Formal Concept Analysis as Mathematical Theory of Concepts and Concept Hierarchies

Rudolf Wille

Technische Universität Darmstadt, Fachbereich Mathematik
Schloßgartenstr. 7, D–64289 Darmstadt
wille@mathematik.tu-darmstadt.de

Abstract. *Formal Concept Analysis* has been originally developed as a subfield of Applied Mathematics based on the mathematization of *concept* and *concept hierarchy*. Only after more than a decade of development, the connections to the philosophical logic of human thought became clearer and even later the connections to Piaget's cognitive structuralism which Thomas Bernhard Seiler convincingly elaborated to a comprehensive *theory of concepts* in his recent book [Se01]. It is the main concern of this paper to show the surprisingly rich correspondences between Seiler's multifarious aspects of concepts in the human mind and the structural properties and relationships of formal concepts in Formal Concept Analysis. These correspondences make understandable, what has been experienced in a great multitude of applications, that Formal Concept Analysis may function in the sense of *transdisciplinary mathematics*, i.e., it allows mathematical thought to aggregate with other ways of thinking and thereby to support human thought and action.

1 Formal Concept Analysis, Mathematics, and Logic

Formal Concept Analysis had its origin in activities of restructuring mathematics, in particular mathematical order and lattice theory. In the initial paper [Wi82], restructuring lattice theory is explained as "an attempt to reinvigorate connections with our general culture by interpreting the theory as concretely as possible, and in this way to promote better communication between lattice theorists and potential users of lattice theory." Since then, Formal Concept Analysis has been developed as a subfield of *Applied Mathematics* based on the mathematization of concepts and concept hierarchies.

Only after more than a decade of development, the connections to *Philosophical Logics* of human thought became clearer, mainly through Charles Sanders Peirce's late philosophy. Even our general understanding of mathematics did improve as pointed out in the recent paper "Kommunikative Rationalität, Logik und Mathematik" ("Communicative Rationality, Logic, and Mathematics") [Wi02b]. The concern of that paper is to explain and to substantiate the following thesis:

B. Ganter et al. (Eds.): Formal Concept Analysis, LNAI 3626, pp. 1–33, 2005.

> *The aim and meaning of mathematics finally lie in the fact that mathematics is able to effectively support the rational communication of humans.*

Here we only recall the key arguments founding this thesis: First, logical thinking as expression of human reason graps actual realities by the main forms of human thought: concepts, judgments, and conclusions (cf. [Ka88], p.6). Second, mathematical thinking abstracts logical human thinking for developing a cosmos of forms of potential realities (see [Pe92], p.121). Therefore, mathematics as a historically, socially and culturally detemined formation of mathematical thinking, respectively, is able to support humans in their logical thinking and hence in their rational communication. Since concepts are also prerequisites for the formation of judgments and conclusions, we can adapt the above thesis to Formal Concept Analysis as follows:

> *The aim and meaning of Formal Concept Analysis as mathematical theory of concepts and concept hierarchies is to support the rational communication of humans by mathematically developing appropriate conceptual structures which can be logically activated.*

2 Concepts and Formal Concepts

Concepts can be philosophically understood as the basic units of thought formed in dynamic processes within social and cultural environments. According to the main philosophical tradition, a concept is constituted by its *extension*, comprising all objects which belong to the concept, and its *intension*, including all attributes (properties, meanings) which apply to all objects of the extension (cf. [Wi95]). Concepts can only live in relationships with many other concepts where the *subconcept-superconcept-relation* plays a prominent role. Being a subconcept of a superconcept means that the extension of the subconcept is contained in the extension of the superconcept which is equivalent to the relationship that the intension of the subconcept contains the intension of the superconcept (cf. [Wa73], p.201).

For a mathematical theory of concepts and concept hierarchies, we obviously need a mathematical model that allows to speak mathematically about *objects*, *attributes*, and relationships which indicate that an object *has* an attribute. Such a model was introduced in [Wi82] by the notion of a "formal context" which turned out to be basic for a new area of applied mathematics: *Formal Concept Analysis*. A *formal context* is defined as a set structure $\mathbb{K} := (G, M, I)$ for which G and M are sets while I is a binary relation between G and M, i.e. $I \subseteq G \times M$; the elements of G and M are called (*formal*) *objects* (in German: *Gegenstände*) and (*formal*) *attributes* (in German: *Merkmale*), respectively, and gIm, i.e. $(g, m) \in I$, is read: the object g *has* the attribute m.

For defining the formal concepts of the formal context (G, M, I), we need the following *derivation operators* defined for arbitrary $X \subseteq G$ and $Y \subseteq M$ as follows:

$$\begin{array}{rcl} X & \mapsto & X^I := \{m \in M \mid gIm \text{ for all } g \in X\}, \\ Y & \mapsto & Y^I := \{g \in G \mid gIm \text{ for all } m \in Y\}. \end{array}$$

The two derivation operators satisfy the following three conditions:

$$(1)\ Z_1 \subseteq Z_2 \Longrightarrow Z_1^I \supseteq Z_2^I,\ \ (2)\ Z \subseteq Z^{II},\ \ (3)\ Z^{III} = Z^I.$$

A *formal concept* of a formal context $\mathbb{K} := (G, M, I)$ is defined as a pair (A, B) with $A \subseteq G$, $B \subseteq M$, $A = B^I$, and $B = A^I$; A and B are called the *extent* and the *intent* of the formal concept (A, B), respectively. The *subconcept-superconcept-relation* is mathematized by

$$(A_1, B_1) \leq (A_2, B_2) :\Longleftrightarrow A_1 \subseteq A_2\ (\Longleftrightarrow B_1 \supseteq B_2).$$

The set of all formal concepts of $\mathbb{K}$ together with the defined order relation is denoted by $\underline{\mathfrak{B}}(\mathbb{K})$.

A general method of constructing formal concepts uses the derivation operators to obtain, for $X \subseteq G$ and $Y \subseteq M$, the formal concepts (X^{II}, X^I) and (Y^I, Y^{II}). For an object $g \in G$, its *object concept* $\gamma g := (\{g\}^{II}, \{g\}^I)$ is the smallest concept in $\underline{\mathfrak{B}}(\mathbb{K})$ whose extent contains g and, for an attribute $m \in M$, its *attribute concept* $\mu m := (\{m\}^I, \{m\}^{II})$ is the greatest concept in $\underline{\mathfrak{B}}(\mathbb{K})$ whose intent contains m. The specific structure of the ordered sets $\underline{\mathfrak{B}}(\mathbb{K})$ of formal contexts $\mathbb{K}$ is clarified by the following theorem:

Basic Theorem on Concept Lattices. [Wi82] *Let $\mathbb{K} := (G, M, I)$ be a formal context. Then $\underline{\mathfrak{B}}(\mathbb{K})$ is a complete lattice, called the* concept lattice *of (G, M, I), for which infimum and supremum can be described as follows:*

$$\bigwedge_{t\in T}(A_t, B_t) = (\bigcap_{t\in T} A_t, (\bigcup_{t\in T} B_t)^{II}),$$
$$\bigvee_{t\in T}(A_t, B_t) = ((\bigcup_{t\in T} A_t)^{II}, \bigcap_{t\in T} B_t).$$

*In general, a complete lattice L is isomorphic to $\underline{\mathfrak{B}}(\mathbb{K})$ if and only if there exist mappings $\tilde{\gamma} : G \longrightarrow L$ and $\tilde{\mu} : M \longrightarrow L$ such that $\tilde{\gamma}G$ is $\bigvee$-*dense *in L (i.e. $L = \{\bigvee X \mid X \subseteq \tilde{\gamma}G\}$), $\tilde{\mu}M$ is $\bigwedge$-*dense *in L (i.e. $L = \{\bigwedge X \mid X \subseteq \tilde{\mu}M\}$), and $gIm \Longleftrightarrow \tilde{\gamma}g \leq \tilde{\mu}m$ for $g \in G$ and $m \in M$; in particular, $L \cong \underline{\mathfrak{B}}(L, L, \leq)$ and furthermore: $L \cong \underline{\mathfrak{B}}(J(L), M(L), \leq)$ if the set $J(V)$ of all $\bigvee$-irreducible elements is $\bigvee$-dense in L and the set of all $\bigwedge$-irreducible elements is $\bigwedge$-dense in L.*

A formal context is best understood if it is depicted by a *cross table* as for example the formal context about bodies of waters in Fig. 1. A concept lattice is best pictured by a *labelled line diagram* as the concept lattice of our example context in Fig. 2 (see the book cover of [GW99a]). In such a diagram the name of each object g is attached to its represented object concept γg and the name of each attribute m is attached to its represented attribute concept μm. By the Basic Theorem, this labelling allows to read the extents, the intents, and the

	natural	artificial	stagnant	running	inland	maritime	constant	temporary
tarn	×		×		×		×	
trickle	×			×	×		×	
rill	×			×	×		×	
beck	×			×	×		×	
rivulet	×			×	×		×	
runnel	×			×	×		×	
brook	×			×	×		×	
burn	×			×	×		×	
stream	×			×	×		×	
torrent	×			×	×		×	
river	×			×	×		×	
channel				×	×		×	
canal		×		×	×		×	
lagoon	×		×			×	×	
lake	×		×		×		×	
mere		×	×		×		×	
plash	×		×		×			×
pond		×	×		×		×	
pool	×		×		×		×	
puddle	×		×		×			×
reservoir		×	×		×		×	
sea	×		×			×	×	

Fig. 1. Formal context partly representing the lexical field "bodies of waters"

underlying formal context from the diagram. Speaking in human logical terms, by the Basic Theorem, each concept is represented by a little circle so that its extension (intension) consists of all the objects (attributes) whose names can be reached by a descending (ascending) path from that circle. In Fig. 2, for instance, the circle vertically above the circle with the label "artificial" represents the formal concept with the extent $\{tarn, lake, pool, sea, lagoon\}$ and the intent $\{natural, stagnant, constant\}$. Furthermore, even all *attribute implications*

$$A \to B :\Longleftrightarrow\ A^I \subseteq B^I \text{ with } A, B \subseteq M$$

can be read from a labelled line diagram; Fig. 2, for instance, shows the attribute implication $\{artificial\} \to \{inland, constant\}$ because there are ascending paths from the circle with the label "artificial" to the circles with the labels "inland" and "constant", respectively. In the case of $M^I = \emptyset$, an implication $A \to M$ is equivalent to $A^I = \emptyset$ wherefore A is then said to be *incompatible*.

The aim of Section 2 is to give an answer to the following basic question: *How adequate is the mathematization of concepts and concept hierarchies used in Formal Concept Analysis?* For answering this question, we have to refer to a comprehensive convincing *theory of concepts*. Such a theory is presented in the book "Begreifen und Verstehen. Ein Buch über Begriffe und Bedeutungen"

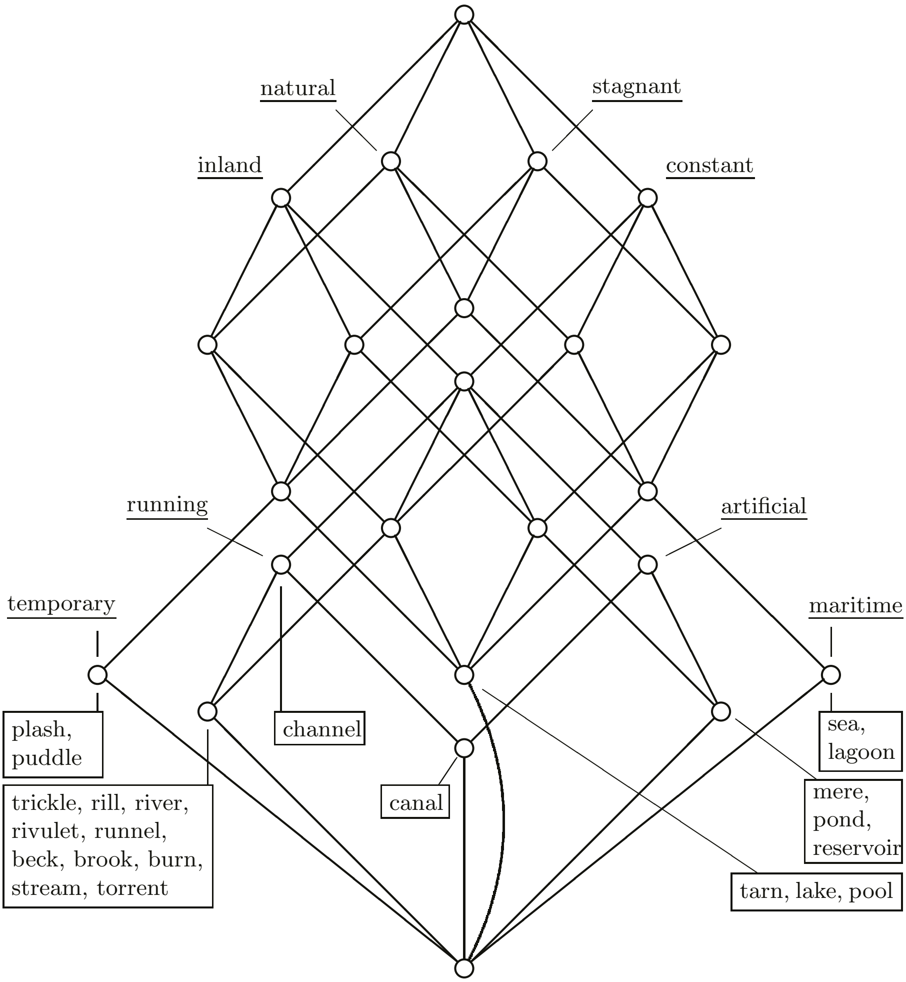

Fig. 2. Concept lattice of the formal context in Fig. 1

("Conceiving and Understanding. A Book on Concepts and Meanings") written by *Thomas Bernhard Seiler* [Se01][1]. In his book, Seiler discusses a great variety of concept theories in philosophy and psychology and concludes with his own theory which extends the concept understanding of Piaget's structure-genetic approach. In his theory, Seiler describes concepts as *cognitive structures* whose development in human mind is constructive and adaptive. Seiler elaborates his approach in twelve aspects which are briefly described in the following twelve subsections and used to review the adequacy of the mathematizations of Formal

[1] It might be desirable to integrate further concept theories in our discussion, but that would exceed the scope of this paper. The connections to those theories may be analyzed later

Concept Analysis. In each subsection, the first paragraph concisely summarizes Seiler's understanding of the corresponding aspect; then related notions and relationships from Formal Concept Analysis are discussed and partly concretized by at least one example. The connections between Seiler's concept theory and Formal Concept Analysis which come apparent in this way are far from being exhaustive. But they already show an astonishing multitude of correspondencies between both theories which may be taken as arguments for the adequacy of the discussed mathematizations.

2.1 Concepts Are Cognitive Acts and Knowledge Units

According to [Se01], concepts are *cognitive acts* and *knowledge units* potentially independent of language. Only if they are used to give meaning to linguistic expressions, they become so-called *word concepts* which are conventualized and incorporated. The meanings of words for an individuum presuppose conceptual knowledge of that individuum which turns linguistic expressions into signs for those concepts. *Personal concepts*[2] exceed conventional meaning with additional aspects and connotations. *Conventional concepts and meanings* are objectified and standardized contents, evolved in recurrently performed discourses. The problem arises how to explain under which conditions which knowledge aspects are actualized.

Formal concepts of formal contexts may mathematize personal and conventional concepts as units of extension and intension independent of specific concept names. They are representable in labelled line diagrams which stimulate individual *cognition acts* of creating personal and conventional concepts and knowledge. Computer programs for drawing labelled line diagrams (like Anaconda [Vo96]) allow to indicate represented *word concepts* by attaching concept labels to the corresponding circles in the diagrams.

Mathematizations of *conventional concepts* are given, for example, through formal contexts of lexical fields in which the conventional meaning of the corresponding words are determined by so-called "noemas" (smallest elements of meaning). The formal concepts depicted in the labelled line diagram of Fig. 2 are mathematizing conventional concepts; they are derived from the formal context in Fig. 1 which originates from a mathematization of lexical fields of bodies of waters performed in [KW87].

An example based on *personal concepts* and their interrelationships is presented in Fig. 3. Its data are taken from psychology research about the development of economic concepts by young persons [Cl90]. The example reports on the outcome of interviews about *price differences* between various articles of commerce. Reasons for those differences were classified by the five characteristics "size, beauty", "use", "rarity", "production costs", and "supply/demand". The personal understanding of price differences of the 48 test persons (16 persons of age 10-11, 15, and 18-19, respectively) is represented in the line diagram of Fig. 3 by the 14 object concepts; the formal concepts with the intents

[2] In [Se01] personal concepts are named "idiosyncratic concepts"

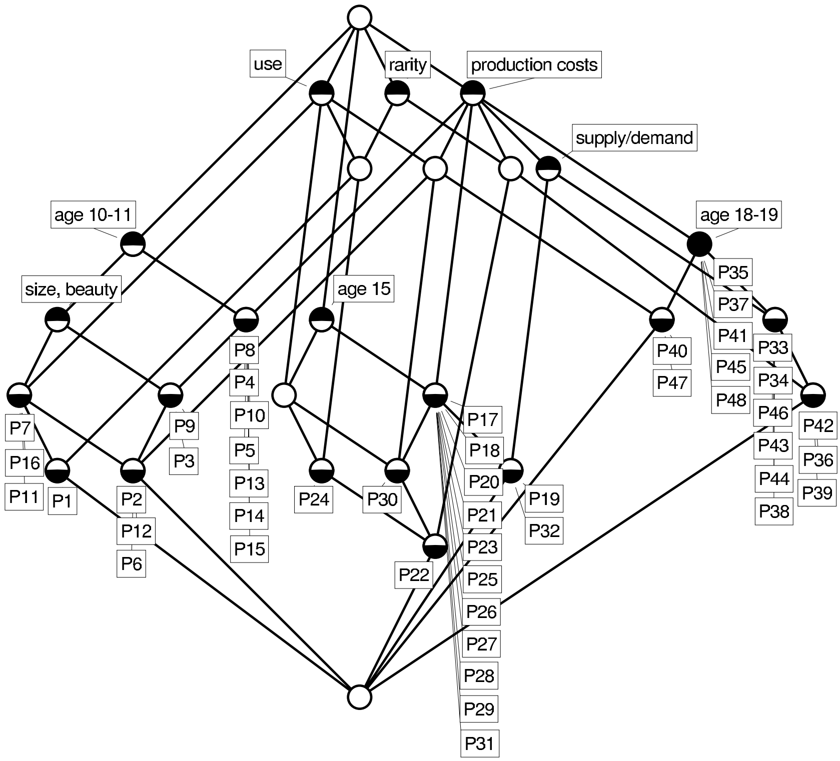

Fig. 3. Concept lattice about economic concepts of young persons

{use,rarity},{rarity, production costs}, and {supply/demand} combine two object concepts with the same characteristics from different age groups, respectively, and {use, production costs} combines even object concepts with the same characteristics from all three age groups. In particular, the development of the personal understanding over the ages becomes transparent by the labelled line diagram. Here we only mention the change from the specific characteristic "size, beauty" in the age group 10-11 to the dominance of the characteristics "production costs" and "supply/demand" in the age group 18-19. This indicates the plausible development towards the conventional meaning of the concept "money".

Concerning the mentioned problem of actualizing knowledge, labelled line diagrams as representations of concept lattices support the *actualization of knowledge aspects*. Especially, the understanding of the concepts represented by the little circles unfolds more and more when the connections of the relevant object and attribute labels with those circles are mentally established.

2.2 Concepts Are Not Categories, but Subjective Theories

According to [Se01], concepts are primarily cognitive structures and therefore elements and subsystems of our understanding and knowledge. As *naive and*

	acute	equiangular	equilateral	isosceles	oblique	obtuse	right-angled	scalene
((0,0),(1,0),(0,1))				×			×	
((0,0),(1,0),(0,2))							×	×
((0,0),(2,0),(3,1))					×	×		×
((0,0),(2,1),(4,0))				×	×	×		
((0,0),(1,2),(2,0))	×			×	×			
((0,0),(1,2),(3,0))	×				×			×
((0,0),(1,root(3)),(2,0))	×	×	×	×	×			

Fig. 4. Formal context with lexical attributes for triangles

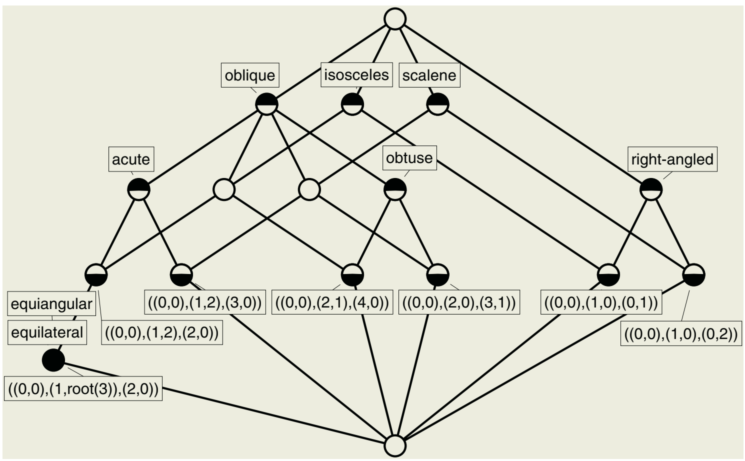

Fig. 5. Concept lattice of the formal context in Fig. 4

subjective theories, concepts contain implicit and explicit assumptions about objects and events, their conditions and causes, their characteristics, relations and functions; they are of an *abstract and idealizing nature*. They are theories which the subject creates and uses to reconstruct and to represent objects, segments, events of the surrounding world. The example in Fig. 3 indicates that young children start with creating subjective theories which slowly adapt intersubjective views and quite lately reach full conventional understanding. For conceptual subjective theories see also the example in Fig. 6 and Fig. 7.

The formal concepts of a formal context live in a *hierarchical network* of a multitude of further formal concepts. They are substructured internally by a network of subconcepts and externally in multi-relationship to further formal concepts. Thus, formal concepts are not only pairs of sets, they are part of a contextual representation of a formal theory which can be linked by inscriptions to *subjective and intersubjective theories* of human beings. As mathematical entities, formal concepts are abstract and of an idealizing nature.

In order to demonstrate a contextual mathematization of an intersubjective theory, we present a *formal theory of the lexical word concepts of triangles*. This theory is based on the formal context in Fig. 4 and conceptually unfolded in the corresponding concept lattice pictured in Fig. 5. The top element of the concept lattice represents the conventional concept of a (plane) triangle. The represented substructure of the general triangle concept is determined by the lexical attributes for triangles: *equilateral, equiangular, scalene, isosceles, oblique, acute, obtuse, right-angled*. Those attributes give rise to exactly seven object concepts for which only one generating triangle is made explicit, respectively (implicitly, there are obviously infinitely many triangles generating each of the seven object concepts). The concept lattice shows that the lexical word concepts of triangles form a *simple-implicational theory* in the sense of [Wi04] which is determined by the implications $equilateral \leftrightarrow equiangular$, $equilateral \rightarrow isosceles$, $equilateral \rightarrow acute$, $acute \rightarrow oblique$, $obtuse \rightarrow oblique$ and the incompatible subsets $\{acute, obtuse\}$, $\{acute, right - angled\}$, $\{obtuse, right - angled\}$, $\{equilateral, scalene\}$, $\{isosceles, scalene\}$, $\{oblique, right - angled\}$. Besides the seven attribute concepts, there are exactly eight consistent word concepts which can be named by combining two of the lexical attributes for triangles, for instance: *scalene obtuse triangle*. It is not surprising that the logic of the lexical word concepts of triangles is determined by implications with one-element-premise and incompatibilities; seemingly, our everyday thinking has *intersubjectively incorporated* the predominant use of logical implications with one-element premise (cf. [Wi04]).

2.3 Concepts Are Not Generally Interlinked in the Sense of Formal Logic

According to [Se01], a one-sided priority of aspects of formal logic leads to view concepts through the conventional perspective and to disregard the primarily *personal nature of concepts*. Conceptual thinking is situation and domain dependent. Personal concepts are not structured in the sense of formal logic, but they are cognitive structures which tend to amalgamate to closed and integrated systems. Although concepts are not of a formal-logic nature, they may form a *basis for logical thinking*.

Formal concepts are *mathematical entities* and not formal-logic constructs; they live in the extremely rich realm of mathematics (that allows applications as in the example of Subsection 2.6). Formal concepts are context dependent and mathematically structured in concept lattices which even tend towards more elaborated integrated systems. Formal concepts (and concept lattices) especially form the basis of *Contextual Logic* [Wi00a], a mathematization of the traditional philosophical logic based on “the three essential main functions of thinking – *concepts*, *judgments*, and *conclusions*” ([Ka88], p.6).

Formal concepts which mathematize *personal concepts* are derived from formal contexts which represent personal views. Fig. 6 yields an example of such a personal context which is an outcome of a *Repertory Grid* examination of an anorectic patient (see [SW93]). Since such a patient is understood to suffer from