$ has been called a *polarity* [1]; it is also the *derivation operator* of formal concept analysis (FCA) [25]. While $[[R]]$

*Co-operation for this paper was supported by EU COST Action 274 "Theory and Applications of Relational Structures as Knowledge Instruments" (TARSKI); http://www.tarski.org/

†The ordering of authors is alphabetical and equal authorship is implied.

has received some prominence via FCA, the operator $\langle R\rangle$ seems to have been largely neglected in the study of object–attribute relations. Interestingly, it is the reverse in logical systems, where $\langle R\rangle$ became the widely studied possibility operator of modal logics associated with Kripke frames $\langle U, R\rangle$, the roots of which go back to the seminal paper by Jónsson & Tarski [12]. The operator $[[R]]$ was introduced to modal logics by Humberstone [11] and Gargov et al. [7], who called it a "sufficiency operator". Their aim was to be able to express "negative" properties of relations such as irreflexivity, which could not be expressed by the common modal operators "possibility" and its dual, "necessity". Recently, Düntsch & Orłowska [6] have investigated the algebraic interplay of $\langle R\rangle$ and $[[R]]$ as well as their (mixed) correspondence theory.

Apart from the sufficiency operator in FCA, modal–style operators have been used in data analysis in connection with rough set approximation [20], where R is an equivalence relation on U; there, $\langle R\rangle$ can be interpreted as an upper approximation, and its dual as a lower approximation, based on the knowledge of the world given by the classification induced by R [e.g. 16, 23, 26–28]. There is a rich literature on binary relations among objects, induced by information systems, and we invite the reader to consult [17] for many examples and details.

Our aim in this note is to explore the possibilities of the $\langle R\rangle$ operator and its dual necessity operator $[R]$ in relational attribute systems, and we shall show that it (quite literally) complements the derivation operator of FCA. Along the way, we will propose a new generalisation for the rough set approximation operators. The paper closes with an application of the concepts to the Morse data set [21].

2 Definitions and notation

Throughout this paper, U and V are nonempty sets, and $R \subseteq U \times V$. For unexplained notation and concepts in lattices and order theory we refer the reader to [4].

A *closure operator* on U is a mapping $\mathrm{cl} : 2^U \to 2^U$ such that for all $X, Y \subseteq U$,

$$\begin{aligned}
&X \subseteq Y \subseteq U \Rightarrow \mathrm{cl}(X) \subseteq \mathrm{cl}(Y), && \text{i.e. cl is monotone},\\
&X \subseteq \mathrm{cl}(X), && \text{i.e. cl is expanding},\\
&\mathrm{cl}(X) = \mathrm{cl}(\mathrm{cl}(X)) && \text{i.e. cl is idempotent}.
\end{aligned}$$

A *closure system on U* is a family of subsets of U which is closed under intersection. It is well known that there are one-one correspondences between closure operators, closure systems, and $\cap$ -congruences, the latter, if U is finite.

Dually, an *interior operator* is a mapping $\mathrm{int} : 2^U \to 2^U$ such that for all $X, Y \subseteq U$,

$$\begin{aligned}
&X \subseteq Y \subseteq U \Rightarrow \mathrm{int}(X) \subseteq \mathrm{int}(Y), && \text{i.e. int is monotone},\\
&\mathrm{int}(X) \subseteq X, && \text{i.e. int is contracting},\\
&\mathrm{int}(X) = \mathrm{int}(\mathrm{int}(X)), && \text{i.e. int is idempotent}.
\end{aligned}$$

An *interior system* is a family of subsets of U which is closed under union.

If L is a lattice and $M \subseteq L$, then M is called *join–dense* (*meet–dense*), if every $x \in L$ is a join (meet) of elements of M.

For each $x \in U$ we let

$$R(x) = \{y \in V : xRy\}$$

be the *R-range of x*, and

$$\operatorname{dom} R = \{x \in U : xRy \text{ for some } y \in V\}$$

is the *domain of R*. Furthermore,

$$R\breve{} = \{\langle y,x\rangle \subseteq V \times U : xRy\}$$

is the *converse of R*. If $f : 2^U \to 2^V$, then the *dual of f* is the mapping $f^\partial : 2^U \to 2^V$ defined by

$$f^\partial(X) = V \setminus f(U \setminus X). \tag{2.1}$$

The operators $2^U \to 2^V$ which we want to consider are the following:

$$\langle R\rangle(X) = \{y \in V : X \cap R\breve{}(y) \neq \emptyset\}, \quad \text{possibility} \tag{2.2}$$

$$[R](X) = \{y \in V : R\breve{}(y) \subseteq X\}, \quad \text{necessity} \tag{2.3}$$

$$[[R]](X) = \{y \in V : X \subseteq R\breve{}(y)\}, \quad \text{sufficiency} \tag{2.4}$$

$$\langle\langle R\rangle\rangle(X) = \{y \in V : (-R)\breve{}(y) \cap (U \setminus X) \neq \emptyset\}, \quad \text{dual sufficiency.} \tag{2.5}$$

Here, $(-R) = \{\langle x,y\rangle \in U \times V : \langle x,y\rangle \notin R\}$ is the complement of the relation R in $U \times V$.

It is well known that these mappings have the following structural properties: Let $\mathfrak{X} \subseteq 2^U, x \in U$.

$$\langle R\rangle(\{x\}) = R(x) = [[R]](\{x\}), \tag{2.6}$$

$$[R](U \setminus \{x\}) = V \setminus R(x) = \langle\langle R\rangle\rangle(U \setminus \{x\}), \tag{2.7}$$

$$\langle R\rangle\left(\bigcup_{X\in\mathfrak{X}} X\right) = \bigcup_{X\in\mathfrak{X}} \langle R\rangle(X), \tag{2.8}$$

$$[R]\left(\bigcap_{X\in\mathfrak{X}} X\right) = \bigcap_{X\in\mathfrak{X}} [R](X), \tag{2.9}$$

$$[[R]]\left(\bigcup_{X\in\mathfrak{X}} X\right) = \bigcap_{X\in\mathfrak{X}} [[R]](X), \tag{2.10}$$

$$\langle\langle R\rangle\rangle\left(\bigcap_{X\in\mathfrak{X}} X\right) = \bigcup_{X\in\mathfrak{X}} \langle\langle R\rangle\rangle(X). \tag{2.11}$$

Hence, the mappings $\langle R\rangle$ and $[[R]]$ are determined by their action on the singleton sets, and $[R]$ as well as $\langle\langle R\rangle\rangle$ are determined by their action on the complements of singletons.

As an example, suppose that U is a set of students, V is a set of problems, and aRb is interpreted as "Student a solves problem b" [8]. If $X \subseteq U$ is a set of students, then for a problem b we have

$$\begin{aligned}
b \in \langle R\rangle(X) &\iff \text{Some student in } X \text{ solves } b,\\
b \in [R](X) &\iff \text{Each student who solves } b \text{ is in } X,\\
b \in [[R]](X) &\iff b \text{ is solved by each student in } X,\\
b \in \langle\langle R\rangle\rangle(X) &\iff \text{Not all students in } U \setminus X \text{ solve } b.
\end{aligned}$$

If we think of qRs as "s is a property which q has", then, for each $Y \subseteq V$, the set $[R\breve{}](Y)$ collects those objects all of whose properties are in Y, and $[[R\breve{}]](Y)$ is the set of objects which possess all properties of Y (and possibly more).

It is not hard to see (and well known) that $\langle R \rangle$ and $[R]$, as well as $[[R]]$ and $\langle\langle R \rangle\rangle$ are dual to each other. Furthermore,

$$[[R]](X) = [(-R)](U \setminus X), \ \langle\langle R \rangle\rangle(X) = \langle(-R)\rangle(U \setminus X). \tag{2.12}$$

We see from (2.12) that each of the four operators, along with the complements on U and $U \times V$ and the converses defines all others. It may be argued that, in principle, everything is already said, when we consider, for example, the sufficiency operator $[[R]]$. As far as the formal Mathematics and the computational aspects go, this may be true; however, the semantic interpretations of the various operators differ widely, and it is useful to start with the other operators if the situation so requires. Indeed, considering complementation on the relational level adds another level to the underlying logic; in order to avoid this, the sufficiency operator was introduced on the language level.

3 Modal operators in data analysis

In the realm of data analysis the sufficiency operator has received the widest attention of all the four modal–style operators via the context of formal concept analysis [25]. There, a *context* is a triple $\langle U, V, R \rangle$, where U, V are sets, and $R \subseteq U \times V$. If $X \subseteq U, \ Y \subseteq V$, the set $[[R]](X)$ is called *intent of* X and $[[R\breve{}]](Y)$ is called the *extent of* Y. Here, we think of Y as a set of properties, and X as the set of objects (of our set U of discourse) which possess these properties. A *concept* is defined as a pair $\langle X, Y \rangle \in 2^U \times 2^V$ such that $[[R]](X) = Y$ and $[[R\breve{}]](Y) = X$. The main theorem of FCA is the following:

Proposition 1. *[25] Let $M = \langle U, V, R \rangle$ be a context, and set*

$$\mathcal{C}_M = \{\langle X, Y \rangle \in 2^U \times 2^V : [[R]](X) = Y, \ [[R\breve{}]](Y) = X\}.$$

Then, $\mathcal{C}_M$ can be made into a complete lattice by setting

$$\sum_{i \in I} \langle X_i, Y_i \rangle = \langle [[R\breve{}]][[R]](\bigcup_{i \in I} X_i), \bigcap_{I \in I} Y_i \rangle,$$
$$\prod_{i \in I} \langle X_i, Y_i \rangle = \langle \bigcap_{i \in I} X_i, [[R]][[R\breve{}]](\bigcup_{i \in I} Y_i) \rangle.$$

Conversely, a complete lattice L is isomorphic to some $\mathcal{C}_M$, if and only if there are mappings $\gamma : U \to L, \ \mu : V \to L$ such that

$$\{\gamma(x) : x \in U\} \text{ is join–dense in } L,$$
$$\{\mu(y) : y \in V\} \text{ is meet–dense in } L,$$
$$xRy \iff \gamma(x) \leq \mu(y) \text{ for all } x \in U, y \in V.$$

$\mathcal{C}_M$ is called the *concept lattice of* M.

Concept lattices have proved to be quite useful in qualitative data analysis, but they are not a panacea, as the following example shows [8]: Let U be a set of problems, V be a set of skills, and $R \subseteq U \times V$ such that qRs is interpreted as

Skill s is necessary to solve q, and the skill set $R(q)$ is minimally sufficient to solve q.

Suppose that $X \subseteq U$ is the set of all problems which some student t has solved in a test. If one assumes that X is a true representation of the student's state of knowledge, then

1. For each $q \in X$, the student has all the skills to solve q (no lucky guesses). By our operationalisation, these are given by $R(q)$; thus, the student possesses all skills in $\bigcup_{q \in X} R(q) = \langle R \rangle(X)$. This is a somewhat conservative interpretation, since the student may possess other skills that are necessary, but not sufficient, to solve an additional problem not in X.

2. The student actually has solved all problems which can be solved with the skills in $\langle R \rangle(X)$ (no careless errors). Thus, for each $q \in U$, $R(q) \subseteq \langle R \rangle(X)$ implies $q \in X$; in other words, $[R\breve{}]\langle R \rangle(X) \subseteq X$.

We shall see in Lemma 2 that $[R\breve{}]\langle R \rangle(X)$ is actually a closure operator; thus, we have $[R\breve{}]\langle R \rangle(X) = X$ in this case, and the true knowledge states are the closed sets with respect to this operator. More generally, we can regard $[R\breve{}]\langle R \rangle(X)$ as an upper bound of the collection of problems which t is capable of solving.

Similarly, if qRs is interpreted as

It is possible to solve problem q with skill s,

then $\langle R\breve{} \rangle[R](X)$ is a lower bound of the collection of problems which t is capable of solving.

The usefulness of $[[R\breve{}]]$ or $[[R]]$ is somewhat limited in this context, since $[[R]](X)$ will be small or empty, in case the student has managed to solve problems which test different skills.

These considerations lead to the following: If $X \subseteq U$ and $Y \subseteq V$ we call $\langle R \rangle(X)$ the *span of* X, and $[R\breve{}](Y)$ the *content of* Y. The span of X is the set of all properties which are related to some element of X, and the content of Y is the set of those objects which can be completely described by the properties in Y. These operators have the following properties:

Lemma 2. *1. $[R\breve{}]\langle R \rangle$ is a closure operator on 2^U.*

2. $\langle R \rangle[R\breve{}]$ is an interior operator on 2^V.

3. $x \in [R\breve{}]\langle R \rangle(\emptyset) \iff x \notin \operatorname{dom} R$.

Proof. 1. Using (2.12) and the fact that $\langle S \rangle$ and $[S]$ as well as $[[S]]$ and $\langle\langle S \rangle\rangle$ are dual to each other, we obtain

(3.1) $$[R\breve{}]\langle R \rangle(X) = [[(-R)\breve{}]](V \setminus \langle R \rangle(X)) = [[(-R)\breve{}]]([R](V \setminus X)) = [[(-R)\breve{}]][[-R]](X).$$

It is well known that $[[S\breve{}]][[S]]$ is a closure operator for any $S \subseteq U \times V$ [25], and thus, so is $[R\breve{}]\langle R \rangle$.

2. follows from the fact that $\langle R \rangle[R\breve{}]$ is the dual of $[R]\langle R\breve{} \rangle$.

3. $x \in [R\breve{}]\langle R \rangle(\emptyset) \iff R(x) \subseteq \langle R \rangle(\emptyset) \iff R(x) \subseteq \emptyset \iff x \notin \operatorname{dom} R$. □

We call $[R\breve{}]\langle R \rangle(X)$ the *upper bound of* X and $\langle R\breve{} \rangle[R](X)$ the *lower bound of* X, both with respect to R.

In related development, Wong et al. [26] define an interval structure as a pair $\langle f, g \rangle$ of mappings between two Boolean algebras which have approximately the properties of the pair of operators $\langle [R], \langle R \rangle \rangle$.

If $U = V$ and R is a transitive relation on U, then $[R\breve{}]\langle R\rangle = \langle R\rangle$ and $\langle R\rangle[R\breve{}] = [R\breve{}]$; therefore, $[R\breve{}]\langle R\rangle$ coincides with the upper approximation operator and $\langle R\rangle[R\breve{}]$ with the lower approximation operator of rough set theory.

For reflexive relations, Slowinski & Vanderpooten [23] propose $\langle R\rangle$ as an upper approximation operator. This definition has the drawback that, unless R is also transitive, $\langle R\rangle$ is not idempotent, so that we may have the situation that $\langle R\rangle(X) \subsetneq \langle R\rangle\langle R\rangle(X)$. On the contrary, $[R\breve{}]\langle R\rangle$ is a closure operator regardless of the properties of R.

Let us denote the set of all pairs $\langle X,Y\rangle$ with $X = [R\breve{}](Y),\ Y = \langle R\rangle(X)$ by SC_M. We now have a fundamental theorem for SC_M, analogous to Proposition 1:

Proposition 3. *SC_M becomes a complete lattice by setting*

$$\sum_{i\in I}\langle X_i,Y_i\rangle = \langle[R\breve{}]\langle R\rangle(\bigcup_{i\in I}X_i),\bigcup_{i\in I}Y_i\rangle, \tag{3.2}$$

$$\prod_{i\in I}\langle X_i,Y_i\rangle = \langle\bigcap_{i\in I}X_i,\langle R\rangle[R\breve{}](\bigcap_{i\in I}Y_i\rangle \tag{3.3}$$

Conversely, a complete lattice L is isomorphic to some SC_M if and only if there are mappings $\gamma : U \to L,\ \mu : V \to L$ such that

$$\begin{aligned}&\{\gamma(x) : x \in U\} \cup \{0\} \text{ is join–dense in } L,\\ &\{\mu(y) : y \in V\} \cup \{1\} \text{ is meet–dense in } L,\\ &xRy \iff \gamma(x) \not\leq \mu(y) \text{ for all } x \in U, y \in V.\end{aligned}$$

Proof. This can be inferred from Proposition 1 and the fact that $[R\breve{}]\langle R\rangle(X) = [[(-R)\breve{}]][[-R]](X)$. $\square$

The result shows that $SC_{\langle U,V,R\rangle}$ is isomorphic to the concept lattice $\mathcal{C}_{\langle U,V,-R\rangle}$. The internal structure, as well as the semantic interpretation of the two lattices are, however, different. Unlike the extent–intent operators of FCA, our construction is asymmetric, since we use one operator into one direction, and its dual in the opposite direction. Furthermore, for $\langle X_0,Y_0\rangle,\ \langle X_1,Y_1\rangle \in SC_M$, we have

$$\langle X_0,Y_0\rangle \leq \langle X_1,Y_1\rangle \iff X_0 \subseteq X_1 \iff Y_0 \subseteq Y_1,$$

so that $\leq$ is isotone in both components.

4 The Morse data

In this Section we will give an example how the modal-style operators can be applied to relations of similarity. For related work in a similar context we invite the reader to consult [2]. The data under investigation, a flagship of multidimensional scaling, were originally collected by Rothkopf [21] in the following context:

> "The S[ubject]s of this experiment were exposed to pairs of aural Morse signals sent at a high tone speed. The signals of each pair were separated by a short temporal interval. The S[ubject]s were asked to indicate whether they thought the signals were the same (or different) by making the appropriate remark on an IBM True–False Answer sheet. Each S[ubject] was asked to respond in this fashion to 351 different pairs of Morse signals."

Table 1: Morse code

a	·–	k	–·–	u	··–	0	–––––
b	–···	l	·–··	v	···–	1	·––––
c	–·–·	m	––	w	·––	2	··–––
d	–··	n	–·	x	–··–	3	···––
e	·	o	–––	y	–·––	4	····–
f	··–·	p	·––·	z	––··	5	·····
g	––·	q	––·–			6	–····
h	····	r	·–·			7	––···
i	··	s	···			8	–––··
j	·–––	t	–			9	––––·

The data is given as a matrix, with rows and columns labeled by the alphanumeric characters (Table 5 on page 12). An entry s in cell $\langle p, q \rangle$ means that s% of subjects regarded the code for p and q as the same signal. In the sequel, we will refer to p as the *first stimulus* or as being *in the first position*, and to q as the *second stimulus*, or as being *in the second position*. We use upper case letters for first stimuli and lower case letters for second stimuli; the numeric characters are prefixed by a $*$, if they occur as second stimuli. We emphasise that these are only notational conveniences, so that e.g. a and A or 1 and $*1$ correspond to the same code sequence. The matrix diagonal corresponds to pairs which are truly the same, the off-diagonal entries correspond to pairs which are truly different. It should be noted that the matrix is not symmetric, and that the entries in the diagonal are always less than 100. Thus, we have an example of a possibly non-reflexive asymmetric relation which expresses some form of similarity.

Shepard [22] describes the data using the dimensions

1. Length of the signal,
2. Distribution of dots and dashes in the signal, going from only dots to only dashes.

The distances between the points in a plane spanned by these dimensions reflect (partially) the ordinal relation among the given proximities, see Figure 1.

For various “cut points ” s we consider the relation

$$R_s = \{\langle p, q \rangle : \text{At least } s\% \text{ of the subjects responded “same”, when } \langle p, q \rangle \text{ was presented}\}.$$

Observe that $R_s \subseteq R_t$ in case $t \leq s$.

As the length of the signal is one of the dimension identified in [22] (and also in [3]), we are interested in the behaviour of the modal–style operators on the sets

$$X_n = \{p : \text{The length of the Morse code for first stimulus } p \text{ is } n\},$$
$$Y_n = \{q : \text{The length of the Morse code for second stimulus } q \text{ is } n\}.$$

which are given in Table 2.

Figure 1: Multi dimensional scaling of the Morse data [from 13, p. 13]

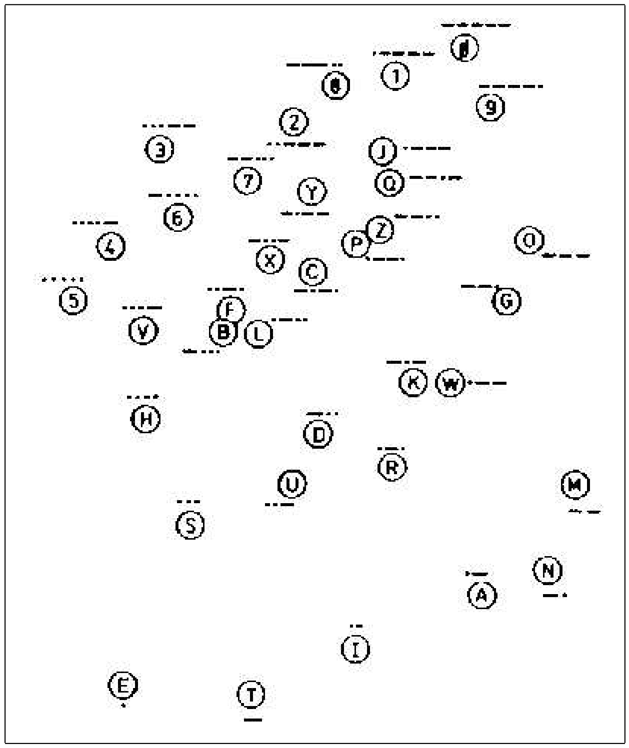

Table 2: Distinguished sets

Stimulus (first position)	Stimulus (second position)
$X_1 = \{E,T\}$	$Y_1 = \{e,t\}$
$X_2 = \{A,I,M,N\}$	$Y_2 = \{a,i,m,n\}$
$X_3 = \{D,G,K,O,R,S,U,W\}$	$Y_3 = \{d,g,k,o,r,s,u,w\}$
$X_4 = \{B,C,F,H,J,L,P,Q,V,X,Y,Z\}$	$Y_4 = \{b,c,f,h,j,l,p,q,v,x,y,z\}$
$X_5 = \{0,1,2,3,4,5,6,7,8,9\}$	$Y_5 = \{*0,*1,*2,*3,*4,*5,*6,*7,*8,*9\}$

Disregarding for the moment the cut off parameter, the equality interpretation of the operators is as follows: If we start with the first stimuli, in particular the sets X_n, then

$$
\begin{aligned}
q \in \langle R \rangle (X_n) &\iff q \text{ was gauged to be the same as some first stimulus of length } n. \\
q \in [R](X_n) &\iff q \text{ was gauged to be the same only as first stimuli of length } n. \\
q \in [[R]](X_n) &\iff q \text{ was gauged to be the same to all first stimuli of length } n, \text{ and possibly others.}
\end{aligned}
$$

If we start with a set Y of stimuli at the second position, we replace R by $R^{\smile}$, which means that “first” is replace by “second” in the definition of the sets. Putting these together, we obtain

$p \in [R^{\smile}]\langle R\rangle(X_n) \iff$ Every signal, which cannot be distinguished from p cannot be distinguished from some stimulus of length n.

$p \in \langle R^{\smile}\rangle[R](X_n) \iff$ Some signals, which cannot be distinguished from p were gauged to be the same only to stimuli of length n.

$p \in [[R^{\smile}]][[R]](X_n) \iff$ Whenever q cannot be distinguished from all stimuli of length n, then q cannot be distinguished from p.

If we consider cut off points s and t, we interpret, for example, for a first stimulus p,

$p \in [R_s^{\smile}]\langle R_t\rangle(X_n) \iff$ Every second stimulus which could not be distinguished from p by at least s% of all subjects could not be distinguished from some first stimulus of length n by at least t% of all subjects.

$p \in \langle R_s^{\smile}\rangle[R_t](X_n) \iff$ There is a second stimulus q such that at least s% of subjects gauged q to be the same as p, and at least t% of subjects gauged q to be the same only as stimuli of length n.

A first impression of the difficulties encountered by the subjects when discriminating the first and second stimuli is offered by the binary relations R_{50} and $R_{50}{}^{\smile}$, which are generated, when the probability cut point $p = 0.5$ is used. Table 3 presents the results of the operators applied to the sets X_n of first and Y_n of second stimuli. One can see that applying the sufficiency operators is not suitable for this situation, since the results are too coarse (see the last column of Table 3). The combination of content and span operators seem to be more promising in either direction.

Each signal can be interpreted in two ways – as confusing the first stimulus with the second one and vice versa–, and we can apply the operators starting with either case.

Stimuli of length 1 or 2 are easily distinguished from those of different length. Starting with second stimuli of length 3, we see that none of d,k,s,u is contained in the lower bound $\langle R\rangle[R^{\smile}](Y_3)$. When we consider these signals as first stimuli, then this is not the case, since $\langle R^{\smile}\rangle[R](X_3) = X_3$. This result is hard to present in geometrical terms, as the scaling proposed by Shepard [22] uses a non-metric MDS approach (Fig. 1).

Signals of length 4 and 5 are harder to distinguish. We observe that the signals H and h of length 4, and *6, *6, 7, *7* of length 5 cause considerable confusion. This cannot be determined from the geometrical representation given in Fig. 1. Indeed, the first stimulus H seems to have the largest distance of any element of set X_4 to the set X_5 in Fig. 1, and thus, according to the model, not much confusion should arise.

Another interesting stimulus seems to be the Morse code ···– of character V, because this code of length 4 is confused with stimuli of length 3 and 5. Therefore, V should be presented in a “bridging position” in a geometrical presentation.

Variation of the cut point offers further insights. In Tab. 4 we present the set differences of lower and upper bound of the signal sets for $s = 80, 70, 60, 40$. The first entry in a cell $\langle Z, s\rangle$ is the set of elements which are in the set Z, but not in the lower bound, and the second entry is the set of those elements which are not in Z, but belong to the upper bound of Z. Inspecting the result in Tab. 4, we observe that the signals *6* and **6* seem to be very hard to distinguish from the signals of length 4, – an effect which is worse for **6*. Furthermore, *5* and **5* frequently appear in one of the differences.

Table 3: Modal-style operators applied to Morse data (Cut point p=0.5)

X_n	$[\breve{R}]\langle R\rangle$	$\langle\breve{R}\rangle[R]$	$[[\breve{R}]][[R]]$
E T	e t E T	e t E T	t E T
A I M N	a i m n A I M N	a i m n A I M N	$\emptyset$ **1**
D G K O R S U W	b d g h k l o p r s u v w x D G K O R S U **V** W	d g o r s u w D G K O R S U W	$\emptyset$ **1**
B C F **H** J **L** P Q **V** X Y Z	b c f h j k l p q v x y z *1 *2 *5 *6 *7 *8 B C F H J L P Q V X Y Z **6 7**	c f j q y C F J P Q X Y Z	$\emptyset$ **1**
1 2 3 4 5 **6 7** 8 9 0	b h v x z *1 *2 *3 *4 *5 *6 *7 *8 *9 *0 **B H V** 1 2 3 4 5 6 7 8 9 0	*3 *4 *9 *0 1 2 3 4 5 8 9 0	$\emptyset$ **1**
Y_n	$[R]\langle\breve{R}\rangle$	$\langle R\rangle[\breve{R}]$	$[[R]][[\breve{R}]]$
e t	E T e t	E T e t	E e t
a i m n	A I M N a i m n	A I M N a i m n	$\emptyset$ **1**
d g **k** o r **s u** w	D G K O R S U W X d g k o r s u w	O R W g o r w	$\emptyset$ **1**
b c f **h** j l p q v x y z	B C D F G H J K L P Q S U V X Y Z 4 5 6 7 8 b c **d** f h j **k** l p q **s u** v x y z ***4 *5 *6 *7**	C F L P V Y b c f j l p q v x y z	$\emptyset$ **1**
*1 *2 *3 ***4 *5 *6 *7** *8 *9 *0	B H J Q X Z 1 2 3 4 5 6 7 8 9 0 *1 *2 *3 *4 *5 *6 *7 *8 *9 *0	1 2 3 9 0 *1 *2 *3 *8 *9 *0	$\emptyset$ **1**

Bold letter in "codes": Letter does not apear in the lower bound.
Bold letter in "$[\breve{R}]\langle R\rangle$", resp. "$[R]\langle\breve{R}\rangle$": Letter is added in the upper bound.
The first line in a second column cell is the result when applying the inner operator to the original set, the second line is the result when applying the outer operator to the first line.

Summarising, we recognise the troublesome first stimuli *5,6,7* of length 5, *B,H,V* of length 4 and their second position counterparts, and, in addition, the second stimuli $\{d,k,s,u\}$.

With respect to the second dimension of the MDS model, namely, the distribution of dots and dashes, we see from Table 1 that, except for *k*, the problematic signals contain more short than long impulses. A geometric representation has to present the data in a "long–short" dimension, but – since the result pattern is asymmetric – the representation cannot deal with the data in an adequate manner. It was shown in [15] that in fact an asymmetric "drift" from short to long can be extracted, when MDS is applied to "residual proximities"; these can be computed by the difference of the original data and the estimated symmetric proximity matrix, which is the base of the classical MDS approach.[1]

Therefore, our operator–based qualitative analysis supports the findings of the MDS model, and offers some additional explanations. These are, in short,

- The signal length is the first determining factor for the discrimination of the stimuli, because:
  - Signals of length 1 or 2 are easy to discriminate from other stimuli.
  - Signals of length 3 are easy to discriminate from other stimuli, if they are located at the first position.

[1] It is interesting to note that the first MDS approach of Shepard [22] was published in 1963 and the "asymmetric extensions" of Möbus [15] appeared 16 years later.

Table 4: Difference of lower and upper bound given varied cut points in Morse Data

Codes	Cut = 0.8	Cut = 0.7	Cut = 0.6	Cut = 0.5	Cut = 0.4
ET	$\emptyset,\emptyset$	$\emptyset,\emptyset$	$\emptyset,\emptyset$	$\emptyset,\emptyset$	$\emptyset,\emptyset$
et	$\emptyset,\emptyset$	$\emptyset,\emptyset$	$\emptyset,\emptyset$	$\emptyset,\emptyset$	$\emptyset,\emptyset$
AIMN	$\emptyset,\emptyset$	$\emptyset,\emptyset$	$\emptyset,\emptyset$	$\emptyset,\emptyset$	$\emptyset,\emptyset$
aimn	$\emptyset,\emptyset$	$\emptyset,\emptyset$	$\emptyset,\emptyset$	$\emptyset,\emptyset$	$\emptyset,\emptyset$
DGKORSUW	$\emptyset,\emptyset$	$\emptyset,\emptyset$	$\emptyset,\emptyset$	–,V	K,–
dgkorsuw	$\emptyset,\emptyset$	$\emptyset,\emptyset$	$\{d,k\},\emptyset$	$\{d,k,s,u\},\emptyset$	$\{d,k,s,u\},\emptyset$
BCFHJLPQVXYZ	$\emptyset,\emptyset$	$\emptyset,\{5,6\}$	$\{B,H\},\{1,2,6\}$	$\{B,H,L,V\},\{6,7\}$	$\{B,F,H,L,V\},\{K,2,3,4,5,6,7\}$
bcfhjlpqvxyz	$\emptyset,\{*6\}$	$\{b,h\},\{*6\}$	$\{h\},\{d,k,*4,*5\}$	$\{h\},\{d,k,s,u,*4,*5,*6,*7\}$	$\{b,f,h,j,l,q,v,x,z\},\{d,k,s,u,*4,*5\}$
1234567890	$\emptyset,\emptyset$	$\{5,6\},\emptyset$	$\{1,2,6\},\{H\}$	$\{6,7\},\{B,H,V\}$	$\{2,3,4,5,6,7\},\{B,H,V\}$
*1*2*3*4*5*6*7*8*9*0	$\{*6\},\emptyset$	$\{*6\},\{b,h\}$	$\{*4,*5\},\{h\}$	$\{*4,*5,*6,*7\},\emptyset$	$\{*4,*5\},\{f,j,q\}$

  - Signals of length 3 in the second position overlap with signals of length 4. Signals of length 4 overlap mainly with signals of length 5.

- The character of the impulses is of less effect because a signal must contain mainly short Morse impulses, and should contain at least 4 (first stimuli) or 3 (second stimuli) Morse impulses to be hard to discriminate.

- Asymmetric features of the data are reflected by the construction. There is no need for an extra analysis of method–dependent "residual matrices".

5 Discussion

The presented modal operator approach offers a complementary view of data with respect to derivation operator of formal concept analysis. In principle, the proposed operators can be derived from concept analysis by applying the intent–extent operators to $-R$, and building complements of the resulting concept sets. This is nice, because the computation of convolutions of possibility and necessity operators can be perfomed by programs for concept analysis, and using the de Morgan rules. Of course, this does not mean that the proposed analysis based on possibility and necessity operators is the same as applying concept analysis, because

- Both proposed operators act asymmetrically, while intent and extent of FCA are symmetric.

- The combination of $\langle R\rangle$ and $[R]$ can be interpreted as a generalisation of rough sets approximations, which are based on equivalence relations.

Comparing the proposed theory with MDS, we observe that it offers comparable results, and that these results are presented in a direct manner: There is no need for a 2-dimensional representation (which is not even adequate for Morse data as Shepard [22] remarks), and the risk of so called *divergence artifacts* [9] is reduced. It should be noted, however, that the proposed theory offers a literally "rough approximation" to the data: Once a cut point p is chosen, all differences below this cut point are neglected: It has to be assumed that these differences are not relevant for further interpretation. This is different to the MDS approach; there, the rank order of the proximities is used, which contains more information than taking a simple cut.

Although the proposed theory is nice, handy and applicable from scratch, there is an observation which opens a box of further questions: Unlike for equivalence relations, the $\subseteq$ ordering on relations is not reflected by the new definition of lower and upper bounds, i.e. $R \subseteq S$ does not necessarily imply $[R\breve{}]\langle R\rangle(X) \subseteq [S\breve{}]\langle S\rangle(X)$. The question arises, which kind of compatibility assumptions must hold in order for the structural properties of the relations to generate comparable properties in the results of the operators.

Table 5: Morse data [22]

	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z	*1	*2	*3	*4	*5	*6	*7	*8	*9	*0
A	92	4	6	13	3	14	10	13	46	5	22	3	25	34	6	6	9	35	23	6	37	13	17	12	7	3	2	7	5	5	8	6	5	6	2	3
B	5	84	37	31	5	28	17	21	5	19	34	40	6	10	12	22	25	16	18	2	18	34	8	84	30	42	12	17	14	40	32	74	43	17	4	4
C	4	38	87	17	4	29	13	7	11	19	24	35	14	3	9	51	34	24	14	6	6	11	14	32	82	38	13	15	31	14	10	30	28	24	18	12
D	8	62	17	88	7	23	40	36	9	13	81	56	8	7	9	27	9	45	29	6	17	20	27	40	15	33	3	9	6	11	9	19	8	10	5	6
E	6	13	14	6	97	2	4	4	17	1	5	6	4	4	5	1	5	10	7	67	3	3	2	5	6	5	4	3	5	3	5	2	4	2	3	3
F	4	51	33	19	2	90	10	29	5	33	16	50	7	6	10	42	12	35	14	2	21	27	25	19	27	13	8	16	47	25	26	24	21	5	5	5
G	9	18	27	38	1	14	90	6	5	22	33	16	14	13	62	52	23	21	5	3	15	14	32	21	23	39	15	14	5	10	4	10	17	23	20	11
H	3	45	23	25	9	32	8	87	10	10	9	29	5	8	8	14	8	17	37	4	36	59	9	33	14	11	3	9	15	43	70	35	17	4	3	3
I	64	7	7	13	10	8	6	12	93	3	5	16	13	30	7	3	5	19	35	16	10	5	8	2	5	7	2	5	8	9	6	8	5	2	4	5
J	7	9	38	9	2	24	18	5	4	85	22	31	8	3	21	63	47	11	2	7	9	9	9	22	32	28	67	66	33	15	7	11	28	29	26	23
K	5	24	38	73	1	17	25	11	5	27	91	33	10	12	31	14	31	22	2	2	23	17	33	63	16	18	5	9	17	8	8	18	14	13	5	6
L	2	69	43	45	10	24	12	26	9	30	27	86	6	2	9	37	36	28	12	5	16	19	20	31	25	59	12	13	17	15	26	29	36	16	7	3
M	24	12	5	14	7	17	29	8	8	11	23	8	96	62	11	10	15	20	7	9	13	4	21	9	18	8	5	7	6	6	5	7	11	7	10	4
N	31	4	13	30	8	12	10	16	13	3	16	8	59	93	5	9	5	28	12	10	16	4	12	4	6	11	5	2	3	4	4	6	2	2	10	2
O	7	7	20	6	5	9	76	7	2	39	26	10	4	8	86	37	35	10	3	4	11	14	25	35	27	27	19	17	7	7	6	18	14	11	20	12
P	5	22	33	12	5	36	22	12	3	78	14	46	5	6	21	83	43	23	9	4	12	19	19	19	41	30	34	44	24	11	15	17	24	23	25	13
Q	8	20	38	11	4	15	10	5	2	27	23	26	7	6	22	51	91	11	2	3	6	14	12	37	50	63	34	32	17	12	9	27	40	58	37	24
R	13	14	16	23	5	34	26	15	7	12	21	33	14	12	12	29	8	87	16	2	23	23	62	14	12	13	7	10	13	4	7	12	7	9	1	2
S	17	24	5	30	11	26	5	59	16	3	13	10	5	17	6	6	3	18	96	9	56	24	12	10	6	7	8	2	2	15	28	9	5	5	5	2
T	13	10	1	5	46	3	6	6	14	6	14	7	6	5	6	11	4	4	7	96	8	5	4	2	2	6	5	5	3	3	3	8	7	6	14	6
U	14	29	12	32	4	32	11	34	21	7	44	32	11	13	6	20	12	40	51	6	93	57	34	17	9	11	6	6	16	34	10	9	9	7	4	3
V	5	17	24	16	9	29	6	39	5	11	26	43	4	1	9	17	10	17	11	6	32	92	17	57	35	10	10	14	28	79	44	36	25	10	1	5
W	9	21	30	22	9	36	25	15	4	25	29	18	15	6	26	20	25	61	12	4	19	20	86	22	25	22	10	22	19	16	5	9	11	6	3	7
X	7	64	45	19	3	28	11	6	1	35	50	42	10	8	24	32	61	10	12	3	12	17	21	91	48	26	12	20	24	27	16	57	29	16	17	6
Y	9	23	62	15	4	26	22	9	1	30	12	14	5	6	14	30	52	5	7	4	6	13	21	44	86	23	26	44	40	15	11	26	22	33	23	16
Z	3	46	45	18	2	22	17	10	7	23	21	51	11	2	15	59	72	14	4	3	9	11	12	36	42	87	16	21	27	9	10	25	66	47	15	15
1	2	5	10	3	3	5	13	4	2	29	5	14	9	7	14	30	28	9	4	2	3	12	14	17	19	22	84	63	13	8	10	8	19	32	57	55
2	7	14	22	5	4	20	13	3	25	26	9	14	2	3	17	37	28	6	5	3	6	10	11	17	30	13	62	89	54	20	5	14	20	21	16	11
3	3	8	21	5	4	32	6	12	2	23	6	13	5	2	5	37	19	9	7	6	4	16	6	22	25	12	18	64	86	31	23	41	16	17	8	10
4	6	19	19	12	8	25	14	16	7	21	13	19	3	3	2	17	29	11	9	3	17	55	8	37	24	3	5	26	44	89	42	44	32	10	3	3
5	8	45	15	14	2	45	4	67	7	14	4	41	2	0	4	13	7	9	27	2	14	45	7	45	10	10	14	10	30	69	90	42	24	10	6	5
6	7	80	30	17	4	23	4	14	2	11	11	27	6	2	7	16	30	11	14	3	12	30	9	58	38	39	15	14	26	24	17	88	69	14	5	14
7	6	33	22	14	5	25	6	4	6	24	13	32	7	6	7	36	39	12	6	2	3	13	9	30	30	50	22	29	18	15	12	61	85	70	20	13
8	3	23	40	6	3	15	15	6	2	33	10	14	3	6	14	12	45	2	6	4	6	7	5	24	35	50	42	29	16	16	9	30	60	89	61	26
9	3	14	23	3	1	6	14	5	2	30	6	7	16	11	10	31	32	5	6	7	6	3	8	11	21	24	57	39	9	12	4	11	42	56	91	78
0	9	3	11	2	5	7	14	4	5	30	8	3	2	3	25	21	29	2	3	4	5	3	2	12	15	20	50	26	9	11	5	22	17	52	81	94

References

[1] Birkhoff, G. (1948). Lattice Theory, vol. 25 of *Am. Math. Soc. Colloquium Publications*. Providence: AMS, 2nd Edn.

[2] Bisdorff, R. & Roubens, M. (2002). Clustering with null kernels for a valued similarity relation: Application to the Morse data. Presented at the l'Aquila workshop of COST Action 274.

[3] Buja, A. & Swayne, D. F. (2001). Visualization methodology for multidimensional scaling. Preprint.

[4] Davey, B. A. & Priestley, H. A. (1990). Introduction to Lattices and Order. Cambridge University Press.

[5] Düntsch, I., Gediga, G. & Orłowska, E. (2001). Relational attribute systems. *International Journal of Human Computer Studies*, **55**, 293–309.

[6] Düntsch, I. & Orłowska, E. (2000). Beyond modalities: Sufficiency and mixed algebras. In E. Orłowska & A. Szałas (Eds.), *Relational Methods in Computer Science Applications*, 277–299, Heidelberg. Physica Verlag.

[7] Gargov, G., Passy, S. & Tinchev, T. (1987). Modal environment for Boolean speculations. In D. Skordev (Ed.), *Mathematical Logic and Applications*, 253–263, New York. Plenum Press.

[8] Gediga, G. & Düntsch, I. (2002). Skill set analysis in knowledge structures. *British Journal of. Mathematical and Statistical Psychology*. To appear.

[9] Gigerenzer, G. (1981). Messung und Modellbildung in der Psychologie. Basel: Birkhäuser.

[10] Godin, R., Missaoui, R. & Alaoui, H. (1995). Incremental concept formation algorithms based on Galois (concept) lattices. *Computational Intelligence*, **11**, 246–267.

[11] Humberstone, I. L. (1983). Inaccessible worlds. *Notre Dame Journal of Formal Logic*, **24**, 346–352.

[12] Jónsson, B. & Tarski, A. (1951). Boolean algebras with operators I. *American Journal of Mathematics*, **73**, 891–939.

[13] Kruskal, J. B. & Wish, M. (1978). Multidimensional Scaling. No. 11 in Quantitative Applications in the Social Sciences. Newbury Park: Sage.

[14] Lipski, W. (1976). Informational systems with incomplete information. In S. Michaelson & R. Milner (Eds.), *Third International Colloquium on Automata, Languages and Programming*, 120–130, University of Edinburgh. Edinburgh University Press.

[15] Möbus, C. (1979). Zur Analyse nichtsymmetrischer Ähnlichkeitsurteile: Ein dimensionales Driftmodell, eine Vergleichshypothese, Tversky's Kontrastmodell und seine Fokushypothese. *Archiv für Psychologie*, **131**, 105–136.

[16] Orłowska, E. (1988). Logical aspects of learning concepts. *Journal of Approximate Reasoning*, **2**, 349–364.

[17] Orłowska, E. (Ed.) (1997). Incomplete Information – Rough Set Analysis. Heidelberg: Physica – Verlag.

[18] Orłowska, E. & Pawlak, Z. (1984). Representation of nondeterministic information. *Theoretical Computer Science*, **29**, 27–39.

[19] Pawlak, Z. (1973). Mathematical foundations of information retrieval. ICS Research Report 101, Polish Academy of Sciences.

[20] Pawlak, Z. (1982). Rough sets. *Internat. J. Comput. Inform. Sci.*, **11**, 341–356.

[21] Rothkopf, E. Z. (1957). A measure of stimulus similarity and errors in some paired-associate learning tasks. *Journal of Experimental Psychology*, **53**, 94–101.

[22] Shepard, R. N. (1963). Analysis of proximities as a technique for the study of information processing in man. *Human Factors*, **5**, 33–48.

[23] Slowinski & Vanderpooten (2000). A generalized definition of rough approximations based on similarity. *IEEE Transactions on Knowledge and Data Engineering*, **12**, 331–336.

[24] Vakarelov, D. (1997). Information systems, similarity relations and modal logics. In [17], 492–550.

[25] Wille, R. (1982). Restructuring lattice theory: An approach based on hierarchies of concepts. In I. Rival (Ed.), *Ordered sets*, vol. 83 of *NATO Advanced Studies Institute*, 445–470. Dordrecht: Reidel.

[26] Wong, S., Wang, L. & Yao, Y. (1995). On modeling uncertainty with interval structures. *Computational Intelligence*, **11**, 406–426.

[27] Yao, Y. Y. (1998). On generalizing Pawlak approximation operators. In L. Polkowski & A. Skowron (Eds.), *Proceedings of the 1st International Conference on Rough Sets and Current Trends in Computing (RSCTC-98)*, vol. 1424 of *LNAI*, 298–307, Berlin. Springer.

[28] Yao, Y. Y. & Lin, T. Y. (1996). Generalization of rough sets using modal logic. *Intelligent Automation and Soft Computing*, **2**, 103–120.