

The Formalization of WordNet by Methods of Relational Concept Analysis¹

Uta E. Priss

School of Library and Information Science, Indiana University
Bloomington, IN 47405, upriss@indiana.edu

1 Introduction

In this chapter a mathematical formalization of WordNet that relies on WordNet's hierarchical and relational structure is described. Conceptual hierarchies are formalized in Formal Concept Analysis (Ganter & Wille, 1996), a theory developed at the Technische Hochschule Darmstadt over the last sixteen years (Section 7.2). This theory can be extended to Relational Concept Analysis by adding further relations (Section 7.4). In a linguistic application the conceptual hierarchies can be interpreted as hypernymy orderings, and other semantic relations such as meronymy and antonymy can be taken as additional relations. The theoretical analysis shows dependencies among semantic relations such as inheritance of relations from subconcepts to superconcepts. It does not provide a complete system of axioms for semantic relations, but it can facilitate the investigation of the logical properties of those relations. For example, it does not answer the question of whether meronymy is transitive in general, but it defines sufficient transitivity conditions to identify those properties that intransitive meronymy relations cannot have.

Conceptual structures are modeled as a hierarchical network in the form of a mathematical lattice. This enables a graphical representation that would be difficult to obtain without using mathematical structures. Some of the concepts of the lattice are lexicalized as they are denominated by words. (In what follows, all lexical units which could appear as an entry of a dictionary are called 'word'.) Other concepts are lexical gaps, which exist in the conceptual structure and can

¹The research for this paper was partially supported by the Zentrum für Interdisziplinäre Technikforschung, Darmstadt

be described by their hypernyms, attributes, and other relations, but are not denominated by words. Two formal contexts are needed for the study of semantic relations, a *denotative context*, which contains the denotational meanings of word forms (denotata) and their conceptual ordering, and a *lexical context*, which has the words as constitutive elements. The words are always assumed to be disambiguated (for example by WordNet sense numbers) to avoid problems of polysemy and homonymy. A denotative context is usually incomplete because it is not possible to write a list of all denotata of a language. But, as semantic relations refer to relations among denotata, they cannot be defined on words without studying the denotata in a denotative context. Examples for lexical contexts are lexical fields (Kipke & Wille, 1987). Furthermore, every dictionary or thesaurus can be interpreted as a lexical context. Words are names for concepts in a denotative context and formal objects in a lexical context. Therefore it has to be investigated whether semantic relations have the same properties in both contexts. WordNet is formalized as such a lexical context, but only the noun synsets are investigated in this chapter (Section 7.3). The other parts of speech and more details can be found elsewhere (Priss, 1996). In Section 7.7, three examples of the meronymy relation in WordNet show how this theoretical framework can be used to find irregularities among the semantic relations in WordNet1.5.

2 Formal Concept Analysis

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 7.1 shows an example of a formal context. It has ‘person’, ‘adult’, and so on as formal objects, and ‘young’, ‘old’, ‘female’, and ‘male’ as formal attributes. It should be noted that this use of ‘context’ must be distinguished from the linguistic use of ‘context’. 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 $\iota A := \{m \in M \mid gIm \text{ for all } g \in A\}$ and, analogously, the set of all common formal objects of a set $B \subseteq M$ of formal attributes is $\varepsilon B := \{g \in G \mid gIm \text{ for all } m \in B\}$. For example, in the formal

context in Figure 7.1, $\iota\{\text{man}\} = \{\text{old, male}\}$ and $\varepsilon\{\text{old}\} = \{\text{adult, woman, man}\}$ hold.

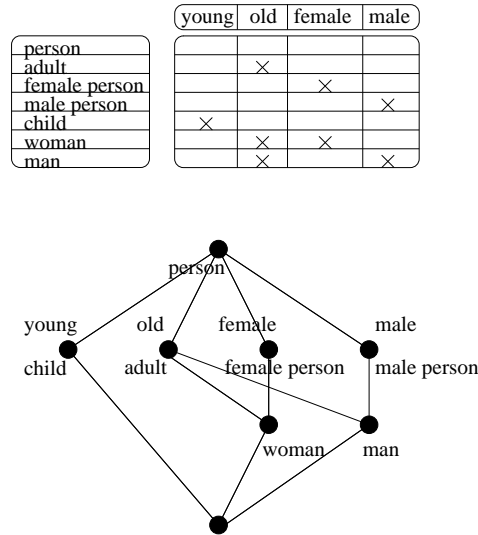


Figure 7.1: A formal context and a line diagram of its concept lattice

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 7.1, $(\{\text{adult, woman, man}\}, \{\text{old}\})$ is a formal concept, because $\iota\{\text{adult, woman, man}\} = \{\text{old}\}$ and $\varepsilon\{\text{old}\} = \{\text{adult, woman, man}\}$. 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, $(\{\text{adult, woman, man}\}, \{\text{old}\})$ as a formal superconcept of $(\{\text{woman}\}, \{\text{old, female}\})$ has more formal objects but less formal attributes than $(\{\text{woman}\}, \{\text{old, female}\})$. 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 excludes 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 $\underline{\mathcal{B}}(G, M, I)$.

Graphically, mathematical lattices can be visualized by line diagrams which represent a formal concept by a small circle. 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 *object concept* and *attribute concept*, 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 7.1 shows the line diagram of the concept lattice of the formal context in Figure 7.1. 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).

3 WordNet as a formal context

As mentioned in the introduction, it is necessary for the formalization of WordNet to define two contexts. In a *denotative context* $\mathcal{K}_D := (D, A_D, I_D)$, denotata $d \in D$ are the formal objects. The set A_D of formal attributes consists of attributes of the denotata. The concepts can additionally be denominated by disambiguated words $w \in W$ via the function $dnt : W \rightarrow \mathcal{B}(\mathcal{K}_D)$. As the words are disambiguated, dnt is really a function. A relational structure consisting of a denotative context \mathcal{K}_D , a set W of words, the function dnt , and further optional relations on the denotata, attributes or concepts, is called *denotative structure* and is denoted by \mathcal{S}_D . A *lexical context* $\mathcal{K}_L := (W, A_L, I_L)$ consists of a set W of disambiguated words as formal objects, a set A_L of attributes, and a relation I_L . The attributes in A_L can be attributes of the denotata of the words, connotative attributes, or formal attributes, such as ‘has four letters’. In many applications the attributes of a word in \mathcal{K}_L are the attributes of its denotata in an underlying \mathcal{K}_D , that means I_L is defined by $w I_L m : \iff (dnt(w) \leq \mu m \text{ in } \underline{\mathcal{B}}(\mathcal{K}_D))$ and therefore $A_L = A_D$. In other

applications, especially in Componential Analysis, which uses a combination of denotational and connotational attributes, the lexical context would not have A_D as the set of attributes. An equivalent representation of $\underline{\mathcal{B}}(\mathcal{K}_D)$ is $\underline{\mathcal{B}}(\mathcal{K}_D^*)$ with $\mathcal{K}_D^* := (D \cup W, A_D, I_D^*)$, where the words are interpreted as a second set of objects joined with the set of denotata. The relation I_D^* is then defined as $I_D \cup I_L$. As \mathcal{K}_L (with set A_D of attributes) is obviously contained in \mathcal{K}_D^* , $\underline{\mathcal{B}}(\mathcal{K}_L)$ is isomorphic to a join-preserving sublattice of $\underline{\mathcal{B}}(\mathcal{K}_D^*) (\cong \underline{\mathcal{B}}(\mathcal{K}_D))$.

WordNet is a lexical context $\mathcal{K}_{\text{WN}} := (W, S, I_{\text{WN}})$ with disambiguated words $w \in W$ as objects. As there is no set A_D of attributes given in WordNet that discriminates the words explicitly, the following construction is needed. An equivalence relation, synonymy SYN, is defined on the set W of disambiguated words via $w_1 \text{ SYN } w_2 : \iff \text{syn}(w_1) = \text{syn}(w_2)$, where the synset of a word w is denoted by $\text{syn}(w)$. Then an order relation, hyponymy HYP, is defined on the set S of synsets and the concept lattice is computed as the Dedekind closure of this ordered set. Therefore, formally each synset is interpreted as an attribute, for example, the synset ‘dog’ is interpreted as the attribute ‘to be a dog.’ The relation I_{WN} is defined by $w I_{\text{WN}} \text{syn}(w_1) : \iff \text{syn}(w) \text{ HYP } \text{syn}(w_1)$. It follows that the extent of a concept consists of all words that belong to the synsets of that concept or lower concepts. The intent of a concept consists of the synsets of that concept or higher concepts. Each concept can be an object concept for at most one synset. An open question concerning WordNet is whether the set S of formal attributes could be replaced by a set A_D of attributes of the denotata or whether that would change the hypernymy ordering of the synsets. Prototypical attributes cannot be used as A_D , because they are not inherited by all subconcepts, for example, not all birds fly. Obviously the underlying context \mathcal{K}_D is not considered at all for the construction of \mathcal{K}_{WN} . But it can be shown (Priss, 1996) that, assuming some axioms hold, \mathcal{K}_{WN} reveals information about a possibly corresponding \mathcal{K}_D .

4 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 or attributes. And it is an extension of Woods’ (1990) quantificational tags and inheritances. In what follows, only binary relations $r \subseteq G \times G$ are considered. These relations are transferred to relations among concepts, i. e., $R \subseteq \mathcal{B}(G, M, I) \times \mathcal{B}(G, M, I)$, according to the following definitions. The quan-

tifiers that are used in the definitions can be natural language quantifiers or mathematical expressions, such as $\|\text{all}\|$, $\|\text{at least } 1\| =: \|\geq 1\|$, or $\|\text{exactly } 1\| =: \|\equiv 1\|$ (for more details on natural language quantifiers see Westerstahl (1989)).

Definition 7.1:

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

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

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

$$c_1 R^r[Q^1, Q^2; Q^3, Q^4] c_2 :\iff c_1 R^r[Q^1, Q^2;] c_2 \text{ and } c_1 R^r[; Q^3, Q^4] c_2 \quad (3)$$

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 leads therefore to several different relations R^r among concepts. The terms ‘quantificational tag’ and ‘relational component’ are taken from Woods’ terminology. The formalization can be best 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 combination 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 7.2:

‘ $R^r[\|\geq 1\|, \|\geq 1\|; \|\geq 1\|, \|\geq 1\|]$ ’ is abbreviated as R^r_0 .

‘ $R^r[\|\text{all}\|, Q^2; \|\text{all}\|, Q^4]$ ’ 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 $R^r_{(Q^4, Q^2)}$

Besides its applications to the modeling of lexical databases, this formalization can be used to describe functions $R^r_{(\geq 0; 1)}$, bijections $R^r_{(1; 1)}$, or Cartesian products

$R_{(\text{all};\text{all})}^r$. It is useful to characterize a conceptual relation by considering the concepts only and not the objects. This leads to a classification of certain relations R^r into different types.

Definition 7.3:

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 $c_1, c_2 \in \mathcal{B}(G, M, I)$

$$c_1 R c_2 \iff Q_{c_{11} \leq c_1}^5 Q_{c_{22} \leq c_2}^6 : c_{11} R c_{21} \quad (4)$$

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

It seems to be impossible to find such quantifiers Q^5 and Q^6 for all $R^r[Q^1, Q^2;]$, but it can be proved (Priss, 1996) that the quantifiers Q^5 and Q^6 fulfill equivalence (4) for the quantifiers Q^1 and Q^2 in Table 7.1 (n is always larger or equal to 1).

Q^1	Q^2	Q^5	Q^6
$ \text{all} $	$ \text{all} $	$ \text{all} $	$ \text{all} $
$ \geq 1 $	$ \geq 1 $	$ \geq 1 $	$ \geq 1 $
$ \text{all} $	$ \geq n $	$ \text{all} $	$ \geq 1 $
$ \geq n $	$ \text{all} $	$ \geq 1 $	$ \text{all} $

Table 7.1

Furthermore, if $Q^1, Q^2 \in \{||> (\text{all} - n)||, ||\geq n||, ||\leq (\text{all} - n)||, ||< n||\}$ ($n \geq 1$) then $Q^5, Q^6 \in \{||\text{all}||, ||\geq 1||\}$. An interpretation of Table 7.1 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’ 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’. From equivalence (4) follows equivalence (5).

$$\begin{aligned} \gamma g_1 R^r[Q^1, Q^2;] \gamma g_2 &\iff \\ Q_{g_{11} \in \text{Ext}(\gamma g_1)}^5 Q_{g_{21} \in \text{Ext}(\gamma g_2)}^6 : \gamma g_{11} R^r[Q^1, Q^2;] \gamma g_{21} &\quad (5) \end{aligned}$$

It is therefore enough to consider object concepts in order to determine the type of relation.

Additional properties of the relation r have consequences for the relations R^r . For example, if r is irreflexive and transitive (and thus, by implication, antisymmetric) and all sets of objects are finite, then $R_{(\leq 0; \geq 1)}^r$, $R_{(\geq 1; \geq 0)}^r$, and $R_{(\geq 1; \geq 1)}^r$ are also irreflexive, antisymmetric, and transitive. 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 holds

$$c_1 R_{(\geq 0; \geq 1)}^= c_2 \iff c_1 \leq c_2 \quad (6)$$

$$c_1 R_{(\geq 1; \geq 0)}^= c_2 \iff c_1 \geq c_2 \quad (7)$$

$$c_1 R_0^= c_2 \iff Ext(c_1) \cap Ext(c_2) \neq \emptyset \quad (8)$$

$$c_1 R_{(\geq 1; \geq 1)}^= c_2 \iff c_1 = c_2 \quad (9)$$

Therefore the conceptual ordering itself results from a relation between objects. For further details and the proofs of the statements above see Priss (1996).

5 Meronymy

Semantic relations, such as meronymy, synonymy and hyponymy, are according to WordNet terminology relations that are defined among synsets. They are distinguished from lexical relations, such as antonymy, which are defined among words and not among synsets. In our terminology a relation $s \subseteq W \times W$ among disambiguated words is called a *semantic relation* if

$$\begin{aligned} dnt(w_1) = dnt(w_2) \implies \\ \forall_{w \in W} (w_1 s w \iff w_2 s w) \text{ and } (w s w_1 \iff w s w_2) \end{aligned} \quad (10)$$

is fulfilled. Relations $s \subseteq W \times W$ that do not fulfill condition (10) are called *lexical relations*. In this section meronymy is formally defined as a semantic relation according to this terminology. In the next section hyponymy and synonymy are defined. While antonymy is a lexical relation, *indirect antonymy* is a semantic relation which holds among ‘antonymous’ concepts.

Although meronymy is a hierarchical relation it should not be modeled as a mathematical lattice. One obvious reason against a modeling of 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’ both 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 7.4:

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 $R_{(Q^4; Q^2)}^m$ where m is a meronymy relation among denotata, i.e.

$$w_1 MER_{(Q^4; Q^2)}^m w_2 : \iff dnt(w_1) R_{(Q^4; Q^2)}^m dnt(w_2)$$

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

In contrast to indirect antonymy whose types are distinguished by the relational components, 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_0^m : facultative-facultative²; 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.
- $MER_{(\geq 0; \geq 1)}^m$: canonical-facultative; for example, all door-handles are parts of doors, but not all doors have to have handles.
- $MER_{(\geq 1; \geq 0)}^m$: facultative-canonical; for example, all ice-cubes consist of water, but not all water is frozen.
- $MER_{(\geq 1; \geq 1)}^m$: canonical-canonical; for example, each bird feather is part of a bird, and each bird has feathers.

The question of transitivity of meronymy, which has been widely discussed (Winston et al., 1987), is not answered here in general. But it can be shown (Priss, 1996) that if m is defined (according to Definition 7.4) to be irreflexive, antisymmetric,

²Cruse uses ‘facultative’ and ‘canonical’ instead of Lyons’ (1977) ‘contingent’ and ‘necessary’.

and acyclic then it follows that $MER_{(\geq 0; \geq 1)}^m$, $MER_{(\geq 1; \geq 0)}^m$, and $MER_{(\geq 1; \geq 1)}^m$ are also irreflexive, antisymmetric, and acyclic. And if m is transitive, then $MER_{(\geq 0; \geq 1)}^m$, $MER_{(\geq 1; \geq 0)}^m$, and $MER_{(\geq 1; \geq 1)}^m$ 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 $MER_{(\geq 0; \geq 1)}^m$ would be contradictory).

Most classifications of semantic relations are based on qualitatively different attributes of the relations. Relational Concept Analysis shows, however, that there may even be differences in the quantificational conditions. These can be discovered in a formalization of the four models of Iris et al. (1988):

- Functional component: m can be described as ‘part of and functional dependency between part and whole’; $MER_{(\geq 0; 1)}^m$ 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. $MER_{(\geq 0; n)}^m$, for $n > 1$, is also possible. For example, for any section, a border belongs to exactly two countries.
- Segmented whole/mass nouns/is-substance-of relation: $MER_{(\text{some}; \leq 1)}^m$ is a typical form, because, for example, each pile of paper consists of several sheets, and each sheet belongs to at most one pile. Using the ||some|| quantifier leaves the actual number of pieces, needed to form the whole, open.
- Membership relation: $MER_{(\geq 0; > 1)}^m$ is possible, because, for example, a human is usually a member of different sets (clubs, family, cultures) at the same time.
- Individual concepts: $MER_{(1; 1)}^m$ holds, because, for example, all objects which are ‘Princeton (NJ)’ are part of the only existing state ‘New Jersey’. (There is no possible confusion with other towns called ‘Princeton’, because of the disambiguation of the words.) This also holds for some abstract nouns: each day has one morning, while each morning is part of one day.

This analysis of the differences in the quantificational conditions is missing in WordNet, which distinguishes only between part-of, substance-of and member-of meronymy relations.

6 Hyponymy and Synonymy

If a relation r is the equality relation, then equivalences 6 - 9 of Section 7.4 show that the corresponding conceptual relations $R^=$ coincide with the conceptual ordering of the lattice. In the linguistic application these relations are given special names.

Definition 7.5:

In a denotative structure \mathcal{S}_D the following semantic relations are defined:

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.

$$w_1 HYP w_2 : \iff dnt(w_1) \leq dnt(w_2) (\iff dnt(w_1) R_{(\geq 0; \geq 1)}^= dnt(w_2))$$

The inverse relation of hyponymy is called *hypernymy*.

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

$$w_1 \neg DISJ w_2 : \iff dnt(w_1) R_0^- dnt(w_2)$$

Two disambiguated words are called *synonyms* if they denote the same concept, i.e.

$$w_1 SYN w_2 : \iff dnt(w_1) = dnt(w_2) (\iff dnt(w_1) R_{(\geq 1; \geq 1)}^= dnt(w_2))$$

These definitions are consistent with the formalization of WordNet in Section 7.3. If an underlying denotative context for WordNet is assumed then a word is a hyponym of another word if its denotata are a subset of the denotata of the other word. Synonymy is defined between two words if their concepts have the same extent (and therefore also the same intent) in \mathcal{K}_D . It should be noted that these definitions of synonymy and hypernymy depend on the context \mathcal{K}_D . It is possible to apply these definitions to any other kind of context, but then, if the underlying context is not a \mathcal{K}_D , these definitions can have a completely different meaning.

Concerning the combination of hyponymy and meronymy, it follows from equivalence (4) that some quantifiers cause inheritance of relations: $MER_{(\geq 0; \geq 1)}^m$ is inherited by hypernyms of the whole and hyponyms of the part; $MER_{(\geq 1; \geq 0)}^m$ is inherited

by hyponyms of the whole and hypernyms of the part; MER_0^m is inherited by hypernyms of part and whole. (This contradicts the statement about transitivity in the paper of Winston et al. (1987, p. 435).) It is possible to have a meronymy and a hyponymy relation at the same time between two concepts. For example, ‘ice’ is a kind of ‘water’ and consists of ‘water’, or ‘musical strings’ are ‘musical supplies’ and part of ‘musical supplies’ at the same time. However, usually it is possible to add additional concepts with single relations. 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 part of ‘musical instruments’, which are a kind of ‘musical supplies.’

7 Identifying irregularities 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 Smalltalk software 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 automatically corrected. 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 automatically correct 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_0^m ’ or ‘ $MER_{(\geq 1, \geq 1)}^m$ ’ and so on. But as there are comparatively few meronymy relations implemented in WordNet, a first approach assumes all of them to be of the strongest kind, $MER_{(\geq 1, \geq 1)}^m$. If irregularities are found, they can be changed to weaker kinds such as ‘ $MER_{(\geq 1, \geq 0)}^m$ ’ or be otherwise repaired.

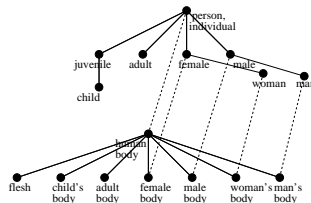


Figure 7.2: An example of the part-of meronymy in WordNet1.5

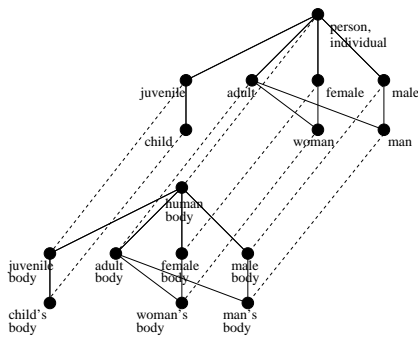


Figure 7.3: A modified version of the example in Figure 7.2

The Figures 7.2, 7.4, and 7.6 show parts of the WordNet1.5 lattice, Figures 7.3, 7.5, and 7.7 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 7.2 ‘human body $MER_{(\geq 1, \geq 1)}^m$ 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 $MER_{(\geq 1, \geq 1)}^m$ 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 7.2 to Figure 7.3 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.

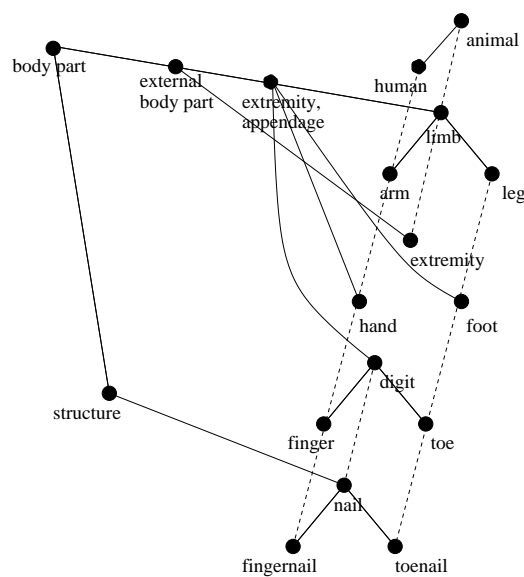


Figure 7.4: Another example of the part-of meronymy in WordNet1.5

The reason for the irregularities in Figure 7.4 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, 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 7.5 shows a more regular pattern than Figure 7.4.

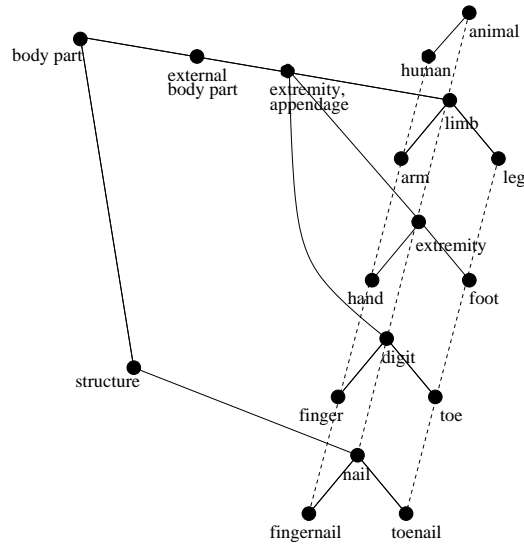


Figure 7.5: A modified version of the example in Figure 7.4

The last example of the substance-of meronymy in Figure 7.6 probably needs improvement, too, as all those fluids should have a common substance – water molecules (see Figure 7.7). 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.

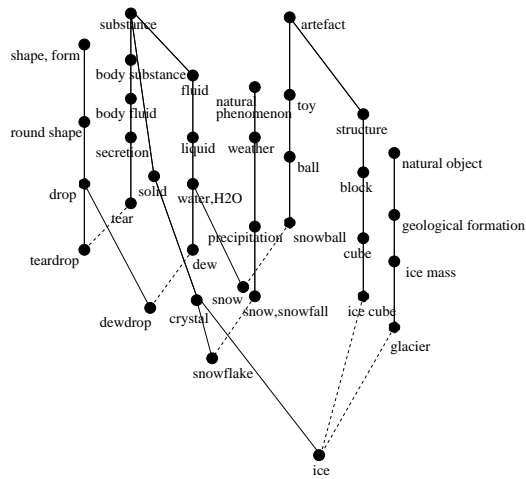


Figure 7.6: An example of the substance-of meronymy in WordNet1.5

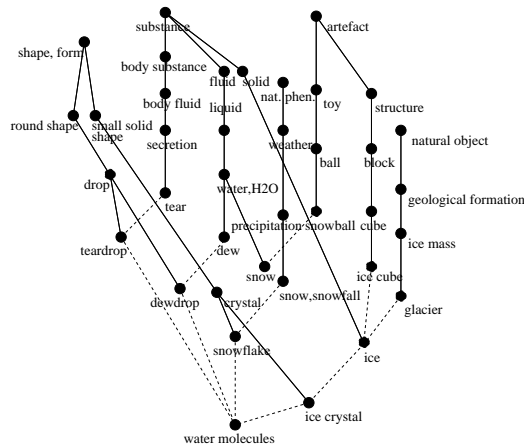


Figure 7.7: A modified version of the example in Figure 7.6

References

- Cruse, D. A. (1986). *Lexical Semantics*. Cambridge, New York.
- Fischer, Dietrich H. (1991). *Consistency Rules and Triggers for Thesauri*. Int. Classif., 18, No. 4.
- Ganter, Bernhard; Wille, Rudolf (1996). *Formale Begriffsanalyse: Mathematische Grundlagen*. Springer-Verlag.
- 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.
- Kipke, Uwe; Wille, Rudolf (1987). *Formale Begriffsanalyse erläutert an einem Wortfeld*. LDV- Forum 5, pp. 31-36.

- Lyons, John (1977). *Semantics*. Cambridge University Press.
- Priss, Uta E. (1996). *Relational Concept Analysis: Semantic Structures in Dictionaries and Lexical Databases*. Dissertation, TH-Darmstadt.
- Westerstahl, Dag (1989). *Quantifiers in Formal and Natural Languages*. In: Gabbay, D.; Guenther F. (ed.). *Handbook of Philosophical Logic*. Vol. 4, Kluwer, Dordrecht.
- Winston, Morton E.; Chaffin, Roger; Herrmann, Douglas (1987). *A Taxonomy of Part-Whole Relations*. *Cogn. Science*. Vol. 11, pp. 417-444.
- Woods, William A. (1990). *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.