Relational Concept Analysis: Semantic Structures in Dictionaries and Lexical Databases[1]

Vom Fachbereich
Gesellschafts- und Geschichtswissenschaften
der Technischen Universität Darmstadt
zur Erlangung des Grades eines
Doktors philosophiae
(Dr. phil.)
genehmigte

Dissertation

von
Dipl.-Math. Uta Priß
aus Braunschweig

Referent:	Prof. Dr. Rudolf Wille
Koreferenten:	Prof. Dr. Rudolf Hoberg
	Prof. Dr. Volker Beeh
Tag der Einreichung:	4. November 1996
Tag der Prüfung:	13. Dezember 1996

Darmstadt 1996
D17

[1]Published: Shaker Verlag, Aachen, 1998, ISBN 3-8265-4099-9

Contents

Introduction

A starting point for this dissertation was the attempt to find mathematical models for the semantic relations in lexical databases such as Roget's International Thesaurus (RIT, 1962) and WordNet. This lead to a detailed analysis of lexical and conceptual structures within linguistic data (Chapter 1), to the creation of Relational Concept Analysis as an extension of Formal Concept Analysis (Chapter 2), and to the formal modeling of semantic relations (Chapter 3). The purpose of this research is to provide a set of formal representation techniques that allow a structural approach to knowledge organisation and representation systems. Applications can be found in computerized systems that customize natural language storage and processing, in relational databases, semantic networks, conceptual knowledge systems as developed by cognitive scientists, library classification systems, thesauri, and others. Chapter 4 demonstrates how Relational Concept Analysis interlinks with some other theories in this area. It further indicates a wide range of applications of Relational Concept Analysis in addition to the lexical databases for which the theory originally was designed.

Chapter 1 presents a short introduction to Formal Concept Analysis, an overview of the linguistic terminology and linguistic or philosophical theories on the subject of 'word', 'concept', 'meaning', and 'denotation'. This leads to the definition of 'disambiguated words' which have 'particular meanings', are contained or stored in a lexical structure, and present the basic units for most of the modelings in this paper. The concepts that correspond to a disambiguated word are differentiated into denotative and connotative word concepts. They always depend on underlying contexts which are given in the denotative and connotative structures. This terminology allows the formalization of several 'linguistic contexts' and 'linguistic lattices' which can be applied to a variety of linguistic datasets for a variety of purposes. For example, the differences between lexical and conceptual structures can be made visible. This can be used to compare implicit conceptual structures of several languages or even improve semi-automatic machine translation systems. Polysemy and synonymy, which evolve from the interrelationship between lexical and conceptual structures and depend on the amount of connotative and denotative features, can also be made visible in graphical representations. The chapter terminates with formalizations of a traditional dictionary (such as Webster (1981)), a natural language thesaurus (such as RIT (1962)), and a lexical database (such as WordNet).

Relational Concept Analysis (Chapter 2) enhances the conceptual structures of Formal Concept Analysis with additional relations that are defined on the object or attribute level and generalized to the conceptual level. Preceding and related theories to Relational Concept Analysis can be found in lexical semantics/linguistics (Cruse, 1986), knowledge representation systems/cognitive science (Woods, 1990), logical quantifiers/philosophy (Westerstahl, 1989) and power relations/mathematics (Brink, 1993). An important feature of Relational Concept Analysis is that implicit quantifiers, which are in other theories often not fully recognized, are made explicit. This together with the lattice formalization of Formal Concept Analysis seems to establish an advantage for Relational Concept Analysis as a tool in knowledge organization and data structuring. The different aspects of Relational Concept Analysis are discussed in detail in Chapter 2: the development of bases for concept relations which allow optimal implementations of relations in lexical databases; the inheritance structures and other formal properties (such as transitivity) of concept relations; auto- and polyrelations; graphical representations; and applications to the linguistic contexts and lattices from Chapter 1 are the major subjects that are discussed.

Chapter 3 concentrates on linguistic aspects of semantic relations. A broad classification of semantic relations is developed based on formal characteristics. The semantic relations, synonymy, hyponymy, hypernymy, cohyponymy, disjointness, meronymy, contrast (antonymy), sequence, cause, backward presupposition, and entailment are formally defined. The major example for this chapter is the meronymy relation which is distinguished from Lesniewski's mereology (compare Luschei (1962)). The modeling with Relational Concept Analysis facilitates the improvement of a theory of transitivity for the meronymy relation that was developed by Winston et al. (1987). Furthermore, a classification of meronymy based on quantifications is developed and compared to the classifications based on content by other authors (Chaffin & Herrmann (1988), Winston et al. (1987), Iris et al. (1988), and Miller et al. (1990)). The version of WordNet as a relational database which was developed in connection with this dissertation leads to a further classification based on content which is also compared to the existing classifications. Relational Concept Analysis can be utilized to discover irregularities in the implementations of semantic relations in lexical databases. This is demonstrated using examples from WordNet. The contrast relations (such as antonymy) are shown to be different from meronymy in that they use only one type of quantification. Finally, verbs and prepositions are interpreted as semantic relations and it is demonstrated that several types of verb entailment (Fellbaum, 1990) are based on quantifications of the meronymy and sequence relations.

Chapter 4 provides applications and extensions of Relational Concept Analysis. Some of the features that Relational Concept Analysis shares with other knowledge representation systems, that are improvements compared to other systems, or that are not as advanced as in other systems are discussed. It is argued that the lexical and conceptual structures of a natural language both form separate systems that interrelate. A graphical representation technique for semantic and lexical relations in a denotative lattice is developed. Formal composition rules for lexical and conceptual items are defined and related to each other. In the second section of Chapter 4 Relational Concept Analysis is revisited within the framework of Relational Algebra (Pratt, 1992). It is proved in this chapter that the quantified relations from Chapter 2 can be computed as binary matrix multiplications (that is, as a relational product). Thus the connection to Relational Algebra provides a simple algorithm for a computerized implementation of Relational Concept Analysis. Unary relations are introduced and used to assign additional attributes to a concept lattice. Such additional attributes can be prototypical or default attributes that are not shared by all objects in the extent of a concept. Finally, Relational Concept Analysis is compared to other systems of knowledge representation or data structuring, such as many-valued contexts, the Entity-Relationship model, terminological logic, and semantic networks.

I wish to thank all those people who helped me during the last three years with comments and discussions about my thesis. Without naming them individually I thank the members of the AG Begriffsanalyse for the constructive discussions during the Mittagseminar and for the supportive atmosphere in the department. I especially thank Prof. Rudolf Wille for his advice and support, for encouraging me to work with Roget's International Thesaurus and for enabling me to spend some time at universities in the United States. I thank Prof. Rudolf Hoberg for accepting my interdisciplinary research. I am further thankful to the Professors Walter and Sally Sedelow (University of Arkansas at Little Rock) and George Miller (Princeton University) for enabling me to study and work in their departments. Discussions with Christiane Fellbaum helped in clarifying some aspects of WordNet and hints from Fritz Lehmann helped in finding some relevant references. And finally I thank John Old for his valuable comments and for correcting my English.

1 Formal Description of Linguistic Terms

1.1 Formal Concept Analysis

The increasing need for computerized knowledge storage and processing requires formal representation techniques. In linguistic applications computerized lexical databases and natural language processing software tools enable a treatment of natural language for the purpose of software development and research goals which has been unimaginable a hundred years ago. The basic condition for all computerized data is formalization. In this paper a mathematical theory called Formal Concept Analysis, which has been especially designed (see Wille, (1992) and (1994)) for the preparation of data in a way that facilitates a critical examination of the data according to the aims of pragmatic philosophy, is introduced and applied to linguistic data. In formalizations, symbolic configurations are assigned meanings that enable substantive interpretations. Nevertheless the distance between formalization and interpretation often causes a loss in content. The basic notions of Formal Concept Analysis are a formalization of 'concept' and of conceptual hierarchies. Based on the mathematical formalization graphical representations, such as concept lattices (see Figure 1.1), are developed which can serve as a tool for scientific communication. To differentiate between formalizations and content, in this paper three sub-languages have to be distinguished: the language of Formal Concept Analysis, the sub-language of scientific linguistic terminology, and the sub-language of everyday communication. Therefore attributes, such as 'formal', 'linguistic', or 'natural language' are used throughout this paper in the cases where ambiguities could occur in order to clarify in which sub-language the terms are used. 'Natural language' is used to refer to the usage of a term in the natural language, whereas 'linguistic' refers to the usage of a term in the linguistic science. In Formal Concept Analysis verbal forms of objects, attributes, or concepts (see below) are elements of sets which get a specific meaning when they are interpreted within the terminology of another sub-language. In the later sections of this chapter, several classes of formal contexts are defined whose formal objects, attributes, or concepts have specific interpretations. For example, the formal objects in one class of formal contexts are interpreted as denotata in a linguistic terminology. Depending on the class of a formal context the structures of the formal conceptual hierarchy also have certain interpretations. For example, for some contexts the formal conceptual hierarchy can be interpreted as linguistic hypernymy. It should be obvious by now, that it is essential to be conscious about which sub-language is used and which interpretation is valid in which context.

Formal Concept Analysis (Ganter & Wille, 1996) starts with the definition of a *formal context* $\mathcal{K}$ as a triple (G, M, I) consisting of two sets G and M and a relation I between G and M (i.e. $I \subseteq G \times M$). The elements of G and M are called *formal objects* (Gegenstände) and *formal attributes* (Merkmale), respectively. The relationship is written as gIm or $(g, m) \in I$ and is read as 'the formal object g has the formal attribute m'. A formal context can be represented by a cross table which has a row for each formal object g, a column for each formal attribute m and a cross in the row of g and the column of m if gIm. The upper half of Figure 1.1 shows an example of a formal context. It has 'person', 'adult', and so on as formal objects, and 'juvenile', 'grown-up', 'female', and 'male' as formal attributes. All formal contexts in this paper are finite. That means they have finite sets of formal objects and attributes because the main application of this paper is *linguistic contexts*. Linguistic contexts are defined in this paper as formal contexts whose formal objects and attributes are elements of linguistic research, such as words, denotata of words, word forms, semantic components of words, and others. Since the basis for linguistic contexts in this paper are dictionaries or lexical database, it is assumed that every language at any time has only a finite number of words, letters, word forms, and so on, therefore all linguistic contexts are finite.

	juvenile	grown-up	female	male
person				
adult		×		
female person			×	
male person				×
child	×			
woman		×	×	
man		×		×
girl	×		×	
boy	×			×

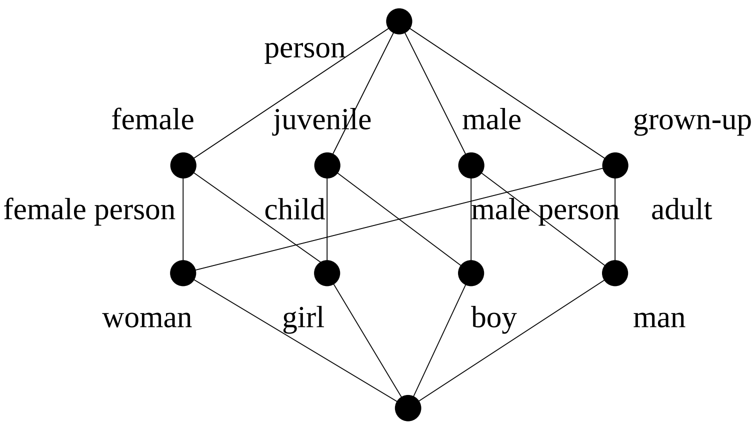

Figure 1.1: A formal context and the line diagram of its concept lattice

In a context (G, M, I) the set of all common formal attributes of a set $A \subseteq G$ of formal objects is denoted by $\iota A := \{m \in M \mid gIm$ for all $g \in A\}$ and, analogously, the set of all common formal objects of a set $B \subseteq M$ of formal attributes is $\varepsilon B := \{g \in G \mid gIm$ for all $m \in B\}$. For example, in the formal context in Figure 1.1, $\iota\{\text{man}\} = \{\text{grown-up, male}\}$ and $\varepsilon\{\text{grown-up}\} = \{\text{adult, woman, man}\}$ hold. The sets $\iota\{\text{man}\}$ and $\varepsilon\{\text{grown-up}\}$ obviously depend on the formal context because only those formal attributes or objects can be retrieved that actually exist in the formal context. A pair (A, B) is said to be a *formal concept* of the formal context (G, M, I) if $A \subseteq G, B \subseteq M, A = \varepsilon B$, and $B = \iota A$. In this paper formal concepts are denoted by c, c_1, c_i and so on. For a formal concept $c := (A, B)$, A is called the *extent* (denoted by $Ext(c)$) and B is called the *intent* (denoted by $Int(c)$) of the formal concept. In the example of Figure 1.1, ({adult, woman, man}, {grown-up}) is a formal concept, because ι{adult, woman, man} = {grown-up} and ε{grown-up} = {adult, woman, man}. The extent of this formal concept is {adult, woman, man}, the intent is {grown-up}. Within a formal context a formal concept can therefore already be uniquely described by either its extent or its intent. This is not necessarily true for natural language concepts, for example, it seems to be impossible to list all attributes of 'adult' which speakers of the English language might intersubjectively assign to the natural language concept 'adult'. Similarly it seems to be impossible to list the class of all objects which might be described by the attribute 'big' in some natural language context.

The set of all formal concepts of (G, M, I) is denoted by $\mathcal{B}(G, M, I)$. The most important structure on $\mathcal{B}(G, M, I)$ is given by the formal subconcept-superconcept relation that is defined as follows: the formal concept c_1 is a *formal subconcept* of the formal concept c_2 (denoted by $c_1 \leq c_2$) if $Ext(c_1) \subseteq Ext(c_2)$, which is equivalent to $Int(c_2) \subseteq Int(c_1)$; c_2 is then a *formal superconcept* of c_1 (denoted by $c_1 \geq c_2$). For example, ({adult, woman, man}, {grown-up}) as a formal superconcept of ({woman}, {grown-up, female}) has more formal objects but fewer formal attributes than ({woman}, {grown-up, female}). It follows from this definition that each formal concept is a formal subconcept of itself in contrast to the natural language use of 'subconcept' which precludes a concept from being a subconcept of itself. The relation '$\leq$' is a mathematical order relation called *formal conceptual ordering* on $\mathcal{B}(G, M, I)$ with which the set of all formal concepts forms a mathematical lattice denoted by $\underline{\mathcal{B}}(G, M, I)$. This means that, for all pairs of concepts, the greatest common subconcept and the least common superconcept exist.

Graphically, mathematical lattices can be depicted as line diagrams which represent a formal concept by a small circle. According to the formal conceptual ordering, formal subconcepts are drawn under their formal superconcepts and connected by a line with their immediate formal superconcepts. For each formal object g the smallest formal concept to whose extent g belongs is denoted by γg. And for each formal attribute m the largest formal concept to whose intent m belongs is denoted by μm. The concepts γg and μm are called the *object concept* of g and the *attribute concept* of m, respectively. In the line diagram it is not necessary to write the full extent and intent for each concept, instead the name (verbal form) of each formal object g is written slightly below the circle of γg and the name of each formal attribute m is written slightly above the circle of μm. The lower half of Figure 1.1 shows the line diagram of the concept lattice of the formal context in Figure 1.1. To read the line diagram, the extent of a formal concept consists of all formal objects which are retrieved by starting with the formal concept and then collecting all formal objects that are written at formal subconcepts of that formal concept. Analogously, the intent is retrieved by collecting all formal attributes that are written at formal superconcepts of the formal concept. More details on Formal Concept Analysis can be found in Ganter & Wille (1996). In contrast to many classification systems that are organized as trees, lattices in Formal Concept Analysis allow for a formal concept to have several immediate formal superconcepts. In the lattice in Figure 1.1, for example, a woman is an adult and a female person at the same time. We consider this feature to be a major advantage over tree based systems.

For further reference, the mappings which are defined for formal contexts and lattices are summarized in Definition 1.1. The mappings ι^+ and ε^+ are added because they are needed in some applications, for example, in Section 1.11. If several formal contexts are involved in one application the mappings $\iota, \varepsilon, \gamma, \mu, \iota^+$, ε^+ can be indexed, such as ι_I, ε_I, and so on. If necessary further indexes can be added.

Definition 1.1:
For a formal context $\mathcal{K} := (G, M, I)$ with set $\mathcal{B}(\mathcal{K})$ of formal concepts the following mappings are defined[2]:

- $\iota : \wp(G) \to \wp(M)$ with $\iota G_1 := \{m \in M \mid \forall_{g \in G_1} : gIm\}$ for $G_1 \subseteq G$
- $\varepsilon : \wp(M) \to \wp(G)$ with $\varepsilon M_1 := \{g \in G \mid \forall_{m \in M_1} : gIm\}$ for $M_1 \subseteq M$
- $\gamma : G \to \mathcal{B}(\mathcal{K})$ with $\gamma g := (\varepsilon\iota\{g\}, \iota\{g\})$
- $\mu : M \to \mathcal{B}(\mathcal{K})$ with $\mu m := (\varepsilon\{m\}, \iota\varepsilon\{m\})$
- $\iota^+ : \wp(G) \to \wp(M)$ with $\iota^+ G_1 := \{m \in M \mid \exists_{g \in G_1} : gIm\}$ for $G_1 \subseteq G$
- $\varepsilon^+ : \wp(M) \to \wp(G)$ with $\varepsilon^+ M_1 := \{g \in G \mid \exists_{m \in M_1} : gIm\}$ for $M_1 \subseteq M$
- $Ext : \mathcal{B}(\mathcal{K}) \to \wp(G)$ with $Ext(c) := \{g \in G \mid \gamma g \leq c\}$
- $Int : \mathcal{B}(\mathcal{K}) \to \wp(M)$ with $Int(c) := \{m \in M \mid \mu m \geq c\}$

1.2 The linguistic terminology

Since linguists and philosophers often use terms, such as ‘meaning’, ‘concept’, ‘word’, ‘denotation’ and others in a slightly different way from their colleagues, there is a need to precisely define these terms for the use of this paper. This section presents the basic terms of our terminology and introduces an example which is used throughout this chapter. The next section compares our terminology to the terminology of other linguists or philosophers. The following sections of this chapter derive a formalization of our terminology to facilitate the application of Formal Concept Analysis to linguistic data. The first question is certainly ”what is a word?” Hofstadter (1996)

[2] $\wp$ denotes the power set of a set, i.e. $\wp(G) := \{G_1 \mid G_1 \subseteq G\}$.

convincingly demonstrates that this question is not easily answered. A survey of approaches to answer this question can be found in Forsgren (1977). In this paper the following is defined:

Definition 1.2:
Anything which can appear as an entry of a dictionary, such as single words ('person'), composite words, or phrases ('female person', 'President of the United States') is called 'word'. A word consists of a word form and a word meaning. The word meaning at a certain time is a cultural unit which contains all intersubjectively associated aspects of the meaning of the word[3]*, such as denotation, connotation, part of speech, syntactic rules, conditions of usage, and so on. These aspects are called 'features' of the word meaning and cannot be defined in general but are subject to underlying linguistic theories. Formally, words are denoted by* v, v_1, v_i, *and so on, word forms by* $form(v)$, $form(v_1)$, $form(v_i)$, *and so on, and word meanings by* $mng(v)$, $mng(v_1)$, $mng(v_i)$, *and so on.*

The following remarks try to clarify this definition: First, the basic application of the theory of this paper are dictionaries, thesauri, and lexical databases which are assumed to represent the words synchronically for one certain time period. Pragmatic aspects are only contained as long as they are part of the word meaning. It would be interesting to extend our theory to include sentences and their meanings, but this could not be achieved in this paper. The dictionary that is used as an example in this paper is Webster's Third New International Dictionary (W3), the thesaurus is Roget's International Thesaurus (RIT), and the lexical database is WordNet.

Second, what we call 'word' is often called 'lexeme' (Lyons, 1977), 'lexical unit' (Cruse, 1986) or 'lexical item'. But since linguistic terms are often used differently among the authors (for example, the French structuralists use 'lexeme' for the morpheme which contains the lexical meaning of a word (Wiegand & Wolski, 1980)) and since compound forms, such as 'meaning of a lexical unit' are not as easy to construct as 'word meaning', we prefer 'word'.

Third, the word form in this paper is always represented in its 'citation form' (Lyons, 1977). It could be interpreted as a paradigm of its tokens (Lutzeier, 1982), but these aspects shall be ignored in this paper.

Fourth, word meanings are considered in this paper to be atomic cultural units. They resemble Eco's (1991) definition of a significat as a cultural unit which is a semantic unit within a system and is denoted by the significant (word). More details on Eco's terminology can be found in the next section. In our definition, the different aspects

[3]This is not a circular definition because 'word meaning' is used as a linguistic term which is defined by this statement and which is used according to this definition in the rest of this paper whereas 'meaning of the word' refers to a not defined notion which is commonly used in the natural language.

of the meaning, such as connotations, denotations, and so on are features of the word meaning. Which features a certain meaning has is not decided in general independently of a certain philosophy or context. For applications, formal contexts (so-called *connotative contexts*) are constructed which have word meanings as formal objects and features of word meanings, such as 'denotes a living being', 'has a derogatory connotation', or 'is an adjective' as formal attributes. These formal contexts can then be used to compare different word meanings according to the features which have been selected for that specific formal context. It seems to be difficult (or impossible) to derive any general statements about the features of word meanings, independent of specific formal contexts, which would be agreed upon by a majority of linguists or philosophers, therefore we choose this approach to represent the word meaning as an atomic constant which has features according to certain contexts.

The following example of 'morning star' and 'evening star', which is selected because of Frege's (1892) famous example, is used throughout this chapter and is based on W3:

> Evening star 1a : a bright planet (as Venus) seen in the western sky after sunset; b: any planet that rises before midnight; c: any of the five planets that may be seen with the naked eye at sunset; 2: a small bulbous plant of Texas (Cooperia drummondii) with grass-like leaves and star-shaped white flowers;
>
> Morning star 1a: a bright planet (as Venus) seen in the eastern sky before sunrise; b: any of the five planets that may be seen with the naked eye if in the sky at sunrise (Venus, Jupiter, Mars, Mercury, and Saturn may be morning stars); c: a planet that sets after midnight; 2: a weapon consisting of a heavy ball set with spikes and either attached to a staff or suspended from one by a chain - called also holy water sprinkler; 3: an annual California herb with showy yellow flowers;

A word meaning can often be divided into more specific word meanings, for example, the word meaning 1 ('celestial body') of the word 'morning star' can be divided into the more specific meanings 'a bright planet (as Venus)', 'any of the five planets ...' and 'a planet that sets after midnight'. The other two meanings 'medieval weapon' and 'California herb' which cannot be further divided are not homographs (which are considered to be different words), but metaphorical uses of 'morning star'. The most specific meanings given in a dictionary indicate that the word is disambiguated and therefore has a particular meaning. Similarly, each location of a word in a thesaurus hierarchy usually denotes one particular meaning of the word and therefore disambiguates it. The distinction between words and disambiguated words is not the same as the distinction between 'virtual' and 'actual' or 'type' and 'token' (for example in Lincke et al. (1994)). Words are virtual and types. Disambiguated words are, so to say, virtual versions of tokens. Consider an example: the type ‘dog’ has six meanings in WordNet (they are called ‘senses’ in WordNet) of which the disambiguated word ‘dog 1’ means the member of the genus Canis. A token of ‘dog 1’ occurs, for example, in the sentence ‘Fido is the dog that ate the sausage’ which can under some circumstances denote a real world dog ‘Fido’. On the other hand in ‘Fido is a dog’, the token of ‘dog’ denotes not only Fido but a concept or a class of dogs. Therefore the denotata of a token (Lyons (1977) calls this ‘referent’) are normally obvious from the natural language context, such as ‘Fido’ or the class of dogs. It is usually agreed upon (Lincke et al., 1994) that the type ‘dog’ has no denotata because as long as it is not disambiguated the denotata could be dogs, villains, or devices. Disambiguated words which are, so to say, in-between types and tokens have denotata (Lyons (1977) calls them ‘denotatum’). The following definition specifies disambiguated words and their denotata based on the general approach of this paper to consider dictionaries (or lexical databases or thesauri) as a basis.

Definition 1.3:
Particular meanings of words are indicated by sense numbers in a dictionary or lexical database or by the location in the hierarchy of a thesaurus. A disambiguated word consists of the word form of the word concatenated with a sense number and a particular meaning of the word which native speakers of a language recognize after reading its definition in a dictionary or after seeing its synonyms in a thesaurus. A denotatum of a disambiguated word is an instance or item which the disambiguated word can denote according to its particular meaning and to the intersubjectively agreed upon rules of the language. Formally, disambiguated words are denoted by w, w_1, w_i, and so on, forms of disambiguated words by $form(w)$, $form(w_1)$, $form(w_i)$, and so on, and particular meanings by $mng(w)$, $mng(w_1)$, $mng(w_i)$, and so on.

Several problems arise from such a definition. First, similarly to what has been said about word meanings, particular meanings are considered in this paper to be atomic units. Within formal contexts the particular meanings are assigned certain features that therefore depend on the formal context. It is not defined in general in this paper how the particular meanings constitute the word meaning because that also depends on underlying linguistic theories and can be modeled in a formal context.

Second, it is assumed that the reader of the dictionary or thesaurus lives in the same time period and social environment which is represented in the dictionary; that he or she fully understands what is written in the dictionary and the dictionary does not contradict the general intersubjective knowledge of the speakers of that language.

Third, the questions of how meanings are represented in the mental lexicon of native speakers of a language, and how the meanings recognized by individual speakers as part of their idiolects differ from the meanings as part of the intersubjective language system, are not discussed in this paper. It is simply assumed that different native speakers of a language recognize approximately the same particular meaning after reading a dictionary definition so that it is possible to treat the particular meaning of a disambiguated word as a constant of a certain natural language at a specific time.

Fourth, the ontological question in what way denotata ‘exist’ is not important for this paper because we attempt to separate the formalization of linguistic terms from philosophical viewpoints. The only application where denotata are used in this paper are so-called *denotative contexts* (see below) which consist of specific sets of denotata and specific sets of attributes of denotata. The person who designs a denotative context has to decide whether or not items which are questionable, such as unicorns, are included in the set of denotata for that specific denotative context. A denotative context of a whole language would probably be too large to ever be constructed. Instead, denotative contexts of partial sets of the vocabulary of a natural language should be considered. Our suggestion for denotative contexts derived from a dictionary is to use prototypical instances or items for the denotata of words which do not denote individuals, such as proper names. The discussion about the existence of denotata inside of possible worlds or the external world does not have to be taken into consideration because even if some denotata of a disambiguated word do not have a physical existence they still have attributes which a native speaker of a language knows if he or she knows the meaning of the disambiguated word. And for the development of formal contexts only a set of denotata, a set of attributes of denotata and their relation are needed. For example, the denotative attributes of unicorns are easy to define inside the intersubjective knowledge of English speakers. For English speakers, unicorns look like white horses and have a long horn growing from their foreheads (according to WordNet). They further have the denotative attribute ‘exists only in imagination’.

Fifth, denotata are always thought to be items or instances. For example, denotata of ‘cow 1’ (according to WordNet) can be all cows to which someone can refer to as ‘cow’. This is in contrast to Lyons (1977) who defines the denotatum (singular) of ‘cow’ as the class of all cows. Denotata of a verb are usually instances, for example, denotata of ‘eat 1’ (WordNet) are instances where it can be said that ‘someone eats something’. Denotata of ‘courage’ are instances which can be called ‘courage’. As said before for a denotative context some specific instances and items, such as the ‘courage of David fighting Goliath’ or ‘the cow Bluebell’ or some prototypical instances or items have to be selected.

Sixth, anaphora is ignored in this paper because it is a phenomenon of the word tokens and not of the disambiguated words. Saying ‘dog’ instead of ‘Fido’ means using a hypernym of ‘Fido’ instead of ‘Fido’. ‘Dog’ is a synonym of ‘Fido’ in this sentence, but ‘dog’ is not defined as a synonym of ‘Fido’ in a thesaurus and ‘Fido’ is not a particular meaning of ‘dog’ in a dictionary. Pronouns and prepositions can denote a variety of denotata, for example, every human being who can speak English can say ‘I’ for her- or himself. This may be a reason why they are not included in WordNet. The denotata of some other classes of words are difficult to describe, for example, what are denotata of gradable adjectives and adverbs? These shall be ignored in this paper.

1.3 Linguistic and philosophical theories on some aspects of meaning

By defining word meaning and particular meaning as atomic constants of natural languages according to a dictionary or lexical database, we tried not to include the complete discussion of meaning among linguists and philosophers in this paper. This section gives a short summary on the treatment of some aspects of meaning – which are needed for the modeling in this paper, such as denotation – among different authors and compares their terminology to our terminology. A more detailed survey on the treatment of reference and meaning among philosophers including Wittgenstein, Frege, Austin, Searle, Grice, and Quine can be found in the philosophical part of Steinberg & Jakobovits (1971). Two uses of 'denotation' in the literature have to be distinguished. First, 'denotation' is used in contrast to 'connotation' in the meaning of 'what the word denotes' versus 'what is intersubjectively associated with the word or the word meaning'. Second, 'denotatum' is used in contrast to 'designatum' or 'significatum' (Morris, 1946) to distinguish between what the word denotes and the characteristics of what the word denotes. This is also called 'extension' and 'intension', 'denotation' and 'connotation' (J. S. Mills according to Lyons (1977)), or 'Bedeutung' ('meaning') and 'Sinn' ('sense') (Frege, 1892). For example, J. S. Mills (1843) is quoted by Lyons (1977) with 'The word white denotes all white things, ..., connotes the attribute whiteness'. Frege (1892) invented his (now famous) example of 'morning star' and 'evening star' which have the same Bedeutung but different Sinn. Furthermore, the intension is sometimes called 'reference' (Ullmann, 1957) as the 'information which the speaker transmits to the hearer'. Some linguists or philosophers (e.g. Saussure (1916) and Eco (1991)) completely exclude the 'external world' because for them language is a system which cannot denote anything outside of itself. For other linguists or philosophers (e.g. Frege) denotation refers to something outside of the language system.

Lyons (1977) distinguishes between the denotation of a token and of a disambiguated word: Reference is the relation between an expression (a series of word tokens uttered by a speaker in a context or written in a text) and a referent that is established by a speaker in a specific context. On the other hand he defines denotation as the relation (independent of any context) between a lexeme (in our terminology: a disambiguated word) and a denotatum which is a class of objects, properties, and so on. For example, the denotatum of 'dog' is the class of 'dogs' and the denotatum of 'canine' is the property of being a dog whereas in 'Fido is the dog that ate the sausage', 'Fido' is the referent of 'dog'. In our terminology Lyons' referents are denotata. His denotatum which he defines as a class of denotata can be modeled according to our theory as the set of all denotata of a disambiguated word in a certain formal context.

Talking about classes of denotata without the restriction to a context (which some linguists do) seems to be not very precise. A different approach is taken by Putnam (1975) and further developed by Lutzeier (1981, see below) who define prototypes and stereotypes (which are lists of prototypical attributes) instead of extension and intension. That means that they exclude reference to items outside of the language system because prototypes and stereotypes are meant to be mental schemas. Furthermore they avoid talking about classes because each disambiguated word corresponds to one stereotype and has only a few prototypes to which it refers.

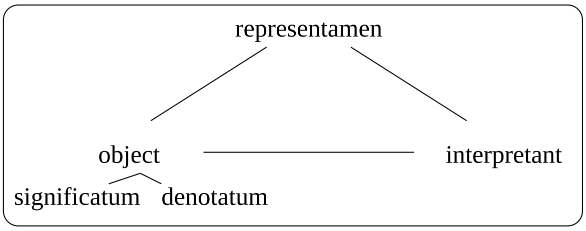

Figure 1.2.a: Peirce's and Morris' conceptual triangle

The relation between a word form and a word meaning has been characterized by Saussure who uses 'signifié' or 'concept' for the meaning of a word and 'signifiant' or 'image acoustique' for the word form. For him words are part of the language system and therefore word meanings are constituted by the opposition of a word to other words inside the system. Peirce (according to Nöth (1987)) invented his 'semiotic triangle' (see Figure 1.2.a) consisting of a representamen which resembles Saussure's 'signifiant', a reference object, and an interpretant. The reference objects can be classified into immediate objects, which depend on the semiotic process and are therefore not outside of the sign system, and dynamic objects, which are independent of the semiotic process and therefore reference objects outside of the sign system. The interpretant of a sign is the effect of the sign in the mind of an interpreter and can be a representamen itself which leads to a theoretically indefinite chain of signs (see Eco's interpretation of Peirce's interpretant below). Peirce originated the distinction between 'type' and 'token' as part of his classification of signs. Morris (1946) adapts Peirce's triangle under a behaviouristic viewpoint. Peirce's 'object' is split into 'significatum' (also called 'designatum') and 'denotatum' similarly to 'intension' and 'extension'. Nevertheless Morris defines a designatum as a class of denotata (compare Lyons' denotatum) and as the characteristics of the denotata so that his designatum actually includes intension and extension. An interpretant is according to his theory an 'interpreter with a disposition to respond' which shows his behaviouristic interpretation of Peirce's philosophical term. From his sign triangle he develops the distinction between the three disciplines: syntax, semantics, and pragmatics. His classification of signs includes 'unambiguous signs' (disambiguated words in our terminology). Ogden & Richards (1923) changed Peirce's triangle to referent, reference, and symbol which has later often been changed to significat, meaning, and significant. While 'symbol' and 'significant' mean approximately the same as Peirce's 'representamen', their 'referent' and 'reference' resemble Morris' 'denotatum' and 'designatum', which means that they split Peirce's 'object' but omit his 'interpretant'.

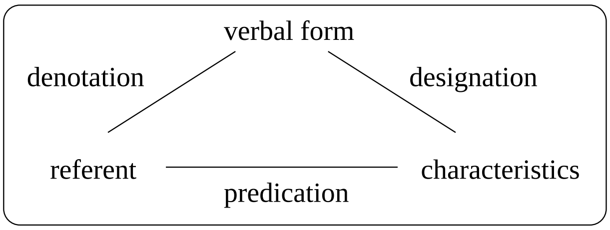

Figure 1.2.b: Dahlberg's conceptual triangle

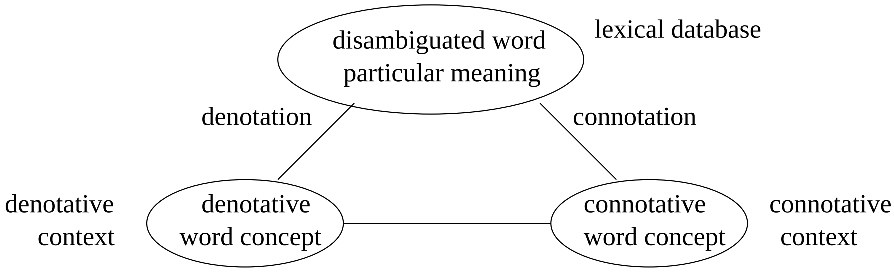

Figure 1.2.c: The word triangle developed in this paper

Dahlberg's (1994) theory concentrates on concepts and ignores problems and features of lexicalizations by ideally considering formal codes as verbal forms of concepts, instead of using natural language words. Her triangle which she uses to define a concept consists of a referent, characteristics, and a verbal form (see Figure 1.2.b). Her 'referent' resembles a denotatum in Lyons' terminology, that means that for example, the concept 'cow' denotes one referent and not a class or set. Her 'characteristics' have features of being a concept themselves. In our theory (see Figure 1.2.c) the word triangle consists of disambiguated words or particular meanings (which are in one-to-one correspondence to each other) as parts of a lexical database and connotative and denotative concepts[4] as parts of connotative and denotative contexts. This approach emphasizes that linguistic analyses are based on underlying theories or intentions therefore all statements about the features of disambiguated words or particular meanings are only achieved within formal contexts. Furthermore statements about disambiguated words can only be produced within the framework of a lexical database. We think that the confusion about whether denotata are outside or inside of the language system can be solved by specifying the attributes of denotative contexts in a corresponding manner. Instead of two terms for extension and intension we distinguish four terms, the extent and intent of denotative concepts and the extent and intent of connotative concepts. For example, connotative concepts replace Saussure's (1916) 'concepts'. Mills' 'denotation' and 'connotation' can be modeled as extent

[4]Connotative and denotative concepts, contexts, and structures are defined in more detail in the next sections.

and intent of a denotative concept. Frege's 'Bedeutung' corresponds to the extent of a denotative concept, whereas his 'Sinn' corresponds to the intent of a connotative concept. The extents and intents of denotative concepts replace Morris' 'denotatum' and 'significatum'. Connotative concepts replace Dahlberg's 'characteristics' which can be concepts themselves according to her theory.

In the rest of this section, some aspects of the terminology of Eco (1991) and Lutzeier (1981) that are relevant for our modeling are explained and compared to our terminology. Lutzeier defines the meaning of a word for a speaker of a language as a structure of stereotypes which the speaker associates with the word. Having a certain intention the word evokes a certain stereotype in the mental lexicon of the speaker (this corresponds to a particular meaning in our terminology). In a certain context and having that intention the speaker refers to a normal or prototypical member of the word (the denotatum in our terminology) called referent. The meaning of a word (independent of a speaker) is defined as a structure of stereotypes which is intersubjectively accepted among the speakers of the language. He adds that these stereotypes are usually formed by a subgroup of speakers who are experts in the relevant subject area of the word. If the meaning is considered independently of a speaker he calls the set of referents (normal or prototypical members of the word) extension. Although we in general agree with Lutzeier's modeling, we think that 'particular meaning' is a better term for the purpose of this paper than 'stereotype' because 'stereotype' has some negative connotations of being the prejudiced opinion about something and it is not clear if 'stereotype' includes all aspects of the meaning of a word or if it is mainly restricted to the prototypical attributes of the denotata.

Eco (1991) considers referents (objects of an external world) as not being part of semiotics and irrelevant for the meaning of a significant (which corresponds to a disambiguated word in our terminology). A significant denotes a significat (particular meaning in our terminology) which is a cultural, semantic unit within a system. Each significat corresponds to exactly one significant and vice versa. The more difficult term is Peirce's 'interpretant' which Eco defines as the as a cultural unit interpreted meaning of a significant which is represented by another significant to show its independence as a cultural unit from the first significant. In our terminology, a particular meaning does always depend on the disambiguated form. For example, the reason that 'morning star' and 'evening star' have different meanings depends on the fact that there are already words 'morning' and 'evening' in the English language which have meanings themselves and that there is a relation between 'morning star' and 'morning' because the word 'morning' is a part of the compound word 'morning star'. The meaning of 'morning star' depends therefore (for this example) on the relation of 'morning star' to another disambiguated word of the English language. However the meaning of 'morning star' cannot be represented as a function of 'morning' and 'star'. Eco states that there is something (the interpretant) which is as a cultural unit independent from a specific significant and therefore facilitates a definition of synonymy.

In our theory this resembles the definition of connotative concepts. While particular meanings are in one-to-one correspondence to words and therefore can never be synonymous to each other, connotative concepts depend on the relation between particular meanings and attributes. Hence two disambiguated words can have the same connotative concept. In Eco's theory, denotation which is not defined for word tokens, but for lexemes (lexical entries, disambiguated words in our terminology), is the semantic valence or position of the lexeme in a semantic field. This definition resembles Saussure's definition of 'meaning' or Lyons (1977) definition of the 'sense of a word' as the place of the word in a system. In our theory it is the place of a particular meaning in a denotative or connotative lattice. Eco defines connotation as the intersubjectively agreed upon union of all cultural units which can be related to the intensional definition of the significant. This corresponds in our terminology to the connotative attributes which a particular meaning can have within a connotative context.

1.4 Denotative contexts and structures

In the following, denotative knowledge and lexical knowledge, which is for this paper contained in a lexical database, dictionary, or thesaurus, are distinguished. This distinction is essential, for example, for the modeling of semantic relations in chapter 2. The denotative knowledge is the common sense or the scientific knowledge of a special social group that is implicitly contained in a reference source, such as a dictionary, encyclopedia, or lexical database. For the purpose of this paper it is formally represented as a relational structure called *denotative structure* (see below) which consists of a formal context called denotative context and some additional sets and relations. First, the denotative context and lattice are defined:

Definition 1.4:
A *denotative context* $\mathcal{K}_D := (D, A_D, I_D)$ is defined as a formal context whose formal objects are denotata and whose formal attributes are attributes of the denotata. The set of denotata is denoted by D, the set of attributes of the denotata by A_D, and a relation that assigns attributes to denotata by I_D. The concept lattice $\underline{\mathcal{B}}(\mathcal{K}_D)$ of a denotative context $\mathcal{K}_D$ is called *denotative lattice*.

Denotative knowledge does not consist of the words and their relations, but it contains all the information about a subject area which is intersubjectively known by a certain social group and does not depend on the verbal representation. Denotative contexts are not thought to be independent from languages (on the contrary, English and Japanese denotative contexts, for example, can be very different), but to be abstracted from connotations or associations which entirely depend on the specific verbal representation or the social context and not on the objects themselves. Denotative and lexical structures are linked by the *denotative concepts*, which are always part of a denotative lattice. If they are also part of a lexical structure, which means they are lexicalized, they are called *denotative word concepts*. (Obviously, some denotative concepts, such as lexical gaps (see Section 1.9), are non-lexicalized.) It is probably not possible to directly construct denotative contexts for a natural language, instead the denotative word concepts have to be selected from a lexical database. From these denotative word concepts and their relations an underlying denotative context can be constructed. In many other applications of Formal Concept Analysis the formal objects can be interpreted as denotata. For example, in a formal context of a library classification system the formal objects might be books (or their contents) which are items to which the titles of the books refer. If their formal attributes entirely depend on the books and not on associations of the titles of the books, such a formal context is a denotative context. For applications to lexical databases, prototypical instances or items are often chosen as denotata.

	seen after sunset	seen before sunrise	one of the 5 planets ...	diameter: 12756 km	diameter: 143650 km	diameter: 6770 km	diameter: 4840 km	diameter: 120670 km	attribute 1	attribute 2
Venus	x	x	x	x						
Jupiter	x	x	x		x					
Mars	x	x	x			x				
Mercury	x	x	x				x			
Saturn	x	x	x					x		
prot. even. star	x								x	
prot. morn. star		x								x

Figure 1.3.a: A denotative context for 'morning star' and 'evening star'

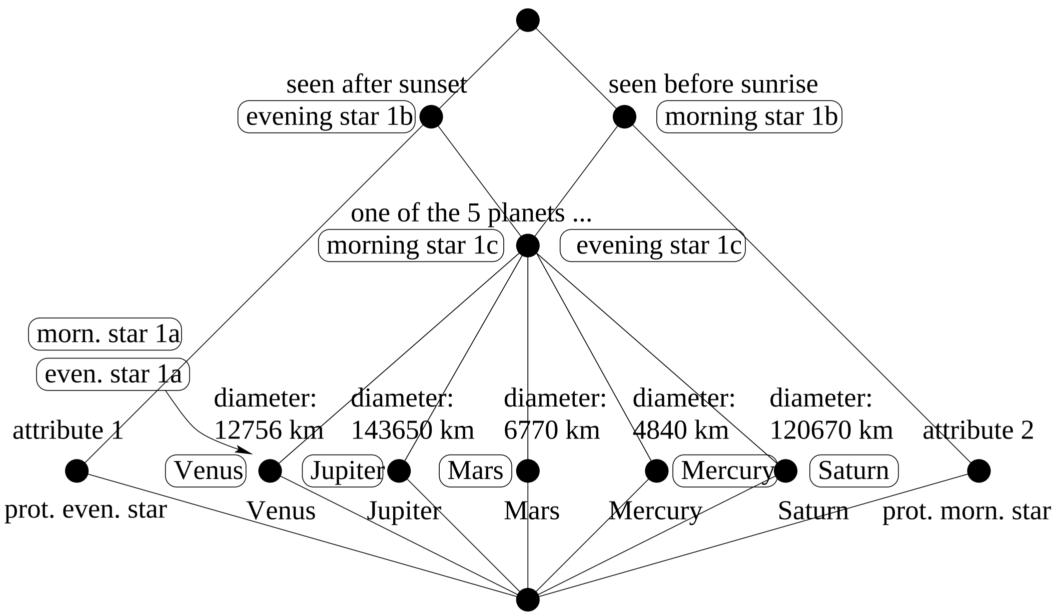

Figure 1.3.b: A line diagram of a denotative lattice for 'morning star'

Figures 1.3.a and 1.3.b show an example of a denotative context and a line diagram of its concept lattice. The denotative knowledge is derived from the definitions of 'morning star' and 'evening star' in W3 and some general background knowledge about planets. The disambiguated words that are notated inside of circles next to the concepts (which are therefore denotative word concepts) are taken from W3. The attributes 'seen after sunset', 'seen before sunrise', and 'one of the five planets' are directly taken from the defining glosses in W3. Since the planets, 'Venus', 'Jupiter', and so on, are usually considered to be different items (their names are not synonyms describing the same item) their diameters are chosen as differentiating attributes. The knowledge about the diameters is not contained in W3. Other attributes, such as 'has two moons' are not chosen in this case because usually 'planet with two moons' is not a denotative word concept in the English language. To represent the implicit denotative structures in a dictionary that directly depend on the disambiguated words it seems to be a good approach not to add too many attributes which lead to non-lexicalized concepts. Obviously, the selection of attributes depends on the purpose of the formal context and lattice. The formal objects, 'Venus', 'Jupiter', and so on, are denotata of the words 'Venus', 'Jupiter', and so on. The other formal objects, 'prototypical morning star' and 'prototypical evening star', are prototypical denotata of the words 'morning star' and 'evening star'. Since W3 distinguishes between the particular meanings '1b' and '1c' it is likely to assume that a prototypical morning star does not have to be a prototypical evening star, hence the addition of the prototypical denotata as formal objects. The attributes, 'attribute 1' and 'attribute 2', are prototypical attributes to differentiate between a prototypical morning and evening star which are not further specified.

The disambiguated words that are notated in the circles next to the concepts are not part of the denotative context. To include them and some other relations into the formalization a denotative structure is defined as an extension of a denotative context.

Definition 1.5:
A *denotative structure* is a relational structure

$$\mathcal{S}_{\mathcal{D}} := (D, A_D, C, W, C(W); I_D, dnt, \mathcal{R}_D, \mathcal{R}_{A_D}, \mathcal{R}_C)$$

with the following conditions:
1) $\mathcal{K}_D := (D, A_D, I_D)$ is a denotative context and C is its set of formal concepts, i.e. $C := \mathcal{B}(D, A_D, I_D)$. The formal concepts in C are called *denotative concepts*. The mappings $\iota, \varepsilon, \gamma, \mu$, and so on are defined according to Definition 1.1.
2) W can be empty. The mapping $dnt : W \to C$ is defined and $C(W) := \{c \in C \mid \exists_{w \in W} : c = dnt(w)\}$. The formal concepts in $C(W)$ are called *denotative word concepts*.
3) $\mathcal{R}_D, \mathcal{R}_{A_D}$, and $\mathcal{R}_C$ are families of relations: $\mathcal{R}_D := (R_{Dj})_{j \in J_D}$ is a family of relations $R_{Dj} \subseteq D \times D$; $\mathcal{R}_{A_D} := (R_{Aj})_{j \in J_A}$ is a family of relations $R_{Aj} \subseteq A_D \times A_D$; and

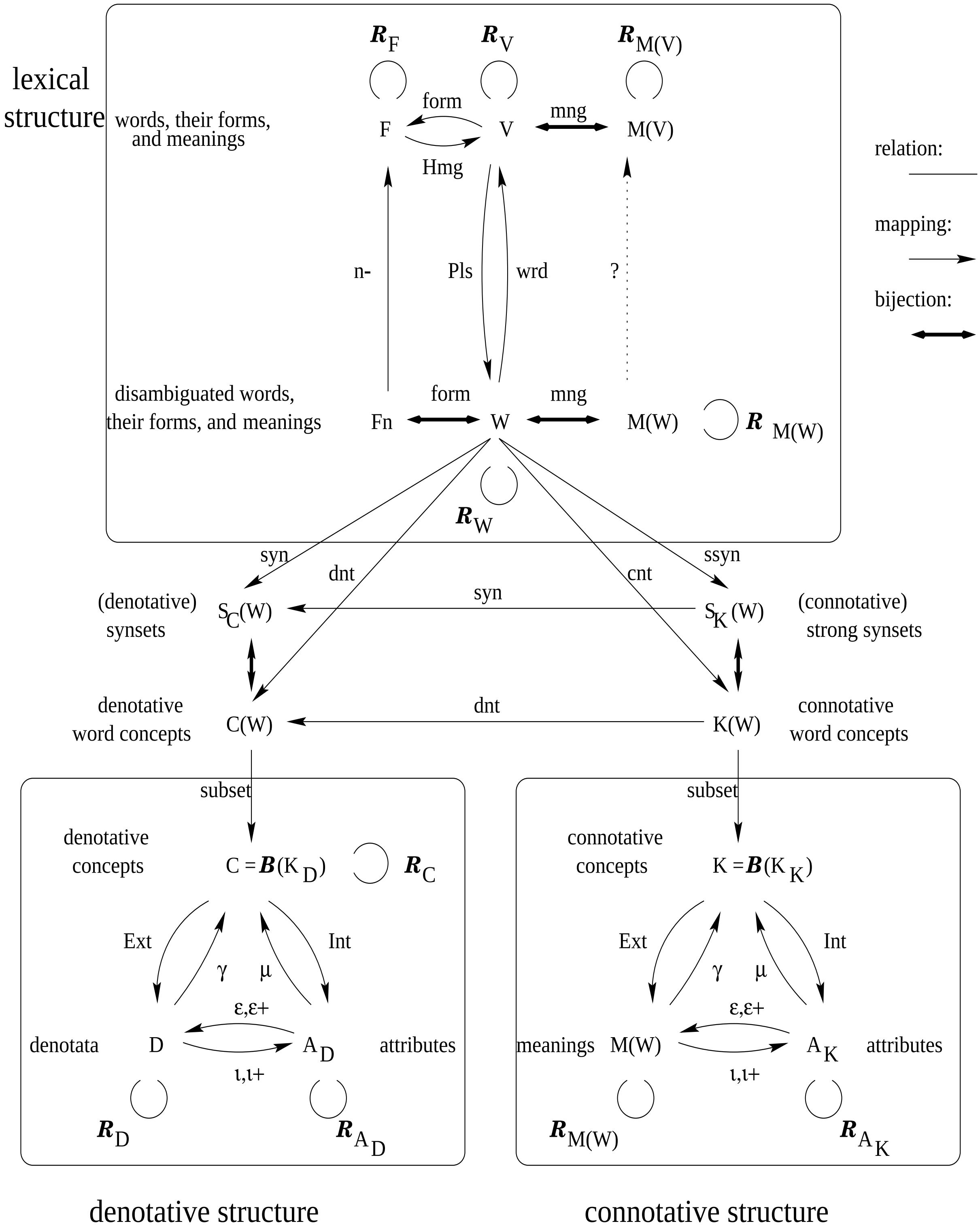

Figure 1.4: The denotative, connotative, and lexical structures

$\mathcal{R}_C := (R_{Cj})_{j \in J_C}$ is a family of relations $R_{Cj} \subseteq C \times C$. Relations of the families $\mathcal{R}_D$ and $\mathcal{R}_{A_D}$ and the relation I_D are called *denotative relations*. Relations of the family $\mathcal{R}_C$ are called *semantic relations*. $\mathcal{R}_D, \mathcal{R}_{A_D}$, and $\mathcal{R}_C$ can be empty.

If no ambiguities can occur some of the indices, such as the D in A_D, can be omitted. On the other hand, if ambiguities can occur, indices, such as I_D can be added to ι, ε, and so on. In the case of several denotative structures, all elements can be labeled with the name of the denotative structure, such as $C^{\mathcal{S}}$, $D^{\mathcal{S}}$, and $A_D^{\mathcal{S}}$. More details about the families of relations in 3) are presented in Section 1.7. The most important feature of denotative contexts is the fact that some concepts may be denoted by disambiguated words $w \in W$, such as the words inside of the circles in Figure 1.3.b. Therefore denotative word concepts can be distinguished from non-lexicalized denotative concepts. The disambiguated words which denote the denotative word concepts provide a further tool of naming concepts: denotative word concepts can either be described by their extents and intents or by the disambiguated words which denote them. A main issue of Section 2.8 is to investigate how some of the information of a denotative lattice is already contained in the subset of denotative word concepts. This is essential because a denotative context of a natural language can be derived from the disambiguated words of a lexical database and the question arises how the relations and structures among the words codify the relations and structures among their denotata.

Figure 1.4 shows a diagram of the sets, relations, and mappings of a denotative structure and of a connotative structure and a lexical structure which are defined in the next section. A short remark on the notation: a relation is represented in Figure 1.4 by a line without arrow which is labeled by the name of the relation family. Mappings are represented as lines with arrows. Except for the mappings that are denoted by small Greek characters (which are standard names in Formal Concept Analysis) mappings are denoted by three or four letter acronyms. They start with a small letter if the arrow in Figure 1.4 points to the set that they map into. They start with a capital letter if they map into the power set of the set to which the arrow in Figure 1.4 points. Therefore $M : X \to Y$ in Figure 1.4 means $M(X) \subseteq \wp(Y)$. Bijections are denoted by boldface lines with two arrowheads.

For some applications it is possible that, for example, a set of denotata and an ordering on the sets of the denotata but no attributes are given. The modeling of WordNet (see Section 1.12) presents such a case. It is then possible to interpret the sets of the denotata as extents of denotative concepts and create a denotative context and lattice from them. This method of constructing denotative contexts is equivalent to the usual method in the following sense: Formal contexts are called equivalent to each other if their reduced versions are isomorphic (Ganter & Wille, 1996). Equivalent versions of (D, A_D, I_D) are (D, C, I_D^{γ}), $(C, C, \leq)$, and (C, A_D, I_D^{μ})[5]. If, for example,

[5]Proof: The first two are equivalent according to the following constructions: starting with $\mathcal{K}_D$

$(C, C, \leq)$ is given, denotata and their attributes which are implicitly contained in the concepts can be explicitly named by using verbal phrases, for example, the denotative concept c with verbal form 'person' leads to the set of denotata 'prototypical persons' and to the set of attributes 'prototypical attributes of a person'.

1.5 Connotative and lexical structures

Connotative contexts and structures are defined analogously to denotative contexts and structures. The main difference (condition 2 in Definition 1.7) is that particular meanings are assumed to correspond to disambiguated words: each disambiguated word has a unique particular meaning because, otherwise, if two disambiguated words had exactly the same meaning one word would be redundant in the language system. On the other hand, each particular meaning is expressed by one word, because it is questionable what meanings that are not expressible by words should be. Particular meanings cannot be denotata because particular meanings contain all connotations of the disambiguated word.

Definition 1.6:
A *connotative context* $\mathcal{K}_K := (M(W), A_K, I_K)$ is defined as a formal context whose formal objects are particular meanings and whose formal attributes are features of the particular meanings. The set of particular meanings is denoted by $M(W)$, the set of features of the particular meanings by A_K, and a relation that assigns features to particular meanings by I_K. The concept lattice $\underline{\mathcal{B}}(\mathcal{K}_K)$ of a connotative context $\mathcal{K}_K$ is called *connotative lattice*.

Definition 1.7:
A *connotative structure* is a relational structure

$$\mathcal{S}_{\mathcal{K}} := (M(W), A_K, K, W, K(W); I_K, cnt, \mathcal{R}_{M(W)}, \mathcal{R}_{A_K})$$

with the following conditions:
1) $\mathcal{K}_K := (M(W), A_K, I_K)$ is a connotative context and K is its set of formal concepts, i.e. $K := \mathcal{B}(M(W), A_K, I_K)$. The formal concepts in K are called *connotative concepts*. The mappings $\iota, \varepsilon, \gamma, \mu$, and so on are defined according to Definition 1.1.
2) A bijection holds between $M(W)$ and W. The mapping $cnt : W \to K$ is defined as $cnt(w) := \gamma mng(w)$ and $K(W) := \{k \in K \mid \exists_{w \in W} : k = cnt(w)\}$. The formal concepts in $K(W)$ are called *connotative word concepts*.
3) $\mathcal{R}_{M(W)}$ and $\mathcal{R}_{A_K}$ are families of relations: $\mathcal{R}_{M(W)} := (R_{Mj})_{j \in J_M}$ is a family of relations $R_{Mj} \subseteq M(W) \times M(W)$ and $\mathcal{R}_{A_K} := (R_{Aj})_{j \in J_A}$ is a family of relations $R_{Aj} \subseteq A_K \times A_K$. Relations of the family $\mathcal{R}_{M(W)}$ are usually similarity relations. $\mathcal{R}_{M(W)}$ and $\mathcal{R}_{A_K}$ can be empty.

---

and $C := \mathcal{B}(\mathcal{K}_D)$ the relation is defined as $dI_D^{\gamma}c :\Longleftrightarrow \gamma d \leq c$ (in $\mathcal{K}_D$). Starting with (D, C, I_D^{γ}) for each meet irreducible concept $c_i \in C$ an additional attribute a_i is defined such that $\mu a_i = c_i$ (in (D, C, I_D^{γ})). Then $dI_D a_i :\Longleftrightarrow dI_D^{\gamma} a_i$. The other equivalences can be similarly shown.

A connotative word concept of a disambiguated word is therefore defined as the smallest connotative concept which contains the particular meaning of a disambiguated word in its extent. An equivalent version of a connotative context $(M(W), A_K, I_K)$ is a context (W, A_K, I_K) because there is a one-to-one correspondence between disambiguated words and their particular meanings and therefore a connotative context can also be modeled using the words as formal objects. The intent of a connotative concept depends on the attributes which it has in a connotative context. These attributes are the features (see Definition 1.3) that a particular meaning has according to the viewpoint of a connotative context. Connotative concepts often represent what linguists (for example Saussure (1916)) mean when they say the meaning of a word is a concept. It can be said that Lyons' (1977) 'sense' of a word which is the place of the word in the language system is described by its place in a connotative lattice. We think that a particular meaning is in the extent of connotative concepts, but should not be modeled as a formal concept itself because the formal objects and attributes of a meaning would be difficult to define. On the other hand, connotative concepts cannot be modeled independent of particular meanings (or disambiguated words) because the connotations depend on the word (or word form) and on its relation to other words. 'Morning star' has a different connotation from 'evening star' because there are words 'morning' and 'evening' in the English language which have meanings themselves. The relation between 'morning star' and 'morning', 'evening star' and 'evening', and so on is an example for a similarity relation that can be in the family $\mathcal{R}_{M(W)}$ of relations.

Figure 1.5.a and 1.5.b show an example of a connotative context and lattice. The attributes 'feline', 'canine', 'domesticated', and 'wild' are attributes of denotata, but the attributes 'common language' and 'biological terminology' depend on the disambiguated words. For example, 'domestic dog' and 'Canis familiaris' have usually the same denotata, but they are used in different language contexts. In this example, a denotative context is part of a connotative context. Often denotative and connotative structures are modeled on the same set of disambiguated words. If they do share the same set of disambiguated words the denotative concepts should be entailed by the connotative concepts because according to Definition 1.3 the particular meaning of a disambiguated word contains all aspects of the word meaning therefore the denotation of the word should be included. Hence the following is defined:

Definition 1.8:
A denotative structure $\mathcal{S}_D$ and a connotative structure $\mathcal{S}_K$ can be *combined* if they contain the same set of disambiguated words, i.e. $W := W^{\mathcal{S}_D}$ and $W = W^{\mathcal{S}_K}$, and a mapping $dnt : K(W) \to C(W)$ can be defined with $dnt(cnt^{\mathcal{S}_K}(w)) = dnt^{\mathcal{S}_D}(w)$ for all $w \in W$.

If $\mathcal{S}_D$ and $\mathcal{S}_K$ can be combined it follows that $cnt^{\mathcal{S}_K}(w_1) = cnt^{\mathcal{S}_K}(w_2) \Longrightarrow dnt^{\mathcal{S}_D}(w_1) = dnt^{\mathcal{S}_D}(w_2)$. The other direction does not have to be true, because each denotative concept can belong to several connotative concepts. For example, 'Canis familiaris' and ‘domestic dog’ can have the same denotation, but they have different connotations. The next definition provides a formal modeling of a lexical database.

	canine	feline	domesticated	wild	common language	biological term
domestic dog	x		x		x	
Canis familiaris	x		x			x
jackal	x			x	x	
Canis aureus	x			x		x
domestic cat		x	x		x	
Felis catus		x	x			x
wildcat		x		x	x	

Figure 1.5.a: A connotative context

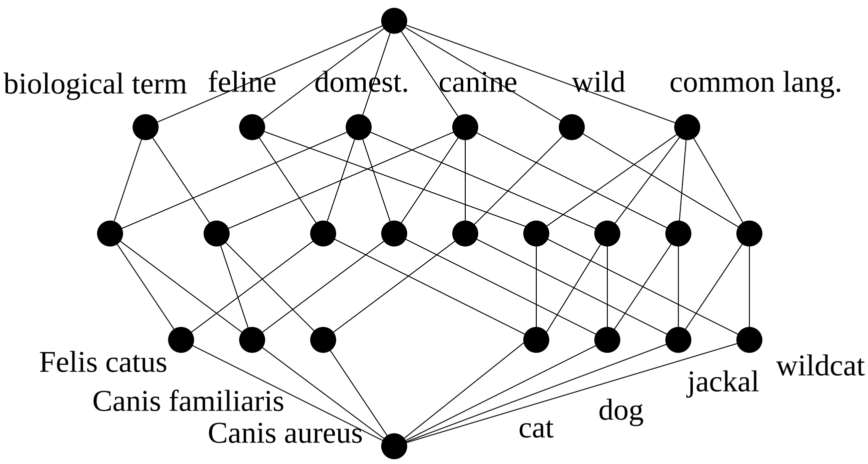

Figure 1.5.b: A line diagram of the connotative lattice of the context in Fig. 1.5.a

Definition 1.9:
A *lexical structure* is a relational structure

$$\mathcal{S_L} := \quad (W, M(W), Fn, V, M(V), F, A_L; I_L, form, mng, Hmg, \\ Pls, n^-, wrd, \mathcal{R}_W, \mathcal{R}_{M(W)}, \mathcal{R}_V, \mathcal{R}_{M(V)}, \mathcal{R}_F,)$$

with the following conditions:
1) The elements of the set V are called *words*, the elements of the set W *disambiguated words*, the elements of the set $M(V)$ *word meanings*, the elements of the set $M(W)$ *particular meanings*, the elements of the set F *word forms*, and the elements of the set Fn *forms of disambiguated words*.
2) A formal context $\mathcal{K}_L := (W, A_L, I_L)$ is called *lexical context*. It has disambiguated words as formal objects and attributes of those disambiguated words as formal attributes. The set of formal attributes is denoted by A_L and the relation between formal objects and attributes by I_L. A lattice $\underline{\mathcal{B}}(\mathcal{K}_L)$ is called *lexical lattice*.

3) $M(V)$ can be empty. $form : W \cup V \to Fn \cup F$ and $mng : W \cup V \to M(W) \cup M(V)$ are mappings with $form(w) \in Fn$ for $w \in W$, $form(v) \in F$ for $v \in V$, $mng(w) \in M(W)$ for $w \in W$, and if $M(V) \neq \emptyset$, $mng(v) \in M(V)$ for $v \in V$. $mng : W \to M(W)$, $form : W \to Fn$, and, if defined, $mng : V \to M(V)$ are bijections.
4) The mappings $n^- : Fn \to F$ and $wrd : W \to V$ must fulfill $n^- form(w) = form(wrd(w))$.
5) The mappings $Pls : V \to \wp(W)$ and $Hmg : F \to \wp(V)$ are defined as $Pls(v) := \{w \in W \mid wrd(w) = v\}$ and $Hmg(f) := \{v \in V \mid form(v) = f\}$ and called *polysemy* and *homography*, respectively.
6) $\mathcal{R}_W$, $\mathcal{R}_{M(W)}$, $\mathcal{R}_V$, $\mathcal{R}_{M(V)}$ and $\mathcal{R}_F$ are families of relations: $\mathcal{R}_W := (R_{Wj})_{j \in J_W}$ is a family of relations $R_{Wj} \subseteq W \times W$; $\mathcal{R}_{M(W)} := (R_{Mj})_{j \in J_M}$ a family of relations $R_{Mj} \subseteq M(W) \times M(W)$; $\mathcal{R}_V := (R_{Vj})_{j \in J_V}$ a family of relations $R_{Vj} \subseteq V \times V$; $\mathcal{R}_{M(V)} := (R_{Mj})_{j \in J_M}$ a family of relations $R_{Mj} \subseteq M(V) \times M(V)$; and $\mathcal{R}_F := (R_{Fj})_{j \in J_F}$ is a family of relations $R_{Fj} \subseteq F \times F$. Relations of the families $\mathcal{R}_{M(W)}$ and $\mathcal{R}_{M(V)}$ are usually similarity relations. Relations of the family $\mathcal{R}_W$ are called *lexical relations* (and under some circumstances (see Section 1.6) *semantic relations*), relations of the family $\mathcal{R}_V$ *morpho-lexical relations*, and relations of the family $\mathcal{R}_F$ *morphological relations*. All families of relations can be empty.

The following should be remarked: Similarly to denotative and connotative structures, which contain denotative or connotative contexts, a lexical structure contains a lexical context. The set of formal attributes of the lexical context can coincide with other sets of the lexical structure, but it can also be disjoint to the other sets. Definition 1.9 is consistent with Definitions 1.2 and 1.3 because according to condition 3 in Definition 1.9 there is a one-to-one correspondence between w and $mng(w)$ and v and $mng(v)$. Therefore a word can be uniquely identified by the tuple $(form(v), mng(v))$ and a disambiguated word by $(form(w), mng(w))$ or $(n^- form(w), mng(w))$. If forms of disambiguated words (elements of the set Fn) are thought to be word forms concatenated with a sense number then the mapping n^- (see condition 4) yields for each form fn the corresponding word form f by deleting the sense number. The mapping $form : W \to Fn$ is a bijection because every word form with a sense number corresponds to exactly one disambiguated word. The mapping wrd from disambiguated words to words obtains the not disambiguated basic word $v := wrd(w)$ for a disambiguated word w. According to condition 5, polysemy is defined as the set of disambiguated words which belong to one word. Or, since disambiguated words correspond to particular meanings, polysemy is represented by the set of particular meanings of a word. Polysemy is not defined on word forms because only disambiguated words that share a word form and have similar particular meanings are normally called polysemous. Modifying conditions 4 and 5, it would be possible to define polysemy based on a similarity relation on the meanings. For example, a similarity of meaning (SIM_{Pls}) could be defined as 'having the same intent' in a connotative structure $\mathcal{S}_K$, i.e. $mng(w_1) SIM_{Pls} mng(w_2) :\Longleftrightarrow \gamma^{\mathcal{S}_K} mng(w_1) = \gamma^{\mathcal{S}_K} mng(w_2)$. Then two disambiguated words would be polysemous if they have the same word form and a similar meaning, i.e. $\exists_{v\in V} : (w_1 \in Pls(v) \text{ and } w_2 \in Pls(v)) :\Longleftrightarrow (mng(w_1) SIM_{Pls} mng(w_2) \text{ and } n^- form(w_1) = n^- form(w_2))$. Since polysemy is a partition of W into equivalence classes, the mapping wrd can also be defined as the mapping which leads from each equivalence class in W to its basic word in V. Homography is defined on word forms as the set of words which share a word form. Hmg creates a partition on V into equivalence classes. As an example for the terminology, the word v_1 with $form(v_1)$ = 'morning star' in W3 has five polysemous disambiguated words ($Pls(v_1) = \{w_1, w_2, w_3, w_4, w_5\}$) with $form(w_1)$ = 'morning star 1a', $form(w_2)$ = 'morning star 1b', and so on. The mapping n^- yields n^-('morning star 1a') ='morning star'. The homography Hmg('morning star') = $\{v_1\}$ is a set with one element v_1.

It is not discussed in this paper if there is a mapping between word meanings and the particular meanings of their disambiguated words. In a connotative structure the meaning of a word could be defined as the meet of the particular meanings in a connotative lattice (the common core meaning, $mng(v) := \bigwedge\{\gamma mng(w) \mid v = wrd(w)\}$) or their join (all the possible usages of a word, $mng(v) := \bigvee\{\gamma mng(w) \mid v = wrd(w)\}$). Or it could be tried to model Wittgenstein's (1971) 'family resemblance', for example, $mng(v) := \bigvee_{w_1,w_2,w_3 \in Pls(v)}(\gamma mng(w_1) \wedge \gamma mng(w_2) \wedge \gamma mng(w_3))$, but this will not be investigated in this paper. In general the Definition 1.9 is thought to be a framework for the modeling of dictionaries, lexical databases, and thesauri. Sections 1.10 to 1.12 illustrate how the definition can be applied to three prototypical examples: W3, RIT, and WordNet. Depending on what is actually given in a lexical database, the sets and relations of the lexical structure might have to be adapted. We excluded synonymy from the definition of a lexical structure, because in a thesaurus it is an important relation which generates the structure, but in a dictionary it is presented only implicitly or it entirely depends on denotative or connotative contexts.

1.6 Synonymy and word equivalence

The term 'synonymy' has probably as many different definitions in linguistic theories as the term 'word'. We mention only two definitions which can be interpreted within our theory: first, in componential semantics (according to Lincke et al. (1994)) two words are called 'synonymous' if they have the same semantic features. In our modeling this corresponds to 'having the same intent in a connotative or denotative context'. Second, another common definition is to call two words synonymous if they can be exchanged for each other in every language context (Ullmann (1957) calls this 'pure synonymy') or if they can be exchanged for each other in some language context (Ullmann (1957) calls this 'pseudo-synonymy'). Since 'pure synonyms' are very rare, we define synonymy for disambiguated words and not for words in general. A distinction can be made whether the disambiguated words share their denotative or their connotative concepts. We call the first one 'synonymy' and the second one 'strong synonymy'. Although it seems that words which are synonymous (but not strong synonymous) to each other for all their particular meanings should be very rare in natural languages, Old (1996) searches for all words in RIT which always occur in the same semicolon-groups (see Section 1.11) and discovers that these 'word equivalents' are not that rare. His lists include words and their abbreviations, words with spelling variations, English words with foreign translations, and so on. Lincke et al. (1994) find similar reasons for synonymy or partial synonymy: a language often has regional, social, and stylistic distinct words which have the same denotation. We adapt the term 'synset' for 'set of synonymous words' from Miller et al. (1990).

Definition 1.10:
Synonymy $SYN^{\mathcal{S}_D}$ is a relation among disambiguated words in a denotative structure $\mathcal{S}_D$. Disambiguated words are called *synonyms* in a denotative structure, if they denote the same denotative word concept, i.e. $w_1 SYN^{\mathcal{S}_D} w_2 :\Longleftrightarrow dnt(w_1) = dnt(w_2)$. The mapping $syn : W \to \wp(W)$ is defined as $syn(w) := \{w_1 \in W \mid dnt(w) = dnt(w_1)\}$ and $S_C(W) := \{A \subseteq W \mid \exists_{w \in W} A = syn(w)\}$. $syn(w)$ is called the *synset* of the disambiguated word w in $\mathcal{S}_D$.
Strong synonymy $SSYN^{\mathcal{S}_K}$ is a relation among disambiguated words in a connotative structure $\mathcal{S}_K$. Disambiguated words are called *strong synonyms* in a connotative structure, if they connote the same connotative word concept, i.e. $w_1 SSYN^{\mathcal{S}_K} w_2 :\Longleftrightarrow cnt(w_1) = cnt(w_2)$. The mapping $ssyn : W \to \wp(W)$ is defined as $ssyn(w) := \{w_1 \in W \mid cnt(w) = cnt(w_1)\}$ and $S_K(W) := \{A \subseteq W \mid \exists_{w \in W} A = ssyn(w)\}$. $ssyn(w)$ is called the *strong synset* of the disambiguated word w in $\mathcal{S}_K$.
Word equivalency $EQV^{\mathcal{S}_D}$ is a morpho-lexical relation in a lexical structure according to a denotative structure $\mathcal{S}_D$. Two words v_1, v_2 are called *word equivalents*, if they share all their synsets, i.e. $v_1 EQV^{\mathcal{S}_D} v_2 \Longleftrightarrow (\forall_{w_1 \in Pls(v_1)} \exists_{w_2 \in Pls(v_2)} : syn(w_1) = syn(w_2)$ and $\forall_{w_2 \in Pls(v_2)} \exists_{w_1 \in Pls(v_1)} : syn(w_1) = syn(w_2))$.

Since dnt and cnt are mappings, $SYN^{\mathcal{S}_D}$ and $SSYN^{\mathcal{S}_K}$ are equivalence relations on W. Lexical relations that hold for all synonyms of the disambiguated word are called semantic relations (see Section 2.8) because they are defined on the denotative word concepts. Therefore synonymy is a semantic relation. In a denotative lattice, synonymous words denote the same denotative word concept. In a connotative lattice the strong synsets are the sets of disambiguated words that connote the same connotative concept which means their object concepts are equal. As an example consider Figure 1.3.b and the disambiguated word w_1 with $form(w_1) =$'morning star 1a'. $S_C(w_1) = \{$morning star1a, evening star1a, Venus$\}$ denotes the denotative word concept 'the planet Venus'. $\{$morning star1b$\}$, $\{$evening star1b$\}$ and $\{$morning star1c, evening star1c$\}$ are further synsets in Figure 1.3.b. None of these are strong synsets according to the lattice in Figure 1.7. Since all disambiguated words in a synset denote the same denotative word concept there is a bijection between synsets and denotative word concepts, and similarly there is a bijection between strong synsets and connotative word concepts. Therefore the denotative and connotative word concepts can be described by their synsets and strong synsets, respectively. But, because of word equivalents and because of synsets which contain only one word, it is usually not possible to uniquely represent a synset as a set of word forms without their sense numbers. For example, {morning star, evening star} is ambiguous in Figure 1.3.b. If $\mathcal{S}_D$ and $\mathcal{S}_K$ can be combined then it follows that $syn(w) = syn(ssyn(w))$.

1.7 Other relations

This section presents a definition of hyponymy, which is another semantic relation besides synonymy, and some details on the classification of relations into families of relations in the Definitions 1.5, 1.7, and 1.9. More details on relations can be found in chapter 3. It should be remarked that hyponymy and synonymy according to Definition 1.10 and 1.11 always depend on the denotative and connotative contexts. That means if the attributes in A_D and A_K are not appropriately selected for the purpose of displaying linguistic synonymy and hyponymy, it is possible that words become synonyms or hyponyms according to the Definitions 1.10 and 1.11 which usually would not be called synonyms or hyponyms in linguistic terminology.

Definition 1.11:
(Denotative) hyponymy $HYP^{\mathcal{S}_D}$ is a relation among disambiguated words in a denotative structure. A disambiguated word is called *(denotative) hyponym* of another disambiguated word, if its denotative concept is a subconcept of the denotative concept of the other word, i.e. $w_1 HYP^{\mathcal{S}_D} w_2 :\Longleftrightarrow dnt(w_1) \leq^{\mathcal{S}_D} dnt(w_2)$.
(Connotative) hyponymy $HYP^{\mathcal{S}_K}$ is a relation among disambiguated words in a connotative structure. A disambiguated word is called *(connotative) hyponym* of another disambiguated word, if its connotative concept is a subconcept of the connotative concept of the other word, i.e. $w_1 HYP^{\mathcal{S}_K} w_2 :\Longleftrightarrow cnt(w_1) \leq^{\mathcal{S}_K} cnt(w_2)$.
Hypernymy is the inverse relation to hyponymy, i.e. if w_1 is a (denotative or connotative) hyponym of w_2 then w_2 is a (denotative or connotative) hypernym of w_1.

The classification of relations in this paper which is based on the sets on which the relations are defined is similar to DIN 2330 (1979) and WordNet terminology. DIN 2330 distinguishes three types of relations: ontological relations, which are based on objects; abstract relations, which are based on intents of concepts; and contextual (syntagmatic) relations, which are based on collocations in a text or on a scheme. WordNet (Miller et al, 1990), on the other hand, distinguishes only semantic relations, which are defined on synsets, from lexical relations, which are defined on disambiguated words. In this paper the ontological relations (DIN 2330) are called denotative relations which are defined on denotata and attributes of denotata and also include the relation I_D between denotata and their attributes. Denotative relations are relations which depend only on the denotata and not on the verbal representations of the denotata and their attributes. These relations include denotative part-whole relations (a certain or prototypical denotatum has a certain or prototypical part), and others. A main issue of chapter 2 is to investigate how these denotative relations can be generalized to semantic relations (for example meronymy) among denotative concepts or disambiguated words via their denotative word concepts. Hence semantic relations in our terminology correspond to abstract relations (DIN 2330), but not to, for example, Rahmstorf's (1991) semantic relations which are defined between expressions and their meanings. Lexical relations (for example antonymy) are defined on disambiguated words. Our distinction between semantic and lexical relations corresponds therefore to the WordNet terminology. Antonymy is a lexical and not a semantic relation, because it depends on certain contrary attributes but also on some intersubjective agreement about which words are considered to be pairs. For example, 'fast' and 'slow' are antonyms, but 'rapid' and 'slow' usually not.

Morpho-lexical relations are relations on words. For example, etymological relations are usually morpho-lexical because they often refer to several particular meanings of a disambiguated word. Since etymological relations do normally not include homographs, they are not morphological relations which entirely depend on the word forms. The alphabetical order of the lemmas in a dictionary is a morphological relation. Contextual or syntagmatic relations (DIN 2330) are only part of our model if they are lexicalized and therefore can be classified as semantic or lexical relations. For example, between 'dog' and 'bark' a functional semantic relation could be defined in a lexical database.

1.8 Linguistic contexts

In this section the terminology is further extended and applied to some examples. It is assumed for this section that a lexical, a denotative, and a connotative structure ($\mathcal{S}_D$, $\mathcal{S}_K$, and $\mathcal{S}_L$) can be combined, that means that they all share the same set W of disambiguated words and the conditions from Definition 1.8 hold. Linguistic contexts are according to Section 1.1 formal contexts whose formal objects and attributes are elements of linguistic research. The most important linguistic contexts are probably lexical contexts (compare Definition 1.9) which have two basic types:

Definition 1.12:
The two basic types of lexical contexts and lattices are: *lexical denotative contexts* $\mathcal{K}_{LD} := (W, A_D, I_{LD})$ (*lexical denotative lattice* $\underline{\mathcal{B}}(\mathcal{K}_{LD})$) and *lexical connotative contexts* $\mathcal{K}_{LK} := (W, A_K, I_{LK})$ (*lexical connotative lattice* $\underline{\mathcal{B}}(\mathcal{K}_{LK})$) whose sets of attributes are taken from a denotative or connotative structure. The relations are defined as $wI_{LD}a :\Longleftrightarrow dnt(w) \leq^{\mathcal{S}_D} \mu a$ and $wI_{LK}a :\Longleftrightarrow cnt(w) \leq^{\mathcal{S}_K} \mu a$, respectively.

From the definition follows that $\underline{\mathcal{B}}(\mathcal{K}_{LD})$ is isomorphic to a join-preserving sublattice of its $\underline{\mathcal{B}}(\mathcal{K}_D)$[6]. A lexical connotative lattice $\underline{\mathcal{B}}(\mathcal{K}_{LK})$ is even isomorphic to its $\underline{\mathcal{B}}(\mathcal{K}_K)$ because of the bijection between $M(W)$ and W and the implied isomorphism between I_{LK} and I_K. Using these isomorphies, in lexical denotative and connotative

[6]Proof: Instead of $\mathcal{K}_D$ a formal context $\mathcal{K}_D^* := (D \cup W, A_D, I_D \cup I_{LD})$ can be defined. Obviously, $\underline{\mathcal{B}}(\mathcal{K}_D) = \underline{\mathcal{B}}(\mathcal{K}_D^*)$. Since all rows of $\mathcal{K}_{LD}$ are contained in $\mathcal{K}_D^*$ it follows that $\underline{\mathcal{B}}(\mathcal{K}_{LD})$ is a join-preserving sublattice of $\underline{\mathcal{B}}(\mathcal{K}_D^*)$ which proves the statement.

lattices exactly the object concepts correspond to the (denotative or connotative) word concepts. The set A_K of connotative attributes A_K can include the set A_D of denotative attributes in a denotative structure. For example, A_K can be the union of A_D and a set of style features, such as {neutral, poetic, vernacular, vulgar, ...} (compare the context in Figure 1.5.a).

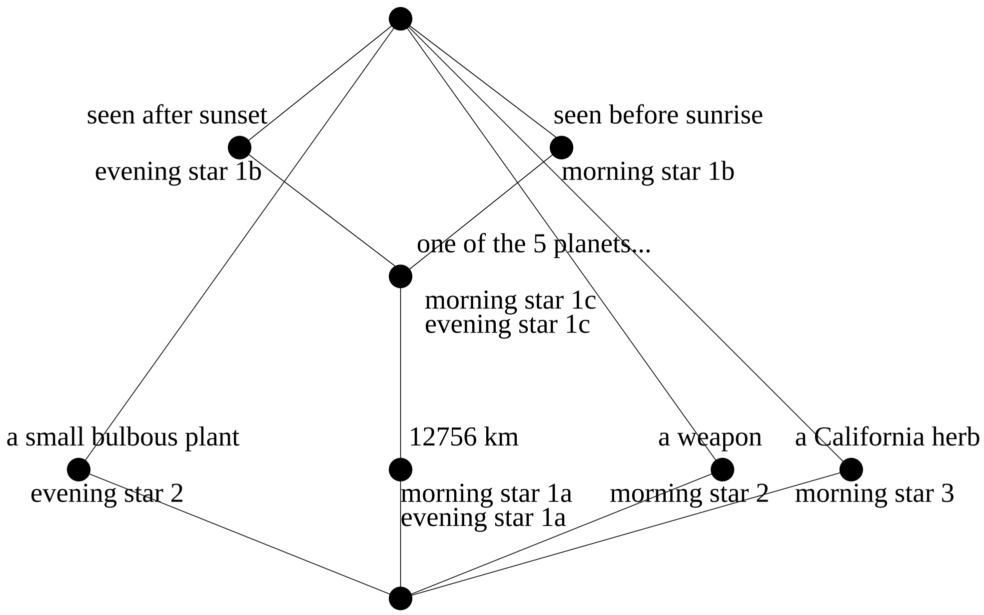

Figure 1.6: A lexical denotative concept lattice of a context (W, A_D, I_{LD})

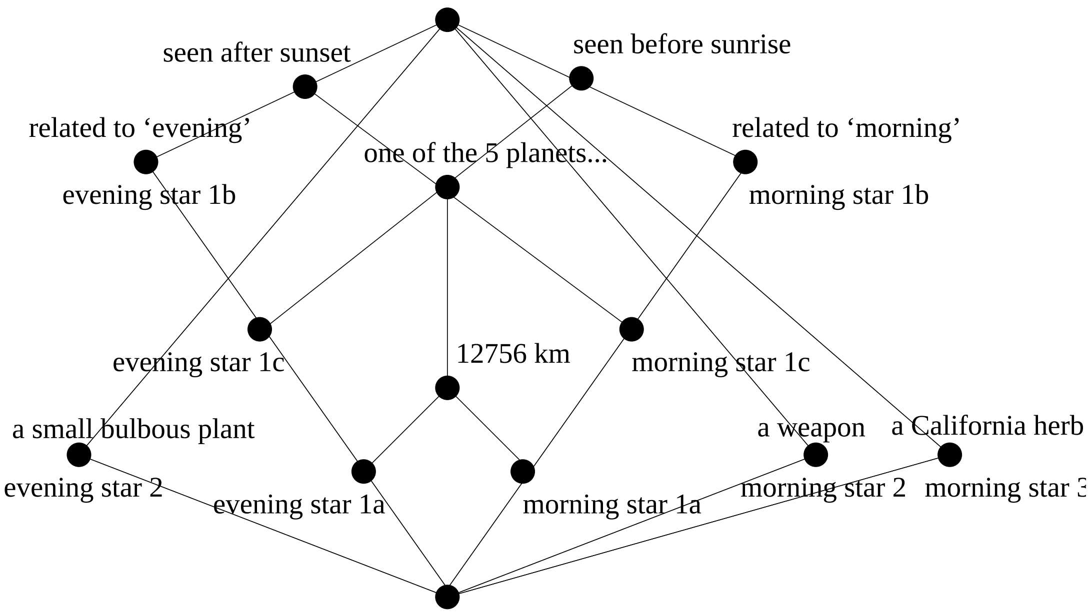

Figure 1.7: A lexical connotative lattice of a context (W, A_K, I_{LK})

Figure 1.6 shows a line diagram of a lexical denotative lattice (W, A_D, I_{LD}), which has the attributes of its denotata in an underlying denotative context $\mathcal{K}_D$ (compare Figure

1.3) as attributes. In many non-linguistic applications of Formal Concept Analysis it is difficult to distinguish between $\mathcal{K}_D$ and $\mathcal{K}_{LD}$, especially, if the disambiguated words denote individual concepts. The main difference is that denotata are always instances or items (see Definition 1.3), therefore 'morning star 3' cannot be a formal object in a denotative context, whereas 'prototypical morning star 3' is an instance and can refer to a denotatum. This distinction is essential for the modeling of relations (see chapter 2) because denotative relations are not quantified whereas semantic relations need quantifiers. Figure 1.7 illustrates a lattice of a lexical connotative context (W, A_K, I_{LK}) with $A_K := A_D \cup \{$related to 'evening', related to 'morning'$\}$, $I_{LD} \subseteq I_{LK}$ and W, A_D and I_{LD} are taken from the denotative lattice in Figure 1.6. The denotative word concepts in Figure 1.6 can be described by their synsets, such as $\{$morning star 1a, evening star 1a$\}$, or by the extents and intents of the formal concepts, such as ($\{$morning star 1a, evening star1a$\}$, $\{$12756 km, one of the 5 planets..., seen after sunset, seen before sunrise$\}$). A connotative word concept in Figure 1.7 is, for example, ($\{$evening star 1c, evening star 1a$\}$, $\{$one of the 5 planets..., seen after sunset, seen before sunrise, related to 'evening'$\}$) which is described by the strong synset $\{$evening star 1c$\}$. It should be noted that the synsets and strong synsets are formed by the disambiguated words which share their object concepts and not by the extents of the formal concepts.

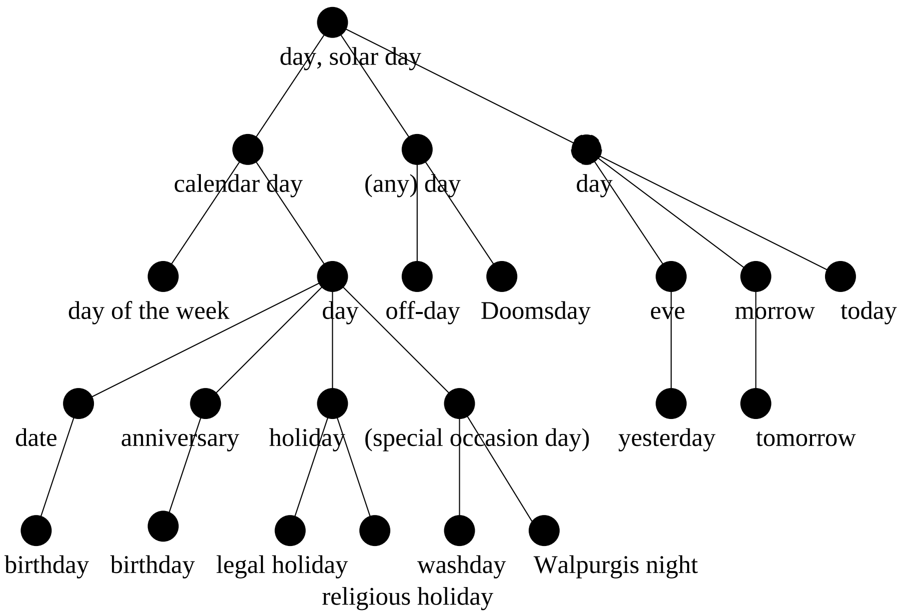

Figure 1.8: A lexical concept hierarchy

Figure 1.8 shows a lexical concept hierarchy (a lattice without bottom concept) which is modeled after WordNet. It is difficult to classify the corresponding lexical lattice. It cannot be a denotative lattice because even a prototypical 'yesterday' should be a day of the week. Therefore the formal objects are not denotata, but disambiguated words. It can be a lexical denotative lattice where the prototypical denotata in an underlying denotative lattice, such as 'prototypical yesterday', are non-lexicalized. It could also be argued that some of the attributes, which are never explicitly mentioned in WordNet, are not purely denotative because every day is a day of the week, a day, a yesterday, a tomorrow, and a today at the same time. Whether a day is a yesterday or a tomorrow does not depend on the day, but on the perspective of the speaker. On the other hand only the word tokens 'yesterday' and 'tomorrow' can refer to the same denotatum, whereas prototypical 'yesterdays' and 'tomorrows' are probably different. Hence, for the modeling of WordNet in Section 1.12, we decided to interpret concept hierarchies, such as the one in Figure 1.8, as lexical denotative.

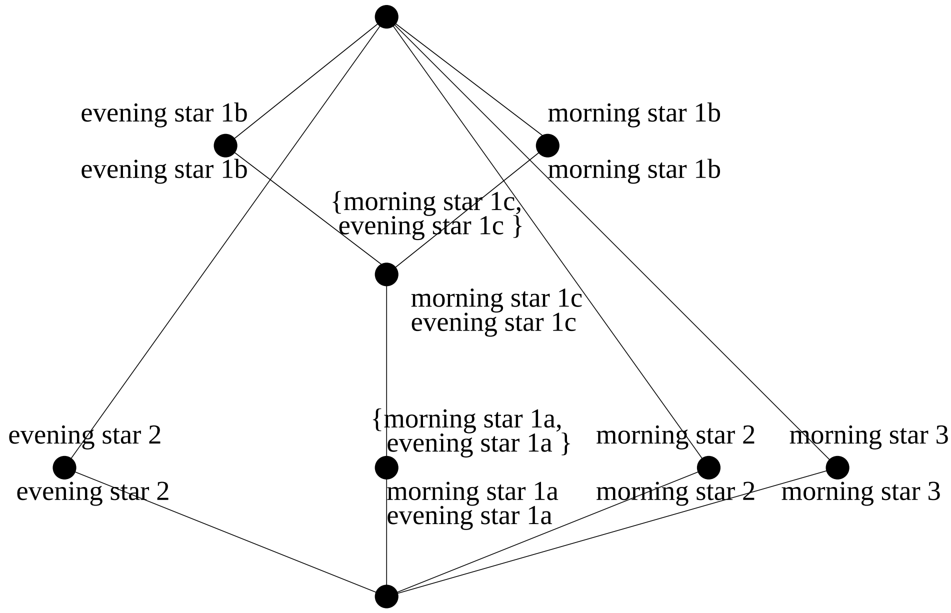

Figure 1.9: The lattice of Figure 1.6 as $(W, S_C(W), HYP^*)$

Similarly to what has been said in Section 1.4 there are different possibilities to construct equivalent versions of formal contexts and lattices. For example, instead of disambiguated words, denotative or connotative word concepts or synsets can be taken as the formal objects of an equivalent version of a lexical context. $(S_C(W), A_D, I^{\mu}_{LD})$ is an equivalent version of $\mathcal{K}_{LD}$ if the relation I^{μ}_{LD} is defined by $syn(w_1)I^{\mu}_{LD}a :\Longleftrightarrow \forall_{w \in syn(w_1)} : wI_{LD}a$. For example, if the lexical denotative concept lattice of morning star in Figure 1.6 is interpreted as a $\mathcal{K}_{LD}$, then the objects are 'morning star 1a', 'evening star 1a', and so on. If it is interpreted as $(S_C(W), A_D, I^{\mu}_{LD})$, then the objects are synsets, such as {morning star 1a, evening star 1a} and so on. Since there is a bijection between $S_C(W)$ and $C(W)$, $(C(W), A_D, I^{\mu}_{LD})$ is another equivalent representation. The object concepts of a lexical denotative lattice are always denotative word concepts. Therefore, since a lattice can be uniquely constructed from its object and attribute concepts and their ordering (Ganter & Wille, 1996), $(C(W), C(W), \leq)$, $(S_C(W), S_C(W), \leq)$, and $(W, S_C(W), HYP^*)$ are further equivalent representations of a $\mathcal{K}_{LD}$ if all attribute concepts are also denotative word concepts. The relation HYP^* is defined by $w_1 HYP^* syn(w_2) :\Longleftrightarrow dnt(w_1) \leq^{LD} dnt(w_2)$. Figure 1.9 shows an example of a lexical lattice for a context $(W, S_C(W), HYP^*)$.

Obviously, it would be sufficient to write either the attributes or the objects in the line diagram. Section 1.12 demonstrates that WordNet can be modeled in the form of $(W, S_C(W), HYP^*)$. Equivalent versions of lexical connotative contexts are $(K(W), A_K, I^{\mu}_{LK})$ and $(S_K(W), A_K, I^{\mu}_{LK})$, and so on.

	god	goddess	planet	plant	weapon
evening star			x	x	
morning star			x		x
Venus		x	x		
Jupiter	x		x		
Aphrodite		x			
Zeus	x				
Cooperia drummondii				x	
holy water sprinkler					x

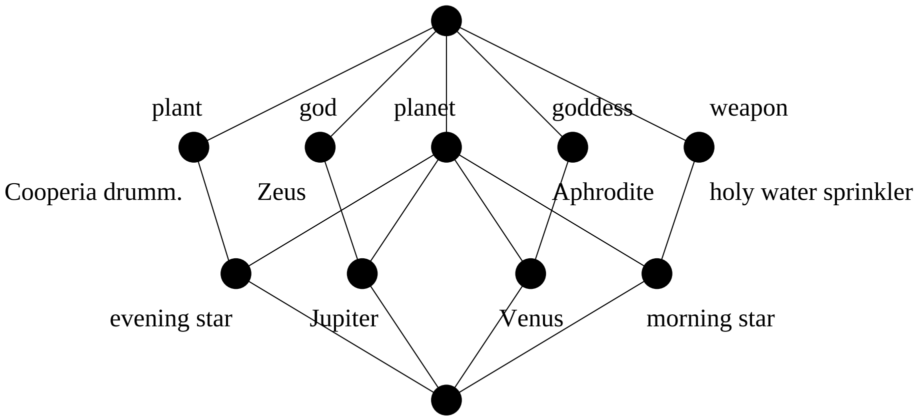

Figure 1.10: A morpho-lexical context (V, A_D, I_{MLD}) and its lattice

To visualize the polysemy, a so-called *morpho-lexical denotative lattice* of a context (V, A_D, I_{MLD}) can be investigated which has words as objects and attributes of the denotata of its particular meanings in an underlying lexical denotative context as attributes. The relation I_{MLD} is usually defined as $vI_{MLD}a :\Longleftrightarrow \exists_{w \in W} : (wrd(w) = v \text{ and } wI_{LD}a)$ that means that all denotative attributes which at least one disambiguated word has are considered[7]. Figure 1.10 illustrates such a morpho-lexical lattice. It gives a survey of the polysemy of ‘morning star’ and ‘evening star’. (‘Morning star 3’ is missing.) Since words do not have denotata (only disambiguated words have denotata according to Definition 1.3), it is not possible to assign denotata to concepts

[7]Alternatively the definition $vI_{MLD}a :\Longleftrightarrow \forall_{w \in W} : (wrd(w) = v \text{ and } wI_{LD}a)$ is possible if only the common denotative attributes of all disambiguated words of a word are considered.

in the lattice in Figure 1.10. For example, ‘evening star’ has the attributes of being a planet and a plant, but no denotatum can be a planet and a plant at the same time. On the other hand, in lexical denotative lattices it is often possible to assign denotata to concepts, especially, if the disambiguated words denote individual concepts. In Section 1.11 so-called neighborhood contexts which are of the form (V, A_D, I_{MLD}) are developed for RIT.

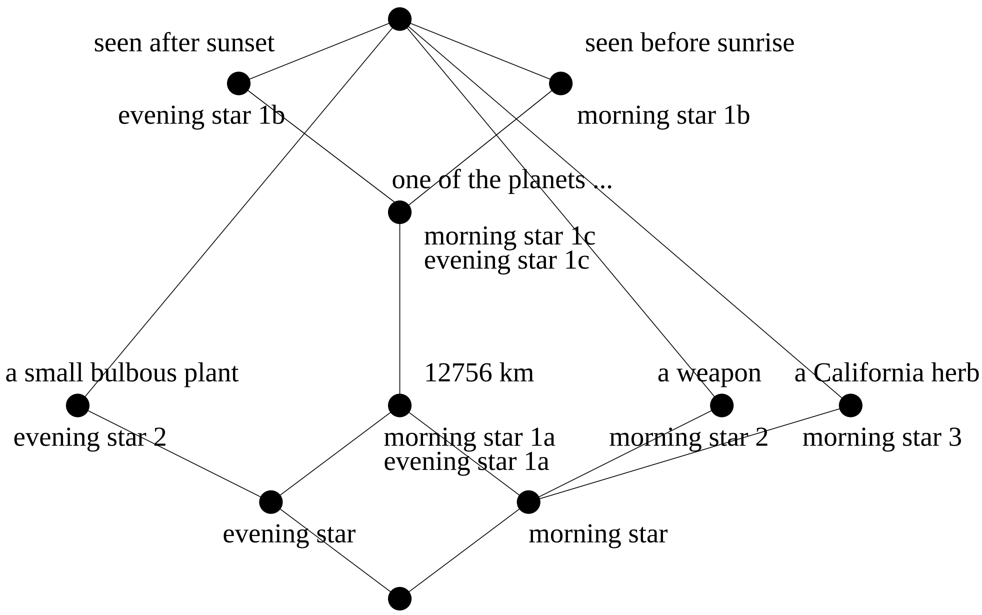

Figure 1.11: A morpho-lexical lattice of a context $(W \cup V, A_D, I_{LD} \cup I_{MLD})$

Figure 1.11 presents another attempt to show the polysemy of ‘morning star’ and ‘evening star’. This time as a morpho-lexical denotative lattice of a formal context $(W \cup V, A_D, I_{LD} \cup I_{MLD})$. It should be obvious by now that many different kinds of linguistic contexts and lattices are possible. The terminology could be extended in many ways. For example, morphological contexts of the form (F, A_F, I_F) can be studied. Phonemes, graphemes and morphemes could be included in the terminology and formal contexts could be defined on them. We think that, for example, phonemes and their features lead to interesting lattices that enable phonological comparisons since phonological data are often already presented as cross-tables. An example for such a lattice of German vowels can be found in Wille (1984), but to our knowledge no further research using Formal Concept Analysis in that area has been achieved. Besides phonology, componential or seme analyses of lexical fields is another application where the data are often presented in cross-tables of disambiguated words and their semantic features. In our terminology they can be interpreted as lexical denotative or connotative contexts and their lattices could be studied. An example (to our knowledge the only example so far) for the application of Formal Concept Analysis to componential analysis is Kipke & Wille (1987). Many other applications are possible.

1.9 Non-lexicalized concepts

Denotative or lexical denotative lattices can be used to visualize how the denotative word concepts are distributed among the denotative concepts. If several languages are to be compared, denotative lattices can be drawn for each language. The lattices visualize how the languages differ in their conceptual systems. As mentioned before, concepts are non-lexicalized if they are denotative concepts but not denotative word concepts. Rahmstorf (1991) defines non-lexicalized expressions as expressions which can be derived from formulas or rules whereas lexicalized expressions cannot be derived from formulas. In a lexical denotative lattice this corresponds in our terminology to the fact that non-lexicalized concepts can be constructed as meets or joins from denotative word concepts in the lattice (since all object concepts in a lexical denotative lattice correspond to denotative word concepts). For a denotative lattice which is, for example, constructed using the scientific knowledge of a scientific sub-language and the set W of disambiguated words of the everyday language, it does not have to be true that every non-lexicalized concept can be constructed as a meet or join of denotative word concepts. But Rahmstorf's definition should probably always hold for lattices that do not combine the background knowledge and vocabulary of different languages or sub-languages.

Since it is difficult to distinguish lexical gaps (for example 'not being thirsty') from complex concepts (for example 'tall woman') by a formal definition both are called *non-lexicalized concepts* for the purpose of this paper. Non-lexicalized concepts can be divided into three different kinds:

First, there are *intra-language non-lexicalized concepts* that are denotative concepts of a language but not denotative word concepts (i.e. the set $C \setminus C(W)$). This means that the denotative knowledge of a language permits the construction of concepts (in English for example 'not being hungry' and 'bald-headed man') which can be expressed in words but do not appear as an entry in the lexical database.

Second, there are *inter-language non-lexicalized concepts* that are denotative concepts of one language L_1, but not denotative concepts of another language L_2 (i.e. the set $C^{L_1} \setminus C^{L_2}$). These concepts are difficult to translate because even concepts for a description might not exist in the target language. For example, the German word 'Kitsch' cannot be translated without a long explanation on style and taste. The American word 'blues' cannot be translated into other languages without background knowledge of the American history.

Third, there are *multi-inter-language non-lexicalized concepts* that can be expressed using terms of several languages, but not of one single language (i.e. the set $C^{L_1 \cup \ldots \cup L_n} \setminus C^{L_1} \cup \ldots \cup C^{L_n}$). This third kind of non-lexicalized concepts is rather theoretical. Concepts, such as 'kitschy blues' (whatever that means) can only be understood by people who have enough knowledge of German and American denotations. These concepts occur if denotative (or lexical) contexts of several languages are combined and the common conceptual structure is computed.

1.10 The formalization of a dictionary, such as W3

Webster's Third New International Dictionary (W3) is used as an example of a traditional dictionary in this paper. It can be modeled as a lexical structure

$$\mathcal{S}_{W3} := (W, M(W), Fn, V, F; form, mng, Hmg, Pls, n^-, wrd)$$

with the following conditions:
A word v is a headword of a paragraph since homographs are not contained in the same paragraph. Meanings of words (in the set $M(V)$) are missing because each gloss belongs to a disambiguated word. A disambiguated word w is given by its form $form(w)$ which consists of the form $form(wrd(w))$ of the basic word and a sense number. The meaning $mng(w)$ of a disambiguated word is indicated by a gloss and sometimes further information in brackets, pictures, and others. For example, information on connotations can be deduced from comments such as 'dialect' or 'slang' and on denotations from the pictures. Semantic relations are implicitly contained in the glosses (Calzolari (1988) extracts semantic relations from such a traditional dictionary to create a thesaurus). But they need not be systematical and cannot be accessed without parsing software. Even a search for synonyms is almost impossible if the synonyms are not explicitly mentioned in the glosses. The Figures 1.3, 1.6, 1.7, 1.9, and 1.11 show different modelings of the example of 'morning star' and 'evening star' in W3 to indicate how implicit structures in dictionaries can be visualized. The formal objects in these figures are denotata of 'morning star' and 'evening star' or the words or disambiguated words themselves. Such formal objects could be selected automatically according to some criteria (such as words from lexical fields, the nouns of a paragraph of a text, or others). Unfortunately, the examples showed that it is not very useful to take the glosses as formal attributes without modification. The glosses do not distinguish properly between denotative and connotative aspects. They are designed to be interpreted by a natural language speaker and not by a machine. In all the examples, the formal attributes where derived by extracting information from the glosses and adding general background knowledge where necessary. Therefore it seems that most types of linguistic contexts cannot automatically be derived from dictionaries, such as W3. It is possible that linguistic contexts consisting of information on phonology or etymology could be automatically derived from such dictionaries, because they are based on word forms or words, but that is not investigated in this paper.

1.11 The formalization of a thesaurus, such as RIT

Roget's International Thesaurus (RIT) has been studied by Sedelow & Sedelow (for example, 1974 and 1986) for more than 20 years. RIT consists of two parts: a classification of categories and an alphabetical index. Words are disambiguated by their grouping together with other words of similar meanings into so-called *semicolon-groups*, the sixth and lowest level of the hierarchy. Semicolon-groups are gathered in paragraphs, the fifth level of the hierarchy. The higher levels are categories or level-4-classes, level-3-classes (indicated by letters), level-2-classes (indicated by Roman numerals), and level-1-classes (indicated by Arabic numerals). For example, one particular meaning of 'summer' is classified as '1.VI.A.105.6.4' where '1' is the class 'abstract relations', '1.VI' is 'time', '1.VI.A' is 'absolute time', and '105' is 'time'. The paragraph and semicolon-group levels do not have names. The other categories of 'summer' are '128 season', '216 shaft', and '327 heat'. Several formalizations have been developed. Bryan ((1973) and (1974)) defines a thesaurus as a triple consisting of entities (disambiguated words in our terminology), word forms, and classes. Since homographs are distinguished in the index, but not in the first part of the RIT and only the first part is contained in the lexical database that the Sedelows use, Bryan's major effort was to distinguish homographs from polysemous words using similarity in meaning. Bryan defines different chains consisting of alternating word forms and semicolon-groups so that each pair '(word form, semicolon-group)' corresponds to an entity in RIT. The strongest chains, so-called 'type-10-chains', form a partition of the entities of RIT. According to Bryan, homographs are entities which occur in separate components of that partition. Talburt & Mooney (1990) investigate the components in more detail and discover that a word often occurs in as many components of RIT as the Oxford English Dictionary has different polysemous senses for it. This questions Bryan's definition of homography. Another interesting modeling of RIT has been achieved by Knuth (1993) who draws graphs of RIT using the cross-references instead of the hierarchy.

A formalization of RIT based on Formal Concept Analysis can be found in Priss & Wille (in preparation). It uses the 'plus' mappings ι^+ and ε^+ (see Definition 1.1): $\iota^+(G_1) := \{m \in M \mid \exists_{g \in G_1} : gIm\}$, which yields for a set G_1 of objects in a context (G, M, I) all attributes that belong to at least one of the objects, and $\varepsilon^+(M_1) := \{g \in G \mid \exists_{m \in M_1} : gIm\}$ for a set M_1 of attributes. If two plus mappings are applied to a set G_1 it results in a set $\varepsilon^+\iota^+(G_1)$ (with $\varepsilon^+\iota^+(G_1) \supseteq G_1$) which is called the *neighborhood* of G_1 under I. A neighborhood of attributes is defined analogously. Wunderlich (1980) mentions neighborhoods (without giving them this name) used with bilingual dictionaries: a neighborhood of a German word is obtained by finding all the English translations of that word, and then all the German translations of the English translations. In general, neighborhoods can be used to select a set of related items from a set that would be too large to oversee. In RIT the neighborhood of a word is used as set of formal objects of a so-called *neighborhood context* $(\varepsilon^+\iota^+(G_1), \iota^+\varepsilon^+\iota^+(G_1), I)$.

In this paper RIT is modeled as a lexical structure

$$\mathcal{S}_{RIT} := (W, M(W), V, F; form, mng, Hmg, Pls, wrd)$$

which is combined with a denotative context

$$\mathcal{K}_{RIT} := (S_C(W), C_{RIT}, \leq)$$

with the following conditions:
A word v is a headword of a paragraph in the index. Since the index is missing in the electronic versions of RIT that are used by the Sedelows and in Darmstadt, in the electronic version homography cannot easily be detected and each word v is in one-to-one correspondence to its form $form(v)$, but hopefully that will be improved in the future. Meanings of words (the set $M(V)$) are missing. A semicolon-group is interpreted as a synset in the denotative context $\mathcal{K}_{RIT}$, i.e. two disambiguated words w_1 and w_2 are synonyms if they occur in the same semicolon-group (or synset) $syn(w_1) = syn(w_2)$ because they denote the same denotative concept in $\mathcal{K}_{RIT}$. A disambiguated word w is therefore given by its form $form(wrd(w))$ and the synset $syn(w)$ in which it occurs. Sense numbers are not denoted but could be generated using the ordering in the index. The meaning $mng(w)$ of a word is according to Definition 1.2 the particular meaning that a native speaker of English recognizes after considering the other words in the same synset and the hierarchy above it. Some information on connotation is explicitly represented in brackets behind the disambiguated word (such as, 'dialect', language origin, author of a quotation). The denotative context $\mathcal{K}_{RIT}$ has the synsets as formal objects and the classes of the RIT hierarchy as formal attributes. Since the synsets as semicolon-groups are level-6-classes and therefore classes themselves, the relation between formal objects and attributes is the conceptual ordering $\leq$. Except of the bottom concept, which has no formal objects in its extent, the concept lattice of $\mathcal{K}_{RIT}$ forms a tree. The atoms of this lattice, which are concepts of the form $(\{syn(w)\}, \{syn(w)\})$, can be interpreted as denotative word concepts, the other concepts are non-lexicalized because they are not represented by synsets. As semantic relations, in RIT only the conceptual ordering $\leq$ and synonymy are presented. In earlier versions (RIT1, 1852) there has also been antonymy. Some other semantic relations, such as meronymy occur un-systematically and implicitly.

Besides the context $\mathcal{K}_{RIT}$ the polysemy of the words can be modeled in a context $(V, S_C(W), HYP_V^*)$[8] with relation HYP_V^* defined as $vHYP_V^*s \iff \exists_{w \in W} : (wrd(w) = v \text{ and } syn(w) = s)$. The relation can be interpreted as 'has polysemous meaning'. Since the synsets correspond to denotative word concepts, the context $(V, S_C(W), HYP_V^*)$ is a morpho-lexical context of the form (V, A_D, I_{MLD}) (compare Figure 1.10). The contexts $\mathcal{K}_{RIT}$ and $(V, S_C(W), HYP_V^*)$ can be composed to one context in different ways, for example, $(V, S_C(W), HYP_V^*) \circ (S_C(W), C_{RIT}, \leq) := (V, C_{RIT}, HYP_V^* \circ \leq)$ with $v(HYP_V^* \circ \leq)c :\iff \exists_{s \in S_C(W)} : (vHYP_V^*s \text{ and } s \leq c)$. Neighborhood contexts and lattices are used to display the relationship between a word and other words that are semantically related to the first word by sharing at

[8]Or $(F, S_C(W), HYP_F^*)$ in the electronic versions of RIT.

least one synset with it. They are essential because the complete contexts, $\mathcal{K}_{RIT}$ and $(V, S_C(W), HYP_V^*)$, are too large to be comprehensible and therefore subcontexts have to be selected. Figure 1.12 shows a line diagram of the neighborhood lattice of 'summer' in RIT, i.e. $(\varepsilon^+\iota^+(\{summer\}), \iota^+\varepsilon^+\iota^+(\{summer\}), HYP_V^*)$ and Figure 1.13 shows a line diagram of the same lattice composed with the three lower levels of the RIT hierarchy. For more details on context constructions in RIT see Priss & Wille (in preparation).

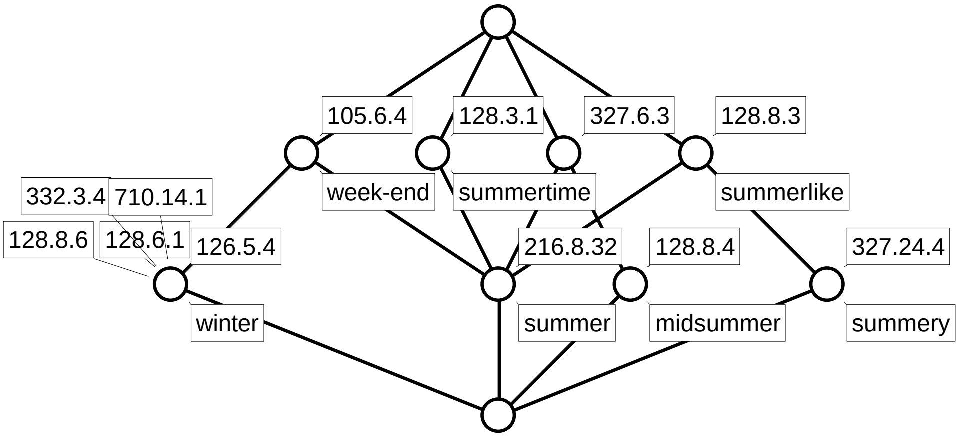

Figure 1.12: A neighborhood lattice of 'summer' in RIT

1.12 The formalization of a lexical database, such as WordNet

WordNet is a lexical database of the English language developed by the Cognitive Science Lab under George Miller (Miller et al., 1990). It was started as a model of the mental lexicon but is actually developed as a tool for computational linguistics. Words are disambiguated by collecting them in synsets. Lexical relations hold among the disambiguated words, and semantic relations hold among the synsets. WordNet can be formalized as a lexical structure

$$\mathcal{S}_{WN} := (W, M(W), V, F; form, mng, Pls, n^-, wrd, \mathcal{R}_W)$$

which contains the lexical denotative context

$$\mathcal{K}_{WN} := (W, S_C(W), HYP_W^*)$$

a family $\mathcal{R}_{S_C(W)}$ of semantic relations, and a family $\mathcal{R}_W$ of lexical relations with the following conditions:
The meaning $mng(w)$ of a disambiguated word w is indicated by its relations to other words, synsets and additional glosses. The sense numbers exist in the non-public versions of WordNet but are not displayed in the public versions. A disambiguated word w is therefore uniquely described by the form $form(wrd(v))$ of its basic word, its

part of speech and its synset $syn(w)$. Homography is not explicitly marked. The context $\mathcal{K}_{WN}$ (compare section 1.8) is the Dedekind closure of the ordered set of the hyponymy ordering that is implemented in WordNet. The semantic relations are specific to parts of speech: synonymy and antonymy are defined for all parts of speech. Among nouns hyponymy, meronymy, and coordination, among verbs troponymy, cause, and entailment are implemented (compare Chapter 3). Figure 1.8 shows an example of the WordNet hyponymy ordering and Chapter 2 illustrates several examples of the meronymy relation.

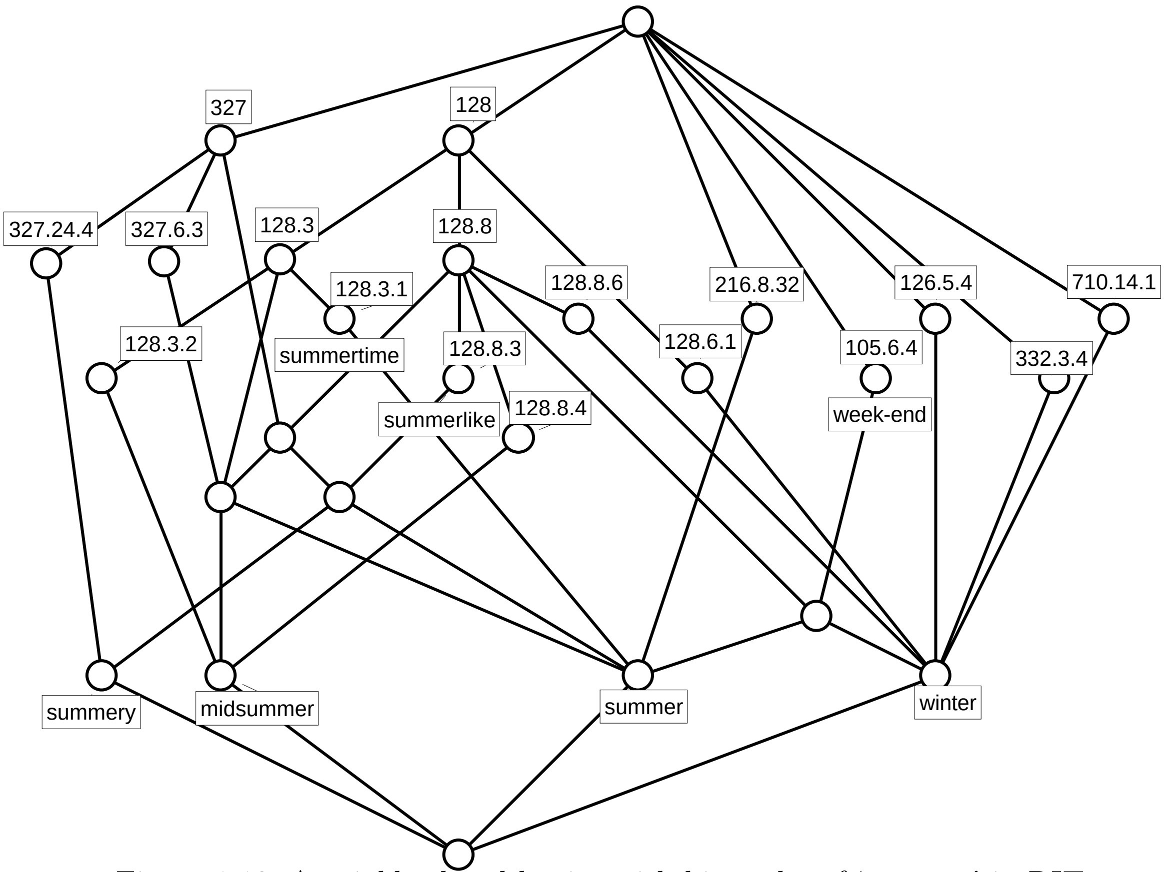

Figure 1.13: A neighborhood lattice with hierarchy of 'summer' in RIT

2 Relational Concept Analysis

Relational Concept Analysis is the extension of Formal Concept Analysis—which provides a conceptual hierarchy—to a more general theory that includes other relations among objects, attributes or concepts. Some of the related questions are: if a relation is given on objects (or attributes) can this relation be extended to a relation among concepts (a concept relation)? If a relation is defined on objects is there a relation on attributes which leads to the same concept relation? Which structures evolve from concept relations? Can relations be inherited from superconcepts to subconcepts, and, if they can, is there a unique basis for a relation? Can concept relations be studied by considering only object (or attribute) concepts? How do properties of relations on objects affect concept relations? How can these relations be effectively visualized? Basic answers to these questions are derived in this chapter for the restriction to binary relations. Unary relations (predicates) are discussed in Section 4.2. A generalization to other relations is yet to be achieved.

The main problem of extending relations among objects to concept relations is that of quantification. It is always necessary to examine whether a concept relation holds for all objects and attributes of the extent and intent of a concept or whether it holds only for a subset. For example, does 'birds fly' mean that all objects of the extent of 'bird' stand in 'ability'-relation to 'fly'? Most natural language statements seem to have an implicit quantification, especially sentences such as 'Women like shopping.' Is the interpretation of this sentence 'All women like some shopping,' (they do not like to shop in hardware stores, therefore they do not like 'all' shopping) or 'All prototypical women like some prototypical shopping,' or 'Prototypically, all women like some shopping'? Such quantifications of verbs are studied by Woods (1991) who separates a quantificational tag of a relation from its relational component and investigates resulting subsumptions. Although he distinguishes between relations among objects (instances) and among concepts, his modeling is not as detailed as the modeling based on Formal Concept Analysis. He uses an equivalence similar to (1) in our Definition 2.1. The detailed analysis of semantic relations by Lyons (1977) is extended by Cruse (1986) who classifies the meronymy relation by considering quantifications. Besides computer scientists (Woods) and linguists (Lyons and Cruse), also mathematicians and philosophers are interested in subjects related to Relational Concept Analysis. A mathematician, Brink (1993), studies relations among sets which are generalized to relations among power sets. Relational Concept Analysis can be interpreted as a special case of his research using extents of concepts instead of arbitrary sets. On the other hand, Relational Concept Analysis is more general since more quantifiers are considered. From a philosophical viewpoint natural language phenomena are a basis for logical theories. Westerstahl (1989), summarizes research on natural language quantifiers. We use some of his rules for the conversion among quantifiers in Section 2.2. Since Woods' paper lacks the details of conceptual modeling which Formal Concept Analysis provides, since Cruse and Lyons do not formalize their ideas and Brink's paper is more general and more specific in some ways, it seems that Relational Concept Analysis is an advancement of its preceding theories.

2.1 The basic definitions

To begin with, all natural language quantifiers are taken into consideration. They are denoted by their word form delimited by two vertical lines, for example $||\text{all}||$. Some can be abbreviated using mathematical notation, such as $||\text{at least } 1|| =: ||{\geq}\,1||$ and $||\text{exactly } 1|| =: ||1||$ (for more details on natural language quantifiers see Westerstahl (1989)). It turns out (see Section 2.2) that only very few of the quantifiers are actually needed for concept relations, but we think it is useful to provide a sufficiently general basic terminology. In what follows, only binary relations among objects (i.e., $r \subseteq G \times G$) are considered. Relations among attributes (i.e., $r \subseteq M \times M$) can be treated analogously (compare Definition 3.4). These relations are transferred to concept relations, i.e., $R \subseteq \mathcal{B}(G, M, I) \times \mathcal{B}(G, M, I)$, according to the following definitions.

Definition 2.1:
For a context (G, M, I), concepts $c_1, c_2 \in \mathcal{B}(G, M, I)$, a relation $r \subseteq G \times G$, and quantifiers Q^i, $1 \leq i \leq 4$, we define

$$c_1 \; R^r[Q^1, Q^2;] \; c_2 :\Longleftrightarrow Q^1_{g_1 \in Ext(c_1)} Q^2_{g_2 \in Ext(c_2)} : g_1 r g_2 \tag{1}$$

$$c_1 \; R^r[; Q^3, Q^4] \; c_2 :\Longleftrightarrow Q^3_{g_2 \in Ext(c_2)} Q^4_{g_1 \in Ext(c_1)} : g_1 r g_2 \tag{2}$$

$$c_1 \; R^r[Q^1, Q^2; Q^3, Q^4] \; c_2 :\Longleftrightarrow c_1 \; R^r[Q^1, Q^2;] \; c_2 \text{ and } c_1 \; R^r[; Q^3, Q^4] \; c_2 \tag{3}$$

r is called the *relational component* and $[Q^1, Q^2;]$, $[; Q^3, Q^4]$, or $[Q^1, Q^2; Q^3, Q^4]$ are called the *quantificational tag* of a relation. If no ambiguities are possible, relational component and quantificational tag can be omitted in the notation of the relation.

Depending on the quantifiers each relation r therefore leads to several different relations R^r among concepts. The terms 'quantificational tag' and 'relational component' are taken from Woods' terminology. The formalization can best be understood through an example: 'all door-handles are parts of doors' states a meronymy relation between door-handles and doors. More precisely it means that all objects that belong to the extent of the concept 'door-handle' have an object in the extent of the concept 'door' so that the meronymy relation holds between them. The variables in equivalence (1) are for this example $Q^1 := ||\text{all}||$, $Q^2 := ||{\geq}\,1||$, c_1 is the concept 'door-handle', c_2 is the concept 'door', and r is the relation 'is part of'. Equivalence (2) could be 'there is at least one door which has a handle', because 'all doors have to have handles' is not true. Equivalence (3) is the conjunction of the first two. For the door-handle example the quantifiers are $Q^1 := ||\text{all}||$, $Q^2 := ||{\geq}\,1||$, $Q^3 := ||{\geq}\,1||$ and $Q^4 := ||{\geq}\,1||$. Abbreviations are used for the more frequently used types of relations:

Definition 2.2:
'$R^r[||\geq 1||, ||\geq 1||; ||\geq 1||, ||\geq 1||]$' is abbreviated as R^r_0 and '$R^r[||\text{all}||, Q^2; ||\text{all}||, Q^4]$' is abbreviated as $R^r_{(Q^4;Q^2)}$. The vertical lines '$||$' can be left out for Q^4 and Q^2 in the subscript of $R^r_{(Q^4;Q^2)}$[9].

Since $Q^1 = Q^2 = ||\geq 1||$ is equivalent to $Q^3 = Q^4 = ||\geq 1||$, we have $R^r[||\geq 1||, ||\geq 1||;]$ $= R^r[; ||\geq 1||, ||\geq 1||] = R^r_0$. This is the minimal relation where at least one pair of objects is in relation r to each other, because if Q^1 or Q^2 equals $||\geq 0||$, then it is not known whether there is a single pair of objects in relation r at all. Similarly, $R^r[||\text{all}||, ||\text{all}||;] = R^r[; ||\text{all}||, ||\text{all}||] = R^r[||\text{all}||, ||\text{all}||; ||\text{all}||, ||\text{all}||]$. It should be noted that not all arbitrary combinations of quantifiers can be chosen because some would lead to empty relations. For example, $R^r[||\text{all}||, ||n||; ||\text{all}||, ||\text{all}||]$ is empty if $Ext(c_2) \neq n$. And there are implications among the quantifiers, for example for $n \geq 1$, from $R^r[||\text{all}||, ||\geq n||;]$ follows $R^r[||\text{all}||, ||\geq (n-1)||;]$ from which follows $R^r[||\text{ all}||, ||\geq 1||;]$. From $R^r[||\geq n||, ||\text{all}||;]$ follows $R^r[; ||\text{all}||, ||\geq n||]$, and so on. In many applications Q^1 and Q^3 equal $||\text{all}||$, therefore the abbreviation $R^r_{(Q^4;Q^2)}$ is useful.[10] The relation $R^r_{(\geq 1;\geq 1)}$ seems to be the most important one for many applications. Brink (1993) calls it the 'power relation' in his modeling of relations among power sets.

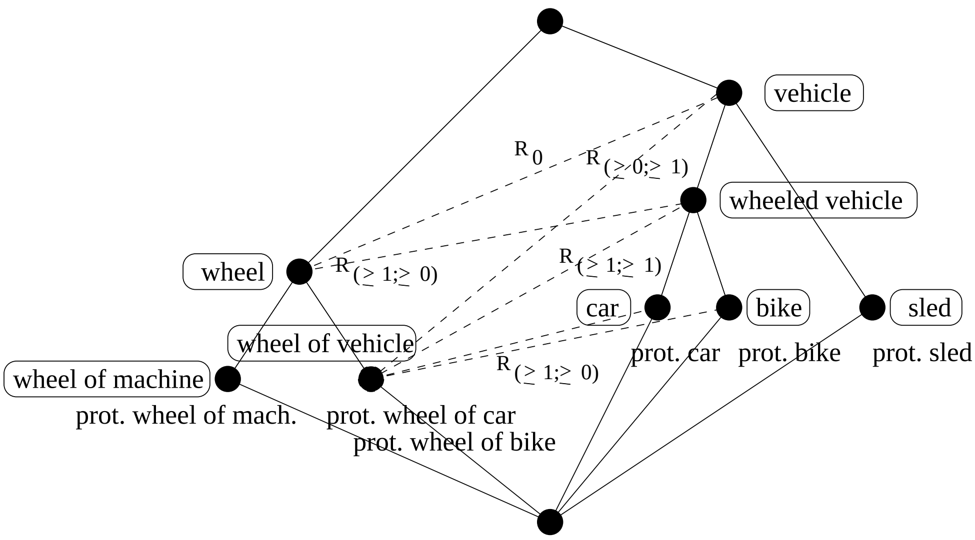

Figure 2.1: Different meronymy relations

Figure 2.1 shows a denotative lattice whose formal objects are prototypes and whose formal attributes are omitted in the diagram. Some of the concepts are denoted by disambiguated words which are surrounded by ellipses in the picture. A relation r is

[9]The inversion of the quantifiers in $(Q^4; Q^2)$ is in analogy to the labeling of relations with 'one-to-many', 'many-to-many', and so on in the Entity-Relationship Model (see Section 4.4).

[10]Besides its applications to the modeling of lexical databases, the formalization can also be used in other disciplines to describe functions $R^r_{(\geq 0;1)}$, bijections $R^r_{(1;1)}$, or Cartesian products $R^r_{(\text{all};\text{all})}$.

defined as a part-whole relation between prototypical wheel of vehicle and prototypical car and bike. The dotted lines represent conceptual meronymy relations which are derived from r. Similarly to the subconcept-superconcept relation, the dotted lines denote a directional relation where the part is below the whole. But, for example, if writing the part below the whole would contradict the conceptual ordering, arrows can be added to the dotted lines to indicate the direction. As an example of how to read the relations in the diagram, the relation $R_{(\geq 1;\geq 0)}$ between 'wheel' and 'wheeled vehicle' means that wheels can be parts of wheeled vehicles, but each wheeled vehicle has at least one wheel. Only between 'wheel of vehicle' and 'wheeled vehicle' does the stronger relation $R_{(\geq 1;\geq 1)}$ hold; wheels of vehicles must be parts of wheeled vehicles and wheeled vehicles must have wheels. Both concepts are lexicalized in the picture, but seem to be non-lexicalized in most dictionaries. As Section 2.3 shows, it is a problem to implement semantic relations in lexical databases without redundancy if the concepts among which the stronger relations hold are not lexicalized. Figure 2.1 demonstrates that relations are sometimes inherited[11] between sub- and superconcepts. This is examined in more detail in Section 2.3.

2.2 Characteristics of concept relations

As it is either possible to define a relation on objects, or on attributes, or on concepts themselves, it is useful to develop a characterization of a concept relation R by considering its structure on the concepts only and ignoring the relational component r. Then investigations can be made to determine which relations on objects or attributes lead to the same concept relation, or, if the concept relation is given, which could be the relational components for it. For simplification this is only demonstrated for relations according to equivalence (1) which lead to the definition of relations of characteristic $[Q^5, Q^6;]$. Definitions and theorems hold similarly for equivalence (2) (relations of characteristic $[;Q^5, Q^6]$). There is no characteristic defined for relations according to equivalence (3), because such relations are characterized by stating whether they have a characteristic $[Q^5, Q^6;]$ and a characteristic $[;Q^5, Q^6]$. A relation $R^r[Q^1, Q^2;]$ either fulfills the following equivalence using two quantifiers Q^5 and Q^6 for all concepts in $\mathcal{B}(G, M, I)$ or there exist no quantifiers Q^5 and Q^6 so that they fulfill the equivalence. Although the characterization is defined for any concept relation independent of r and $[Q^1, Q^2;]$, the rest of this section uses this characterization only to classify certain relations $R^r[Q^1, Q^2;]$ into four characteristics.

Definition 2.3:
A relation $R \subseteq \mathcal{B}(G, M, I) \times \mathcal{B}(G, M, I)$ for which there exist quantifiers Q^5, Q^6 so that for all $c_1, c_2 \in \mathcal{B}(G, M, I)$

$$c_1 \; R \; c_2 \iff Q^5_{c_{11} \leq c_1} Q^6_{c_{21} \leq c_2} : c_{11} \; R \; c_{21} \tag{4}$$

[11] Generalization from subconcepts to superconcepts is also called 'inheritance' for the rest of this paper.

holds is called *of characteristic* $[Q^5, Q^6;]$. Relations of characteristic $[;Q^5, Q^6]$ are defined analogously.

Relations can have several characteristics because some quantifiers entail other quantifiers and therefore several quantifiers can fulfill (4). If there are no quantifiers Q^5 and Q^6 for a relation R that fulfill (4), the relation does not have a characteristic. Relations without characteristics usually do not have simple rules for their inheritance between subconcepts and superconcepts. For example, ||half|| cannot in general be inherited because it is not possible to predict the differences in number of objects between a concept and its sub- or superconcept. A large number of natural language quantifiers is of the kind ||at least all but one||, ||at least one||, ||less than five|| and so on. Figure 2.2 shows the formal expressions (with natural number $n > 0$) of these quantifiers (see Westerstahl (1989), pp. 69-71). The quantifiers on the opposite ends of the lines are logical negations of each other (Q and $\neg Q$; $Q^d := \neg Q\neg$ and $Q\neg$), such as 'every' and 'not every', 'more or equal one' and 'less than one' in the case $n = 1$. The quantifiers on the same horizontal level are inner negations of each other, i.e. the expressions behind the quantifier are negations of each other, such as 'every man snores' versus 'every man does not snore', which equals 'no man snores'. The quantifiers on the same vertical level are dual to each other. The basic rules for *conversion* among quantifiers follow from $Q^d = \neg Q\neg = (\neg Q)\neg = \neg(Q\neg)$ and from the fact that double negation means no negation at all, i.e. $\neg\neg Q = Q = Q\neg\neg$ (see Westerstahl (1986)).

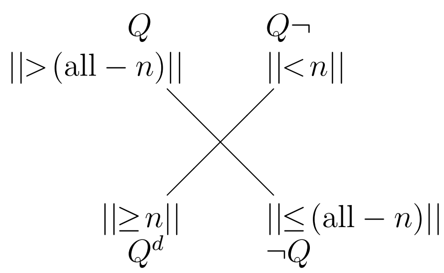

Figure 2.2: A quantifier, its dual, negation and inner negation

The following theorem demonstrates that although different quantifiers can be used on the object level, on the conceptual level basically only the ||all||- and $||\geq 1||$-quantifiers are needed. Therefore exactly four characteristics of relations $[Q^5, Q^6;]$ and analogously four characteristics of relations $[;Q^5, Q^6]$ exist if quantifiers according to Figure 2.2 are involved.

Theorem 2.1:
The relations $R^r[Q^1, Q^2;]$ with quantifiers Q^1, Q^2 of the form $|| > (\text{all} - n)||$, $|| \leq (\text{all} - n)||$, $|| > m||$, or $|| \leq m||$ (n and m are natural numbers larger 0) are of four characteristics according to Table 2.1.

Q^1	Q^2	Q^5	Q^6
$\|\|>(\text{all}-n)\|\|$	$\|\|>(\text{all}-m)\|\|$	$\|\|\text{all}\|\|$	$\|\|\text{all}\|\|$
$\|\|>(\text{all}-n)\|\|$	$\|\|\geq m\|\|$	$\|\|\text{all}\|\|$	$\|\|\geq 1\|\|$
$\|\|\geq n\|\|$	$\|\|>(\text{all}-m)\|\|$	$\|\|\geq 1\|\|$	$\|\|\text{all}\|\|$
$\|\|\geq n\|\|$	$\|\|\geq m\|\|$	$\|\|\geq 1\|\|$	$\|\|\geq 1\|\|$
$\|\|>(\text{all}-n)\|\|$	$\|\|<m\|\|$	$\|\|\text{all}\|\|$	$\|\|\text{all}\|\|$
$\|\|>(\text{all}-n)\|\|$	$\|\|\leq(\text{all}-m)\|\|$	$\|\|\text{all}\|\|$	$\|\|\geq 1\|\|$
$\|\|\geq n\|\|$	$\|\|<m\|\|$	$\|\|\geq 1\|\|$	$\|\|\text{all}\|\|$
$\|\|\geq n\|\|$	$\|\|\leq(\text{all}-m)\|\|$	$\|\|\geq 1\|\|$	$\|\|\geq 1\|\|$
$\|\|<n\|\|$	$\|\|\leq(\text{all}-m)\|\|$	$\|\|\text{all}\|\|$	$\|\|\text{all}\|\|$
$\|\|<n\|\|$	$\|\|<m\|\|$	$\|\|\text{all}\|\|$	$\|\|\geq 1\|\|$
$\|\|\leq(\text{all}-n)\|\|$	$\|\|\leq(\text{all}-m)\|\|$	$\|\|\geq 1\|\|$	$\|\|\text{all}\|\|$
$\|\|\leq(\text{all}-n)\|\|$	$\|\|<m\|\|$	$\|\|\geq 1\|\|$	$\|\|\geq 1\|\|$
$\|\|<n\|\|$	$\|\|\geq m\|\|$	$\|\|\text{all}\|\|$	$\|\|\text{all}\|\|$
$\|\|<n\|\|$	$\|\|>(\text{all}-m)\|\|$	$\|\|\text{all}\|\|$	$\|\|\geq 1\|\|$
$\|\|\leq(\text{all}-n)\|\|$	$\|\|\geq m\|\|$	$\|\|\geq 1\|\|$	$\|\|\text{all}\|\|$
$\|\|\leq(\text{all}-n)\|\|$	$\|\|>(\text{all}-m)\|\|$	$\|\|\geq 1\|\|$	$\|\|\geq 1\|\|$

Table 2.1: Characteristics of relations

The proof (see Proof 1[12]) uses the conversion of quantifiers in a way that only the first four rows of Table 2.1 have to be proved, the others follow from conversion. The first eight rows are even equal to the last eight rows. For example, 'all rabbits fear all snakes' (row 1 with $n = m = 1$) is equal to 'no rabbit fears less than all snakes' (row 9 with $n = m = 1$). An interpretation of Table 2.1 is that if a specific number occurs on the object level, for example, all hands have five fingers, it does not occur on the conceptual level. For a concept 'hand' there is one concept 'finger' so that each object of 'hand' has five parts among the objects of 'finger'; and not: for a concept 'hand' there exist five concepts 'finger' with that property. A linguistic example where this is even reflected in the language is that 'having two shoes' can also be expressed as 'having a pair of shoes'. Some other quantifications can be achieved by intersection of the relations, for example, the $||n||$-quantifier (as Q^1 or Q^2) results from a relation that fulfills the $|| \leq n||$-quantifier and the $|| \geq n||$-quantifier. It seems that these relations have two characteristics, a characteristic $[Q^5, Q^6;]$ for $||\leq n||$ and another characteristic $[Q^5, Q^6;]$ for $||\geq n||$, but not a single characteristic which entails these two characteristics in the way $||n||$ entails $||\leq n||$ and $||\geq n||$, because these relations cannot in general be inherited in the lattice. For example, the statement that each hand has exactly five parts in the concept 'finger' is not inherited to all hypernyms of 'finger' because in the hypernym 'body parts' the hand may have 'palm' as an additional part. It is also not inherited to hyponyms of 'finger' because 'hand' has only one part in the hyponym 'ring finger'. Relations should therefore be modeled using $||\geq n||$- and $||\leq n||$-quantifiers, for example, 'each hand has at least five parts', and 'each hand has less than or equal to five fingers'.

[12]The proofs of this chapter are in Section 2.10.

2.3 Bases of concept relations

An $||\text{all}||$-quantifier for Q^5 or Q^6 in equivalence (4) obviously causes inheritance of the relation to all subconcepts, whereas an $||\geq 1||$-quantifier causes inheritance to all superconcepts. For example, if all bird feathers are part of birds then all sparrow feathers (as special bird feathers) are parts of birds and all bird feathers are parts of animals (a generalization of bird). For an implementation in a lexical database it is desirable to deduce all the relations which follow from inheritance from a set of rules and a basis for each relation. The Theorem 2.2 (see Proof 2) shows that bases exist and are unique for the four characteristics of concept relations from Table 2.1. Relations of characteristic $[;Q^5, Q^6]$ can be treated analogously.

Definition 2.4:
A *basis* $\mathcal{R}$ of a relation R of characteristic $[||\text{all}||, ||\geq 1||;]$ is defined as a relation $\mathcal{R} \subseteq \mathcal{B}(G,M,I) \times \mathcal{B}(G,M,I)$ satisfying, for all $c_1, c_2 \in \mathcal{B}(G,M,I)$,

$$\text{a) } c_1 R c_2 \iff \exists_{(c_1^\circ, c_2^\circ) \in \mathcal{R}} : (c_1 \leq c_1^\circ \text{ and } c_2 \geq c_2^\circ) \tag{5}$$

and b) $\mathcal{R}$ has the minimal number of elements among all relations that fulfill equivalence (5). Bases of characteristics $[||\text{all}||, ||\text{all}||;]$, $[||\geq 1||, ||\geq 1||;]$ and $[||\geq 1||, ||\text{all}||;]$ are defined respectively using ($c_1 \leq c_1^\circ$ and $c_2 \leq c_2^\circ$), ($c_1 \geq c_1^\circ$ and $c_2 \geq c_2^\circ$), and ($c_1 \geq c_1^\circ$ and $c_2 \leq c_2^\circ$).

Theorem 2.2:
Bases as defined in Definition 2.4 are unique. A relation $R \subseteq \mathcal{B}(G,M,I) \times \mathcal{B}(G,M,I)$ is of characteristic $[||\text{all}||, ||\geq 1||;]$, $[||\text{all}||, ||\text{all}||;]$, $[||\geq 1||, ||\geq 1||;]$ or $[||\geq 1||, ||\text{all}||;]$ if and only if it has a basis according to Definition 2.4.

Equivalence (5) uses the fact that $||\text{all}||$-quantifiers cause inheritance to subconcepts (therefore $c_1 \leq c_1^\circ$) and $||\geq 1||$-quantifiers cause inheritance to superconcepts (therefore $c_2 \geq c_2^\circ$). A basis can thus be used to define a characteristic of a concept relation which is defined on the concepts and not derived from a relation among objects or attributes. Theorem 2.3 (see Proof 3) shows that for all concept relations (of certain characteristics and not defined on objects or attributes) an equivalent context with an equivalent concept relation exists, i.e. a context whose reduced version is isomorphic to the reduced version of the first context (Ganter & Wille, 1996) and whose concept relation holds among the corresponding concepts, so that its concept relation is derived from a relation on objects or attributes. That means that if for some reason a concept relation is given in a line diagram, objects and attributes can (so to say) be added to or deleted from the diagram so that the concept relation can be derived from relations among objects or attributes. If it is impossible to add or delete objects or attributes and still keep the same line diagram, then concept relations cannot always be derived from relations among objects or attributes (if the common characteristics of relations according to Table 2.1 are considered). It follows that if a concept relation is derived from a relation on objects there does not have to exist a relation on attributes which leads to the same concept relation.

Theorem 2.3:
A concept relation of a characteristic according to Table 2.1 cannot always be derived from a relation among objects or attributes. But for each concept lattice on which a concept relation R_1 of characteristic $[||\text{all}||, ||\geq 1||;]$, $[||\text{all}||, ||\text{all}||;]$, $[||\geq 1||, ||\geq 1||;]$ or $[||\geq 1||, ||\text{all}||;]$ is defined there exists an equivalent context with an equivalent concept relation (i.e. with order-preserving isomorphism $ism : \mathcal{B}(G_1, M_1, I_1) \rightarrow \mathcal{B}(G_2, M_2, I_2)$ between the concepts satisfying $c_1 R_1 c_2 \Longleftrightarrow ism(c_1) R_2 ism(c_2)$) such that the concept relation R_2 can be derived from a relation r on objects or attributes according to Definition 2.1.

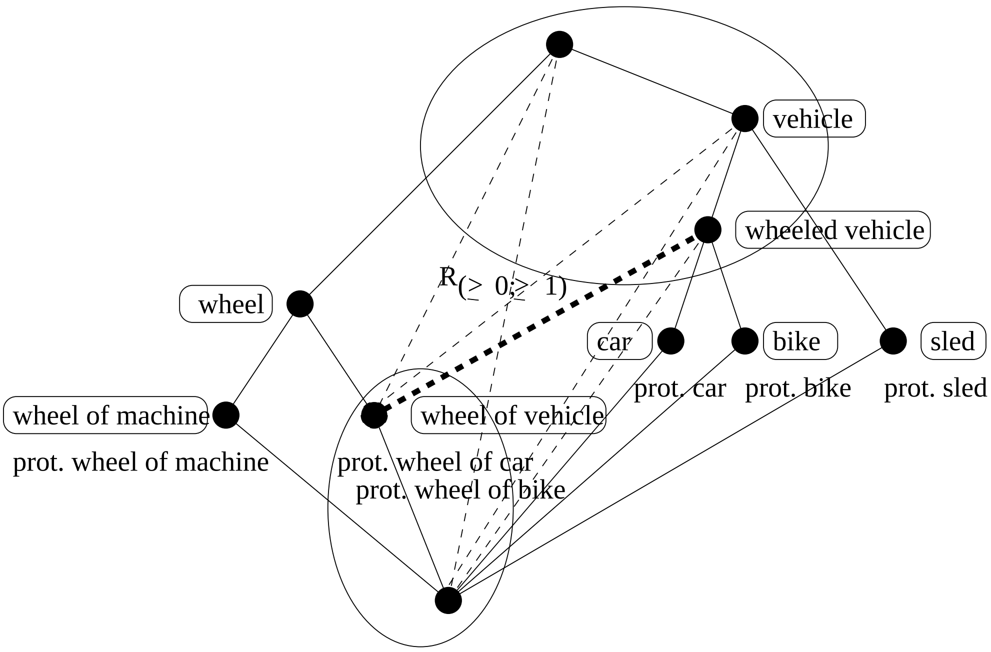

Figure 2.3: A basis for meronymy relations

Figure 2.3 shows the same example as Figure 2.1. The boldface dotted line is the basis for the relations $R_{(\geq 0;\geq 1)}$, which is represented by the dotted lines in the figure, and $R_{(\geq 1;\geq 0)}$, which is not shown in the figure. There is no basis for a relation $R_{(\geq 1;\geq 1)}$ which always must be represented as an intersection of relations $R_{(\geq 0;\geq 1)}$ and $R_{(\geq 1;\geq 0)}$. All the concepts in the lower ellipsis are in relation $R_{(\geq 0;\geq 1)}$ to all concepts in the upper ellipsis. 'Wheel of vehicle' is the most general part of 'wheeled vehicle' which is itself the most specific whole of 'wheel of vehicle'. A basis can consist of several elements, which means a concept can have several most general parts or most specific wholes. For example, 'wheeled vehicle' could also have 'engine of wheeled vehicle' as another most general part. Of course, then there could be a concept 'parts of wheeled vehicle' generated which would be a hypernym for 'engine of wheeled vehicle' and 'wheel of vehicle'. Neither 'engine of wheeled vehicle' nor 'wheel of vehicle' nor 'parts of wheeled vehicle' are usually lexicalized. The sets of objects which such concepts would have as extents can be generated, but adding their concepts to a lattice might be complicated. For example, if $c_1 R_{(\geq 0;\geq 1)} c_2$ then $c_1 R_{(\geq 1;\geq 1)} c_1^*$ with $Ext(c_1^*) := \{g_2 \in Ext(c_2) \mid \exists_{g_1 \in Ext(c_1)} : g_1 r g_2\}$. But for $c_1 R_{(\geq 0;\geq 2)} c_2$ it is only possible to generate $c_1 R_{(\geq 1;\geq 2)} c_1^*$ with $Ext(c_1^*) := \{g_2 \in Ext(c_2) \mid \exists_{g_1 \in Ext(c_1)} : g_1 r g_2\}$ because $c_1 R_{(\geq 2;\geq 2)} c_1^*$ may be contradictory. For the implementation of meronymy it is helpful if these concepts exist, therefore the question arises whether and how non-lexicalized concepts should be added to a lexical database to ease the implementation of semantic relations — especially, since a verbal description of these non-lexicalized concepts is easy to obtain using 'part of...', 'most general part of ...', and so on.

As another example the hyponymy relation itself is considered. Each lattice can be interpreted as a relation of characteristic $[||\text{all}||, ||\geq 1||;]$, if r is the equality relation '=' (see Section 2.4) because each object g of a concept has at least one object (itself) in any superconcept so that it is equal to it ($g = g$). Each concept is then a maximal hyponym and a minimal hypernym of itself. Therefore the basis of the hyponymy relation is $\{(c, c) \mid c \in \mathcal{B}(G, M, I)\}$.

An application of Theorem 2.3 is the determination of an underlying denotative context for a lexical denotative context. Semantic relations are usually defined on words (denotative word concepts or synsets) while the underlying denotata are not known. For example, 'wheels are parts of bikes' might be implemented in a lexical denotative context, but the sets of all prototypical bikes and wheels exist only implicitly. Therefore the question arises as to which axioms must be fulfilled so that a relation among disambiguated words actually could be caused by a relation among denotata (see Section 2.7). Since not all concepts are lexicalized a relation among denotative word concepts is usually only a subset of a concept relation. It is not necessarily a basis but a generating system. Lemma 2.1 (see Proof 4a) demonstrates how a generating system is used to generate a relation.

Definition 2.5:
A relation $\mathcal{R} \subseteq \mathcal{B}(G, M, I) \times \mathcal{B}(G, M, I)$ satisfying, for all $(c_1^\circ, c_2^\circ) \in \mathcal{R}$,

$$c_1^\circ R c_2^\circ \Longleftrightarrow \exists_{(c_{11}^\circ, c_{21}^\circ) \in \mathcal{R}} : (c_1^\circ \leq c_{11}^\circ \text{ and } c_2^\circ \geq c_{21}^\circ) \tag{6}$$

is called a *generating system* of a relation of characteristic $[||\text{all}||, ||\geq 1||;]$. (Analogously $c_1^\circ \leq c_{11}^\circ$ and $c_2^\circ \leq c_{21}^\circ$, $c_1^\circ \geq c_{11}^\circ$ and $c_2^\circ \geq c_{21}^\circ$, and $c_1^\circ \geq c_{11}^\circ$ and $c_2^\circ \leq c_{21}^\circ$ for the other characteristics, $[||\text{all}||, ||\text{all}||;]$, $[||\geq 1||, ||\geq 1||;]$, and $[||\geq 1||, ||\text{all}||;]$, respectively.)

Lemma 2.1:
A generating system of characteristic $[||\text{all}||, ||\geq 1||;]$ generates a relation of the same characteristic if it is extended to all $c_1, c_2 \in \mathcal{B}(G, M, I)$ according to

$$c_1 R c_2 :\Longleftrightarrow \exists_{(c_1^\circ, c_2^\circ) \in \mathcal{R}} : (c_1 \leq c_1^\circ \text{ and } c_2 \geq c_2^\circ) \tag{7}$$

(Analogously for the other characteristics.)

Therefore any concept relation on lexicalized concepts which fulfills equivalence (6) can be extended to a relation on all concepts. But as Theorem 2.3 shows this does not ensure that there is an underlying relation on denotata with quantifiers Q^1 and Q^2 that would lead to the same concept relation of characteristic $[Q^5, Q^6;]$. Part of the problem is the fact that equivalence (7) is not the only possible method for generalizing a relation on a subset of concepts to a relation on the complete lattice because equivalence (8) can also be used.

$$c_1 \; R \; c_2 :\Longleftrightarrow Q^5_{c_1^\circ \le c_1} Q^6_{c_2^\circ \le c_2} : c_1^\circ \; \mathcal{R} \; c_2^\circ \tag{8}$$

A relation generated according to equivalence (7) is a subset of a relation generated according to equivalence (8) from the same generating system. For example, if Q^5 and Q^6 are $||$all$||$-quantifiers then according to equivalence (8) the relation is also inherited to superconcepts if all the basis elements under them are in relation to each other. This is not the case for equivalence (7). On the other hand, both equivalences generate relations of the same characteristic using the same generating system. In lexical denotative contexts the relations are usually defined on the subset of lexicalized concepts. Which equivalence, (7) or (8), has to be used to generate a relation on the complete lattice, assuming a relation with certain quantifiers Q^1 and Q^2 arises from the denotata (according to Definition 2.1), depends on Q^1 and Q^2. Theorem 2.4 (see Proof 4b) demonstrates that equivalence (8) has to be used under certain conditions instead of equivalence (7) if $Q^1 = Q^2 = ||\text{all}||$ or $Q^1 = Q^2 = ||{\geq 1}||$ is assumed, but not if $Q^1 = ||\text{all}||$ and $Q^2 = ||{<n}||$ is assumed. Section 2.7 provides further axioms and conditions which ensure that a relation among words could represent a relation of a certain characteristic among denotata.

Theorem 2.4:
1) If a relation on the set $\{\gamma g \mid g \in G\}$ of object concepts is a generating system of characteristic $[Q^5, Q^6;]$, and the relation is extended to a relation R on all concepts according to equivalence (8), and
a) $Q^5 = Q^6 = ||\text{all}||$ holds then R is a relation $R^r[Q^1, Q^2;]$ with $Q^1 = Q^2 = ||\text{all}||$ and $g_1 r g_2 :\Longleftrightarrow \exists_{c_1 \in \mathcal{B}(G,M,I)} \exists_{c_2 \in \mathcal{B}(G,M,I)} : (g_1 \in Ext(c_1)$ and $g_2 \in Ext(c_2)$ and $c_1 R c_2)$.
b) $Q^5 = Q^6 = ||{\geq 1}||$ holds then R is a relation $R^r[Q^1, Q^2;]$ with $Q^1 = Q^2 = ||{\geq 1}||$ and $g_1 r g_2 :\Longleftrightarrow \forall_{c_1 \in \mathcal{B}(G,M,I)} \forall_{c_2 \in \mathcal{B}(G,M,I)} : (g_1 \in Ext(c_1)$ and $g_2 \in Ext(c_2)$ and $c_1 R c_2)$.
2) If a relation on the set $\{\gamma g \mid g \in G\}$ of object concepts is a generating system of characteristic $[Q^5, Q^6;]$ and if $Q^1 = ||\text{all}||$ and $Q^2 = ||{<n}||$ then equivalence (8) does not have to be true for $Q^5 = Q^6 = ||\text{all}||$.

2.4 Special properties of relations

So far not much has been said about the relational component r. Some mathematical properties of relations on objects or attributes or of concept relations are important for applications. An example is the long standing discussion as to whether meronymy is transitive or not (see Winston et al. (1987) for a summary). Our solution to this question is to distinguish between properties of relations that can be proved mathematically and axioms that cannot be proved. This distinction facilitates a scientific discussion about features at a basic level. For example, if r is transitive then $R^r_{(\geq 0;\geq 1)}$ is also transitive. Therefore the question is not whether conceptual meronymy is transitive, but whether meronymy is transitive on the object level. It would not be correct to say that the meronymy relation between specific door-handles, specific doors, and specific houses is transitive, but that the corresponding meronymy relation $R^r_{(\geq 0;\geq 1)}$ between the concepts 'door-handle', 'door', and 'house' is not. If r is just the local containment, such as 'door-handle physically attached to door', 'door physically contained in house' then 'door-handle physically contained in house' holds. As Iris et al. (1988) point out, the meronymy relation often involves a function between the part and the whole instead of being the pure physical containment. It follows from our theory that this function must already exist on the object level and not only on the conceptual level, such as 'this specific door-handle is a part of and has a function for this specific door', and so on. Relational Concept Analysis can therefore not produce general statements (such as 'a certain relation is transitive') about semantic relations, but it helps to show precisely where mathematical properties should be defined, what consequences they have, and whether there are certain contradictions within a linguistic modeling.

Table 2.2 summarizes (see Proof 5) some of the cases in which a special mathematical property[13] of the relational component causes properties of the resulting concept relations R^r. The properties marked with '1' hold only if the extents of the concepts are not empty. The properties marked with '2' require that the extents are finite.

r	$R^r_{(\geq 0;\geq 1)}$	R^r_0	$R^r_{(\text{all};\text{all})}$	$R^r_{(\geq 1;\geq 1)}$
reflexive	refl.	refl.[1]	—	refl.
symmetric	—	sym.	sym.	sym.
transitive	trns.	—	trns.	trns.
irreflexive	—	—	irrefl.	—
irrefl. and trns.	irrefl.[2], trns.	—	irrefl., trns.	irrefl.[2], trns.
refl., sym., trns.	refl., trns.	refl., sym.	sym., trns.	refl., sym., trns.
irrefl., acycl.	irrefl.[2], acycl.[2]	—	irrefl., acycl.	irrefl.[2], acycl.[2]

Table 2.2: Properties of relations

Furthermore the following statements can be concluded:

[13]A relation r is reflexive if $\forall_{g\in G} : grg$, irreflexive if $\forall_{g\in G} : \neg grg$, symmetric if $\forall_{g_1,g_2\in G} : g_1rg_2 \Longrightarrow g_2rg_1$, transitive if $\forall_{g_1,g_2,g_3\in G} : g_1rg_2, g_2rg_3 \Longrightarrow g_1rg_3$, antisymmetric if $\forall_{g_1,g_2\in G} : g_1rg_2, g_2rg_1 \Longrightarrow g_1 = g_2$, acyclic if $\forall_{n>1}\forall_{g_1,g_2,\ldots g_n\in G} : g_1rg_2, g_2rg_3, \ldots, g_{n-1}rg_n \Longrightarrow \neg(g_nrg_1)$.

[1] If $R^r_{(\geq 0;\geq 1)}$ is reflexive then $R^r_{(\geq 0;\geq 1)}$ includes the subconcept-superconcept relation of the lattice, i.e. $c_1 \subseteq c_2 \Longrightarrow c_1 R^r_{(\geq 0;\geq 1)} c_2$. If $R^r_{(\geq 1;\geq 0)}$ is reflexive then $c_1 \supseteq c_2 \Longrightarrow c_1 R^r_{(\geq 1;\geq 0)} c_2$. If $R^r_{(\text{all};\text{all})}$ is reflexive then all objects (of all concepts) are in relation with all objects which is obviously a trivial case.

[2] If $R^r_{(\geq 0;\geq 1)}$ is irreflexive then '$c_1 \supseteq c_2$ and $c_1 R^r_{(\geq 0;\geq 1)} c_2$' is not possible because $c_2 R^r_{(\geq 0;\geq 1)} c_2$ would follow. If $R^r_{(\geq 1;\geq 0)}$ is irreflexive then '$c_1 \subseteq c_2$ and $c_1 R^r_{(\geq 1;\geq 0)} c_2$' is not possible.

[3] If a relation is irreflexive and transitive then it is also antisymmetric. Therefore, if r is irreflexive and transitive and all sets of objects are finite, then $R^r_{(\geq 0;\geq 1)}$, $R^r_{(\geq 1;\geq 0)}$, and $R^r_{(\geq 1;\geq 1)}$ are also irreflexive, antisymmetric, and transitive.

[4] If r is antisymmetric and transitive and the extents of the concepts are finite then $R^r_{(\geq 0;\geq 1)}$ does not have to be antisymmetric. Only ($c_1 R^r_{(\geq 0;\geq 1)} c_2$ and $c_2 R^r_{(\geq 0;\geq 1)} c_1$) $\Longrightarrow Ext(c_1) \cap Ext(c_2) \neq \emptyset$ can be proved (see Proof 5).

The table shows that only $R^r_{(\geq 1;\geq 1)}$ must be an equivalence relation (reflexive, symmetric, and transitive) if r is an equivalence relation. This may be another reason why $R^r_{(\geq 1;\geq 1)}$ seems to be important for many applications (compare Section 2.1). [1] demonstrates that if, for example, $R^r_{(\geq 0;\geq 1)}$ is an order relation (reflexive, antisymmetric, and transitive) then it includes the subconcept-superconcept order relation. Therefore if relations are to be modeled which have features of order relations but do not include the subconcept-superconcept relation it is possible to study $R \setminus \subseteq$ instead of R (according to [1]). [1] and [4] show that even if r is an order relation the corresponding R's do not have to be order relations. It is questionable whether some types of meronymy are order relations. Meronymy is usually defined as an irreflexive relation, but it could be asked whether it has the other features of order relations: antisymmetry and transitivity. [4] shows that antisymmetry and transitivity of r do not lead to strong properties of R, but [3] demonstrates that irreflexivity, antisymmetry, and transitivity of r can be useful. Therefore transitive meronymy can be modeled according to [3]. If meronymy is not transitive, it usually should at least be acyclic which is a weaker property than antisymmetry and transitivity which together entail acyclicity. Meronymy is usually acyclic because if a is a part of b and b is part of c then c should not be a part of a. Since acyclicity entails antisymmetry, intransitive meronymy can be modeled as a relation with an irreflexive and acyclic relational component r (see Section 3.4). If r is the equality relation '=' then $R^=_{(\geq 0;\geq 1)}$ is an order relation, $R^=_{(\geq 1;\geq 0)}$ is the dual order, and $R^=_{(\geq 1;\geq 1)}$ is an equivalence relation, and the following equivalences hold (see Proof 6a).

$$c_1 R^=_{(\geq 0;\geq 1)} c_2 \iff c_1 \leq c_2 \tag{9}$$

$$c_1 R^=_{(\geq 1;\geq 0)} c_2 \iff c_1 \geq c_2 \tag{10}$$

$$c_1 R^=_0 c_2 \iff Ext(c_1) \cap Ext(c_2) \neq \emptyset \tag{11}$$

$$c_1 R^=_{(\geq 1;\geq 1)} c_2 \iff c_1 = c_2 \tag{12}$$

Thus the conceptual ordering itself can be derived from a relation among objects. Transitivity that involves different relations can also be investigated (see Proof 6b): If r is transitive and all extents are not empty then

$$c_1 R_i c_2 \text{ and } c_2 R_{(\text{all};\text{all})} c_3 \implies c_1 R_i c_3 \text{ for } i \in \{0, (\geq 0; \geq 1), (\geq 1; \geq 0)\} \quad (13)$$

$$c_1 R_{(\text{all};\text{all})} c_2 \text{ and } c_2 R_i c_3 \implies c_1 R_i c_3 \text{ for } i \in \{0, (\geq 0; \geq 1), (\geq 1; \geq 0)\} \quad (14)$$

$$c_1 R_i c_2 \text{ and } c_2 R_{(\geq 0; \geq 1)} c_3 \implies c_1 R_0 c_3 \text{ for } i \in \{0, (\geq 1; \geq 0)\} \quad (15)$$

$$c_1 R_{(\geq 1; \geq 0)} c_2 \text{ and } c_2 R_0 c_3 \implies c_1 R_0 c_3 \quad (16)$$

An example for (16) with $i = (\geq 1; \geq 0)$, c_1 = 'eggs', c_2 = 'pasta', c_3 = 'lasagne' is: if all pasta contains some eggs and some pasta is part of some lasagne then it follows that some eggs are part of some lasagne. Some common sense 'implications', such as 'some spices contain pesticides' and 'some food contains spices' therefore 'some food contains pesticides' cannot be dealt with although they seem to be very likely to be true in many applications (see Section 2.9).

2.5 Auto- and polyrelations

It is possible to have polyrelations (i.e. several relations between two concepts) or autorelations (i.e. a relation of a concept with itself). Concept autorelations must be distinguished from lexical autorelations which involve a relation between two different polysemous senses of the same word form. For example, the two senses of agonize in WordNet, {agonize, agonise, cause agony in} and {agonize, agonise, suffer anguish}, are in lexical autorelation 'cause' to each other. Thus agonizing causes agonizing. Lexical autorelations have been studied by Horn (1989). Miller et al. (1994) distinguish word-'cousins', -'twins', and -'sisters' which are systematically occurring lexical polyrelations, such as when the same word is used for a language, country and citizen of that country, or for a tree and the wood of that tree. It is not known if there are any concept autorelations that are not trivial, such as 'some things are parts of some things.' Concept polyrelations also often do not seem to carry much information.

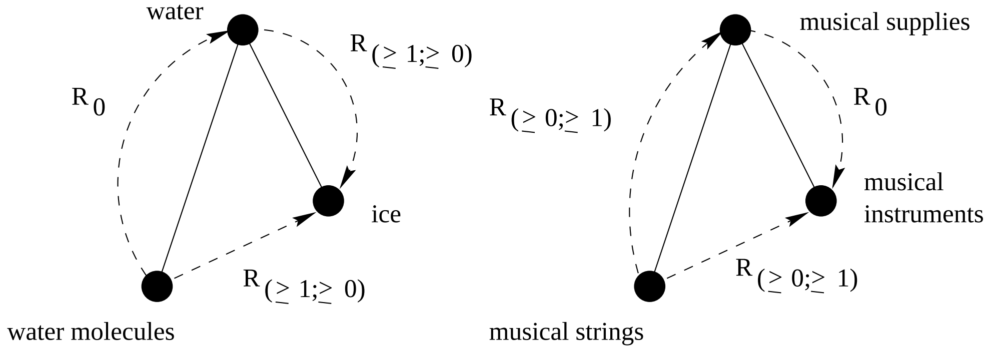

Figure 2.4: Concept polyrelations

Examples of concept polyrelations between meronymy and hyponymy are: 'ice' is a kind of 'water' and consists of 'water', or 'musical strings' are 'musical supplies' and parts of 'musical supplies' at the same time (compare Figure 2.4). Usually, however, it is possible to add additional concepts so that the bases of the relations are single relations and the polyrelations follow from inheritance. In the examples, 'water molecules' and 'musical instruments' can be added because 'ice' consists of 'water molecules' and is a kind of 'water' and 'musical strings' are parts of 'musical instruments', which are a kind of 'musical supplies.'

Figure 2.5 shows all types of concept polyrelations that can occur for irreflexive relations $R^r_{(\geq 1;\geq 0)}$, $R^r_{(\geq 0;\geq 1)}$, $R^r_{(\geq 1;\geq 1)}$, and R_0 (compare statement [2] in the last section). The arrows in Figures 2.4 and 2.5 indicate the direction of the meronymy relations. The R_0 relations in the upper half of Figure 2.5 hold only if the extents of 'bike' and 'wheel of vehicle' are not empty. The figure illustrates that concept polyrelations for meronymy and hyponymy occur if wholes are too general (all wheels of vehicles are things and parts of things) or parts are too general (all bikes are things and have things as parts). Therefore implementing relations as specifically as possible should reduce polyrelations.

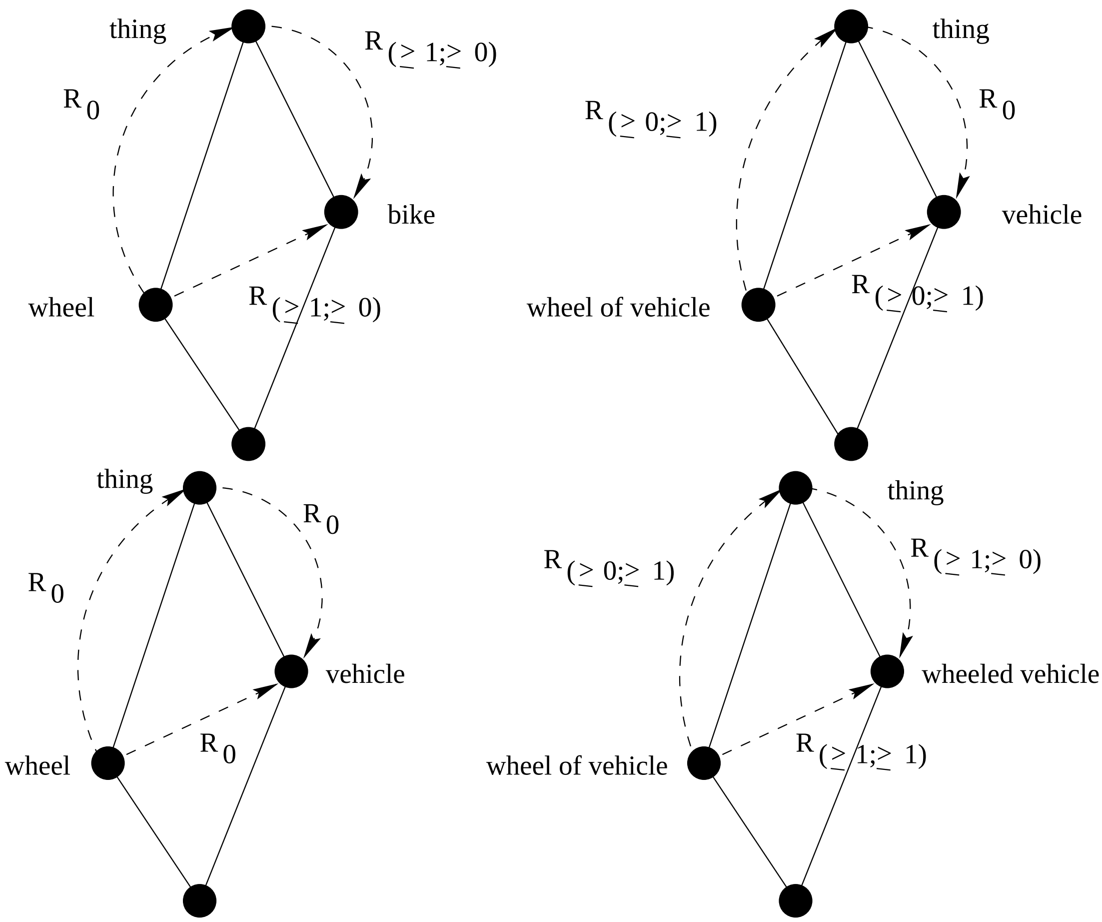

Figure 2.5: The main types of concept polyrelations

2.6 Graphical representations

It is often the case that the terminology of naming parts resembles the terminology of naming wholes. For example (see Figure 2.6[14]), body parts often have different names related to the different names of animals. If the structures are completely regular it is possible to use a representation technique similar to 'nested line diagrams' (Wille, 1984) where instead of parallel lines between two (or more) sets of concepts these concepts are surrounded by a box or an ellipsis and a boldface line is drawn between the boxes or ellipses. In Figure 2.6 all concepts in the lower ellipsis are in $R_{(\geq 1;\geq 1)}$ relation to the concepts in the upper ellipsis that are in the corresponding positions, such as 'hoof' and 'hoofed animal', and so on. The lower half of Figure 2.6 shows the complete version. The sets of $R_{(\geq 0;\geq 1)}$ and $R_{(\geq 1;\geq 0)}$ relations, which are represented by the dotted lines, form bases. Unfortunately the structures in lexical databases are usually not that regular because often the corresponding concepts are not lexicalized or because the relations are not very regularly implemented (compare Section 3.7).

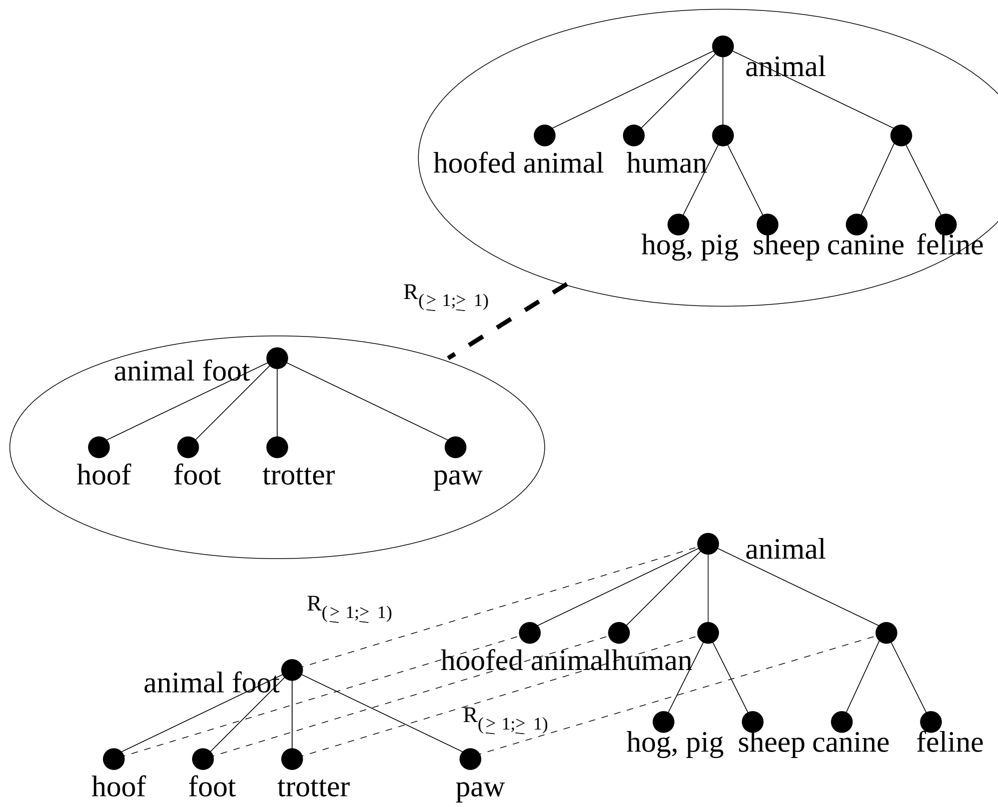

Figure 2.6: One representation of a nested line diagram

[14]Figures 2.6 and 2.7 show parts of a lattice instead of a complete lattice to indicate that they are parts of a larger lexical database, such as WordNet. Attributes are omitted. Accidently most of the figures of this chapter are tree hierarchies. This is due to the fact that WordNet is modeled as a tree and some of the figures are based on WordNet. This does not mean that the theory only holds for trees or that trees are preferable to other ordered sets.

A different representation technique is chosen in Figure 2.7. Here the double boldface dotted lines between the boxes mean that all concepts in the lower box stand in the relation that labels the dotted line to all concepts in the upper box. A basis for the relations consists of $R_{(\geq 0;\geq 1)}$ relations between 'hour, time of day' and 'day of the week', 'day', '(any) day', 'yesterday', 'tomorrow', and 'today'; and $R_{(\geq 1;\geq 0)}$ relations between 'morning', 'noon', and so on and 'day, solar day'. The reason for this very different situation compared to Figure 2.6 is that every day is a day of the week, a day, a yesterday, a today, and a tomorrow at the same time (this is possible in a lexical lattice, compare Section 1.8). Furthermore each day has all the parts, such as morning, evening, night, and so on, whereas in Figure 2.6 no animal can be a hog and a feline at the same time, and every animal has only one kind of feet. The hyponymy relations in Figure 2.7 follow WordNet. The meronymy relations are not as systematically implemented in WordNet as in the figure because as the figure demonstrates their structure is relatively complicated.

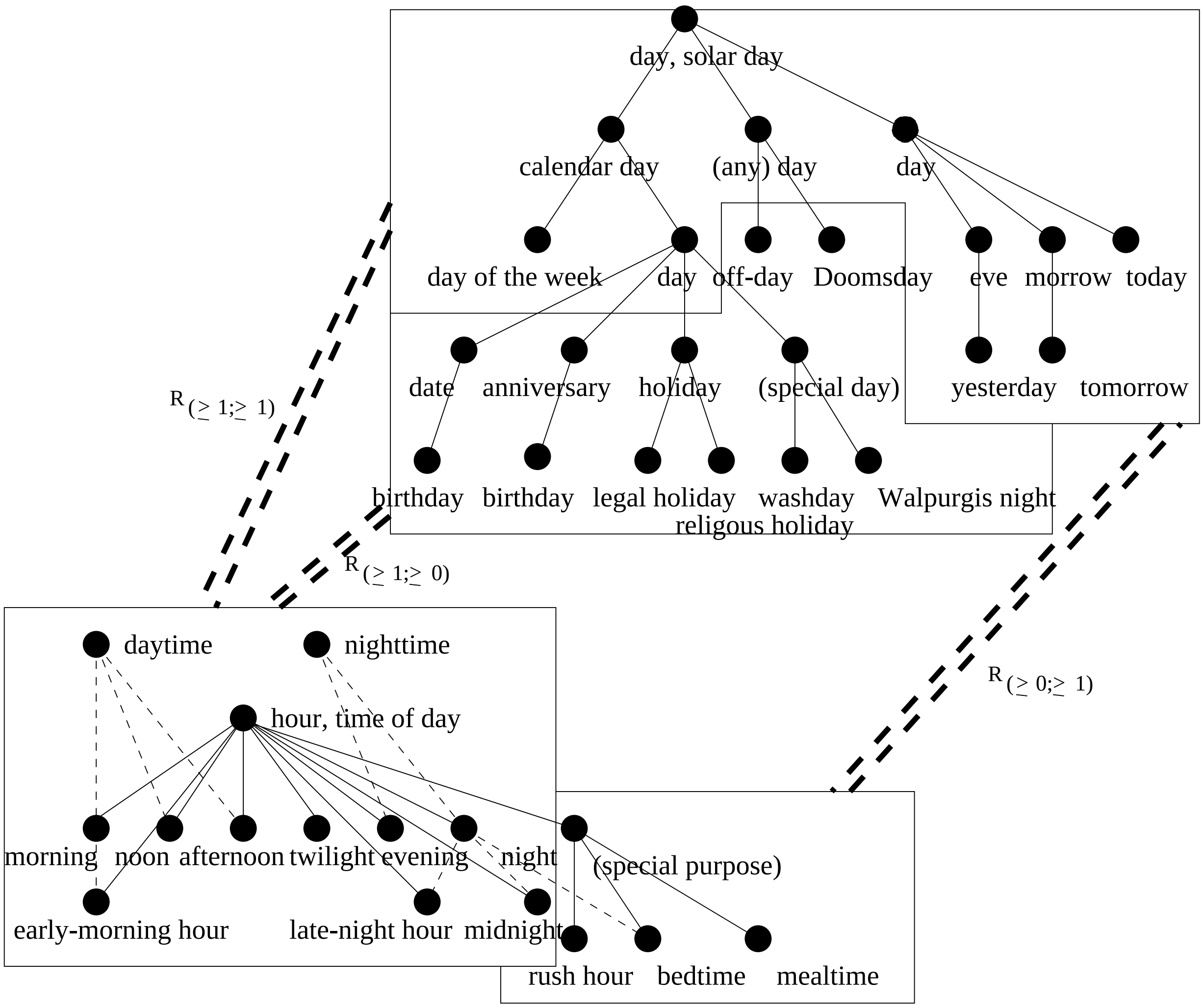

Figure 2.7: Another representation of a nested line diagram

2.7 Relations among words as concept relations - the context $\mathcal{K}_D$

In what follows, the application of Relational Concept Analysis to linguistic contexts and lattices is further investigated. At first a denotative structure $\mathcal{S}_D$ with denotative context $\mathcal{K}_D$, defined as (D, A_D, I_D), for which disambiguated words $w \in W$ denominate concepts via the mapping $dnt : W \longrightarrow \mathcal{B}(\mathcal{K}_D)$ is considered. As Section 1.8 presents, $\underline{\mathcal{B}}(\mathcal{K}_{LD})$, defined as $\underline{\mathcal{B}}(W, A_D, I_{LD})$, is isomorphic to a join-preserving sublattice of $\underline{\mathcal{B}}(\mathcal{K}_D)$. Two viewpoints can be investigated. First, a denotative context $\mathcal{K}_D$ can be given and a lexical denotative context $\mathcal{K}_{LD}$ can be derived from it. This is investigated in this section. Or, second, a lexical denotative context can be given and questions about an underlying hypothetical denotative context can be asked. This is investigated in the next section. In the first case the following is defined:

Definition 2.6:
In a denotative structure $\mathcal{S}_D$, a semantic relation $R^r \subseteq \mathcal{B}(\mathcal{K}_D) \times \mathcal{B}(\mathcal{K}_D)$ is said to be transferred to a *semantic relation* $S^r \subseteq W \times W$ of the same characteristic *among disambiguated words* if the following is defined

$$w_1 S^r w_2 \;:\Longleftrightarrow\; dnt(w_1) R^r dnt(w_2) \tag{17}$$

Since in lexical denotative lattices object concepts correspond to denotative word concepts, it follows from Definition 2.6 that $w_1 S^r w_2 \iff \gamma w_1 R^r \gamma w_2$ holds for a lexical denotative lattice derived from a denotative lattice. According to Section 2.3, $R^r \subseteq C(W) \times C(W)$ is a generating system (i.e. fulfills equivalence (6)) and generates a subset of $R^r \subseteq \mathcal{B}(\mathcal{K}_D) \times \mathcal{B}(\mathcal{K}_D)$, but does not necessarily generate $R^r \subseteq \mathcal{B}(\mathcal{K}_D) \times \mathcal{B}(\mathcal{K}_D)$ itself. Axioms are required to ensure that the relation on denotative word concepts generates the complete relation on the lattice. A first approach is to define that only information about denotata which are non-lexicalized can be lost in lexical denotative contexts in comparison to denotative contexts. That means that at least all relations which are, so to say, above the level of lexicalization in $\underline{\mathcal{B}}(\mathcal{K}_D)$ should exist in $\underline{\mathcal{B}}(\mathcal{K}_{LD})$. This is true for the hyponymy relation because $\underline{\mathcal{B}}(\mathcal{K}_{LD})$ is isomorphic to a join-preserving sublattice of $\underline{\mathcal{B}}(\mathcal{K}_D)$. A sufficient condition for other semantic relations is that the equations (7) or (8),

$$c_1 \; R^r \; c_2 :\Longleftrightarrow \exists_{(dnt(w_1), dnt(w_2)) \in R^r} : (c_1 \leq dnt(w_1) \text{ and } c_2 \leq dnt(w_2)) \tag{18}$$

(analogously for the other characteristics) or

$$c_1 \; R^r \; c_2 :\Longleftrightarrow Q^5_{dnt(w_1) \leq c_1} Q^6_{dnt(w_2) \leq c_2} : dnt(w_1) \; R^r \; dnt(w_2) \tag{19}$$

are fulfilled at least for all $c_1, c_2 \in \mathcal{B}(\mathcal{K}_D)$ with $\exists_{w_1, w_2 \in W} : (dnt(w_1) \leq c_1$ and $dnt(w_2) \leq c_2)$. The following axioms ensure for some characteristics of relations that the relations in $\underline{\mathcal{B}}(\mathcal{K}_D)$ can be generated from relations among denotative word concepts. For practical applications it may not be necessary to consider these axioms at all.

The extension of generating systems to relations on the whole lattice according to equivalence (18) which ensures that they have the correct characteristic although they may not be complete is often sufficient. The axioms in Theorems 2.5 and 2.6 present sufficient conditions for generating systems of some major types of relations (see Proof 7 for Theorem 2.5 and Proof 8 for Theorem 2.6).

Theorem 2.5:
If in $\mathcal{S}_D$ with $\mathcal{K}_D := (D, A_D, I_D)$ the *covering axiom*

$$\forall_{d \in D} \forall_{w \in W} \exists_{w_1 \in W} : \gamma d \leq dnt(w_1) \leq (\gamma d \vee dnt(w)) \tag{20}$$

holds then relations $R^r[||\text{all}||, ||\text{all}||;]$ and $R^r[||{\geq}1||, ||{\geq}1||;]$ fulfill equivalence (19).

The axiom ensures that a denotative concept that is a hypernym of a denotative word concept and of an object concept of a denotata has a denotative word concept between itself and the object concept of the denotata. In other words, the lowest concept at which an object concept of a denotata reaches the level of lexicalization has to be a denotative word concept. Or, below the level of lexicalization in a denotative lattice nothing is known, but in the sublattice which is generated by the disambiguated word concepts and therefore, so to say, above the level of lexicalization, everything is known. Unfortunately, more axioms are needed if other relations $R^r[Q^1, Q^2;]$ are to be dealt with, for example:

Theorem 2.6:
If the covering axiom and the axiom

$$dnt(w_1) R^r_{(\geq 0; \geq 1)} c_2 \iff \exists_{dnt(w_2) \leq c_2} : dnt(w_1) R^r_{(\geq 0; \geq 1)} dnt(w_2) \tag{21}$$

hold then relations of characteristic $[||\text{all}||, ||{\geq}1||;]$ fulfill equivalence (19).

This axiom ensures that the most specific whole of a relation $R^r_{(\geq 0; \geq 1)}$ is lexicalized (compare Section 2.3).

2.8 Semantic relations - the context $\mathcal{K}_{LD}$

In the last section the conditions of restricting a denotative lattice to a lexical denotative lattice are investigated. In this section a lexical denotative context $\mathcal{K}_{LD}$ is given and nothing is explicitly known about an underlying denotative context $\mathcal{K}_D$. While in $\mathcal{K}_D$ the semantic relations are transferred to the words from the concepts, in $\mathcal{K}_{LD}$ semantic relations are defined on the words and then transferred to the concepts. If it assumed that the axioms defined in the last section hold it follows that the semantic relations in $\mathcal{K}_{LD}$ could arise from corresponding relations in an underlying $\mathcal{K}_D$.

Definition 2.7:
A relation $s \subseteq W \times W$ among disambiguated words is called a *semantic relation among disambiguated words* if

$$\gamma w_1 = \gamma w_2 \implies \forall_{w \in W} \ (w_1 s w \Leftrightarrow w_2 s w) \ \text{ and } \ (w s w_1 \Leftrightarrow w s w_2), \tag{22}$$

is fulfilled. The relation is a *semantic relation of characteristic* $[Q^5, Q^6;]$ if it fulfills equivalence (4).
A semantic relation s among disambiguated words is transferred to a relation R among concepts according to

$$\gamma w_1 R \gamma w_2 \ :\Longleftrightarrow \ w_1 s w_2 \tag{23}$$

The condition (22) ensures that equivalence (23) is well-defined. If a relation on the object concepts fulfills equivalence (6) it is a generating system of that characteristic and generates a relation in the complete lattice according to (7) or (8). There are two possible ways to handle synonymy and hyponymy. First, synonymy and hyponymy can be defined according to

$$w_1 \ syn \ w_2 \ :\Longleftrightarrow \ \gamma w_1 = \gamma w_2 \tag{24}$$

$$w_1 \ hyp \ w_2 \ :\Longleftrightarrow \ \gamma w_1 \leq \gamma w_2. \tag{25}$$

It follows that they are semantic relations because they fulfill equivalence (22). Second, synonymy and hyponymy can be defined as semantic relations $syn, hyp \subseteq W \times W$ with the properties that syn is an equivalence relation and hyp is an order relation on the equivalence classes of syn. A concept lattice can be developed from them (see Section 1.8) and with $\mathcal{R}_W^{sem}$ as family of semantic relations on disambiguated words even

$$\gamma w_1 = \gamma w_2 \iff \forall_{s \in \mathcal{R}_W^{sem}} \forall_{w \in W} \ (w_1 s w \Leftrightarrow w_2 s w) \ \text{ and } \ (w s w_1 \Leftrightarrow w s w_2), \tag{26}$$

holds. In contrast to semantic relations lexical relations, such as antonymy, are defined on disambiguated words, but cannot be generalized to relations among concepts. For example, ‘nasty’ and ‘nice’ are antonyms in WordNet, but ‘awful’ which is a synonym to ‘nasty’ is not an antonym of ‘nice’.

Definition 2.8:
Relations $s \subseteq W \times W$ that do not fulfill condition (22) are called *lexical relations.*

2.9 Future research

Not all the problems of the implementation of semantic relations in lexical databases are solved. Fischer (1991) has found some algorithms that check whether relations are contradictory implemented. His solution is to implement consistency rules as conditions within the object-oriented programming language Smalltalk. He checks the consistency of relations concerning inverse relations, implicit relations, circularity, and other basic properties. It would be interesting to study whether his approach can be extended to cover the inheritance and consistency rules that follow from Relational Concept Analysis.

The extension of Relational Concept Analysis to unary relations (logical predicates), such as 'all objects of a concept have an attribute', 'most objects of a concept have an attribute', 'some prototypical objects of a concept have an attribute', and so on is developed in Section 4.2. But ternary or other n-ary relations are left to future research. The problem of higher n-ary relations is that the number of involved quantifiers and therefore the number of characteristics of relations rises. Another open problem is that in linguistic applications many relations use quantifiers of the kind ||almost all|| or ||all typical||. Therefore Relational Concept Analysis should be extended to incorporate prototype theory and heuristic rules.

2.10 Proofs

Proof 1: (Theorem 2.1)
First, the first four rows must be proved by showing two directions of the equivalence (4).The direction '$\Longrightarrow$' of the first row is true because c_{11} and c_{21} are subconcepts of c_1 and c_2, respectively. The number of exceptions m and n cannot be larger for subconcepts than for the original concepts. The direction '$\Longleftarrow$' of the first row is true because if anything is said about all subconcepts of a concept (including the concept) it has to be true for the concept itself. The next three rows are similarly proved or follow from the quantifier conversion rules: Q^2 in row 1 and 2 are dual to each other, so are Q^6 in row 1 and 2, and so on.

Rows 5 to 8 are inner negations (such as Q^1Q^2 and $Q^1Q^2\neg$) of rows 1 to 4. An inner negation changes the relation r into its negation $\neg r$, but obviously the quantifiers stay the same. Rows 9 to 16 are equal to rows 1 to 8, because $Q^1Q^2 = Q^1\neg\neg Q^2 = (Q^1\neg)(\neg Q^2)$.

Proof 2: (Theorem 2.2)
First, it has to be shown ('$\Longrightarrow$') that a relation of one of the four characteristics always has a unique basis. Second, it has to be shown ('$\Longleftarrow$') that any relation with basis according to Definition 2.4 is of such a characteristic. The proof is only demonstrated for relations of characteristic $[||\text{all}||, ||\geq 1||;]$. The other characteristics are analogously.

To prove '$\Longrightarrow$' the following construction yields a basis. If a relation $R \subseteq \mathcal{B}(G, M, I) \times \mathcal{B}(G, M, I)$ is given then for concepts c_1 and c_2 we define $(c_1R)_{min} := min_{\leq}\{c \in \mathcal{B}(G, M, I) \mid c_1Rc\}$ and $(Rc_2)_{max} := max_{\leq}\{c \in \mathcal{B}(G, M, I) \mid cRc_2\}$ as sets of minimal second components or maximal first components of the relations of a concept. A basis is then $\mathcal{R} := \{(c_1, c_2) \mid c_2 \in (c_1R)_{min}$ and $c_1 \in (Rc_2)_{max}\}$. It can be proved that $\mathcal{R}$ is a basis of R: for c_1Rc_2 there are c_3 and c_4 with $c_4 \in (c_1R)_{min}, c_3 \in (Rc_4)_{max}$, $c_4 \leq c_2$, and $c_3 \geq c_1$. Since R is of characteristic $[||\text{all}||, ||\!\geq 1||;]$, we also have $c_4 \in (c_3R)_{min}$ and hence $(c_3, c_4) \in \mathcal{R}$. $\mathcal{R}$ is minimal and unique because, if $\mathcal{R}_1$ is a basis of R and if $(c_1, c_2) \in \mathcal{R}$, then there must exist $(c_3, c_4) \in \mathcal{R}_1$ with $c_1 \leq c_3$, $c_2 \geq c_4$ which implies $c_1 = c_3$, $c_2 = c_4$ and so $\mathcal{R} = \mathcal{R}_1$.

For the other direction '$\Longleftarrow$', it has to be shown that the generated relation is of characteristic $[||\text{all}||, || \geq 1||;]$ (equivalence (4)). If c_1Rc_2, there is a basis element (c_1°, c_2°) with $c_1 \leq c_1^\circ$ and $c_2 \geq c_2^\circ$ according to equivalence (5) '$\Longrightarrow$'. From equivalence (5) '$\Longleftarrow$' it follows that $||\text{all}||_{c_{11} \leq c_1^\circ} : c_{11}Rc_2^\circ$ and because of $c_1 \leq c_1^\circ$ and $c_2 \geq c_2^\circ$ follows $||\text{all}||_{c_{11} \leq c_1}\ ||\!\geq 1||_{c_{21} \leq c_2} : c_{11}Rc_{21}$. On the other hand, if $||\text{all}||_{c_{11} \leq c_1} ||\!\geq 1||_{c_{21} \leq c_2} : c_{11}Rc_{21}$ then $|| \geq 1||_{c_{21} \leq c_2} : c_1Rc_{21}$ and equivalence (5) '$\Longrightarrow$' implies that there is a basis element (c_1°, c_2°) so that $c_1 \leq c_1^\circ$ and $c_2^\circ \leq c_{21} \leq c_2$. Equivalence (5) '$\Longleftarrow$' entails c_1Rc_2.

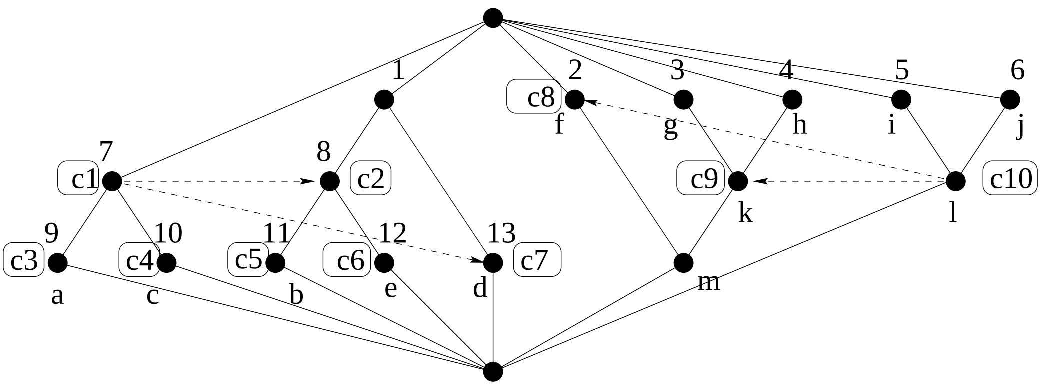

Figure 2.8: A counter example

Proof 3: (Theorem 2.3)
The relation R of characteristic $[||\text{all}||, ||\!\geq 1||;]$ with basis $\{(c_1, c_2), (c_1, c_7), (c_{10}, c_9), (c_{10}, c_8)\}$ in Figure 2.8 cannot be generated from a relation among objects or attributes. To prove this it is enough to demonstrate that the relation in the left half of the lattice, $\{(c_1, c_2), (c_1, c_7)\}$, cannot be generated from a relation on objects. That implies that the relation in the right half of the lattice, $\{(c_{10}, c_9), (c_{10}, c_8)\}$, cannot be generated from a relation among attributes. If even parts of the relation cannot be constructed from relations among objects or attributes the whole relation cannot be constructed.

It has to be shown that it is not possible to find quantifiers Q^1 and Q^2 so that equivalence (1) is fulfilled. According to Theorem 2.1 only $Q^1 = ||\!> (\text{all} - n)||$ and $Q^2 = ||\!\geq m||$ or $Q^1 = ||\!> (\text{all} - n)||$ and $Q^2 = ||\!\leq (\text{all} - m)||$ must be considered. Since $|Ext(c_1)| = |Ext(c_2)| = 2$ only $n, m \in \{1, 2\}$ are possible. (The quantifiers $Q^1, Q^2 \in \{||\!\geq 0||, ||\!\leq 0||, ||\!\leq \text{all}||, ||\!\geq \text{all}||\}$ can be ignored because they would involve trivial relations with every concept as first or second component.)

If $Q^2 = ||\geq 1||$ then, for example, (c_3, c_5) or (c_3, c_6) would also have to be element of R because (c_3, c_2) is always element of the relation and the condition in Q^2 involves only one element. On the other hand (c_3, c_5) or (c_3, c_6) cannot be generated from the basis therefore they cannot be in R.
$Q^2 = ||\geq 2||$ is impossible because then (c_1, c_7) could not be in R.
$Q^2 = ||\leq(\text{all}-1)||$ is equivalent to $Q^2 = ||\neg\text{all}||$ therefore either $\neg arb$, $\neg are$, $\neg crb$, or $\neg cre$. It follows that either c_3 or c_4 is in relation with c_5 or c_6 which is a contradiction to the basis.
$Q^2 = ||\leq(\text{all} - 2)||$ is impossible because then (c_1, c_7) could not be in R.

Now, the first half of the theorem has to be proved: Instead of $\underline{\mathcal{B}_1}(G_1, M_1, I_1)$ a lattice $\underline{\mathcal{B}_2}(G_2, M_2, I_2)$ is chosen which has an equivalent context and concept relation, such that all its concepts are object concepts of exactly one object ($c \in \underline{\mathcal{B}_2}(G_2, M_2, I_2) \Longrightarrow ||1||_{g\in G_2} : c = \gamma g$). The relation $\{(c_{11}, c_{21}), \ldots, (c_{1n}, c_{2m})\} = \mathcal{R}$ is given on $\underline{\mathcal{B}_2}(G, M, I)$ as the equivalent to the concept relation on $\underline{\mathcal{B}_1}(G_1, M_1, I_1)$. With $g_1 r g_2 :\Longleftrightarrow \exists_{c_{11}\geq\gamma g_1} : c_{11} R \gamma g_2$ a relation is defined on the objects. Equivalence (1) has to be shown for this relation (with $Q^1 = ||\text{all}||$ and $Q^2 = ||\geq 1||$) to prove the theorem. If $c_1 R c_2$ then $||\text{all}||_{g_1\in Ext(c_1)} ||\geq 1||_{g_2\in Ext(c_2)} : g_1 r g_2$ follows from the definition of r. On the other hand, if $c_1 := \gamma g_1$, $c_2 := \gamma g_2$, and $||\text{all}||_{g_1\in Ext(c_1)} ||\geq 1||_{g_2\in Ext(c_2)} : g_1 r g_2$ then $||\text{all}||_{g_1\in Ext(c_1)} ||\geq 1||_{g_2\in Ext(c_2)} ||\geq 1||_{c_{11}\geq\gamma g_1} : c_{11} R \gamma g_2$ follows from the definition of r. From $c_{11} R \gamma g_2$ with $c_{11} \geq \gamma g_1$ follows $\gamma g_1 R \gamma g_2$ because R is inherited to subconcepts of the first component. It follows: $c_1 R c_2$.

Proof 4a: (Lemma 2.1)
The proof is only demonstrated for relations of characteristic $[||\text{all}||; ||\geq 1||;]$. It has to be shown that the generated relation is of characteristic $[||\text{all}||; ||\geq 1||;]$ using equivalence (4). If $c_1 R c_2$ then equivalence (7) '$\Longrightarrow$' entails $c_1^\circ R c_2^\circ$ and equivalence (7) '$\Longleftarrow$' entails $||\text{all}||_{c_{11}\leq c_1^\circ} : c_{11} R c_2$. It follows that $||\text{all}||_{c_{11}\leq c_1} ||\geq 1||_{c_{21}\leq c_2} : c_{11} R c_{21}$. If $||\text{all}||_{c_{11}\leq c_1} ||\geq 1||_{c_{21}\leq c_2} : c_{11} R c_{21}$ then $\exists_{c_{21}\leq c_2} : c_1 R c_{21}$ and equivalence (7) '$\Longrightarrow$' implies $c_1^\circ R c_2^\circ$ and equivalence (7) '$\Longleftarrow$' implies $c_1 R c_2$.

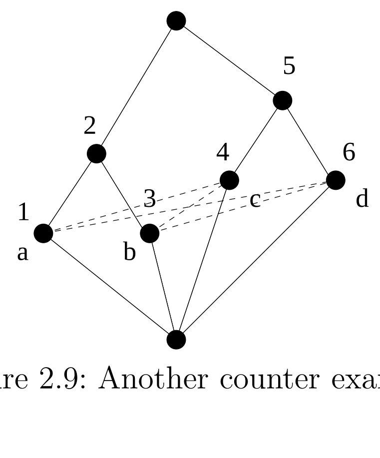

Figure 2.9: Another counter example

Proof 4b: (Theorem 2.4)
The definition of r in a) and b) together with the fact that the characteristic of relation follows from Q^1 and Q^2 entails $g_1 r g_2 \iff \gamma g_1 R^r \gamma g_2$ in both cases. Therefore in a) and b) holds $c_1 R^r c_2 \iff Q^1_{g_1 \in Ext(c_1)} Q^2_{g_2 \in Ext(c_2)} : g_1 r g_2 \iff Q^1_{g_1 \in Ext(c_1)} Q^2_{g_2 \in Ext(c_2)} : \gamma g_1 R^r \gamma g_2$. This entails equivalence (8) for the generating system of object concepts which proves the statement.

Equivalence (8) does not hold for the relation in Figure 2.9 if $Q^1 = ||\text{all}||$ and $Q^2 = ||<2||$ because γa is in relation to γc and γd, and so on, but $\mu 2$ and $\mu 5$ are not in relation to each other.

Proof 5: (Table 2.2)
Reflexivity:
If $\forall_{g \in G} : grg$ then $\forall_{c \in \mathcal{B}(G,M,I)} \forall_{g \in Ext(c)} : grg$ and therefore $R^r_{(\geq 0; \geq 1)}$ is reflexive. And if the extent of c is not empty $\exists_{g \in Ext(c)} : grg$ holds which entails that R^r_0 is reflexive.

Symmetry:
If r is symmetric then R^r_0 and $R^r_{(\text{all;all})}$ are symmetric because $\forall$- and $\exists$-quantifiers are interchangeable, respectively. From $c_1 R^r_{(\geq 0; \geq 1)} c_2 \iff \forall_{g_1 \in Ext(c_1)} \exists_{g_2 \in Ext(c_2)} : g_1 r g_2 \iff \forall_{g_1 \in Ext(c_1)} \exists_{g_2 \in Ext(c_2)} : g_2 r g_1 \iff c_2 R^r_{(\geq 1; \geq 0)} c_1$ follows the symmetry of $R^r_{(\geq 1; \geq 1)}$.

Transitivity:
If r is transitive then $\forall_{g_1 \in Ext(c_1)} \exists_{g_2 \in Ext(c_2)} : g_1 r g_2$ and $\forall_{g_2 \in Ext(c_2)} \exists_{g_3 \in Ext(c_3)} : g_2 r g_3$ implies $\forall_{g_1 \in Ext(c_1)} \exists_{g_3 \in Ext(c_3)} : g_1 r g_3$. Therefore $R^r_{(\geq 0; \geq 1)}$ and $R^r_{(\geq 1; \geq 1)}$ are transitive. Similarly, $R^r_{(\text{all;all})}$ is transitive.

Irreflexivity:
If r is irreflexive then $c_1 R^r_{(\text{all;all})} c_2 \Longrightarrow Ext(c_1) \cap Ext(c_2) = \emptyset$.

Irreflexive and transitive:
If r is irreflexive, transitive and the extents of all concepts are finite then $c_1 R^r_{(\geq 0; \geq 1)} c_1$ is impossible because the transitivity implies for any chain $g_1 r g_2; g_2 r g_3; \ldots; g_{n-1} r g_n$ that $g_1 r g_n$. Since each object g_i has to have a follower in such a chain and there are only finitely many objects, $g_i r g_i$ follows for at least one g_i. This is in contradiction to the irreflexivity of r.

Reflexive, symmetric, and transitive:
The statement follows from the other rows of the table.

Irreflexive and acyclic:
Similarly to the proof of 'irreflexive and transitive', the finiteness implies that if $R^r_{(\geq 0; \geq 1)}$ or $R^r_{(\text{all;all})}$ had a cycle or were not irreflexive, r could not be acyclic or irreflexive.

Antisymmetric and transitive:
Similarly to the proof of 'irreflexive and transitive', the finiteness implies for any chain between objects of c_1 and c_2 that it contains a circle and therefore it follows from the transitivity and antisymmetry of r that $Ext(c_1)$ and $Ext(c_2)$ must have an object in common.

Proof 6a: ((9)-(12))
The first row is proved by $c_1 R^{=}_{(\geq 0;\geq 1)} c_2 \iff ||\text{all}||_{g_1 \in Ext(c_1)} || \geq 1||_{g_2 \in Ext(c_2)} : g_1 = g_2 \iff Ext(c_1) \subseteq Ext(c_2)$. The others are analogously.

Proof 6b: ((13)-(16))
For example, $c_1 R_0 c_2$ and $c_2 R_{(\text{all};\text{all})} c_3 \Longrightarrow \exists_{g_1 \in Ext(c_1)} \exists_{g_2 \in Ext(c_2)} : g_1 r g_2$ and $\forall_{g_2 \in Ext(c_2)} \forall_{g_3 \in Ext(c_3)} : g_2 r g_3 \Longrightarrow \exists_{g_1 \in Ext(c_1)} \exists_{g_3 \in Ext(c_3)} : g_1 r g_3$. The other implications are proved analogously.

Proof 7: (Theorem 2.5)
A relation $R^r[||\text{all}||, ||\text{all}||;]$ always fulfills '$\Longrightarrow$' in equivalence (19). The other direction '$\Longleftarrow$' was not fulfilled if c_1 or c_2 were object concepts, $c_1 = \gamma d_1$ or $c_2 = \gamma d_2$ and not lexicalized. But for $c_1 = \gamma d_1$ and $\neg \exists_{w \in W} : c_1 = dnt(w)$ the covering axiom entails $\neg \exists_{w \in W} : c_1 \geq dnt(w)$ and therefore (19) is not applicable.
A relation $R^r[||\geq 1||, ||\geq 1||;]$ always fulfills '$\Longleftarrow$' in equivalence (19). The other direction '$\Longrightarrow$' was not fulfilled if there were non-lexicalized c_1, c_2 with $\neg dnt(w_1) R^r dnt(w_2)$ for all $dnt(w_1) < c_1$ and $dnt(w_2) < c_2$, $d_1 r d_2$, $\gamma d_1 \leq c_1$, $\gamma d_2 \leq c_2$. This is impossible according to the covering axiom.

Proof 8: (Theorem 2.6)
The axiom in equivalence (21) ensures that the direction '$\Longleftarrow$' in equivalence (19) is fulfilled. The other direction follows from the covering axiom similarly to Proof 7.

3 Semantic Relations

3.1 Introduction

Semantic relations are studied in many disciplines, such as linguistics, logic, cognitive science, psychology, anthropology, and artificial intelligence. Furthermore they are explicitly or implicitly used in many applications, such as the storage of information in classification systems, knowledge bases, bibliographic thesauri and natural language thesauri (for example RIT). This chapter presents a summary of some current research on and applications of semantic relations and develops formalizations based on Relational Concept Analysis. Meronymy which seems to be a very important type of relation in many applications but which is often not properly distinguished from other relations and whose properties are often not agreed upon, is the main example for this chapter. More detailed surveys on semantic relations in lexical databases in general can be found in Evens (1988) and on semantic relations in anthropology, linguistics, psychology, and computer science in Evens et al. (1980).

Semantic relations can be divided into four classes: conceptual orderings, meronymy, functional relations, and contrast or sequence relations. Conceptual orderings (also called 'taxonomy', 'hyponymy', 'IS-A relation', 'class inclusion', or 'super-ordination') and meronymy (also called 'part-whole relation') form hierarchical orderings. They occur in knowledge bases, thesauri, biological taxonomies, and library classification systems. Conceptual ordering and meronymy are sometimes not properly distinguished because they are very similar in some cases (compare Winston et al. (1987)). For example, should 'algebra' be called a part of mathematics or a kind of mathematics? Phenomena of natural languages, such as collective nouns, increase the difficulties. For example, is a tomato a part of the groceries, or a kind of grocery? Since the extent of a concept is a subset of the extent of a superconcept of that concept, the conceptual ordering is in some way similar to some types of meronymy. For example, the set of poodles is a subset of the set of dogs – the area of a city is a subset of the area of a country. The same holds for the instance relation between denotata and concepts and the membership meronymy. A denotatum is a 'kind of' the concept to which extent it belongs, for example, Fido is a kind of dog. But it is also a member of the set of elements in the extent of that concept, for example, Fido is a member of the class of dogs. Other relations such as synonymy (equality in a conceptual hierarchy), coordination (between cohyponyms, which have the same immediate superconcept), attribution (for example, the relation between objects and attributes in a formal context), and some kinds of cross-references evolve from hierarchies. Fischer at al. (1996) call them 'virtual relations'. Cross-references are related to hierarchies if they are based on hierarchical relations, such as cross-references among the cohyponyms of a concept.

Functional relations, such as the case relations of a verb (agent, instrument, object, and so on) do not form hierarchical orderings. They are usually visualized as semantic nets, conceptual graphs (Sowa, 1984) or Entity-Relationship-Diagrams. Contrast relations (such as antonymy) are binary sequences (such as yesterday, today, tomorrow) which are linear orderings. If a conceptual ordering is represented as a vertical hierarchy, contrast and sequence relations are usually horizontal chains among cohyponyms. For example, yesterday, today, and tomorrow are hyponyms of 'day'; a dwarf and a giant are both hyponyms of 'person from fairy tale', and so on. The classification of semantic relations into the four basic types, conceptual ordering, meronymy, contrast and sequences, and functional relations, can be found in Dahlberg (1994) and DIN 32705 (1987). Non-hierarchical relations occur in the analyses of lexical fields, semantic networks and in the Entity-Relationship-Model of database theory. Unfortunately many of the existing relational models concentrate either on the hierarchical or on the non-hierarchical relations. For example, before the invention of Relational Concept Analysis, Formal Concept Analysis studied only hierarchical relations. On the other hand, the Entity-Relationship-Model and conceptual graphs (Sowa, 1984) concentrate on non-hierarchical relations. In applications often both types of relations occur. For example, library classification systems contain see-also-references and scientific thesauri contain related terms (RT) as non-hierarchical relations. While hierarchical relations in thesauri are usually organized within a consistent system which is reflected by a formal notation, the cross-references or related terms often do not seem to be systematic. They can be coordinate terms, meronyms or others, and they are usually not indicated by the formal notation of a classification system. It seems that the question of how to integrate cross-references into a hierarchy as systematically and consistently as possible is still not solved. In Formal Concept Analysis cross-references often become redundant since lattice orderings allow several immediate superconcepts. For example, in a tree hierarchy 'biological applications of computer science' has to be classified either under biology or under computer science and a cross-reference should be established. In a lattice any concept can have two or more immediate superconcepts therefore this kind of cross-reference is redundant.

There are different approaches to the investigation of semantic relations. Linguists, such as Lyons (1977) and Cruse (1986), study the lexicalization of semantic relations. They define relations by 'linguistic tests'. For example, Cruse defines the meronymy relation as a relation which is expressed by 'A Y has Xs/an X' or 'An X is part of a Y' in the language. Different types of relations are then differentiated by different expressions in the language. Psychologists, such as Chaffin & Herrmann (1988) design psychological experiments to determine how and which relations are represented in the mental lexicon. Logicians and philosophers, such as Dahlberg (1994), often try to completely ignore the lexical component of semantic relations and investigate their logical features on a conceptual level. Dahlberg, for example, states that Winston et al.'s (1987) six types of meronymy relations differ only on the lexical level but not on the conceptual. We think that it is essential to distinguish between concept relations and their lexicalizations. Concept relations can often be modeled according to mathematical or logical rules whereas lexicalizations of relations tend to have many exceptions. It seems to be a good approach to base a model on concept relations and then to compare which concepts are lexicalized in which manner.

A distinction can be made as to whether a model includes only a few basic relations and claims that other relations are based on these basic relations or whether it includes many relations. Models which use only basic relations, such as WordNet, seem to be concept based, whereas models that use many relations, such as the Explanatory Combinatorial Dictionary (according to Frawley, (1988)), which uses 53 relations, or Rahmstorf's analysis of German noun phrases (1983), which uses 38 relations, are oriented in the lexical surface. Since our research concentrates on semantic relations as concept relations (see Definition 3.1), our model claims that there are only a few basic types of semantic relations based on a few basic properties. We define semantic relations as relations in a denotative or connotative structure among the denotative or connotative concepts or among the disambiguated words via their denotative or connotative word concepts. In what follows only denotative structures are considered. Connotative structures and lexical connotative or denotative contexts can be treated similarly using the definitions from Section 2.8. The following definition is partly a repetition of Definitions 2.6.

Definition 3.1:
In a denotative structure $\mathcal{S}_D$, the set of relations among concepts is denoted by $\mathcal{R}_C$. The elements of $\mathcal{R}_C$ are called semantic relations. A relation among disambiguated words is called a *lexical relation*. A lexical relation S^r is called a *semantic relation (among disambiguated words)* if it fulfills

$$w_1 S^r w_2 \Longleftrightarrow dnt(w_1) R^r dnt(w_2)$$

for a semantic relation R^r and all $w_1, w_2 \in W$.

3.2 Conceptual ordering

The semantic relations, synonymy, hyponymy, cohyponymy, attribution, and disjunction follow directly from the relations in a concept lattice. Except for cohyponymy which seems to be not equivalent to a semantic relation of the form R^r, the other relations are of the form $R^=$ (compare Section 2.4). The relational tag '=' denotes the equality relation among denotata and represents a trivial case because each denotatum is equal only to itself. In contrast to meronymy, where a denotatum is in relation to another denotatum, synonymy and so on entirely depend on the conceptual ordering and not on non-trivial relations among denotata. The Definition 3.2 summarizes the relevant definitions from Sections 1.6 and 1.7 and in addition explains cohyponymy and disjunction. Attribution could be defined between formal objects and attributes using ι and ε. A distinction could be made between 'classifying attributes' that a concept shares with its cohyponyms and differentiating attributes that distinguish a concept from its superconcepts. Furthermore, while the attributes in A_D are 'essential attributes' because they are needed for the classification in the concept lattice, 'accidental attributes' (Dahlberg, 1994) can be added as unary relations of the denotata (compare Section 4.2) which are not formal attributes of the concept lattice. Essential attributes are inherited from concepts to subconcepts, whereas accidental attributes are not inherited.

Definition 3.2:
In a denotative structure $\mathcal{S}_D$ the following semantic relations are defined:
Synonymy: Two disambiguated words are called *synonyms* if they denote the same concept, i.e.

$$w_1 \ SYN \, w_2 :\Longleftrightarrow dnt(w_1) = dnt(w_2) (\Longleftrightarrow dnt(w_1) R^{=}_{(\geq 1;\geq 1)} dnt(w_2))$$

Hyponymy: A disambiguated word is a *hyponym* of another word if the concept it denotes is a subconcept of the concept the other word denotes, i.e.

$$w_1 \ HYP \, w_2 :\Longleftrightarrow dnt(w_1) \leq dnt(w_2) (\Longleftrightarrow dnt(w_1) R^{=}_{(\geq 0;\geq 1)} dnt(w_2))$$

The hyponymy of a disambiguated word to a denotative concept (which does not have to be lexicalized) is denoted by $HYP^\star$, i.e.

$$w_1 \ HYP^\star c :\Longleftrightarrow dnt(w_1) \leq c$$

Hypernymy is the inverse relation of hyponymy.
Cohyponymy: Two disambiguated words are *cohyponyms* if they denote immediate subconcepts[15] of the same concept and are not synonyms, i.e.

$$w_1 \ COH \, w_2 :\Longleftrightarrow \exists_{c \in \mathcal{B}(\mathcal{K}_D)} : dnt(w_1) \prec c \text{ and } dnt(w_2) \prec c \text{ and } \neg(w_1 \ SYN \, w_2)$$

Disjointness: Two disambiguated words are *disjoint* if they do not have a common object in their extents, i.e.

$$w_1 \ DISJ \, w_2 :\Longleftrightarrow dnt(w_1) \neg R^{=}_{0} dnt(w_2) (\Longleftrightarrow dnt(w_1) R^{\neq}_{(\text{all};\text{all})} dnt(w_2))$$

The *instance relation*, which is not a semantic relation, is defined as

$$d \ INST \, w :\Longleftrightarrow d \in Ext(dnt(w))$$

From the definition follows that synonymy is an equivalence relation (reflexive, symmetric, transitive) and hyponymy is an ordering relation (reflexive, antisymmetric, transitive). Cohyponymy in a lattice is symmetric, but not transitive because, for example, 'piano' can be a cohyponym to 'chair' and 'violin' if it is classified as furniture and musical instrument, but 'chair' and 'violin' are not cohyponyms. (Cohyponymy is transitive in a tree structure.) Cohyponyms are called 'contrast sets' by Kay (1971). Disjointness is also a symmetric, but not transitive relation. Disjointness is distinguished from antonymy: while antonymy is defined using the intents of concepts, disjointness is defined using the extents of concepts. Hyponymy among verbs is sometimes called *troponymy* (Miller et al., 1990).

[15] $\prec$ denotes the relation '$<$' between a concept and its immediate superconcept.

3.3 Meronymy: Lesniewski's mereology

Meronymy serves as the main example for this chapter. This section and Section 3.5 provide information on existing theories of meronymy and explain why these existing theories may be replaced or improved by methods of Relational Concept Analysis. Section 3.4 formally defines meronymy and shows an application of the formal definition that is continued in Section 3.6 to obtain a classification of meronymy. Finally, Section 3.7 investigates how Relational Concept Analysis can facilitate avoiding irregularities in implementations of relations in lexical databases. The philosopher Lesniewski (Luschei (1962)) developed a theory called 'mereology' which is based on part-whole relations instead of set theory. A summary of mereology can be found in Rescher (1975). Mereology has never been applied to natural language research for possible reasons that are explained in this section. Influenced by Russel's paradoxon of set theory, Lesniewski developed mereology which uses instead of the element relation of set theory the part-whole relation as basic relation. His terminology includes 'part', 'disjoint' (two items do not share a part), and 'sum', in analogy to the terms 'element', 'disjoint' (having no intersection), and 'union' of set theory. Linguists only criticize his axiom of transitivity (his part-whole relation is transitive), which according to their opinion does not hold for all kinds of meronymy (see below). It seems that also some of Lesniewski's other axioms and theorems are not valid for natural language meronymy. He states that an item is uniquely defined by its parts. This may be true for the membership meronymy (a tennis club consists exactly of its members), but it does not hold for the so called 'functional meronymy' (Iris et al. 1988), because an item which consists of all parts of a car does not have to be a car. It could be assembled in a completely unusual manner and be a work of art in a museum. According to Lesniewski the sum of two items is always another item. This is at least not valid for the lexicalization of geographical meronymy relations. For example, the geographical unit that contains two cities is usually a county, state or continent, and so on. That means that the mereological sum of two cities should contain more than just the areas of the cities. Furthermore this kind of sum would not be unique. Different pairs of cities can be part of the same geographical unit. Therefore it seems that Lesniewski's sum should be replaced by a closure operator which yields the next smallest unit that contains the parts. But even using a closure operator, the sum need not be uniquely described by its parts in geographical meronymy relations because, for example, Monaco is a city and a country at the same time. Therefore the area of the city Monaco equals the area of the country Monaco, but the country has the city as a part, hence has more parts than the city. Mereology seems to be appropriate for portion-mass meronymy, for example every two lumps of mud can be joined to a bigger lump of mud. But even this is not reflected in the language: there are no distinct words for 'small lump of mud' and 'bigger lump of mud'. For these reasons Lesniewski's mereology does not seem to be a good modeling of natural language meronymy relations.

3.4 A formal definition of meronymy

Although meronymy is a hierarchical relation it should not be modeled as a mathematical lattice. One obvious reason for not modeling meronymy as a concept lattice using denotata as formal objects and attributes and meronymy as the relation between them is that, for example, the formal attributes 'ketchup' and 'pizza' share the formal objects 'sugar' and 'salt' as parts. Therefore a formal concept 'salt, sugar' would evolve, but 'salt, sugar' is usually only a mixture and not a denotative word concept itself in the English language. Such a concept lattice would therefore provide an embedding of meronymy, but not all concepts would have useful interpretations. A better solution is therefore to use part-whole relations as attributes, such as 'has handle as part' which would, for example, differentiate a cup from a glass. A third option is to interpret meronymy as an additional relation besides the conceptual ordering. This is done in the following definition.

Definition 3.3:
In a denotative structure $\mathcal{S}_D$ the semantic relation meronymy is defined as follows: Two disambiguated words are in *meronymy relation* if their denotative word concepts are in relation $R^m_{(Q^4;Q^2)}$ where m is a meronymy relation among denotata, i.e.

$$w_1 \ MER^m_{(Q^4;Q^2)} w_2 :\Longleftrightarrow dnt(w_1) R^m_{(Q^4;Q^2)} dnt(w_2)$$

and the meronymy relation m is irreflexive, antisymmetric, and acyclic.

From the definition follows that $MER^m_{(\geq 0;\geq 1)}$, $MER^m_{(\geq 1;\geq 0)}$, and $MER^m_{(\geq 1;\geq 1)}$ are also irreflexive, antisymmetric, and acyclic (see Section 2.4). And if m is transitive, then $MER^m_{(\geq 0;\geq 1)}$, $MER^m_{(\geq 1;\geq 0)}$, and $MER^m_{(\geq 1;\geq 1)}$ are also transitive (Section 2.4). In contrast to antonymy whose types are distinguished by the relational components (Section 3.8), many types of meronymy differ in their quantificational tags which can therefore be used for a rough classification of meronymy. For example, the four kinds of meronymy relations described by Cruse (1986) consist of combinations of the basic quantifiers $||\geq 1||$, $||\geq 0||$, and $||\text{all}||$:

- MER^m_0: facultative-facultative[16]; for example, a child can be a member of a tennis-club, but not all children are members of tennis-clubs, nor do all tennis-clubs have children as members.

- $MER^m_{(\geq 0;\geq 1)}$: canonical-facultative; for example, all door-handles are parts of doors, but not all doors have to have handles.

- $MER^m_{(\geq 1;\geq 0)}$: facultative-canonical; for example, all ice-cubes consist of water, but not all water is frozen.

- $MER^m_{(\geq 1;\geq 1)}$: canonical-canonical; for example, each bird feather is part of a bird, and each bird has feathers.

[16]Cruse uses 'facultative' and 'canonical' instead of Lyons' (1977) 'contingent' and 'necessary'.

3.5 Transitivity and inheritance of meronymy

The question of transitivity of meronymy has been widely discussed (Winston et al., 1987). Logicians often claim that meronymy is transitive because they use models such as Lesniewski's mereology which are defined to be transitive. Some linguists (Iris et al. 1988) decide that certain types of meronymy are transitive whereas others are not transitive. Functional meronymy is according to them not transitive. We are not going to investigate which types of meronymy might be transitive and which are not, but from Relational Concept Analysis follows that if m is transitive then certain meronymy relations R^m are also transitive. This means that if the relation m is transitive on the denotative level, then meronymy is also transitive on the conceptual level for these kinds of meronymy relations. As the pure spatial inclusion on the object level seems to be always transitive (a particular door-handle is part of a particular door which is part of a particular house), it follows that meronymy is often not a conceptual extension of the spatial inclusion, but that it has other features, such as, for example, functional dependencies, which are not transitive. (The concept of 'door-handle' does not include 'in general having a function for a house.') Relational Concept Analysis facilitates a more detailed analysis of where features occur (on the denotative or on the conceptual level) and can help to show inconsistencies in a linguistic model, such as to assume m is the spatial inclusion for an intransitive relation $MER^m_{(\geq 0;\geq 1)}$ is contradictory.

Winston et al. (1987) claim to have discovered the solution of the transitivity question for meronymy. The rest of this paragraph, however, shows that their theory is not entirely correct. They claim that syllogisms of the kind 'a is part of b, b is member of c therefore a is part of c' are valid if the three meronymy relations are of the same type. The example syllogism is not valid because 'part of' and 'member of' are not the same type of meronymy relation. Furthermore 'mixed inclusion relation syllogisms are valid if and only if the conclusion expresses the lowest relation appearing in the premises, where the ordering of relations is: class inclusion > meronymy > spatial inclusion'. For example, wings are parts of birds, birds are creatures, therefore wings are parts of creatures (and not: wings are creatures). We think it is pure coincidence that Winston et al.'s statement is actually true for all the examples they use in their paper. All syllogisms contain implicit quantifications and from these quantifications, rules follow which seem to be more substantial than Winston et al.'s theory. A sentence 'a is part of b' is usually interpreted as 'all a are parts of some b' since it can be observed in general that the subject of a sentence tends to imply the $||$all$||$-quantifier whereas other parts of the sentence tend to imply $|| \geq 1||$-quantifiers if quantifiers are not explicitly mentioned. From 'all a's are parts of some b and b's are c's' follows 'all a's are parts of some c' according to the rule that a $|| \geq 1||$-quantifier causes inheritance to superconcepts (Section 2.3). Winston et al.'s other examples are of the following types:

(All) pies are desserts.	All a are b.
(All) desserts are partly sugar.	All b have a part c.
(All) pies are partly sugar.	All a have a part c.

Wings are part of birds.	Socrates is in Athens.	All a are related to some b.
Birds are creatures.	Athens is a city.	All b are c.
Wings are part of creatures.	Socrates is in a city.	All a are related to some c.

The wheel is part of the bike.	All a are part of b.
The bike is in the garage.	All b are in c.
The wheel is in the garage.	All a are in c.

The first syllogism is true because ||all||-quantifiers cause inheritance to subconcepts. The second syllogism is the type we already discussed. Individual concepts, i.e. concepts which have only one denotatum, always fulfill the ||all||-quantifier. The third is based on the assumption that this type of meronymy implies spatial inclusion (which is not true in general) and that spatial inclusion is transitive. The following counter examples show that Winston et al.'s examples are correct, but their reasoning is faulty. First, since they use the meronymy relation in both directions as 'is part of' and 'is partly', their theory should also be valid if class inclusion is used in both directions. 'Includes' expresses the inverse relation to the class inclusion 'is a'. Therefore from Winston et al.'s statement would follow: Creatures include birds, birds have wings, therefore all creatures have wings. This is obviously not true. Second, other quantifiers can be used: All pies are kinds of desserts, some kinds of desserts are partly honey, therefore all pies are partly honey. This is also not true. Third, a meronymy relation that does not entail spatial inclusion can be used: The character is part of the book, the book is in the garage, therefore the character is in the garage? Obviously, analysing the implicit quantifications and using the rules for inheritance among relations according to Relational Concept Analysis (Chapter 2) provides a better solution to the transitivity questions than Winston et al.s theory.

3.6 The classification of meronymy

Classifications of semantic relations are usually based on qualitatively different attributes of the relations. The most detailed classifications of meronymy are probably achieved by Winston et al. (1987), Chaffin & Herrmann (1988), and Chaffin (1992). They use the method of 'Relation Elements' which characterizes a relation in terms of its defining features, the so-called 'relation elements'. Certain dependencies among relation elements can be expressed and statistical methods allow a computation of the similarity of meronymy relations to each other. Using Formal Concept Analysis, the meronymy relations can be taken as formal objects, and the relation elements as formal attributes of a formal context. The dependencies among relation elements and the similarity of the types of meronymy relations can be studied in a concept lattice. Typical relation elements are: functional (there is a function between part and whole), homogeneous (the parts are similar to each other), homeomerous (the parts are similar to each other and to the whole according to their substance, such as a piece of cake and a cake. 'Homeomerous' is a special case of 'homogeneous'), separable (the part can be separated from the whole), attached (the part is attached to the whole, a special case of separability), locative or spatial inclusion (the part is 'in' the whole), social (the parts form the whole according to a social agreement), and so on. Cruse (1986) distinguishes between constituents and ingredients. These can also be used as relation elements. Constituents are parts that are found by analysing the whole (in German: Inhaltsstoffe), whereas ingredients are used to synthesize the whole (in German: Zutaten). Collections are usually synthesized, for example, a forest would not exist without trees.

	inclus.	part.	attach.	compon.	prop.	possess.	homog.	soci.	locat.
funct. object	X	X	X	X	X	X			
collection	X	X			X	X	X		
group	X	X			X	X	X	X	
ingredient	X	X		X	X	X			X
funct. location	X	X	X	X	X	X			
organization	X	X	X	X	X	X		X	
measure	X	X					X		
place	X	X				X			X

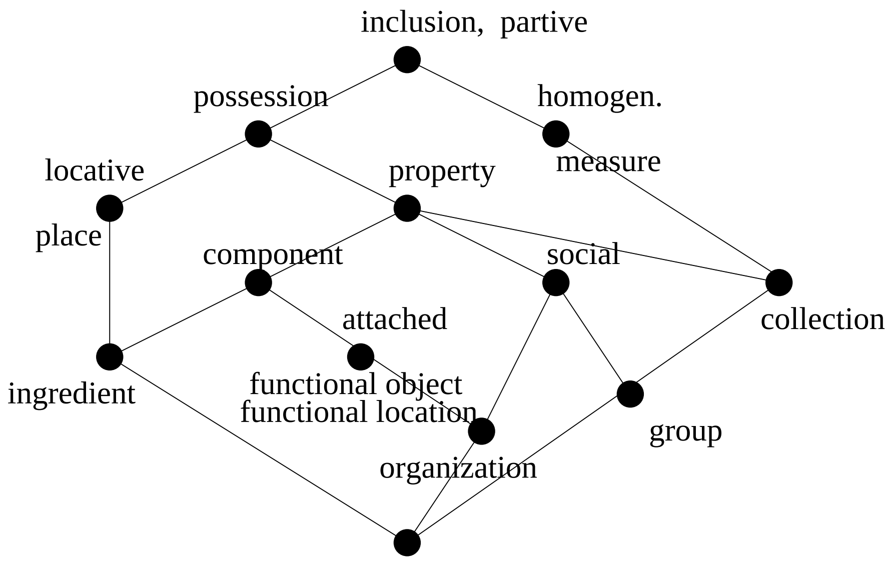

Figure 3.1: Chaffin & Herrmann's classification of meronymy

Chaffin & Herrmann (1988) base their classifications not primarily on lexicalizations of the meronymy relation (such as 'is partly' expresses a substance meronymy), but on psychological experiments, such as letting a person classify, compare, or distinguish several tokens of meronymy (and other) relations. Their experiments result in several similar classifications. In Figure 3.1 the classification of meronymy via relation elements which Chaffin & Herrmann base on Stasio et al. is modeled as a concept lattice. Prototypical relation elements of meronymy relations seem to be 'inclusion', 'partive', 'property', and 'possession'. 'Place' and 'measure' do not appear to be typical meronymy relations. In the lattice the typical meronymy relations are subconcepts of the attribute concept of 'property'. 'Place' and 'measure' are not subconcepts of this concept. 'Group' is a social 'collection', since it shares all attributes with 'collection' and has the differentiating attribute 'social'. It seems that 'functional object' and 'functional location' are not properly distinguished. Maybe a relation element 'functional' should be added and the element 'locative' should be assigned to 'functional location'. A similar classification is given in Winston et al. (1987) as: The most dominant meronymy relation is the component/integral object or functional object meronymy (Iris et al. (1988) call this 'functional component'), such as 'cup/handle'. The other classes are feature/activity (such as 'bride/wedding'), member/collection ('tree/forest'), place/area ('Everglades/Florida'), portion/mass ('slice/pie'), stuff/mass ('lettuce/salad'), and stuff/object ('aluminium/bike'). According to Chaffin & Herrmann (1988) these basic classes can be further divided, for example, member/collection meronymy includes unit/organization ('delegation/UN') and member/group ('cow/herd') meronymy.

It seems that a universal classification based on relation elements cannot be obtained because it seems to be difficult for linguists to agree upon a common set of relation elements. For example, the classifications in the papers by Winston, Chaffin & Herrmann (1987) and Chaffin & Herrmann (1988) contradict each other. According to Winston at al. the place/area meronymy is homeomerous, but according to Chaffin & Herrmann it is not homogeneous, therefore it cannot be homeomerous. An experimental analysis using all tokens of part-whole relations from WordNet and assigning relation elements to them (instead of assigning the relation elements to types of relations) could lead to a more convincing classification and could be the basis for future research. As an example for such an approach, in this paper the substance meronymy which has only about 400 tokens in WordNet is analysed using relation elements (see below). Classifications of meronymy relations based on quantificational differences (or quantificational tags) as provided by Relational Concept Analysis appear to be easier to be constructed than classifications based on qualitative differences (which are contained in the relational component m). The rest of this section shows that these quantificational differences often coincide with qualitative differences, although they are more elaborated in some cases and less detailed in other cases. It seems that quantificational tags have the advantage of being less subjective because it is easier to decide which quantifiers a relation needs instead of deciding which relational component it has. Figure 3.2 shows an attempt to develop such a classification of the meronymy relation. The table is probably not complete, but five major classes can be observed. In addition to the quantifiers from chapter 2, the $||\text{some}_1||$-quantifier is used for mass nouns and denotes the singular sense of 'some', such as 'there is some bread' to distinguish it from the plural sense, such as 'there are some people'. In German these two senses of 'some' have different translations: 'some$_1$' means 'etwas' whereas 'some$_2$' means 'einige'. In English this distinction is reflected in 'much' and 'many'. $|| \leq \text{some}_1||$ means there are none or some, but never all ('a pizza contains some meat or no meat at all'). $||\text{some}_1||$ means there are exactly some, and not none or all. $|| \geq \text{some}_1||$ means there are some or all, but this quantifier is not used for meronymy. ||several|| denotes the analogous quantifiers for a collection of objects, such as 'a book contains several chapters'. The objects of a ||several||-quantifier are always interchangeable, therefore this quantifier corresponds to Chaffin & Herrmann's (1988) relation element 'homogeneous'. It should be noted that all examples are always to be understood in a prototypical sense: 'a prototypical sausage contains some meat', and so on.

relational component	tag	example
stuff/object	(some$_1$; $\leq$1)	meat/sausage
	(some$_1$; 1)	sausage meat/sausage
	($\leq$ some$_1$; $\leq$1)	meat/pizza
stuff/mass	(some$_1$; $\leq$ some$_1$)	salt/seawater
	(some$_1$; some$_1$)	sea salt/seawater
	($\leq$ some$_1$; $\leq$ some$_1$)	sage/tea
	($\leq$ some$_1$; some$_1$)	sausage meat/food
element/mass	(several; $\leq$ some$_1$)	body cells/skin
	(several; some$_1$)	body cells/body tissue
element/mass; portion/mass	($\leq$ several; some$_1$)	skin cells/body tissue; slice/bread
member/set	(several; $\leq$1)	tree/forest
	(several; $\geq$0)	human/citizenship
member/set; section/object	(several; 1)	human/sex; chapter/book
unit/measure; memb./set; obj./obj.	(n; 1)	sec./hour; card/deck; finger/hand
object/object	($\geq$0; 1)	refrigerator/kitchen
	(1; $\geq$0)	melody/song
obj./obj; individual/individual	(1; 1)	punch line/joke; Princeton/NJ

Figure 3.2: A classification of meronymy based on quantificational tags

The distinctions within the five classes show that the quantificational tags depend on the level of abstraction of the concepts. For example, some meat can be contained in a sausage, whereas 'sausage' meat (at least prototypically) has to be contained in a sausage, otherwise it would not be 'sausage' meat. Substituting a general part or whole by a more specific part or whole often changes the quantifiers. For a basic classification only the type of quantifier is essential, such as ||some|| or ||several||, and not the modifier, such as 'at least', 'exactly', or 'at most'. Although meronymy relations with different relational components can share the same tags, each class of relational components tends to prefer a prototypical tag. Therefore the tags can be used as the basis of a classification. For example, for the object/object meronymy, $MER^m_{(\geq 0;1)}$ seems to be the dominant form, because each object is part of exactly one whole. For example, a door-handle belongs to exactly one door. $MER^m_{(\geq 0;n)}$, for $n > 1$, is also possible, but this quantifier is rare. For example, any section of a border belongs to exactly two countries. A prototypical feature of the member/set meronymy is that $MER^m_{(\geq 0;>1)}$ is possible, because, for example, a human is usually a member of different sets (clubs, family, cultures) at the same time. For individual concepts $MER^m_{(1;1)}$ always holds, because the ||all||-quantifier becomes trivial if the extent of a concept has only one element. Our classification coincides partly with the classifications of Chaffin & Herrmann (1988), for example as already mentioned, the ||several||-quantifier corresponds to the relation element 'homogeneous'. The relation element 'attached' requires that the whole is a physical object (quantifiers '||1||' or '||n||'), and the relation element 'separable' does not apply to parts that use the ||some||-quantifier. Obviously, for the object/object relations, the classification based on quantification seems to be the most unsatisfactory. For example, Chaffin & Herrmann's component/integral object, topological part/object, time/time, and place/area are all subsumed under object/object.

	artific.	natur.	const.	ingred.	hom.	f,st,mo	funct.
cell/tissue		X	X		X		
chalk/limest.		X	X		X		
ice/ice cube		X	X		X	X	
water/ice		X	X		X	X	
air/wind		X	X		X	X	
caff./coffee	X		X				
flour/dough	X			X			
flour/bread	X			X			
wheat/flour	X			X	X	X	
dough/cake	X			X	X	X	
gold/dental g.	X			X	X		X

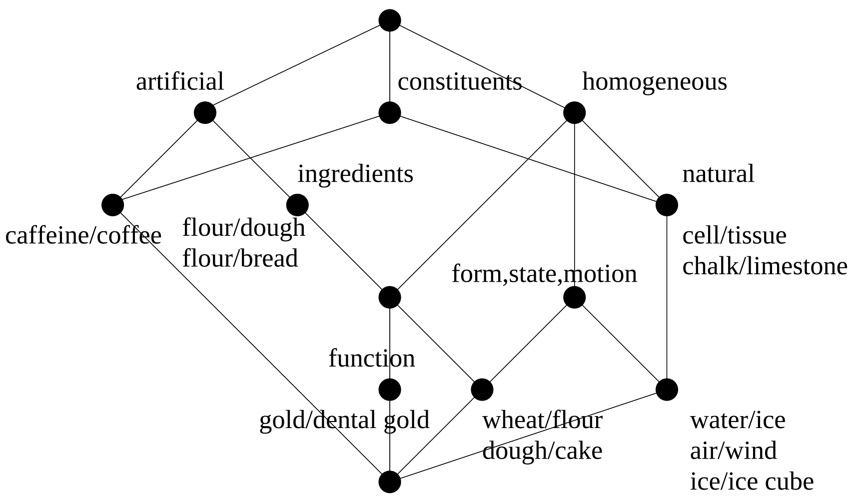

Figure 3.3: A classification of substance meronymy

Iris et al., (1988) develop four models of meronymy which also partly coincide with our classification. Their first model is the functional component meronymy which corresponds to our first and fifth class. Their second model is the member/collection meronymy which is contained in our fourth class. Their third model is the segmented whole meronymy which partly coincides with our second and third class. Their fourth model, the subset/set relation seems to be not a proper meronymy relation. WordNet (Miller et al., 1990) distinguishes only three types of meronymy relations: substance, membership, and part. Our first three classes are subsumed under substance meronymy, our fourth class is WordNet's membership meronymy, and our fifth class is their is-part-of meronymy. Under a pragmatic viewpoint the WordNet classification is efficient since there are less tokens in WordNet that would belong to our first three classes (about 400 tokens), than tokens in the last class alone (about 5000 tokens). The WordNet numbers can be misleading because the is-part-of meronymy contains many geographical proper names and the membership meronymy (about 12000 tokens) contains the biological classification (a species is a member of its genus), which would usually not be considered a meronymy relation.

As mentioned above, we studied a classification based on relational components for the substance meronymy in WordNet. All tokens of this meronymy relation in WordNet have been automatically grouped according to their hypernyms and then each group was manually assigned relation elements. Substance meronymy exclusively occurs among concrete nouns. It dominates in the areas: building materials, fabrics, chemical products, human body, nutrition, medicine, weather and geological formations. The following relation elements seem useful: natural (the whole is a natural item), artificial (the whole is manufactured), homogeneous, change of form, state, or motion (the whole is the same substance as its part but has a different form, such as 'ice/ice cube', is in a different state, such as 'water molecules/ice', or motion, 'air/wind'), change in function (the whole is the same substance as its part but has a different function, 'ascorbic acid/vitamin C'), ingredients, and constituents. After grouping the tokens of the substance meronymy in WordNet according to these relation elements, one token that seemed to be prototypical was chosen for each group. These tokens are the formal objects of the formal context and the concept lattice in Figure 3.3. Six basic classes can be observed. The parts of natural wholes are always constituents. Parts can be found by analysing the whole (class 1: chalk/limestone, cell/tissue) or they are distinguished from the whole by a change of form, state or motion (class 2). Parts of artificial wholes can be constituents (class 3: caffeine/coffee) or they are ingredients. Ingredients are non-homogeneous (class 4: flour/bread, flour/dough) or homogeneous. Homogeneous parts of artificial wholes are distinguished from the whole by a change in function (class 5: gold/dental gold), or a change in form, state or motion (class 6). The classifications in Figures 3.2 and 3.3 can be combined. For example, the elements of class 1 in Figure 3.3 are element/mass- (cell/tissue) and stuff/mass-meronymy relations (chalk/limestone). It is not claimed that the classification in Figure 3.3 is complete. Obviously, other authors could select other relation elements or tokens of the meronymy relation. But using Formal Concept Analysis seems to be a very efficient method of obtaining, graphically representing, and comparing classifications.

The other two meronymy relations in WordNet have not been analysed as detailed as the substance meronymy. The membership meronymy in WordNet contains mainly humans (or other individuals) and their groups (members of religious groups, countries, organizations, families, geographical areas, committees). Furthermore, it contains sections of institutions ('school/university'), collections of similar items for some purpose ('card/deck', 'restaurant/chain', 'star/constellation'), geographical or political units ('Belgium/NATO') and social units as members of associations. The is-part-of meronymy in WordNet consists of time units ('month/year'), geographical or political units ('city/country'), body parts or parts of plants, other units ('cent/dollar'), and others.

3.7 Irregularities in the implementation of meronymy in WordNet

Properties of semantic relations can be used to identify irregularities in the relations of a lexical database or thesaurus. Rules can be implemented as a computer program and then be automatically tested. Fischer (1991) has written a Smalltalk program system to check some mathematical properties of semantic relations, such as inverse relations, circularity, implicit relations, and so on. It would be possible to implement the rules which are implied by Relational Concept Analysis in a similar way, but this has not been undertaken so far. Some irregularities can be corrected automatically. For example, if Fischer's software detects a relation which should be symmetric, but is implemented as a unidirectional pointer, the other direction can simply be added. In many cases, however, it is not possible to correct automatically the irregularities. Irregularities can be detected, but then lexicographers are needed to decide which concepts or relations have to be added or changed to solve the problems. Three examples of the meronymy relation in WordNet are chosen to demonstrate the possibilities of Relational Concept Analysis. WordNet distinguishes only 'part-of', 'substance-of', and 'member-of' meronymy, but not the quantificational tags, such as 'MER^m_0' and '$MER^m_{(\geq 1;\geq 1)}$'. But because comparatively few meronymy relations are implemented in WordNet, a first approach assumes all of them to be of the strongest kind, $MER^m_{(\geq 1;\geq 1)}$. If irregularities are found, they can be changed to weaker kinds such as '$MER^m_{(\geq 1;\geq 0)}$' or be otherwise repaired.

Figures 3.4, 3.6, and 3.8 show parts of the WordNet1.5 lattice, Figures 3.5, 3.7, and 3.9 demonstrate how they could be improved. The examples from WordNet are not complete, because some relations are omitted and only one or two representative words are selected for each synset. The dotted lines represent meronymy, the others hyponymy. In the first example in Figure 3.4 'human body $MER^m_{(\geq 1;\geq 1)}$ person' holds, therefore a child's body and an adult body must also be part of a person. 'Flesh', which does not follow that pattern, seems to be misplaced as a subconcept to 'human body'. If, furthermore, 'female body $MER^{m}_{(\geq 1;\geq 1)}$ female' holds, then, likewise, a woman has to have a female body; therefore 'woman's body' should be a subconcept of 'female body.' It should be noted, though, that not all of the changes from Figure 3.4 to Figure 3.5 can be derived from the theoretical properties of the relations only. In most cases additional semantic knowledge is needed that can be provided only by lexicographers.

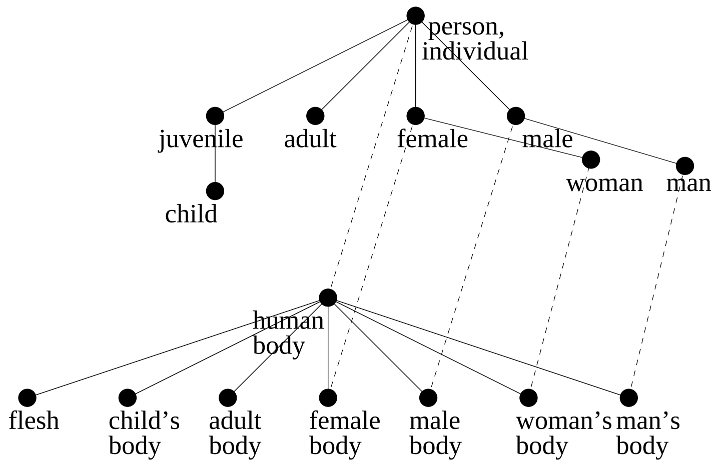

Figure 3.4: An example of the part-of relation in WordNet

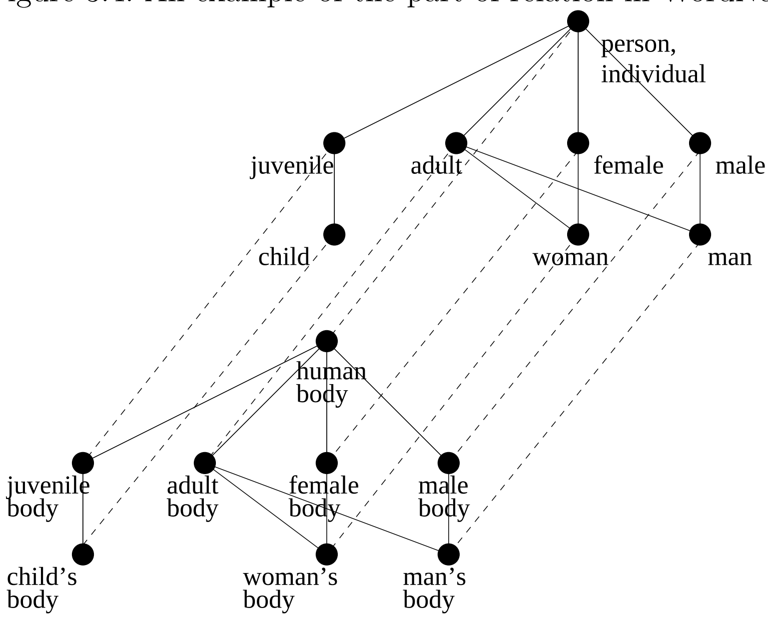

Figure 3.5: A modified version of the example in Figure 3.4

The reason for the irregularities in Figure 3.6 is probably the polysemy of 'extremity', because 'hand' and 'foot' are subconcepts of the wrong 'extremity' concept. The 'extremity' concept with the meaning 'hand and foot' should be a subconcept of 'extremity, appendage'. The meronymy relation is irregular in this example, because,

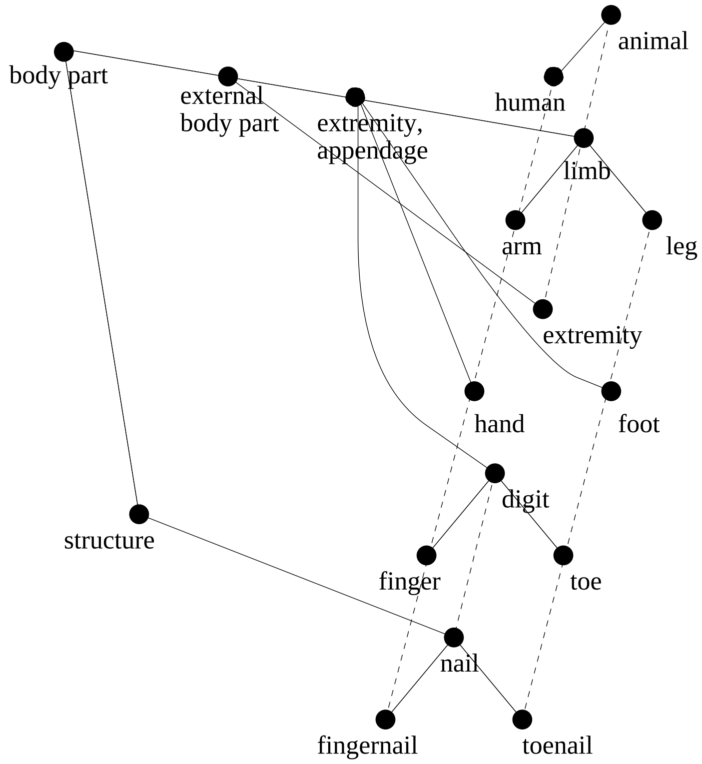

Figure 3.6: Another example of the part-of relation in WordNet

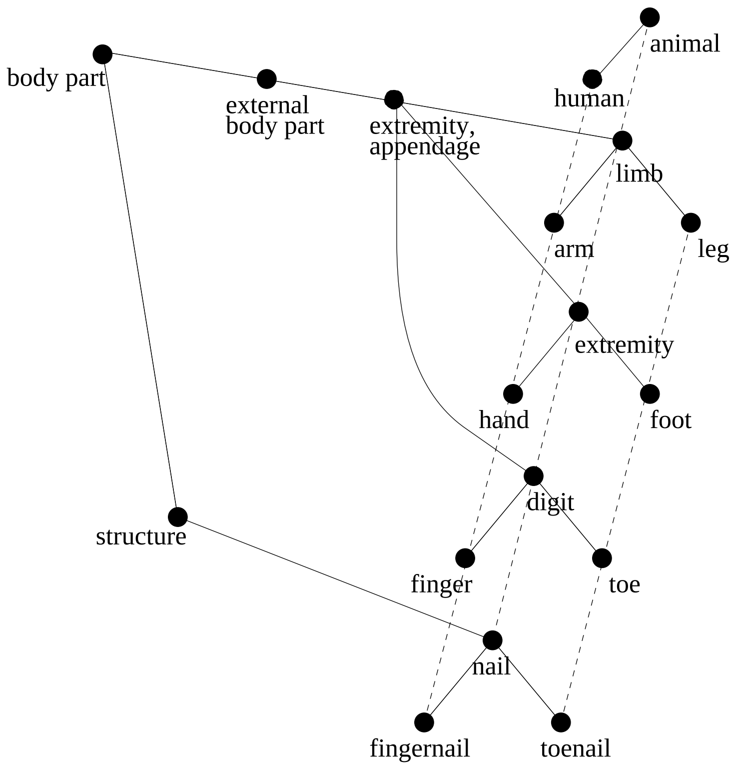

Figure 3.7: A modified version of the example in Figure 3.6

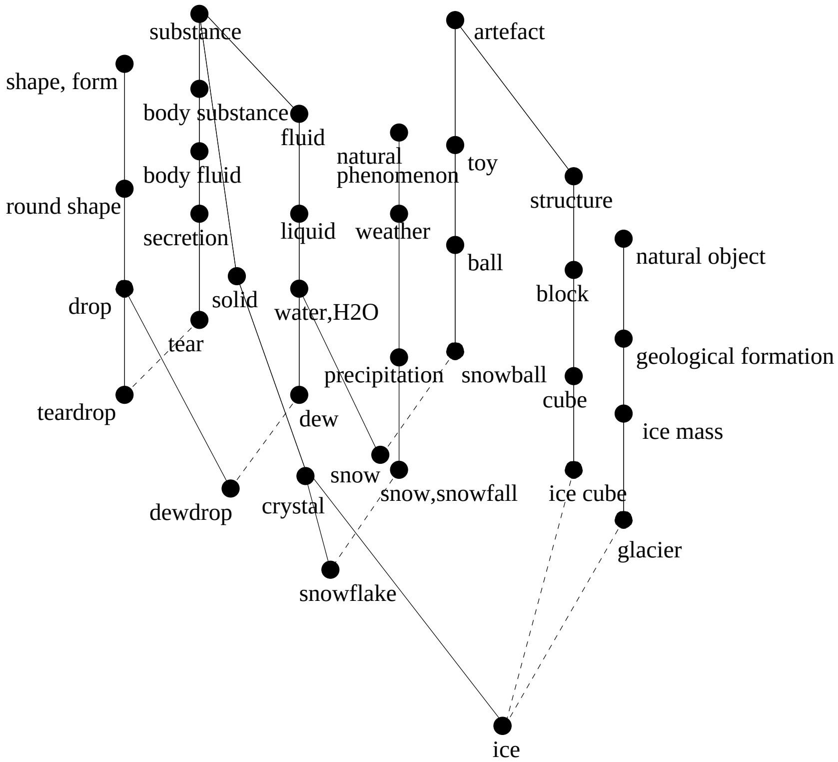

Figure 3.8: An example of the substance-of relation in WordNet

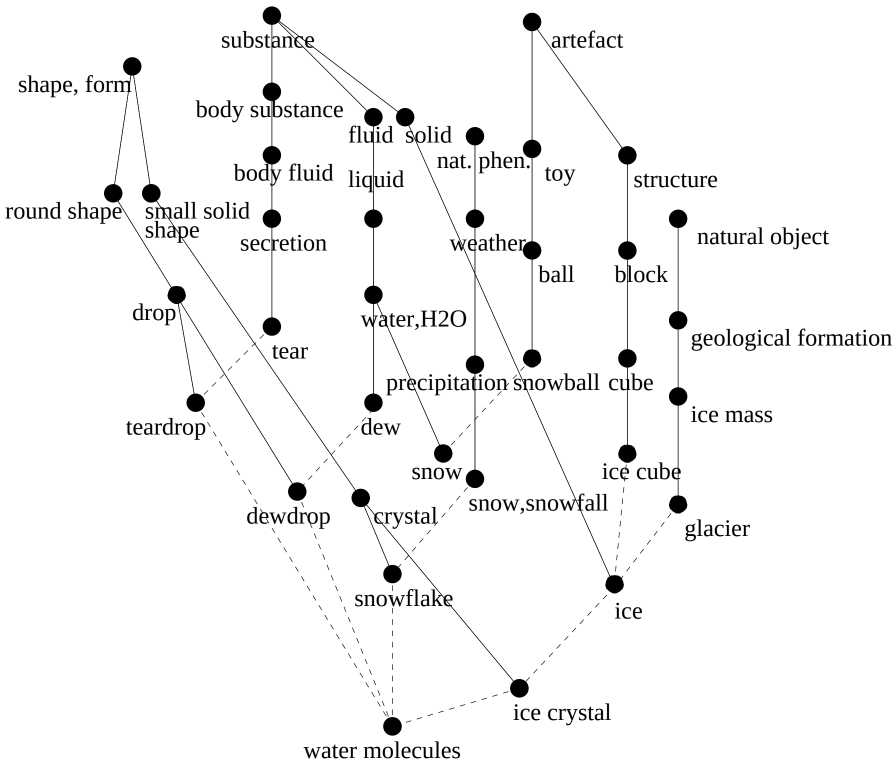

Figure 3.9: A modified version of the example in Figure 3.8

if there are no digits other than fingers or toes, and if all fingers or toes are part of some concepts that have a common hypernym 'extremity', then digits should be part of extremities. The corrected version in Figure 3.7 shows a more regular pattern than Figure 3.6.

The last example of the substance-of meronymy in Figure 3.8 probably needs improvement, too, as all those fluids should have a common substance – water molecules (see Figure 3.9). Similarly a distinction should be made between ice crystals and ice. This example does not contain further irregularities, but it shows a certain pattern that can be discovered by comparing the hypernyms at different levels. It seems to be a property of the substance meronymy that shapeless and shaped forms alternate. Drops and crystals are small shapes. On the next level, 'tear', 'dew', 'snow' and 'ice' are shapeless nouns with the hypernym 'substance'. On the last level, objects are shaped again, but this time the shape is formed by humans ('artefact') or nature ('geological formation'). This last example shows how meronymy and hyponymy may form a regular pattern in some areas of the vocabulary. A more complete analysis of all the relations in WordNet will probably reveal more patterns, which can ultimately be formalized as properties of special relations.

3.8 Other semantic relations: contrast and sequences

Contrast relations are special cases of sequence relations because they are 'binary sequences', i.e. sequences of two elements. Chaffin & Herrmann (1988) distinguish the following kinds of contrast relations: contrary or gradable antonyms ('tiny/huge', a tiny item cannot be huge and vice versa, but there are steps in-between: 'tiny/small/-medium/large/huge'); contradictory antonyms ('dead/alive', an item is either dead or alive); asymmetric contraries ('dry/wet', contraries that are not equally distant from the medium value); incompatible terms ('happy/morbid'; an item cannot be happy and morbid at the same time, but happy and morbid are not usually considered to be antonyms); reverse terms ('attack/defend'); directional terms ('front/back'); pseudo-antonyms ('popular/shy'); and 'terms with similar attributes' ('painting/movie'). Miller at al. (1990) emphasize that antonymy is a lexical relation and not a concept relation, because it is a relation between disambiguated words which does not have to hold for the synonyms of the words. For example, 'happy/sad' are antonyms but not 'happy/melancholic'. Miller et al. call the concept relation that holds between the synonyms of antonyms 'indirect antonymy'. Antonymy is clearly a relation among disambiguated words (a lexical relation) and not among words (a morpho-lexical relation), because the different polysemous senses of a word can correspond to different antonyms. In Roget's International Thesaurus (RIT, 1962) adjacent categories are often antonymous to each other showing the different antonyms of a word according to its polysemous senses. For example, the categories: '665 misuse', '666 disuse', '667 uselessness' are antonyms to different senses of '663 use', such as 665 use right/use wrong, 666 continue using/stop using, 667 useful/useless. Contrasts can occur in all parts of speech, but verbs and prepositions represent a special case. Verbs and prepositions can be formal objects of a lexical context and therefore can have the same types of semantic relations among each other (including antonymy) as nouns. For example, 'grow' is in contrast to 'shrink' in the same way as 'big' is to 'small'. But verbs and prepositions can also be semantic relations themselves. Some of these semantic relations can be inverted (such as 'A teaches B' versus 'B learns from A'). The contrast relation between 'teach' and 'learn from' is a relation between relations represented by verbs. As another example, 'before' can be a relation between 'morning' and 'evening' with the inverse relation 'after' and a contrast relation between 'before' and 'after'. Therefore there are two types of relations among verbs and prepositions: semantic relations among the denotative word concepts of verbs and prepositions and 'meta-relations' among semantic relations that are denoted by verbs or prepositions. (In rare occasions nouns can also be semantic relations themselves such as 'A is brother of B' and 'B is sister of A'.) Chaffin & Herrmann's types 'reverse' and 'directional' represent 'meta-relations' whereas their other types are usually ordinary semantic relations. Prepositions and verbs that are interpreted as semantic relations themselves are studied in the next section.

Contrast relations are different from meronymy in that they seem to be based on attributes instead of objects. In analogy to Definition 2.1 concept relations based on denotative relations among attributes are defined in Definition 3.4. It is not the purpose of this paper to discuss antonymy, the prototypical contrast relation, in detail. Several definitions of antonymy can be found in the literature (compare, for example, Miller et al. (1990)). Usually the differences between the types of antonyms are expressed by logical formulas. For example, a and b are defined to be contrary if $a \Rightarrow \neg b$ and $b \Rightarrow \neg a$; a and b are defined to be contradictory if $a = \neg b$ and $b = \neg a$, and so on. Nevertheless these logical properties are usually not sufficient conditions for antonymy. In a concept lattice a negation operator '$\neg$' among attributes, such as in $a = \neg b$ for attributes a and b would not have to correspond to the '$\neg$'-operator in a Boolean lattice. Otherwise a contradictory antonymy relation between attributes a and b would require that every object that does not have the attribute a has the attribute b. The status 'a and b are both irrelevant for an object' could not be expressed. An antonymy relation among attributes in a formal context can therefore not be deduced from the relations in the lattice, but has to be defined as a denotative relation among attributes. This is done in Definition 3.5 for a general contrast relation. The definition is not always appropriate for indirect antonymy since word concepts can be included which usually are not considered to be indirect antonyms. Fischer et al. (1996) recommend the following definition of antonymy: 'When two concepts are opposed, there will be a common ancestor with respect to hypernymy, and the concepts found in the different chains up to this nearest common ancestor will be opposed to each other'. It is true that antonyms often have more attributes in common than distinguishing attributes, but in a formal context each pair of concepts has a common hypernym and defining a 'nearest common' hypernym would require some kind of measure on the lattice. Therefore Fischer's definition cannot easily be realized in a concept lattice. So there are still open questions concerning the modeling of antonymy.

Definition 3.4:
For a context (G, M, I), concepts $c_1, c_2 \in \mathcal{B}(G, M, I)$, a relation $r \subseteq M \times M$, and quantifiers Q^i, $1 \leq i \leq 4$,

$$c_1 \ \overline{R}^r[Q^1, Q^2;] \ c_2 :\Longleftrightarrow Q^1_{m_1 \in Int(c_1)} Q^2_{m_2 \in Int(c_2)} : m_1 r m_2 \tag{27}$$

$$c_1 \ \overline{R}^r[; Q^3, Q^4] \ c_2 :\Longleftrightarrow Q^3_{m_2 \in Int(c_2)} Q^4_{m_1 \in Int(c_1)} : m_1 r m_2 \tag{28}$$

$$c_1 \ \overline{R}^r[Q^1, Q^2; Q^3, Q^4] \ c_2 :\Longleftrightarrow c_1 \ \overline{R}^r[Q^1, Q^2;] \ c_2 \text{ and } c_1 \ \overline{R}^r[; Q^3, Q^4] \ c_2 \tag{29}$$

are defined as concept relations.

Definition 3.5:
In a denotative structure $\mathcal{S}_D$ the following semantic relation is defined: Two disambiguated words are *in contrast* if there is a contrast relation x defined on attributes (i.e. $x \subseteq A_D \times A_D$) that holds among at least one pair of attributes of their denotative word concepts, i.e.

$$w_1 \ CTR^x w_2 :\Longleftrightarrow dnt(w_1) \overline{R}^x_0 dnt(w_2)$$

and the contrast relation x is a binary, symmetric, denotative relation. Contrast relations are *contrary*, *contradictory*, *incompatible*, or *asymmetric contrary* if the relation $x \subseteq A_D \times A_D$ is contrary, contradictory, incompatible, or asymmetric contrary[17].

In contrast to meronymy where the different kinds of relations depend on the quantifiers, the different kinds of contrasts depend solely on the relational component x. All contrast relations are of the type $\overline{R}_0$. This means that one pair of attributes is enough to create a contrast relation among disambiguated words. Obviously, a contrast relation is always inherited to subconcepts. Sequence relations can be defined analogously to Definitions 3.5 or 3.3 depending on whether they require relations on objects or attributes for their construction.

3.9 Verbs and prepositions as semantic relations

Verbs and prepositions can be semantic relations themselves. For example, 'teach' is a relation $TEACHTO_{(\geq 1; \geq 1)}$ between the concept 'teacher' and the concept 'student'. Its inverse relation (formally defined in Definition 3.6) is $LEARNFROM_{(\geq 1; \geq 1)}$. The inverse relation of a verb does not have to be lexicalized. Often the inverse relation of a verb is its passive voice ($TAUGHTBY_{(\geq 1; \geq 1)}$). Formally inverse relations are defined as follows:

---

[17]Sometimes necessary and/or sufficient conditions can be given to decide whether a relation x is contrary, contradictory, incompatible, or asymmetric contrary. But that depends on the context. The following condition is usually necessary: $a_1 x a_2 \Longrightarrow \forall_{d \in D} \neg (d I_D a_1 \text{ and } d I_D a_2)$ for $a_1, a_2 \in A_D$.

Definition 3.6:
Two semantic relations $R^1_{(Q^4;Q^2)}$ and $R^2_{(Q^2;Q^4)}$ are inverse to each other (denoted by $R^1 = (R^2)^{-1}$) if

$$\forall_{w_1,w_2 \in W} : w_1 R^1_{(Q^4;Q^2)} w_2 :\Longleftrightarrow w_2 R^2_{(Q^2;Q^4)} w_1$$

Teaching and learning usually involve a teacher, a student and a subject. Either the student or the subject can be missing for 'teach': 'She teaches him. She teaches English. She teaches him English.' The teacher or the subject can be missing for 'learn': 'He learns English. He learns from her. He learns English from her.' The best modeling of these two verbs would probably include ternary relations and three quantifiers. In our binary modeling, '$LEARNFROM_{(\geq 1;\geq 1)}$' and '$TEACHTO_{(\geq 1;\geq 1)}$' are inverse to each other, because they both require a teacher and a student and have the same quantifiers Q^4 and Q^2. '*LEARN*' as a relation between the concepts 'student' and 'teacher' has the quantificational tag $(\geq 1; \geq 0)$ and is therefore not inverse to '*TEACHTO*'. It is also not inverse to '$TEACH_{(\geq 1;\geq 0)}$' because the quantifiers of inverse relations are inverted according to Definition 3.6. Often semantic relations are inverse to other semantic relations only in a prototypical sense. For example, teaching someone does not always result in learning.

Whereas meronymy and antonymy represent a few types of relations (depending on how they are counted: about three to twenty types of meronymy or antonymy relations) which have several thousands of tokens (according to WordNet), verbs and prepositions represent a large set of types of relations which have only a few tokens each. Therefore it may not be useful to define each type of verbal or prepositional relation separately. Among such relations other relations, such as inversion, can hold. Even hierarchies of relations can be built (the classification of meronymy in Figure 3.1 is a hierarchy of meronymy relations). Other relations among verbs and prepositions are semantic relations because they do not depend on the relational character of verbs or prepositions. Semantic relations among verbs that usually not occur in other parts of speech are sequence, cause, backward presupposition and entailment (the last two according to Fellbaum (1990)). The next paragraph demonstrates that cause and backward presupposition are special kinds of sequence relations.

Definition 3.7:
A *sequence relation* $SEQ_{(Q^4;Q^2)}$ is an antisymmetric semantic relation among denotative word concepts of verbs.
A *cause relation* is a sequence relation of the type $SEQ_{(\geq 0;\geq 1)}$. A *backward presupposition* is a sequence relation of the type $SEQ_{(\geq 1;\geq 0)}$.
A semantic relation among verbs with quantifier $|| \geq n||$, $n \geq 1$, for Q^4 or Q^2 is called *entailment relation*.

Fellbaum (1990) distinguishes four types of entailment among verbs. A troponym can entail its hypernym (limp/walk or lisp/talk). This kind of entailment is usually 'co-extensive' because limping and walking occur at the same time. Actions that are properly included in other actions can entail the basic action (snore/sleep or buy/pay). The other two types of entailment are backward presuppositions (succeed/try or untie/tie) and cause (raise/rise or give/have). Fellbaum's first kind of entailment is the normal inheritance of attributes from a hypernym to its hyponym (or troponym). The same as 'being a dog' entails 'being an animal', the act of 'limping' entails the act of 'walking'. Her second kind of entailment represents a feature of meronymy relations and their quantificational tags. Since 'snoring' is always part of 'sleeping' it entails sleeping ('snore $MER_{(\geq 0;\geq 1)}$ sleep'). Since 'buying' has always the part 'paying', 'buying' entails 'paying' ('pay $MER_{(\geq 1;\geq 0)}$ buy').

Her other two types of entailment depend on quantificational tags of the sequence relation. For example, 'succeed' presupposes 'try' which means that succeeding follows trying in the sequence relation. Always when someone succeeds in something he or she must have tried it before. Between 'try' and 'succeed' holds therefore the backward presupposition 'try $SEQ_{(\geq 1;\geq 0)}$ succeed'. Verbs which stand in cause relation to each other are in sequence relation to each other with inverted quantification compared to backward presupposition. 'Give' causes 'have', but 'have' does not presuppose 'give'. 'Give' precedes 'have' in the sequence relation, and 'give $SEQ_{(\geq 0;\geq 1)}$ have'. The sequence relation can in some rare occasions be a partially ordered set instead of a linear ordering. For example, 'aim' has two successors 'hit' and 'miss' which are antonymous to each other. WordNet itself has no verbs implemented which stand in cause relation and backward presupposition to each other at the same time (that would be the relation $SEQ_{(\geq 1;\geq 1)}$), probably because such verbs are usually highly specific and usually not lexicalized, such as 'sleep $SEQ_{(\geq 1;\geq 1)}$ wake up from sleep'. This last section shows that Relational Concept Analysis can be effectively applied to semantic relations among verbs. Obviously many questions and problems are left for future research, such as the modeling of ternary relations and the modeling of case relations (object, agent, instrument, and so on).

4 Applications and extensions of Relational Concept Analysis

In the preceeding chapters the basic features of Relational Concept Analysis have been defined and applied to semantic relations. In this chapter Relational Concept Analysis is compared to other theories or models of knowledge representation, such as semantic networks, the Entity Relationship Model, and terminological logic. A study is made as to whether Relational Concept Analysis can express or can be extended to express the same statements and relations as the other theories. It seems likely that all models and theories of knowledge representation and data structuring have advantages and disadvantages. The strategy of Relational Concept Analysis is to use the formal representation of formal contexts and concept lattices as provided by Formal Concept Analysis as its main component. The formal conceptual approach has been proven to be very reliable and efficient in many applications (Ganter & Wille, 1996). Other relations, attributes and features that cannot be modeled according to the mathematical properties of lattices are then formalized as additional relations and additional attributes that follow rules which are not as strict. It seems that this approach of extending the strict formal kernel of Formal Concept Analysis with less formalized structures fulfills two goals: first, it provides a highly structured environment for most of its elements; second, it is extended to also contain less structured elements and can therefore be applied to a wider range than Formal Concept Analysis by itself. Future research is needed to investigate how far Formal and Relational Concept Analysis can be extended and whether it is even possible to use them for all applications for which other systems of knowledge representation and data structuring are used.

4.1 Lexical structures versus conceptual structures

A major difficulty for knowledge representation systems seems to be the difference between lexical items (with their lexical or syntactic relations) and concepts (with their semantic relations). Pustejovsky (1993) presents an extensive collection of attempts at combining them. In natural languages lexical and conceptual structures are usually dependent on and complementary to each other. For example, polysemy, which is a more or less regular method (compare Kilgarriff (1995)) of applying one word to several concepts, allows the expression of a large number of concepts using a fixed lexicon of words. Sowa (1993) even speaks of infinitely many concepts that can be composed from a finite number of words. Polysemy, metaphor, and metonymy allow speakers, for example, to be creative, to relate concepts, and to connote associations to concepts. For a machine and for knowledge representation purposes, polysemy, metaphor, and metonymy provide difficulties: does each polysemous meaning or each metaphorical use of a word represent a new concept which should be represented separately? Obviously, a purely lexical syntactic representation cannot cope with semantic disambiguation. To quote Cruse's example, 'topless dress', 'topless dancer', and ‘topless bar’ have the same syntactic structures, but the meaning of ‘topless’ cannot be parsed as ‘having no top’ in all cases. On the other hand, a purely conceptual representation in some kind of metalanguage denoting ‘concept1: dress without top’, ‘concept2: dancer who wears no top’ and ‘concept3: bar with topless dancers’ does not demonstrate the subtle joke that results from the polysemy of ‘topless’.

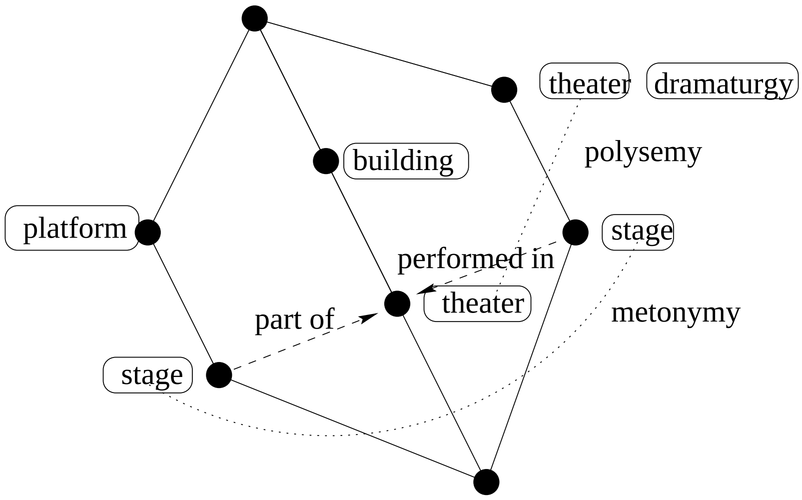

Figure 4.1: Semantic and lexical relations in a denotative lattice

Another discrepancy between lexical and conceptual structures are lexical irregularities which are not represented in the concepts. For example, a fork (in its primary sense) is a kind of cutlery, but ‘fork’ is a singular term whereas ‘cutlery’ is a collective noun. Therefore according to the lexical structure it might be more adequate to say that a fork is a part of cutlery. On the other hand, a fork as an agricultural tool is a kind of tool and not a ‘part of tool’. Although this distinction can also be made on the conceptual level: ‘cutlery’ is a set of fixed elements whereas ‘tool’ is a generic term for many different items, this distinction is not necessary for a conceptual modeling and does not have to occur in other natural languages. ‘Cutlery’ is also a collective noun in German (‘Besteck’), but the plural nouns ‘glasses’ and ‘trousers’ are singular in German (‘Brille’, ‘Hose’). To summarize: lexical items and conceptual items both follow certain rules and have certain features. But not all features of lexical items and relations are represented in the conceptual items and relations, nor are all features of conceptual items and relations represented in the lexical items and relations. On the other hand, some features of lexical items do influence features of concepts and vice versa. Therefore lexical and conceptual items and relations form two separate systems that influence each other. It seems that in order to present the complete content of a natural language expression, both lexical and conceptual structures should be considered. Sowa (1993) claims that different Conceptual Graphs can be used for conceptual and lexical structures, but he does not explain how the conceptual and the lexical representations can be related to each other. Relational Concept Analysis presents a method of demonstrating conceptual and lexical structures and the relationship between them by drawing lattices for the underlying conceptual structures and using additional relations to display lexical relations. In the example in Figure 4.1 the solid lines denote a denotative lattice describing the hyponymy relation. The dashed lines denote other semantic relations, such as 'part of' or 'performed in'. They can be quantified according to Chapter 2 and 3. The dotted lines represent lexical relations, i.e. they connect disambiguated words instead of concepts. Two polysemous senses of 'theater' connected by the polysemy relation occur in the figure. The two senses of 'stage' are in metonymy relation to each other. This kind of representation shows the differences in the lexical and semantic relations. It can also be used to demonstrate 'lexical gaps' (concepts that are not denoted by a disambiguated word) or differences between several languages by drawing and comparing diagrams for similar concepts in several languages.

In the rest of this section, a suggestion is made as to how to relate lexical syntactic composition rules with conceptual composition rules. A more detailed and complete discussion of this subject would require the definition of some kind of 'formal grammar' and would be too comprehensive for this paper. Definition 4.1 and Lemma 4.1 intend to present a suggestion of how Formal Concept Analysis can be used in this area. A lexical expression is, for example, 'a man is a male, mature person' or, formalized as 'man IS A male person AND mature person'. A corresponding conceptual expression in a denotative structure is '$c_1 = c_2 \wedge c_3$' with dnt(man) $= c_1$, dnt(male person) $= c_2$, and dnt(mature person) $= c_3$. The same conceptual expression could also be represented by a Conceptual Graph or, for example, by classes of the Universal Decimal Classification (Rowley, 1992) that are combined by '+', ':' '/', and so on. Lexical and conceptual terms and expressions follow different composition rules: Lexical items can be combined to form lexical terms, such as 'tall woman'. Concepts can be combined according to rules of formal representation systems, such as the mathematical operations in a concept lattice or the composition rules of Conceptual Graphs or the Universal Decimal Classification. In the case of denotative lattices in denotative structures this is formalized in Definition 4.1. This is not the only possibility of assigning a natural or formal language interpretation to a concept lattice. For example, Prediger (1996) interprets the formal attributes as attribute names of a formal language in terms of terminological logic (Baader, 1992).

Definition 4.1:
Let $\mathcal{S}_D$ be a denotative structure with denotative context $\mathcal{K}_D := (D, A_D, I_D)$ and a set W of words. Let the formal attributes in A_D be English adjectives, and let the words in W be English nouns. Let the top concept of the lattice be a denotative word concept denoted by the word $w_\top \in W$.
A *lexical term* is a concatenation of adjectives, nouns, 'AND' and 'OR' according to one of the following rules[18] ($\overline{L}$ denotes the set of lexical terms):

$$w \in \overline{L} \quad \text{for all} \quad w \in W$$

[18]This terminology could be extended to include 'NOT'.

$$
\begin{aligned}
al \in \overline{L} \quad & \text{for all} \quad a \in A_D, l \in \overline{L} \\
(l_1 \text{ AND } l_2) \in \overline{L} \quad & \text{for all} \quad l_1, l_2 \in \overline{L} \\
(l_1 \text{ OR } l_2) \in \overline{L} \quad & \text{for all} \quad l_1, l_2 \in \overline{L}
\end{aligned}
$$

A lexical term is a *meaningful lexical term* in a denotative structure if it denotes a denotative concept other than the bottom concept '$\perp$' according to the judgement of native language speakers, i.e. the mapping $dnt : W \to C$ is extended to $dnt : \overline{L} \to C \cup \{\emptyset\}$ with $dnt(l) = \emptyset$ for meaningless lexical terms. The set $L_{\mathcal{S}_D}$ of meaningful lexical terms in a denotative structure is defined as $L_{\mathcal{S}_D} := \{l \in \overline{L} \mid \exists_{c \in C \setminus \{\perp\}} : dnt(l) = c\}$.
Lexical expressions are of the form

$$
\begin{aligned}
d \text{ IS A } l \quad & \text{for all} \quad d \in D, l \in L_{\mathcal{S}_D} \\
l_1 \text{ IS A } l_2 \quad & \text{for all} \quad l_1, l_2 \in L_{\mathcal{S}_D}
\end{aligned}
$$

Meaningful lexical expressions are lexical expressions with d IS A $l \iff d$ INST l, and l_1 IS A $l_2 \iff l_1$ HYP l_2 according to Definition 3.2.
A *conceptual term* is of the form $c_1 \wedge c_2$ or $c_1 \vee c_2$ with $c_1, c_2 \in C$. *Conceptual expressions* are of the form $d \in Ext(c)$ and $c_1 \leq c_2$ with $d \in D$, $c, c_1, c_2 \in C$.

According to this definition the most general concept in a concept lattice is always denoted by a disambiguated word. This word can, for example, be 'entity' or 'item'. The bottom concept of a concept lattice often has no denotata in its extent. Lexical terms that denote the bottom concept, such as 'male female person' are thus not considered to be meaningful. By using INST and HYP according to Definition 3.2 it is insured that meaningful lexical expressions are always true within a denotative structure. Conceptual expressions can be true or false.

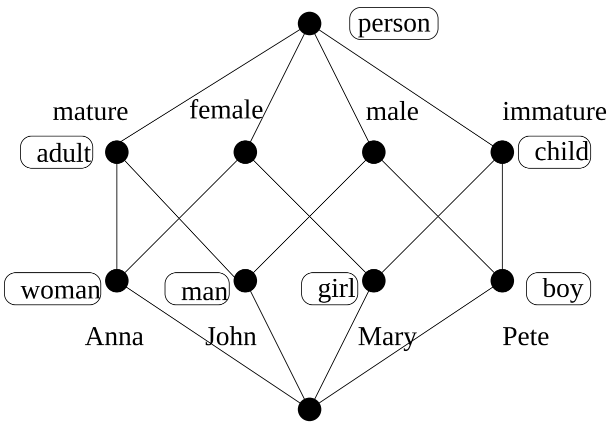

Figure 4.2: A denotative lattice

With Definition 4.1 lexical terms and expressions can be derived from the denotative lattice in Figure 4.2. 'Mngful' in brackets indicates that we consider the term to be meaningful according to Lemma 4.1.

woman (mngful), female woman, female man, female person (mngful)
female person AND mature person (mngful), mature boy AND immature child
female person OR mature person, woman OR girl (mngful)

John IS A man (mngful). Mary IS A girl (mngful). John IS A woman.
A woman IS A adult (mngful). A child IS A person (mngful).

The following paragraph including Lemma 4.1 explains why we consider some of the terms and expressions meaningful or meaningless. Obviously, rules can be developed to decide whether certain expressions must be meaningful if their components are meaningful. But since meaningfulness is defined with respect to native language speakers, there may not be a general agreement about all rules. The following lemma seems to be acceptable.

Lemma 4.1:
For $l, l_1, l_2 \in L_{\mathcal{S}_D}$, $a, a_1, a_2 \in A_D$:

$$\begin{aligned} dnt(aw_\top) &= \mu a & & \\ \mu a \geq dnt(l) \text{ or } \mu a \leq dnt(l) & \implies & al \notin L_{\mathcal{S}_D} \\ dnt(l_1) \wedge dnt(l_2) \neq \bot, l_1 \not\leq l_2 \text{ and } l_2 \not\leq l_1 & \implies & (l_1 \text{ AND } l_2) \in L_{\mathcal{S}_D} \\ Ext(dnt(l_1) \vee dnt(l_2)) \neq Ext(dnt(l_1)) \cup Ext(dnt(l_2)) & \implies & (l_1 \text{ OR } l_2) \notin L_{\mathcal{S}_D} \end{aligned}$$

$$\begin{aligned} (l_1 \text{ AND } l_2) \in L_{\mathcal{S}_D} & \implies & dnt(l_1 \text{ AND } l_2) = dnt(l_1) \wedge dnt(l_2) \\ (l_1 \text{ OR } l_2) \in L_{\mathcal{S}_D} & \implies & dnt(l_1 \text{ OR } l_2) = dnt(l_1) \vee dnt(l_2) \\ al \in L_{\mathcal{S}_D} & \implies & dnt(al) = (\mu a \wedge dnt(l)) \\ a_1 a_2 l \in L_{\mathcal{S}_D} & \implies & dnt(a_1 a_2 l) = (dnt(a_1 l) \wedge dnt(a_2 l)) \end{aligned}$$

'Female woman' is not meaningful according to the lemma since women are always female i.e. μfemale $\geq dnt$(woman) in Figure 4.2. 'Female man' denotes the bottom concept and is therefore not meaningful. Compositions with 'OR' are not meaningful in the given denotative structure if the concept $dnt(l_1) \vee dnt(l_2)$ has more denotata in its extent than the union of $Ext(dnt(l_1))$ and $Ext(dnt(l_1))$. For example in Figure 4.2, 'female person OR mature person' includes women, adults, men, and girls, but dnt(female person) $\vee$ dnt(mature person) also has the boy Pete in its extent who is neither 'female' nor 'mature'. This phenomenon is in opposition to 'lexical gaps'. Lexical gaps are concepts that can be constructed according to conceptual compositions, but cannot be described by a disambiguated word. 'Female person OR mature person' is constructed according to lexical composition rules, but does not denote a concept in the denotative lattice of Figure 4.2. Obviously, meaningfulness depends on the underlying denotative structure. In another lattice, a concept 'female person OR mature person' could exist. Stumme (1994) investigates concept lattices that include all concepts that can be constructed by using 'OR'.

It is not attempted to develop a comprehensive theory of lexical and conceptual expressions in this paper. Many questions remain open: for example, plural nouns behave differently from singular nouns. ‘Woman AND man’ is usually not a meaningful term, but ‘women AND men’ denotes the extent of ‘dnt(woman OR man)’. Furthermore compositions of adjectives require a more elaborate formalization. In Lemma 4.1 several adjectives that modify the same noun are interpreted in a sense that is often indicated in English by inserting a comma between the adjectives. According to Lemma 4.1 ‘female mature person’ is synonymous to ‘mature female person’ or ‘mature, female person’, a person that is mature and female. Lemma 4.1 is not always adequate for grading adjectives: ‘tall female persons’ is a set of tall persons selected from a set of women. On the other hand, ‘female tall persons’ is a set of women selected from a set of tall persons. If ‘tall’ is interpreted as ‘tall for a woman’ in the first case and ‘tall for a person’ in the second case, the set of ‘tall persons’ might not contain any woman whereas ‘tall female persons’ might contain women. Similarly, in some languages composite nouns can be more or less systematically built from single nouns, such as ‘pet dog’ which is synonymous to ‘pet AND dog’. Composite nouns, however often have a slightly different meaning than the combined single meanings and the order is not arbitrary (‘dog pet’ does not exist in the English language). Besides ‘AND’ and ‘OR’, ‘NOT’ could be included into the formalization in Definition 4.1. It would then have to be investigated how ‘NOT’ and the Boolean lattice operation ‘$\neg$’ interrelate. All these problems are left for future research.

4.2 Relational Algebra

This section demonstrates how the relations of Relational Concept Analysis can be defined within Relational Algebra (Pratt, 1992). One advantage of the modeling in Relational Algebra is that the evaluation of quantified expressions is reduced to the computation of binary matrix multiplications which are easily implemented in a computer program. Only quantifiers $||\text{all}||$ and $|| \geq 1||$ and their negations are so far implicitly considered in Relational Algebra. While this section formalizes some aspects of Relational Concept Analysis in terms of Relational Algebra, it might also be possible to extend Relational Algebra to include other quantifiers (which has not yet been achieved according to our knowledge) using the ideas of this section. In Relational Algebra operations such as ‘$^-$’, ‘$^\cup$’, ‘$+$’, ‘$\cdot$’, and ‘$;$’ are defined, among others. We use $\overline{I}$ instead of I^- for the complement of a relation, I^d instead of $I^\cup$ for the dual of a relation, and $I \circ J$ instead of $I;J$ for the relational product. The operations ‘$+$’ and ‘$\cdot$’, union and intersection of relations, are not needed in our modeling.

Definition 4.2
For a binary relation $I(\subseteq G_1 \times G_2)$, the *complementary relation* $\overline{I}(\subseteq G_1 \times G_2)$, and the *dual relation* $I^d(\subseteq G_2 \times G_1)$ are defined as follows. For $g_1 \in G_1, g_2 \in G_2$:

$$g_1 \overline{I} g_2 \quad :\Longleftrightarrow \neg(g_1 I g_2)$$

$$g_2 I^d g_1 \quad :\Longleftrightarrow g_1 I g_2$$

The *(relational) product* $I \circ J$ of two binary relations $I(\subseteq G_1 \times G_2)$ and $J(\subseteq G_2 \times G_3)$ is defined for all $g_1 \in G_1, g_3 \in G_3$ as

$$g_1(I \circ J)g_3 :\Longleftrightarrow \exists_{g_2 \in G_2} : g_1 I g_2 \text{ and } g_2 J g_3$$

The *identity* $Id(\subseteq G \times G)$ is defined for all $g_1, g_2 \in G$ as

$$g_1 \; Id \; g_2 :\Longleftrightarrow g_1 = g_2$$

Using Relational Algebra some of the concept relations from Chapter 2 can be expressed as a relational product. We develop the formalizations in this section only with respect to certain applications: It is often useful not to consider all formal objects and attributes individually, but to group them into classes. A relation between objects and attributes can then be transformed into a relation between object and attribute classes. Or, if the formal objects and attributes are denotata and the classes are concepts in another formal context, relations among denotata can be generalized to relations among concepts as in Chapter 2. An example where classification of objects and attributes is needed is in psychological experiments that are interested in statements about certain groups, such as women between the ages 50 and 60, derived from a questionnaire given to individuals. After classing the objects into classes it is then necessary to decide whether a class has an attribute if all its objects, most of its objects, or some of its objects have the attribute. This decision requires quantifiers, such as $||\text{all}||$, $||\text{most}||$ or $|| \geq 1||$. In many applications these quantifiers are not explicitly named and therefore valuable information can be lost. Modelings with Relational Concept Analysis always make quantifiers explicit. Definition 4.5 provides a formalization of a context schema consisting of objects, attributes, and classes of objects. Some of the attributes in the context schema are not the formal (essential) attributes in a concept lattice, but they are additional (accidental) attributes that are assigned to concepts in the same way as semantic relations, such as meronymy, are additional relations to the lattice structure. Whereas semantic relations are binary relations, a set of attributes can be interpreted as a set of unary relations. Definition 4.3 formalizes therefore unary concept relations in analogy to Definition 2.1. Definition 4.4 formalizes the relations between concepts and accidental attributes.

Definition 4.3:
For a context (G, M, I), a concept $c \in \mathcal{B}(G, M, I)$, a unary relation $r_a \subseteq G$, and a quantifier Q a unary concept relation $R^{ra}[Q] \subseteq \mathcal{B}(G, M, I)$ is defined as

$$c \, R^{ra}[Q] \quad :\Longleftrightarrow \quad Q_{g \in Ext(c)} : g r_a$$

The context (G, M, I) in Figure 4.3. provides an example. Tomato, cucumber, apple, and banana are classed into fruit and vegetable[19]. The classes 'fruit' and 'vegetable'

[19]'Vegetable' and 'fruit' are here used in their natural language sense, not in the biological sense where tomatoes and cucumbers would be fruit.

are the formal attributes of (G, M, I). The denotative attributes that are used to distinguish between fruit and vegetable are not known. A second context (G, A, r) assigns other attributes (color attributes) to tomato, cucumber, and so on. In the sense of (G, M, I) the formal attributes of (G, A, r) are additional, accidental attributes that cannot be assigned to concepts in the concept lattice of (G, M, I) without quantification. But if r_{red} :='has color red', r_{yel} :='has color yellow', and r_{gre} :='has color green', are defined and if Q is the $|| \geq 1||$-quantifier then $c_1 R^{r_{red}}[|| \geq 1||]$ and $c_1 R^{r_{gre}}[|| \geq 1||]$. If Q is the $||\text{all}||$-quantifier then $c_1 R^{r_{gre}}[||\text{all}||]$. In other words, the concept of 'vegetable' in the concept lattice of (G, M, I) can get the additional attributes 'all denotata can be green' or 'some denotata are green or red' assigned. The advantage of this approach is that attributes can be added to a concept lattice with a preceding quantifier. For example, the attribute 'flying' can be assigned to the concept 'bird' if it is modified by the quantifier $||\text{almost all}||$. Instead of defining three unary relations r_{red}, r_{yel}, and r_{gre} a binary relation $r(\subseteq G \times A)$, which is 'has color' in our example, can be defined (Definition 4.4). The difference from Definition 2.1 is that this binary relation is not defined among two sets of concepts but between concepts and attributes, therefore only one quantifier is required.

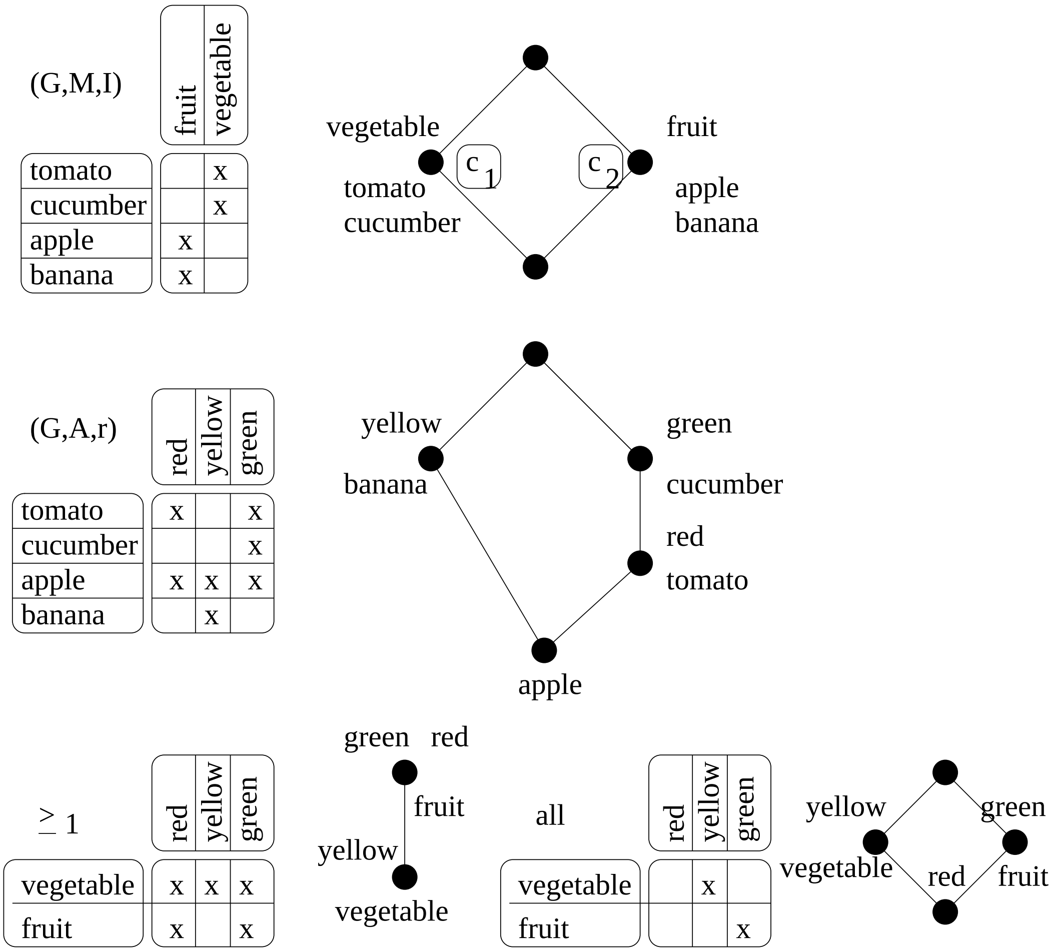

Figure 4.3: Different quantifiers in a combination of contexts

The two formal contexts in Figure 4.3 could also be interpreted in the sense of ‘multicontexts’ (Wille, 1996) for which several methods of composition exist. Our modeling differs from multicontexts in that a concept lattice is computed only for one formal context ((G, M, I) in the example) whereas the other formal context ((G, A, r) in the example) provides additional information for the first context.

Definition 4.4:
For formal contexts (G, M, I) and (G, A, r), a concept $c \in \mathcal{B}(G, M, I)$, and a quantifier Q, binary relations $R^r[Q;] \subseteq \mathcal{B}(G, M, I) \times A$ and $R^r[;Q] \subseteq A \times \mathcal{B}(G, M, I)$ are defined as

$$\begin{aligned} c\, R^r[Q;]\, a \quad &:\Longleftrightarrow \quad Q_{g \in Ext(c)} : gra \\ a\, R^r[;Q]\, c \quad &:\Longleftrightarrow \quad Q_{g \in Ext(c)} : ar^d g. \end{aligned}$$

In the example of Figure 4.3, $c_1 R^{\text{has color}}[||\geq 1||;]$‘green’, $c_1 R^{\text{has color}}[||\geq 1||;]$‘red’, and $c_1 R^{\text{has color}}[||\text{all}||;]$‘green’ hold. The example is further formalized in the next definition.

	tomato	cucumber	apple	banana	red	yellow	green
tomato	x				x		x
cucumber		x					x
apple			x		x	x	x
banana				x		x	
fruit			x	x	R*		
vegetable	x	x					

Figure 4.4: The example from Figure 4.3 as relational context schema

Definition 4.5:
Two formal contexts (G, M, I) and (G, A, r) and a relation $R^\star := R^{\star r}[Q;](\subseteq M \times A)$ which is defined as $mR^\star a :\Longleftrightarrow \mu m R^r[Q;]a$ combined according to the schema in Figure 4.5.a form a *relational context schema with one quantifier* denoted by $(G, M, A, I, r, R^{\star r}[Q;])$.
Two formal contexts (G, M, I) and (G, A, r) and a relation $R^\star := R^{\star r}[;Q](\subseteq A \times M)$ which is defined as $aR^\star m :\Longleftrightarrow aR^r[;Q]\mu m$ combined according to the schema in Figure 4.5.b form a *relational context schema with one quantifier* denoted by $(G, M, A, I, r, R^{\star r}[;Q])$.

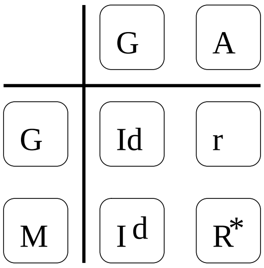

Figure 4.5.a

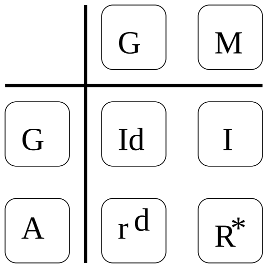

Figure 4.5.b

While R is a relation between concepts and attributes (Definition 4.4), $R^\star$ is a relation between two sets of formal attributes. Theorem 4.1 shows that, for some quantifiers, R (or $R^\star$) can be constructed as a product of the relations I and r (or their complements) according to Relational Algebra. Figure 4.4 presents the relational context schema for the example in Figure 4.3.

Theorem 4.1:
If Q is an element of $\{||\text{all}||, ||\geq 1||, ||0||, ||\neg\text{all}||\}$ then the following equations hold[20].

$$
\begin{aligned}
\mu m\, R^r[||\text{all}||;]\, a &\iff m\,(\overline{I^d \circ \overline{r}})\, a &&\iff ||\text{all}||_{g\in G} : (gIm \implies gra)\\
\mu m\, R^r[||\geq 1||;]\, a &\iff m\,(I^d \circ r)\, a &&\iff ||\geq 1||_{g\in G} : (gIm \text{ and } gra)\\
\mu m\, R^r[||0||;]\, a &\iff m\,(\overline{I^d \circ r})\, a &&\iff ||0||_{g\in G} : (gIm \text{ and } gra)\\
\mu m\, R^r[||\neg\text{all}||;]\, a &\iff m\,(I^d \circ \overline{r})\, a &&\iff ||\geq 1||_{g\in G} : (gIm \text{ and } \neg(gra))
\end{aligned}
$$

Definition 4.5 and Theorem 4.1 demonstrate that already the combination of two contexts that share their objects (or attributes) can lead to different relational products depending on the selection of the quantifier. The other possible products, $\overline{I^d} \circ \overline{r}$ $\overline{\overline{I^d} \circ \overline{r}}$, $\overline{I^d} \circ r$, and $\overline{\overline{I^d} \circ r}$, are not useful in Relational Concept Analysis since they would lead to quantifiers $Q_{g\notin Ext(\mu m)}$. Other natural language quantifiers Q could be chosen, but they cannot be expressed within Relational Algebra. Definition 4.6 extends the relational context schemata to three formal contexts with two quantifiers.

Definition 4.6:
Three formal contexts (G_1, M_1, I_1), (G_2, M_2, I_2), (G_1, G_2, r) and two relations $R_1^\star := R_1^{\star r}[Q^1;](\subseteq M_1 \times G_2)$ and $R^\star := R^{\star r}[Q^1, Q^2;](\subseteq M_1 \times M_2)$, which is defined as $m_1 R^\star m_2 :\iff \mu m_1 R^r[Q^1, Q^2;]\mu m_2$, combined according to the schema in Figure 4.6.a form a *relational context schema with two quantifiers* denoted by
$(G_1, M_1, G_2, M_2, I_1, I_2, r, R_1^{\star r}[Q^1;], R^{\star r}[Q^1, Q^2;])$.
Three formal contexts (G_1, M_1, I_1), (G_2, M_2, I_2), (G_1, G_2, r) and two relations $R_2^\star := R_2^{\star r}[;Q^3](\subseteq G_1 \times M_2)$ and $R^\star := R^{\star r}[;Q^3, Q^4](\subseteq M_1 \times M_2)$, which is defined as $m_1 R^\star m_2 :\iff \mu m_1 R^r[;Q^3, Q^4]\mu m_2$, combined according to the schema in Figure 4.6.b form a *relational context schema with two quantifiers* denoted by
$(G_1, M_1, G_2, M_2, I_1, I_2, r, R_2^{\star r}[;Q^3], R^{\star r}[;Q^3, Q^4])$.

[20] Proof: From $gIm \iff g \in Ext(\mu m)$ follows $||\geq 1||_{g\in G} : (gIm \text{ and } gra) \iff ||\geq 1||_{g\in Ext(\mu m)} : gra$. The other quantifiers are treated similarly.

	G_1	G_2	M_2
G_2		Id	I_2
G_1	Id	r	
M_1	I_1^d	R_1^*	R^*

Figure 4.6.a

	G_1	G_2	M_2
G_2		Id	I_2
G_1	Id	r	R_2^*
M_1	I_1^d		R^*

Figure 4.6.b

The two cases of Definition 4.6 represent the two aspects of concept relations (first quantifying the first set or first quantifying the second set) according to Definition 2.1. Theorem 4.2 demonstrates how Definition 2.1 is related to Definition 4.6 by showing how R is related to $R^\star$. Similarly to that which has been stated about Theorem 4.1, some relational products that are possible according to Relational Algebra are irrelevant to Relational Concept Analysis because they involve quantifiers about elements 'not' in a set and vice versa: natural language quantifiers which can be used in Relational Concept Analysis are not used in Relational Algebra.

Theorem 4.2:
With the terminology of Definition 4.6 and quantifiers Q^1, $Q^2 \in \{||\text{all}||, || \geq 1||, ||0||, ||\neg\text{all}||\}$ the following holds[21]

$$
\begin{aligned}
m_1\, \overline{I_1^d \circ \overline{r} \circ I_2}\, m_2 &\iff \mu m_1\, R^r[||\text{all}||, ||\text{all}||;]\, \mu m_2 \\
m_1\, \overline{I_1^d \circ \overline{\overline{r \circ I_2}}}\, m_2 &\iff \mu m_1\, R^r[||\text{all}||, || \geq 1||;]\, \mu m_2 \\
m_1\, \overline{I_1^d \circ r \circ I_2}\, m_2 &\iff \mu m_1\, R^r[||\text{all}||, ||0||;]\, \mu m_2 \\
m_1\, \overline{I_1^d \circ \overline{\overline{r} \circ I_2}}\, m_2 &\iff \mu m_1\, R^r[||\text{all}||, ||\neg\text{all}||\, \mu m_2 \\
m_1\, I_1^d \circ \overline{\overline{r} \circ I_2}\, m_2 &\iff \mu m_1\, R^r[|| \geq 1||, ||\text{all}||;]\, \mu m_2 \\
m_1\, I_1^d \circ r \circ I_2\, m_2 &\iff \mu m_1\, R^r[|| \geq 1||, || \geq 1||;]\, \mu m_2 \\
m_1\, I_1^d \circ \overline{r \circ I_2}\, m_2 &\iff \mu m_1\, R^r[|| \geq 1||, ||0||;]\, \mu m_2 \\
m_1\, I_1^d \circ \overline{r} \circ I_2\, m_2 &\iff \mu m_1\, R^r[|| \geq 1||, ||\neg\text{all}||;]\, \mu n_2
\end{aligned}
$$

The equivalences for $Q^1 \in \{||\neg\text{all}||, ||0||\}$ are built analogously.

The rest of this section shows an example of a relational context schema with two quantifiers. The two attribute sets M_1 and M_2 are interpreted as classes of objects.

[21]Proof: The first of these equations is proved by $m_1\, \overline{I^d \circ \overline{r} \circ I}\, m_2 \iff \neg\exists_{g_1 \in Ext(\mu m_1)} \exists_{g_2 \in Ext(\mu m_2)} : \neg(g_1 r g_2) \iff \forall_{g_1 \in Ext(\mu m_1)} \neg\exists_{g_2 \in Ext(\mu m_2)} : \neg(g_1 r g_2) \iff \forall_{g_1 \in Ext(\mu m_1)} \forall_{g_2 \in Ext(\mu m_2)} : g_1 r g_2$. The others are proved similarly. Parenthesis are not needed for the $\circ$ operation since it is associative, i.e. $I \circ J \circ K = (I \circ J) \circ K = I \circ (J \circ K)$.

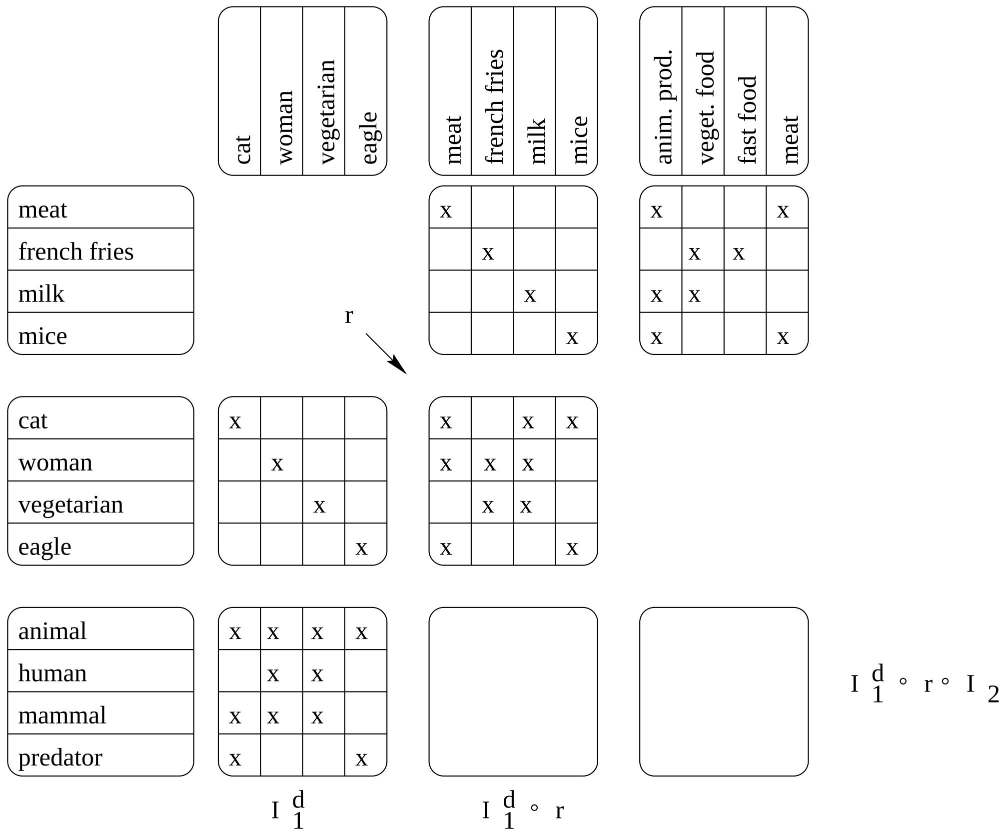

Figure 4.7

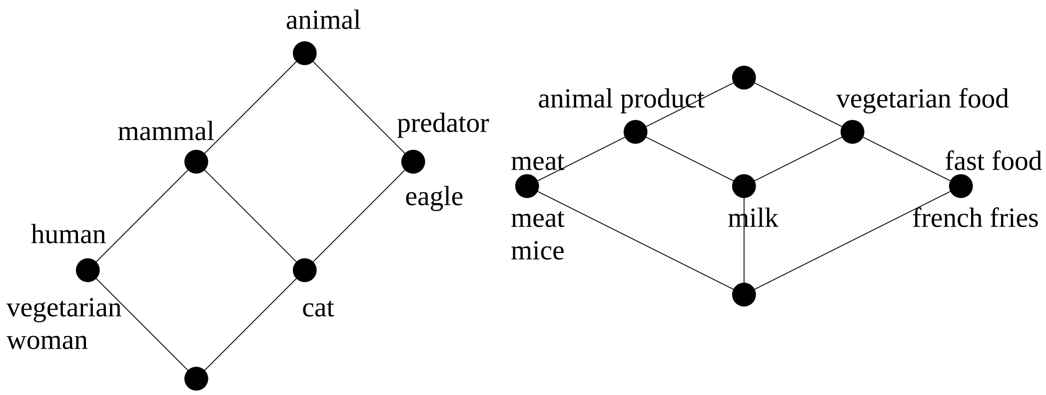

Figure 4.8

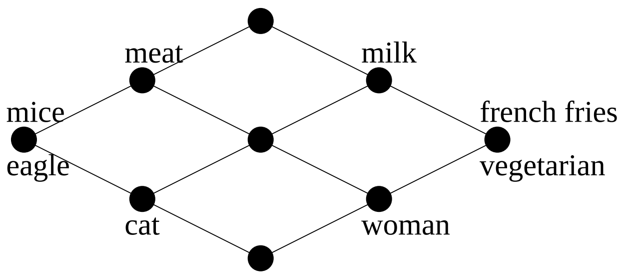

Figure 4.9

The set G_1 of objects consists of 'cat', 'woman', 'vegetarian', and 'eagle'. The set M_1 presents a set of classes for the objects in G_1: 'animal', 'human', 'mammal', and 'predator'. The objects should be interpreted as prototypical objects. The other set G_2 of objects consists of nutrition 'meat', 'French fries', 'milk', and 'mice' and M_2 are classes of food: 'animal products', 'vegetarian', 'fast food', and 'meat'. 'Meat' is in one case considered as an object (prototypical object or one specific piece of meat), in the other case as the class 'meat'. The relation r between the objects of G_1 (living beings) and G_2 (kinds of food) is 'to eat'. According to the purposes, the quantifiers can be selected to form the relation $R^\star$ between classes of living beings and classes of food. Figure 4.7 shows the relational context of the example. Figure 4.8 demonstrates the classifications on the two sets of objects: the lattices $\underline{\mathcal{B}}(G_1, M_1, I_1)$ and $\underline{\mathcal{B}}(G_2, M_2, I_2)$. Figure 4.9 shows the relation 'eat' between the objects.

all eat some	animal prod.	veget. food	fast food	meat
animal	x			
human	x	x	x	
mammal	x	x		
predator	x			x

all eat all	animal prod.	veget. food	fast food	meat
animal				
human		x	x	
mammal				
predator				x

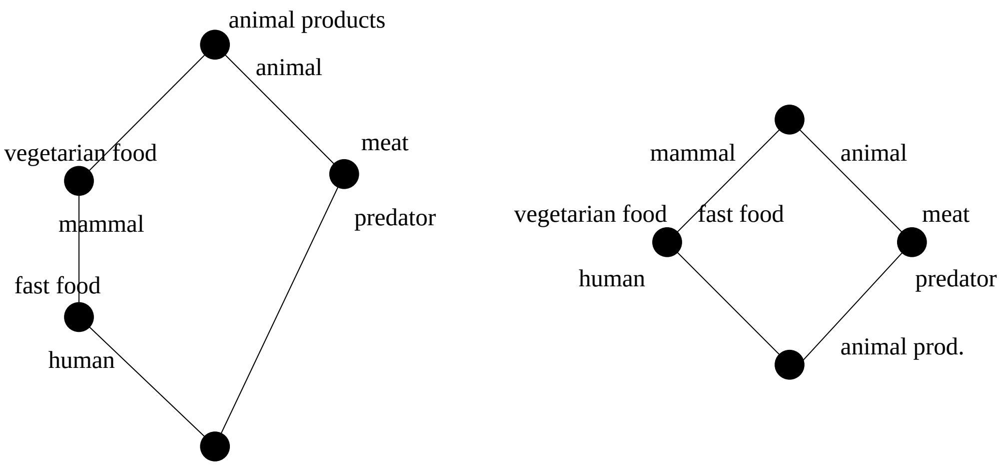

Figure 4.10

Figure 4.10 shows the lattice $\underline{\mathcal{B}}(M_1, M_2, R^\star)$ for the examples $R^\star = \overline{I_1^d \circ \overline{r \circ \overline{I_2}}}$ and $R^\star = \overline{I_1^d \circ \overline{r} \circ I_2}$. The relation 'EAT$_{(\geq 0;\geq 1)}$' represented by the dotted lines in Figure 4.11 generates the relation 'all eat some' in Figure 4.10. Since humans, mammals, and predators are animals, the relation 'eat' between them and 'animal products' follows from the relation between 'animal' and 'animal products'. The same holds for 'humans eat vegetarian food' and 'mammals eat vegetarian food'. Nevertheless the relation in Figure 4.11 is not a basis of the original relation between objects since the relation 'all mammals eat some milk' is missing. This element of the relation is missing because the object concept 'milk' is not an attribute concept. On the more abstract level of classes 'animal products', 'fast food', and so on, there is not a single class for animal products which are vegetarian food at the same time. The information about this special relation between 'mammals' and 'milk' can therefore not be expressed in terms of mammals and classes of foods. This demonstrates that classes have to be chosen carefully if information is to be transferred from an object level to a more abstract level.

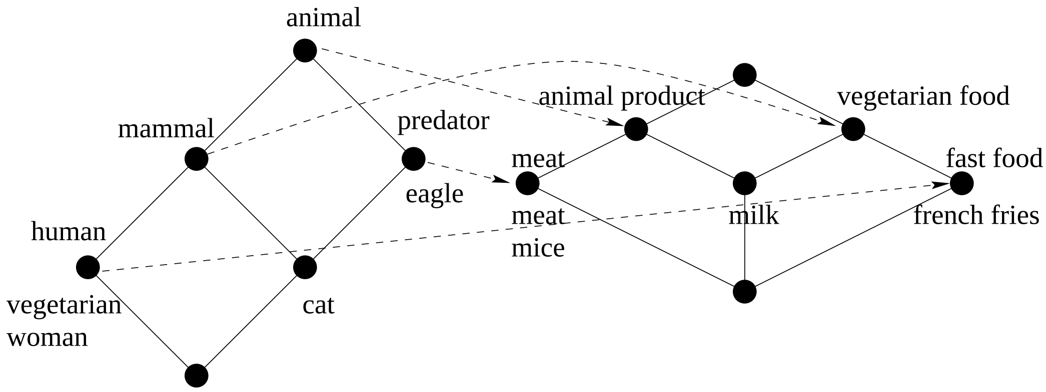

Figure 4.11

4.3 Many-valued contexts

This section shows that Relational Concept Analysis provides a generalization of the many-valued contexts (Ganter & Wille (1996)) of Formal Concept Analysis. A many-valued context is defined as a context (G, M, W, I)[22] consisting of three sets G, M, and W and a ternary relation $I(\subseteq G \times M \times W)$ with the condition $(g, m, w) \in I$ and $(g, m, v) \in I \implies w = v$. The elements of G are called formal objects, the elements of M many-valued attributes, and the elements of W values. (g, m, w) is read as 'the attribute m has the value w for the object g. This is also denoted by $m(g) = w$. The first context (context $\mathcal{K}_A := (G, M, W, I)$) in Figure 4.12 shows an example: 'keyword 1', 'keyword 2', and 'published in' are many-valued attributes for books, such as 'keyword 1(book1)= catalogs'. Since the attributes of many-valued contexts can have only one value for each object, keyword 1 and keyword 2 are not joined into one attribute 'keyword' that has several values. Another solution to this problem would be to allow elements of the power set of keywords to be values of 'keyword'. In relational databases (relational tables can be interpreted as many-valued contexts as explained in the next section) this second solution is usually considered to be in contradiction to database normalization. Another example of a

[22]The letters G, M, W, I are taken from Ganter & Wille (1996). Therefore in this section W does not denote the set of disambiguated words.

problematic attribute is ‘telephone-number’ which has to be split into ‘first telephone-number’, ‘second telephone-number’, and so on. This is problematic because adding a new object to the context may require the addition of a new attribute (for example, if the new object has more telephone numbers than expected).

K_A	keyword 1	keyword 2	published in
book1	catalogs	classification	1995
book2	hardware		1980
book3	hardware	Internet	1995
book4	Internet		1996

K_B	comp. sci.	inform. sci.	library sci.
catalogs			x
classific.		x	x
hardware	x		
Internet	x	x	

K_C	recent	old
1980		x
1995	x	
1996	x	

D	catalogs	classific.	hardware	Internet	comp. sci.	inform. sci.	library sci.
catalogs	x						x
classific.		x				x	x
hardware			x		x		
Internet				x	x	x	
book1	x	x					x
book2			x		x		
book3			x	x	x		
book4				x	x	x	

E	1980	1995	1996	recent	old
1980	x				x
1995		x		x	
1996			x	x	
book1		x		x	
book2	x				x
book3		x		x	
book4			x	x	

Figure 4.12: An example

In Formal Concept Analysis, many-valued contexts are usually conceptually scaled into single-valued contexts. A *conceptual scale* for an attribute m is a single-valued context $S_m = (G_m, M_m, J_m)$ with $m(G) \subseteq G_m$. Context $\mathcal{K}_C := (G_{m3}, M_{m3}, J_{m3})$ in Figure 4.12 is a conceptual scale for the years of publishing in context $\mathcal{K}_A$. In context $\mathcal{K}_B := (G_{m1} \cup G_{m2}, M_{m1,m2}, J_{m1,m2})$ the values of ‘keyword 1’ and ‘keyword 2’ are unioned into one set of formal objects. They are ‘scaled’ into the disciplines ‘computer science’, ‘information science’, and ‘library science’. Conceptual scales are usually developed depending on abstract knowledge of the conceptual properties of the attributes in the formal context and not depending on the existing objects. For example, the decision about ‘recent’ and ‘old’ in $\mathcal{K}_C$ should depend on knowledge about the age of books in scientific disciplines and not on the actual age of the books in $\mathcal{K}_A$. A *realized scale* is a formal context (G, M_m, J_m)

with $\forall_{g\in G, n\in M_m} : (gJ_m n :\Longleftrightarrow \exists_{w\in m(G)} : (m(g) = w \text{ and } wJ_m n))$. The context $(G, M_{m3}, R^{\star J_{m3}}[|| \geq 1||;])$ in the right lower corner of the relational context schema $E(:= (G_{m3}, G, M_{m3}, I_{m3}, J_{m3}, R^{\star J_{m3}}[|| \geq 1||;]))$ is an example of a *realized scale*. According to E, books are 'recent' or 'old' if the years in which they were published are 'recent' or 'old' according to the conceptual scale in $\mathcal{K}_C$. Realized scales are combined to a *derived context* of a many-valued context (G, M, W, I) as a context (G, N, J) with $N := \bigcup_{m\in M}\{m\} \times M_m$ and $gJ(m, n) :\Longleftrightarrow \exists_{w\in m(G)} : (m(g) = w \text{ and } wJ_m n)$. In *plain scaling* the derived context consists of appositions of realized scales. A software tool TOSCANA (Vogt & Wille, 1994) allows the combination of scales and navigation through them.

The definitions of 'realized scales' and 'derived contexts' in the last paragraph show that an $|| \geq 1||$-quantifier is involved. Since in Formal Concept Analysis many-valued attributes have only one value for each object, the $|| \geq 1||$-quantifier is the only useful quantifier. If the values of an attribute are set into relation with the objects (as in context $\mathcal{K}_B$), an object can relate to several values of an attribute. Therefore different quantifiers can be used, and realized scales can be interpreted as contexts $(M, A, R^\star)$ (or $(A, M, R^\star)$) in a relational context schema according to Definition 4.5. In the relational context schema $D := (G_{m1} \cup G_{m2}, G, M_{m1,m2}, I_{m1,m2}, J_{m1,m2}, R^{\star J_{m1,m2}}[||\text{all}||;])$ the $||\text{all}||$-quantifier is utilized: 'all keywords of book 1 belong to library science therefore book 1 belongs to library science'. If the $|| \geq 1||$-quantifier was used, book 1 would also belong to information science since one of its keywords belongs to information science. A relational context schema (such as D in the example) could also be interpreted according to Figure 4.5.b. An $||\text{all}||$-quantifier would then produce statements such as 'all keywords of a discipline are assigned to a book', but that does not seem to be very convincing for this application. Other quantifiers, such as $||\text{almost all}||$, might be useful for other applications.

In summary: relational context schemas are useful for attributes that have several values for an object. Furthermore they allow a variety of quantifiers to be applied. Unfortunately, for an implementation, the query algorithm of the software tool TOSCANA would have to be modified. Furthermore, the lattice of $(M, A, R^\star)$ (compare Definition 4.5) is not easily constructed from the lattices of (G, M, I) and (G, A, r) if other quantifiers than the $|| \geq 1||$-quantifier are involved. Thus, solving the problems of an implementation is left for future research. Instead of the apposition of the realized scales to form a derived context, other combinations could be utilized. For example, the values of several many-valued attributes could be scaled into {good, medium, bad} and it could be asked whether an object has 'good' values for all its attributes which would be the relational intersection (Pratt, 1992) of the realized scales. This approach would be similar to Prediger's (1996) interpretation of formal attributes as attribute names of a terminological description language. A comparison of Relational Concept Analysis and Prediger's theory is left to future research.

4.4 The Entity-Relationship Model

Many-valued contexts, such as context $\mathcal{K}_A$ in Figure 4.12, can represent tables of a relational database. Since the Entity-Relationship Model (ER Model, compare Chen & Knoell (1991)) is often utilized to display the relationships among the objects and tables it is of interest to compare Relational Concept Analysis to the ER Model. A comprehensive comparison would exceed the scope of this paper therefore only some differences are mentioned in this section. The ER Model provides a graphical representation of the internal structure of a relational database. ER Diagrams are mainly utilized in the design stage of a relational database, but they can also be helpful in forming queries for the database. Attempts have been made to design user friendly interfaces for relational databases which graphically display the ER Diagrams (compare for example Burg & van de Riet (in prep.)). As practical implementations of Relational Concept Analysis do not yet exist, no comparison of the user-friendliness of both theories can be made. On the theoretical level Relational Concept Analysis has the advantage of treating the cardinality of relations in more detail. The ER Model (Mannila & Räihä, 1994) distinguishes only 'one to many', 'one to one', and 'many to many' relations which are defined (in our terminology) as follows: $R_{(\leq 1;\geq 0)}$ (one to many), $R_{(\leq 1;\leq 1)}$ (one to one) and $R[|| \geq 1||, || > 1||; || \geq 1||, || > 1||]$ (many to many). Other quantifications cannot be expressed. Instead, if other quantifications are needed they have to be added separately as constraints (Burg & van de Riet, in prep.).

Although the ER Model can facilitate the formation of queries to a database by showing the relationships among the elements of the database, it does not provide a graphical display for the query itself, or for the answer. Queries still have to be constructed using SQL statements which have to be carefully matched to the names of entities and objects and the layout of the database. They produce answers in the form of tables. Consider context $\mathcal{K}_A$ in Figure 4.12 as an example database table called 'library catalog'. The SQL query

```
select book, keyword_1, keyword_2 from library_catalog where
keyword_1="Internet" OR keyword_2="Internet" AND published_in > 1985
```

is answered by

```
book     keyword_1   keyword_2
-----------------------------------
book3    hardware    Internet
book4    Internet

2 row(s) retrieved
```

The system does not give the user any information as to whether other books are contained in the database that deal with similar subjects. Nor does it help the user in correcting spelling mistakes which would result in an empty query result, despite the fact that the desired objects are contained in the table. Nor does the system structure the answer so that the user can get hints about the relevance of retrieved documents. Relevance becomes important if, for example, hundreds of rows are retrieved. Using Formal Concept Analysis the system could answer with a graphical display, such as Figure 4.13 which helps the user by showing part of an internal classification of the keywords. Each keyword has the number of documents that have that keyword attached. Using Relational Concept Analysis a variety of queries, such as 'find books that have at least 2 keywords in the area of computer science and are written after 1985' can be answered by presenting appropriate concept lattices. The details of an implementation are again left for future research.

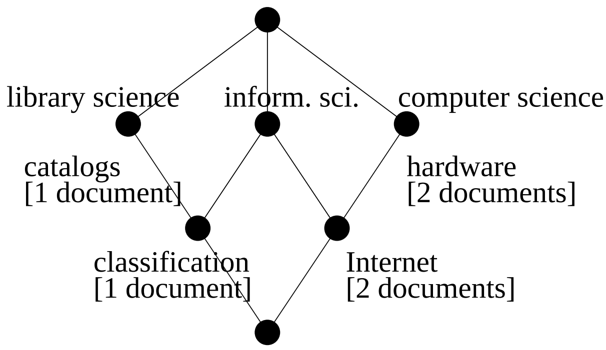

Figure 4.13: A lattice for a database query

4.5 Terminological representation systems and semantic networks

An attempt to compare and combine Relational Concept Analysis with all aspects of terminological representation systems (compare Baader (1992)) or semantic networks (compare Kilgarriff (1995)) would be too comprehensive for this paper. Therefore only some crucial aspects are mentioned in this section. Prediger (1996) shows how Formal Concept Analysis and terminological representation systems can be combined. Burkert (1995) explains how terminological representation systems can be applied to lexical databases. The most apparent difference between Relational Concept Analysis and terminological representation systems is the fact that Relational Concept Analysis describes semantic relations in lexical databases whereas in terminological representation systems relations are used to define concepts. For example, a mother can be defined as a woman that has at least one child. In Relational Concept Analysis this means that there is a relation 'HAS CHILD$_{(\geq 0; \geq 1)}$' between the concepts 'mother' and 'child' and no other subconcept of 'woman' that is not a subconcept of 'mother' is in relation 'HAS CHILD$_0$' to 'child'. Obviously a quantification of relations themselves is involved here that cannot be formalized in Relational Concept Analysis so far. In terminological logic this quantifier is denoted by $\exists R.C$ (Baader, 1992). The formal definition of 'mother' would be denoted by 'mother := woman $\cap$

$\exists$ HAS-CHILD.human'. The other basic quantifier of terminological logic is denoted by $\forall R.C$ and can be used, for example, to define a 'community' as 'community := set $\cap$ $\forall$ HAS-MEMBER.human'. This quantifier can be interpreted as a value restriction for a relation: the values of a HAS-MEMBER-relation from 'community' are in 'human'. In Relational Concept Analysis this would have to be expressed by the conditions that a concept '$\neg$human' exists, no 'HAS MEMBER$_0$'-relation holds between 'community' and '$\neg$human', and all other subconcepts of 'set' that are not subconcepts of 'community' are in 'HAS MEMBER$_{(\geq 0;\geq 1)}$ relation to '$\neg$human'. Both quantifiers, $\exists R.C$ and $\forall R.c$, cause problems for lexical databases modeled according to Relational Concept Analysis since they cannot be implemented as quantifiers of concept relations. They could be realized as conditions which would have to be evaluated each time a relation is added to or deleted from a lexical database. The defining character of quantified relations in terminological representation systems allows furthermore the following example. A quartet is defined as a music group with exactly four musicians as members (Burkert, 1995): 'quartet = music group $\cap$ ||exactly-four|| HAS-MEMBER.musician'. Again in Relational Concept Analysis a relation HAS MEMBER$_{(\geq 0;4)}$ can be defined between 'quartet' and 'musician', but that does not imply that this relation defines the concept 'quartet'. Therefore, the only solution to overcome these problems seems to be to use relations, such as 'has exactly four members', as formal attributes and not as relations, such as HAS MEMBER$_{(\geq 0;4)}$. This approach is similar to the attribute logic in Prediger's (1996) work. Future research might answer all these questions.

Other differences between Relational Concept Analysis and terminological representation systems include the larger variety of quantifiers in Relational Concept Analysis and the fact that episodic (or assertional) knowledge is integrated into terminological systems. The combination of a description language with an assertion language is also a common feature of knowledge representation systems, such as KL-ONE (Brachman & Schmolze, 1985) and semantic networks, such as DATR (according to Kilgarriff (1995)). Inheritance rules (for example, 'term subsumption languages' (Burkert, 1995)) are usually used to store attributes as high as possible in the hierarchy and to inherit them from superconcepts to subconcepts. A major difficulty is posed by default attributes that hold for most of the objects, but not for all of them. A discussion of these problems and a proposed solution can be found in Touretzky (1986). For example, the attribute 'CAN FLY' can be stored as a default attribute of the concept 'bird'. It is then valid for all subconcepts of 'bird' if it is not explicitly negated for a concept, such as 'penguin'. Our solution is to model 'CAN FLY' as a accidental attribute of the concept 'bird' according to Definition 4.4 using the quantifier ||almost all|| or ||all typical||. The highest subconcepts of 'bird' that denote 'flying birds' are assigned 'CAN FLY$_{(||\text{all}||;)}$' whereas the highest subconcepts of 'bird' that denote 'not flying birds' are assigned 'CAN FLY$_{(0;)}$'. This approach allows correct attribution for all concepts. Some attributes must be stored in several places with different quantifiers, but we think that this does not pose a problem since computer storage is becoming cheaper and database management software is becoming faster. Only accidental attributes need to be stored for several concepts, essential attributes that are used as formal attributes in a concept lattice are stored only with their attribute concepts.

As a final conclusion it can be said that some of the differences between Relational Concept Analysis and other knowledge representation systems seem to be in favor of Relational Concept Analysis whereas others seem to be in favor of other systems. Formal Concept Analysis has already been utilized for many applications (Ganter & Wille, 1996). This dissertation demonstrates that Relational Concept Analysis extends the scope of applications. Future research has to investigate whether the remaining problems concerning implementations and applications of Relational Concept Analysis can be solved.

5 References

5.1 The lexical databases, dictionaries or thesauri

RIT1 (1852). Peter M. Roget. *A Thesaurus of English words and Phrases.* 1st Ed., London.

RIT (1962). *Roget's International Thesaurus.* 3rd Ed. Th. Y. Crowell Co., New York.

Webster (1981). *Webster's Third New International Dictionary of the English Language Unabridged.* Merriam Company, Springfield, Massachusetts.

WordNet: a lexical database developed by George Miller and his team at the Cognitive Science Laboratory, Princeton University. Different versions have been used in this paper, such as WordNet 1.4, WordNet 1.5 and some in-official versions in between.

5.2 References

Baader, Franz (1992). *Logische Methoden in der Wissensrepräsentation am Beispiel der terminologischen Repräsentationssprachen.* RWTH Aachen.

Brachman, R. J.; Schmolze, J. G (1985). *An Overview of the KL-ONE Knowledge Representation Systems.* Cogn. Sci., 9 (2), pp. 171-216.

Brink, Chris (1993). *Power Structures.* Algebra Univ., 30, No. 2, pp. 177-216.

Bryan, Robert (1973). *Abstract Thesauri and Graph Theory Applications to Thesaurus Research.* In: Sedelow, Sally Y. (ed.). *Automated Language Analysis.* Report on research 1972-73, University of Kansas, Lawrence.

Bryan, Robert (1974): *Modeling in Thesaurus Research.* In: Sedelow, Sally Y. (ed.). *Automated Language Analysis.* Report on research 1973-74, University of Kansas, Lawrence.

Burg, J. F. M.; van de Riet, R. P. (in prep.). *Color-X: Using Knowledge from WordNet for Conceptual Modeling.* In: Fellbaum, Christiane (ed.). *WordNet: an Electronic Lexical Database and some of its Applications.* MIT Press.

Burkert, Gerrit (1995). *Lexical semantics and terminological knowledge representation.* In: Saint-Dizier, P.; Viegas, E.: *Computational lexical semantics.* Cambridge Univ. Press.

Calzolari, Nicoletta (1988). *The Dictionary and the Thesaurus can be Combined.* In: Evens, Martha W. (ed.). *Relational Models of the Lexicon.* Cambridge University Press.

Chaffin, Roger; Herrmann, Douglas J. (1988). *The nature of semantic relations: a comparison of two approaches.* In: Evens, Martha W. (ed.). *Relational Models of the Lexicon.* Cambridge University Press.

Chaffin, Roger (1992). *The Concept of a Semantical Relation.* In: Adrienne Lehrer, Eva Feder Kittay: *Frames, Fields, and Contrasts.* New Essays in Semantical and Lexical Organisation.

Chen, Peter P. S.; Knöll, Heinz-Dieter (1991). *Der Entity-Relationship Ansatz zum logischen Systementwurf.*

Cruse, D. A. (1986). *Lexical Semantics.* Cambridge, New York.

Dahlberg, Ingetraud (1994). *Conceptual Structures and Systematization.* In: Negrini, G. (ed.). *Categorie, Ogetti e Strutture della Conoscenza.* Roma.

DIN 2330 (1979). *Begriffe und Benennungen: Allgemeine Grundsätze.* Deutsches Institut für Normung, Berlin-Köln, Beuth Verlag.

DIN 32705 (1987). *Klassifikationssysteme.* Deutsches Institut für Normung, Berlin-Köln, Beuth Verlag.

Eco, Umberto (1991). *Einführung in die Semiotik.* München, Fink.

Evens, Martha W.; Litowitz, B. E.; Markowitz, J. A.; Smith, R. N.; Werner O. (1980). *Lexical-Semantic Relations: A Comparative Survey.* Edmonton, Alberta, Canada.

Evens, Martha Walton (ed.) (1988). *Relational Models of the Lexicon.* Cambridge University Press.

Fellbaum, Christiane (1990). *English Verbs as a Semantic Net.* In: G. A. Miller et al. *WordNet an Online Lexical Database.* Int. Journ. of Lexicology, Vol. 3, Number 4.

Fellbaum, Christiane (1996). *WordNet: An Electronic Lexical Database and Some of Its Applications.* MIT Press. (to appear)

Fischer, Dietrich H. (1991). *Consistency Rules and Triggers for Thesauri.* International Classification, 18, No. 4, pp. 212 - 225.

Fischer, Dietrich H.; Moehr, Wiebke; Rostek, Lothar (1996). *A Modular, Object-Oriented and Generic Approach for Building Terminology Maintenance Systems.* Galinski & Schmitz (eds.): *TKE'96, Terminology and Knowledge Engineering.* Frankfurt, INDEKS-Verlag.

Forsgren, K. (1977). *Wortdefinition und Feldstruktur. Zum Problem der Kategorisierung in der Sprachwissenschaft.* Göteburger Germanistische Forschungen 16.

Frawley, W. (1988). *New Forms of Specialized Dictionaries.* Internat. Journ. of Lexicography, Vol. 1, Nr. 3.

Frege, G. (1892). *Über Sinn und Bedeutung.* Zeitschr. f. Phil. und phil. Kritik 100. pp. 25-50.

Ganter, Bernhard; Wille, Rudolf (1996). *Formale Begriffsanalyse: Mathematische Grundlagen.* Berlin-Heidelberg, Springer-Verlag.

Hofstadter, Douglas (1996). *Speechstuff and Thoughtstuff: Musings on the Resonances Created by Words and Phrases via the Subliminal Perception of their Buried Parts.* Presentation at the MAICS, Bloomington, Indiana.

Horn, L. (1989).*Morphology, Pragmatics, and the Verb.* In: Powers, J., and de Jong, K. (eds.). Proceedings of the Fifth Eastern State Conference on Linguistics. Columbus, Ohio: Ohio State University.

Iris, Madelyn; Litowitz, Bonnie; Evens, Martha (1988). *Problems of Part-Whole Relations.* In: Evens, Martha W. (ed.). *Relational Models of the Lexicon.* Cambridge University Press, pp. 261 - 288.

Kay, Paul (1971). *Taxonomy and Semantic Contrast.* In: Language, Vol. 47, 4.

Kilgarriff, Adam (1995). *Inheriting polysemy.* In: Patrick Saint-Dizier, Evelyne Viegas (eds.). *Computational lexical semantics*, Cambridge Univ. Press.

Kipke, Uwe; Wille, Rudolf (1987). *Formale Begriffsanalyse erläutert an einem Wortfeld.* LDV- Forum 5, pp. 31-36.

Knuth, D. E. (1993). *The Stanford GraphBase: a Platform for Combinatorial Computing.* New York, Addison Wesley.

Lincke, Angelika; Nussbaumer, Markus; Portmann, Paul (1994). *Studienbuch Linguistik.* Niemeyer, Tübingen.

Luschei, Eugene C. (1962). *The logical systems of Lesniewski.* Amsterdam, North-Holland Pub. Co.

Lutzeier, Peter Rolf (1981). *Wort und Feld. Wortsemantische Fragestellungen mit besonderer Berücksichtigung des Wortfeldbegriffs.* Tübingen, Niemeyer.

Lyons, John (1977). *Semantics.* Cambridge University Press.

Mannila, Heikki; Räihä, Kari-Jouko (1994). *The Design of Relational Databases.* Addison Wesley.

Miller, G. A.; et al. (1990). *WordNet an Online Lexical Database.* Int. Journ. of Lexicology, Vol. 3, Number 4.

Miller, G. A.; et al. (1994). *WordNet: A Lexical Database for English.* Status Report No. 12, Princeton University.

Morris, C. W. (1946). *Signs, Language, and Behaviour.* Prentice Hall.

Nöth, W. (ed.) (1987). *Handbuch der Semeotik.* Mehler, Stuttgart.

Ogden, C. K.; Richards, I. A. (1923). *The Meaning of Meaning.* New York, Harcourt.

Old, L. John (1996). *Synonymy and Word Equivalence.* In: Online Proceedings of the 1996 Midwest Artificial Intelligence Conference, Bloomington, Indiana. (URL: http://www.cs.indiana.edu/event/maics96/Proceedings/old.html)

Pratt, V. R. (1992). *Origins of the calculus of binary relations.* Proc. 7th Annual IEEE Symp. on Logic in Comp. Sci., Santa Cruz, CA, pp 248-254.

Prediger, Susanne (1996). *Symbolische Datenanalyse und ihre begriffsanalytische Einordnung.* Staatsexamensarbeit, Technische Hochschule Darmstadt, Fachbereich Mathematik.

Priss, Uta; Wille, Rudolf (in prep.). *A Formalization of Roget's International Thesaurus.*

Pustejovsky, James (ed.) (1993). *Semantics and the Lexicon.* Kluwer, Dordrecht.

Putnam, H. (1975). *The meaning of 'meaning'.* In: Gunderson, K. (ed.). *Language, mind, and knowledge.* Minneapolis, Univ. of Minnesota Press.

Rahmstorf, Gerhard (1983). *Die semantischen Relationen in nominalen Ausdrücken des Deutschen.* Dissertation, Joh. Gutenberg Universität, Mainz.

Rahmstorf, Gerhard (1991). *Compositional Semantics and Concept Representation.* In: Bock, H. -H.; Ihm, P. (eds.). *Classification, Data Analysis, and Knowledge Organization.* Springer-Verlag.

Rescher, N. (1975). *Applied Logic.* Encyc. Brit., 15 ed., Vol. 11.

Rowley, J. E. (1992). *Organizing knowledge : an introduction to information retrieval.* Brookfield, Ashgate.

Saussure, F. de (1916). *Cours de Linguistique Générale.* Paris, Payot.

Sedelow, Sally Y. (ed.) (1974). *Automated Language Analysis.* Report on research 1973-74, University of Kansas, Lawrence.

Sedelow, Sally Y.; Sedelow, Walter A. (1986). *Thesaural Knowledge Representation.* In: *Advances in Lexicology.* Univ. of Waterloo Centre for the New Oxford English Dictionary-Proceedings.

Sowa, John (1984). *Conceptual Structures: Information Processing in Mind and Machine*. Addison-Wesley.

Sowa, John (1993). *Lexical Structure and Conceptual Structures.* In: Pustejovsky, James (ed.). *Semantics and the Lexicon.* Kluwer, Dordrecht.

Steinberg, Danny D.; Jakobovits, Leon A. (1971). *Semantics. An Interdisciplinary Reader in Philosophy, Linguistics and Psychology.* Cambridge University Press.

Stumme, Gerd (1994). *Boolesche Begriffe*. Diplomarbeit, TH Darmstadt.

Talburt, John R.; Mooney, Donna (1990). *An Evaluation of Type-10-Homograph Discrimination at the Semi-Colon Level in RIT.* Proc. 1990 ACM Symposium On Personal and Small Computers, Arlington, Virgina, pp 156-159.

Touretzky, David S. (1986). *The Mathematics of Inheritance Systems.* Pitman, London.

Ullmann, Stephen (1957). *The Principles of Semantics.* Oxford.

Vogt, Frank; Wille, Rudolf (1994). *TOSCANA - A Graphical Tool for Analyzing and Exploring Data.* In: Tamassia; Tollis (eds.). *Graph Drawing.* Heidelberg, Springer-Verlag, pp. 226-233.

Westerstahl, Dag (1989). *Quantifiers in Formal and Natural Languages.* In: Gabbay, D.; Guenthner F. (ed.). *Handbook of Philosophical Logic.* Vol. 4, Kluwer, Dordrecht, pp. 1 - 131.

Wiegand, H. E.; Wolski, W. (1980). *Lexikalische Semantik.* In: Althaus, Hans Peter; Henne, Helmut; Wiegand, Herbert E. (eds.). *Lexikon der Germanistischen Linguistik.* Tübingen, Niemeyer.

Wille, Rudolf (1984). *Line diagrams of hierarchical concept systems.* Int. Class., 11, pp 77-86.

Wille, Rudolf (1992). *Begriffliche Datensysteme als Werkzeuge der Wissenskommunikation.* In: H. H. Zimermann; H. -D. Luckhardt; A. Schulz (eds.). *Mensch und Maschine – Informationelle Schnittstellen der Kommunikation.* Universitätsverlag Konstanz.

Wille, Rudolf (1994). *Plädoyer für eine philosophische Grundlegung der Begrifflichen Wissensverarbeitung.* In: R. Wille; M. Zickwolff (eds). *Begriffliche Wissensverarbeitung – Grundfragen und Aufgaben.* B.I.-Wissenschaftsverlag, Mannheim.

Wille, Rudolf (1996). *Conceptual Structures of Multicontexts.* In: Eklund; Ellis; Mann (eds.). *Conceptual Structures: Representation as Interlingua.* Berlin-Heidelberg-New York, Springer-Verlag, pp. 23-39.

Winston, Morton E.; Chaffin, Roger; Herrmann, Douglas (1987). *A Taxonomy of Part-Whole Relations*. Cogn. Science. Vol. 11, pp. 417-444.

Wittgenstein, Ludwig (1971). *Philosophische Untersuchungen*. Frankfurt. (First Edition: 1958).

Wunderlich, Dieter (1980). *Arbeitsbuch Semantik*. Königstein/Ts: Athenaeum.

Woods, William A. (1991). *Understanding Subsumption and Taxonomy: A Framework for Progress*. In: Sowa, John (ed.). *Principles of Semantic Networks: Explorations in the Representation of Knowledge*. M. Kaufmann, California, pp. 45 - 94.