The Logical Structure of Semantic Networks and its Role in Natural Language Processing

Hermann Helbig, Ingo Phoenix

WG Intelligent Information and Communication Systems
University of Hagen, Germany

Abstract. Semantic Networks (SN) are a knowledge representation paradigm especially suited for the meaning representation of natural language expressions. Whereas SN formalisms often do not give a clear logical specification of their basic constructs, logical formalisms, on the other hand, do seldom bother about a connection to natural language or about the intensional interpretation of the non-logical symbols used in the calculus. In order to bridge this gap, the relations and functions used in a semantic network must be given a logical characterization. Moreover, the logical descriptions must be provided with a richer inner structure compared to traditional logical knowledge bases. The paper exemplifies this strategy for Multilayered Extended Semantic Networks (the so-called MultiNet paradigm), but the basic approach should also be of importance for other models of natural language semantics and knowledge representation. In particular, it is argued that the axioms defining the logical properties of the expressional means of a knowledge base have to be classified according to well-defined criteria. The typology of axioms and the assignment of concepts to different semantic layers derived from this classification have an important influence on the inferences carried out over such a knowledge base. This is in contrast to logical knowledge bases where one normally deals with an axiomatic apparatus having a very flat structure or no structure at all.

1 Introduction

1.1 General Remarks

Semantic networks [1,2] belong to the knowledge representation formalisms well-suited for the meaning representation of natural language (NL) expressions. However, there is often no clear understanding of the nodes and arcs contained in an SN [3]. Logically oriented systems, like the different kinds of Description Logics [4], by contrast, do seldom bother about a close connection to NL, especially regarding lexical semantics and the description of large computational lexica. In addition, they do not define a structure on their axiomatic apparatus which is typically considered a "*flat*" list of well-formed formulas assumed to express valid assertions. In contrast to that it is argued in this paper that axioms in a real-life knowledge base have to be classified according to different criteria, where each class or axiom type has a special influence on the inference process. However, the semantic network paradigms which have a logical underpinning, like SNePs [5], Conceptual Structures [6] or Structured Inheritance Networks [7], do not support a classification of expressional means (and in particular of axioms) which could be used for controlling the inference process. With Multilayered Extended Semantic Networks (the MultiNet paradigm [8]), this situation has changed. MultiNet is a very comprehensively described semantic formalism and successfully used in real NLP applications. It is employed for describing large semantically based computational

lexica [9], for representing the results of syntactic-semantic analysis working over large text corpora [10], and for the implementation of NL interfaces [11]. Its application in the InSicht question-answering system [12] has been successfully evaluated in the CLEF contest [13]. InSicht uses a MultiNet knowledge base automatically generated from 4,9 Mio sentences [14].
MultiNet extends simple semantic networks by the following features (see also Sect. 2): a) Every node is labeled by a sort from a predefined ontology of sorts and by bundles of layer attributes. b) MultiNet admits functions and relations of arbitrary arity. c) The arcs (relations) are formally characterized by associated axioms which are structured according to a well-defined typology. This paper will emphasize this typology and its role for controlling the inference processes. d) Subnetworks can be encapsulated to form concepts of higher order which can be connected to other concepts by relations and functions. e) The relationships in the network are assigned knowledge types with regard to their arguments which discern categorical (strict) knowledge from prototypical (default) knowledge and modal restrictions.

1.2 Related Work

Semantic Networks have a long tradition as a paradigm for representing cognitive structures, starting with Quillian [1]. As to the formal specification and logical underpinning of relations and functions used in such networks, we find a logically oriented and a more linguistically oriented approach. The works in the first line, like Shapiro's SNePS [5], Sowa's Conceptual Structures [6], and Brachmann's KL-ONE [7], have a clear logical foundation, but none of them give a systematic and complete description of the relations and functions constituting an SN. By contrast, linguistically oriented work normally discusses selected semantic relations (so-called cognitive roles or theta-roles) in greater detail [15,16]. However, these verbal discussions often lack a logical underpinning, and the proposed relations and roles are not contrasted with each other to form a balanced system of expressional means. In SNePS [5], for example, the guideline for choosing the appropriate relations is deliberately given the status of a recommendation only. To meet the criteria to be fulfilled by a knowledge representation formalism useful for NLP (especially the universality, homogeneity, and interoperability requirements [8, Chap. 1]), one must have a clear verbal definition of the representational means and a corresponding formal specification. This is necessary for consistent use of the formalism when characterizing a large number of lexemes.
With MultiNet we have developed both, a semantic formalism and a linguistic theory, by fixing a set of expressional means and formalizing axioms which describe these functions and relations in a consistent and all-embracing framework. The MultiNet paradigm is also well-supported by software tools, like the workbench MWR for the knowledge engineer and the workbench LIA for the computer lexicographer (see [8, Chapt. 14]). Figure 1 shows a the semantic representation of a sample sentence generated by means of MWR (exhibiting also the layer information of a selected node).

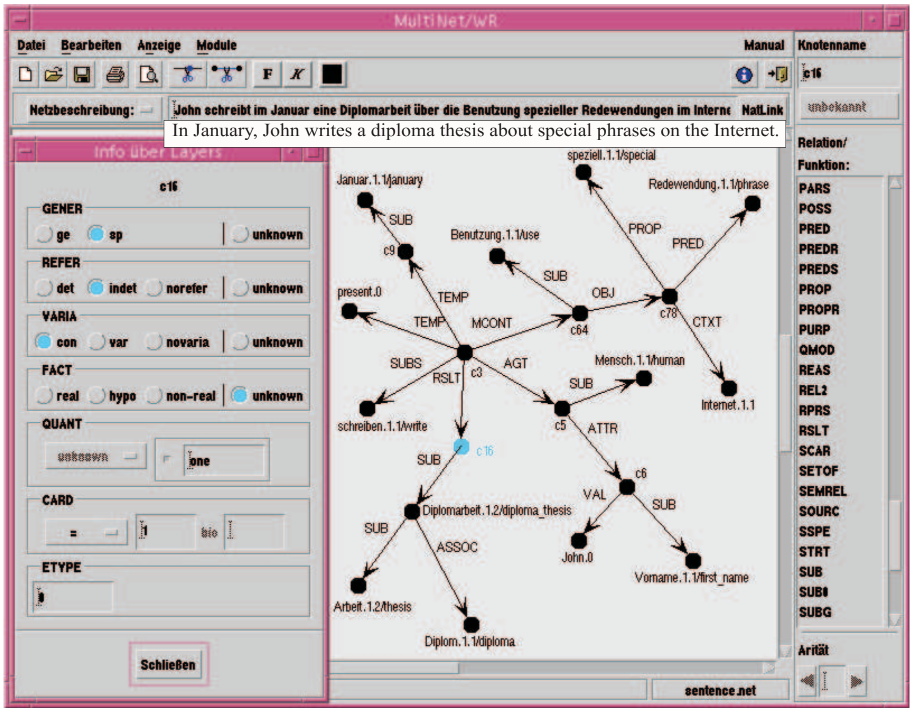


Fig. 1. The workbench for the knowledge engineer MWR

2 The Expressional Means of MultiNet

2.1 Sorts

Sorts are needed for characterizing the algebraic properties of relations and functions of a knowledge representation system. They describe the domains and ranges of these expressional means, i.e. their signatures. MultiNet distinguishes forty-five classes of conceptual entities [8, Sect. 17.1]. Table 1 shows that part of the sortal hierarchy relevant to this paper. Sorts do not only serve to define the signatures of relations and functions, they even have an influence on the applicability of axioms (see Sect. 3) and on inferences (Sect. 4).

2.2 Layer Attributes

Facticity. We discern three possible kinds of facticity expressed by the attribute FACT: [FACT=*real*] for existing entities (*Eiffel tower*), [FACT=*non*] for non-existing entities (*the light ether*), and [FACT=*hypo*] for hypothetical entities (*quarks*). The assignment of facticity values induces a stratification of the conceptual world into layers of existing, non-existing and hypothetical entities. In addition to the *extensional negation* expressed

Table 1. Detail from the hierarchy of ontological sorts

entity [*ent*]	
object [*o*]	
concrete object [*co*]	*house*, *milk*
abstract object [*ab*]	
attribute [*at*]	*height*, *weight*, *length*
situational object [*abs*]	*race*, *robbery*, *movement*
...	
situation [*si*]	
dynamic situation [*dy*]	*write*, *explode*
static situation [*st*]	*stand*, *be ill*
situational descriptor [*sd*]	
time [*t*]	*yesterday*, *Monday*, *tomorrow*
location [*l*]	*here*, *there*
modal situational descriptor [*md*]	*impossible*, *necessary*, *desirable*
quality [*ql*]	
property [*p*]	
total quality [*tq*]	*dead*, *empty*, *green*
gradable quality [*gq*]	*friendly*, *expensive*
...	
...	
quantificators and measurements [*qn*]	*all*, *many*, *two litres*
formal entity [*fe*]	(meta level entities like figures and tables)
...	

by a non-existing situation with [FACT=*non*], MultiNet also supports the *intensional negation* of a situation *s*, expressed by the relation $(s\ \mathrm{MODL}\ {*}\mathrm{NON})$. Both types occur in the example "*It is not true that* [*Peter didn't*$_{(MODL*NON)}$ *drive to Boston with his car*]$_{[FACT=non]}$", whose semantic representation is shown in Fig. 2. The law of double negation can be applied to ascertain the truth value of the sentence "*Peter drove to Boston with his car*". However, from the perspective of NL semantics and pragmatics, the latter sentence is not equivalent to the original one.

Genericity. The GENER attribute (degree of generality) divides the world of concepts into generic objects with [GENER=*ge*] (*house*) and specific objects with [GENER=*sp*] (⟨*my house*⟩). In this way, assertions about the generic concept can be clearly separated from assertions about instances of that concept. Generic concepts are also needed to model prototypical knowledge. Consider the example "*Lions feed on antelopes*". A formalization by a universal quantifier ranging over all lions would be inadequate because the sentence expresses only default knowledge. Layer features are useful to model the combinatorics of quantification and determination in natural language. The interplay of these operators with lexemes designating concepts can be captured by a method called 'layer unification' [17].

Table 2. Strongly abbreviated description of relations used in this paper

Relation	Signature	Short Characteristics
AFF	$si \times [o \cup si]$	C-Role – Affected object
AGT	$si \times o$	C-Role – Agent
ANTE	$[t \cup si] \times [t \cup si]$	Temporal successorship
ATTR	$o \times at$	Specification of an attribute
AVRT	$si \times o$	C-Role – Averting/Turning away from an object
CAUS	$si' \times si'$	Relation between cause and effect (Causality)
CIRC	$si \times [si \cup abs]$	Relation between situation and circumstance
COMPL	$p \times p$	Complementarity relation
DIRCL	$[si \cup o] \times l$	Relation specifying a direction
FIN	$si \times [t \cup si]$	Relation between a situation and its temporal end
LOC	$[o \cup si] \times l$	Relation specifying the location
MIN	$qn \times qn$	Smaller-than relation
MODL	$si \times md$	Relation specifying a restricting modality
OBJ	$si \times [o \cup si]$	C-Role – Neutral object of a situation
ORNT	$si \times o$	C-Role – Orientation of a situation toward something
PARS	$co \times co$	Part-whole relationship
PROP	$o \times p$	Relation between object and property
SCAR	$st \times o$	Relation between state and carrier of the state
SSPE	$st \times ent$	Relation between state and state-specifying entity
SUB	$o \times \overline{o}$	Relation of conceptual subordination (for objects)
SUBS	$si \times \overline{si}$	Relation of conceptual subordination (for situations)
TEMP	$si \times [t \cup si]$	Relation specifying the temporal embedding of a situation
VAL	$at \times [o \cup qn \cup p \cup fe]$	Relation between an attribute and its value

2.3 Relations and functions

To characterize an SN, we need a precise specification of the relations corresponding to the arcs (links). To this end, MultiNet provides about 140 relations and functions described on the basis of a uniform schema. For every relation or function R, this schema comprises: (1) a mnemonic remark, (2) a signature based on the sorts introduced in Sect. 2.1, (3) a verbal definition, (4) a set of question patterns aiming at R in the question-answering game, and (5) an explanation establishing connections to other expressional means and distinguishing R from other relations and functions. The formal definition of relations and functions is given by means of logical axioms (see Sect. 3). Table 2 sketches the relations and functions needed for our discussion. Here, the notation si' demands [FACT $=$ *real*], and the notation $\overline{si}$ demands [GENER $=$ *ge*].

The main characteristics of an SN in MultiNet format are shown by an example in Fig. 2, where every arc is classified with respect to its first and second argument as belonging to the categorical valid knowledge (c), situationally bounded knowledge (s), or semantically restrictive knowledge (r).

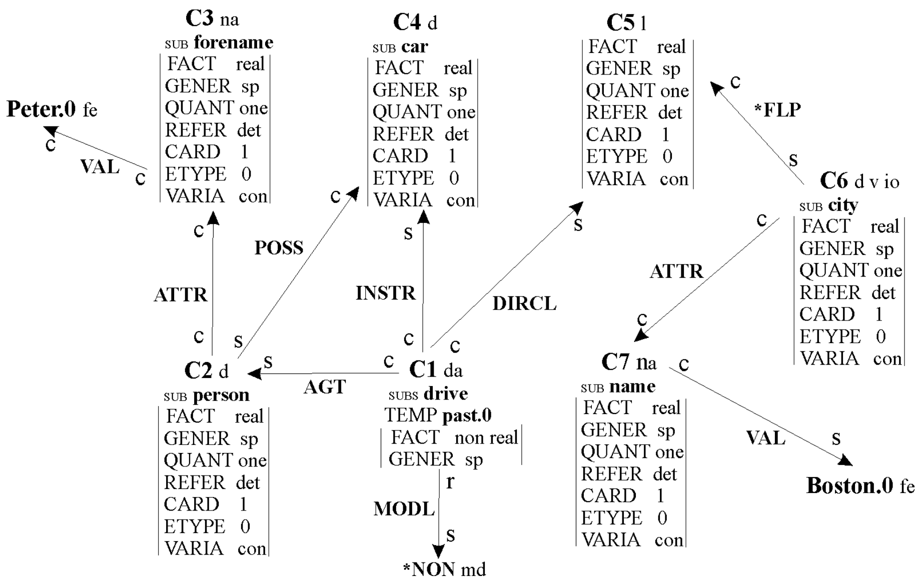


Fig. 2. *"It is not true that Peter didn't drive to Boston with his car"*

3 A Typology of Axioms for Inferences over an SN

In the following, we use a first-order language to formulate axioms. The axioms are written as implicational rules or equivalences of the general form

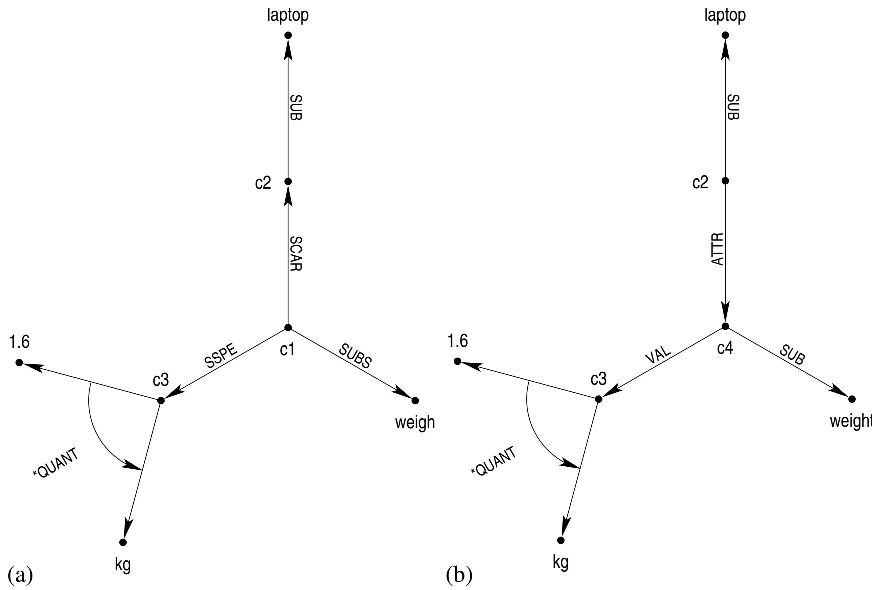
$$L_1 \wedge \cdots \wedge L_n \rightarrow (\exists x_1, \ldots, x_k) R_1 \wedge \cdots \wedge R_m \quad \text{or}$$
$$(\exists x_1, \ldots, x_k) L_1 \wedge \cdots \wedge L_n \rightarrow (\exists y_1, \ldots, y_\ell) R_1 \wedge \cdots \wedge R_m$$

where the L_i and R_j are (possibly negated) literals which correspond to the edges of the subnetworks described by premise and conclusion. Layer features and sorts can also be used in these axioms to further constrain the admissible variable bindings. Notice that variables not explicitly bound by an existential quantifier will be regarded universal.

The way in which such axioms will be used in inferences over MultiNet representations is demonstrated by an example of linking *"The laptop weighs 1.6 kilograms"* and *"The laptop has a weight of 1.6 kilograms"*. Consider the MultiNet representation of these sentences shown in Fig. 3 below.[1]

Given the question *"What is the weight of the laptop?"*, a question pattern will be generated from the MultiNet representation which comprises the conjunction of edge literals, i.e. $(x\,\mathrm{SUB}\,\mathit{laptop}) \wedge (x\,\mathrm{ATTR}\,y) \wedge (y\,\mathrm{SUB}\,\mathit{weight}) \wedge (y\,\mathrm{VAL}\,v)$. The pattern can be directly matched with the network shown in Fig. 3.b, based on the substitution

[1] Layer features are not displayed for simplicity. The function *QUANT serves to construct a measurement (*QUANT 1.6 *kg*) from the given numerical quantificator 1.6 and measurement unit *kg*.



(a) (b)

Fig. 3. MultiNet representation of sentences (a) "The laptop weighs 1.6 kilograms" and (b) "The laptop has a weight of 1.6 kilograms"

$\{x/c_2, y/c_4, v/c_3\}$. However, an axiom is needed to prove the question from the network shown in Fig. 3.a.

The relationship between weighing and having a weight can be expressed as follows,

$$(w\,\text{SUBS}\,\mathit{weigh}) \wedge (w\,\text{SCAR}\,k) \wedge (w\,\text{SSPE}\,q) \rightarrow (\exists a)(w\,\text{ATTR}\,a) \wedge (a\,\text{SUBS}\,\mathit{weight}) \wedge (a\,\text{VAL}\,q)\,.$$

In order to prove "*x* has a weight of *v*" from the semantic network representation of "*x* weighs *v*", a backward chaining step must then be carried out which reduces the representation of the former sentence to the representation of the latter sentence. Following that, the constraints expressed by the layer features must be checked in order to ensure the correctness of the result on a more fine-grained level of meaning analysis.

From a linguistic view, the standard first-order logic (FOL) is too rigid with regard to the validity of expressions. While a logical expression is either true or false in FOL, a semantic formalism dealing with NL cannot assume this even for the basic assertions. Moreover, logical calculi normally do not give a clue how to use the axioms in an effective inference strategy. Additional information about the axioms is needed to achieve this. These considerations suggest a cross-classification of axioms according to several criteria (detailed in Sect. 3.1–3.4), which result in eight basic types of axioms shown in Fig. 4.

Though our discussion is based on MultiNet expressions, the axiom types and their effect on inferences should be of relevance to other frameworks, too. In the following

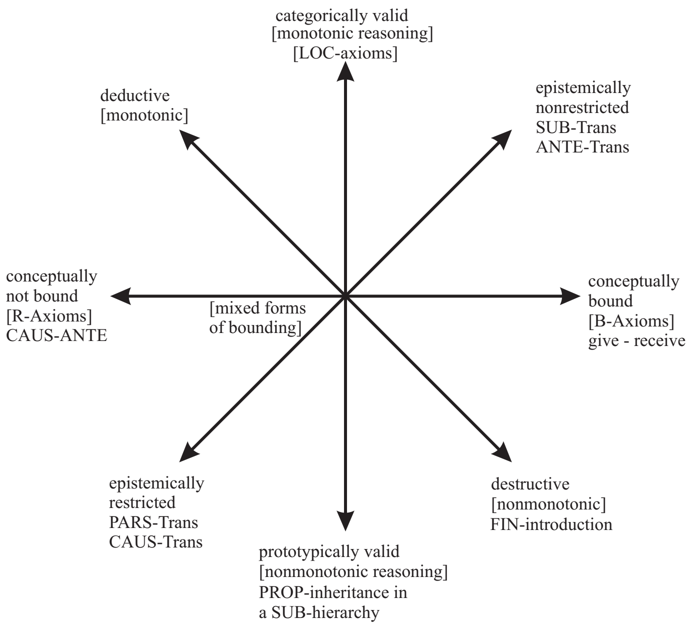


Fig. 4. The Classification of Axioms

Sections 3.1 through 3.4 we give at first a description of the different types of axioms, while the influence of these types on the inference process is discussed in Sect 4 yielding also a motivation for the typology of axioms presented in this paper.

3.1 Conceptually Bound vs. Conceptually Non-bound Axioms

R-axioms. From a syntactical point of view, there are two types of expressions describing axiomatic knowledge. The first type contains no lexical constants but only relation and function symbols (apart from logical signs). These expressions are called *conceptually non-bound* or *R-Axioms*. The following R-Axiom connects causality and time, saying that the effect never takes place before the cause:

$$(x\,\mathrm{CAUS}\,y) \rightarrow \neg(y\,\mathrm{ANTE}\,x) \tag{1}$$

An example is “Since temperatures fell below 0°C, the river froze.” The axiom lets us conclude that the freezing of the river did not take place before temperatures fell below 0°C. Other examples of conceptually non-bound axioms are given by (3), (5), and (6) below. The property of conceptual boundedness affects the inference strategy, since R-axioms (which also comprise the transitivity, symmetry, and reflexivity properties of certain relations) have a global effect on the inference process (see Sect. 4.1).

B-axioms. Axioms containing the representative of at least one concept are called *conceptually bound* or B-axioms. An important and quantitatively large source of axioms in NL semantics is given by meaning postulates expressing entailments and presuppositions connected with lexicalized concepts (especially with the meanings of verbs).[2] Thus, with every selling act s there is a buying act b entailed by s. The corresponding relationship is given by the following axiom:

$$(s\,\mathrm{SUBS}\,\mathit{sell}) \wedge (s\,\mathrm{AGT}\,a) \wedge (s\,\mathrm{OBJ}\,o) \wedge (s\,\mathrm{ORNT}\,d) \rightarrow$$
$$\exists b (b\,\mathrm{SUBS}\,\mathit{buy}) \wedge (b\,\mathrm{OBJ}\,o) \wedge (b\,\mathrm{AVRT}\,a) \wedge (b\,\mathrm{AGT}\;d) \qquad (2)$$

Such a B-axiom has only a local effect, i.e. it is applied only in those cases where one concept has to be connected to another during the inference process. Another example of a B-axiom is (4) which contains only one concept.

3.2 Categorically vs. Prototypically Valid Axioms

Categorically Valid Axioms. It seems to be a contradiction to speak of axioms which are restricted in their validity. But, if we want to formalize natural language semantics, we must also account for *prototypical* regularities.

The following two axioms express knowledge which is categorically valid:

$$(p_1\,\mathrm{COMPL}\,p_2) \wedge \neg(o\,\mathrm{PROP}\,p_1) \rightarrow (o\,\mathrm{PROP}\,p_2) \qquad (3)$$
$$(k_1\,\mathrm{PARS}\,k_2) \wedge (k_2\,\mathrm{ATTR}\,m_2) \wedge (m_2\,\mathrm{SUB}\,\mathit{weight}) \wedge (m_2\,\mathrm{VAL}\,q_2) \rightarrow$$
$$\exists m_1 \exists q_1 [(k_1\,\mathrm{ATTR}\,m_1) \wedge (m_1\,\mathrm{SUB}\,\mathit{weight}) \wedge (m_1\,\mathrm{VAL}\,q_1) \wedge (q_1\,\mathrm{MIN}\,q_2)] \qquad (4)$$

Axiom (3) states that one from two complementary properties (if applicable at all) must hold (for example, if a man is not ‘married’, then he is ‘unmarried’). Axiom (4) asserts that a part must always weigh less than the whole. For example, if the weight of a car is $q_1 = 1,500\,kg$, then the weight q_2 of the car’s engine must be less than $1,500\,kg$, as expressed by the smaller-than relation MIN. It is obvious that there is no exception from these rules.

Prototypically Valid Axioms. By contrast, the rules (5) and (6), governing the inheritance of the part-whole relationship and of properties within the SUB hierarchy, have only the status of default (or prototypically valid) knowledge:

$$(d_1\,\mathrm{SUB}\,d_2) \wedge (d_3\,\mathrm{PARS}\,d_2) \rightarrow \exists d_4 [(d_4\,\mathrm{SUB}\,d_3) \wedge (d_4\,\mathrm{PARS}\,d_1)] \qquad (5)$$
$$(o_1\,\mathrm{SUB}\,o_2) \wedge (o_2\,\mathrm{PROP}\,p) \wedge p \in \mathrm{tq} \rightarrow (o_1\,\mathrm{PROP}\,p) \qquad (6)$$

Axiom (5) expresses, for example, that a specific car will have wheels given that cars have wheels. (6) lets us conclude from the knowledge that bears are dangerous that a specific bear will also be dangerous. It is a good assumption that a conceptual object subordinated to a generic object inherit known parts and properties from the latter. However, there are exceptions. For example, although individual bears normally inherit the property *dangerous* from the generic concept *bear*; there are also circus bears, bears being ill etc. which should not be regarded dangerous.

3.3 Deductive Axioms vs. Destructive Axioms

Deductive Axioms. Many axioms, like (1) through (6), can be used in a deductive process to derive new knowledge, given by the conclusion, provided that the premise be fulfilled. The important feature of monotonic deduction is that no piece of knowledge in the knowledge base must ever be retracted.

[2] We cannot discuss the phenomenon of lexical ambiguity here, but we actually discern three phenomena in our computational lexicon: polysemy, homography, and meaning molecules [9].

Destructive Axioms. There are also axiomatic regularities which not only generate new knowledge but also cancel earlier knowledge. Into this class of 'destructive' axioms we number the derivation of the temporal end of a situation s:

$$(e\,\mathrm{SUBS}\,\mathit{end}) \wedge (e\,\mathrm{AFF}\,s) \wedge (e\,\mathrm{TEMP}\,t) \rightarrow (s\,\mathrm{FIN}\,t) \,|\, \mathrm{DEL}\,(s\,\mathrm{TEMP}\,\_) \tag{8}$$

Thus if an activity e ends a situation s at time t, then a new relation FIN for s must be established and the earlier specification of s by the relation TEMP must be deleted. Similarly, the transfer of negation from its extensional form to its intensional equivalent must be written as a mutual substitution (indicated by the dot over the arrow) and not as a logical equivalence:

$$[\mathrm{FACT}(s) = \mathit{non}] \wedge (s\,\mathrm{MODL}\,\mathit{nil}) \stackrel{\cdot}{\longleftrightarrow} (s\,\mathrm{MODL}\,{*}\mathrm{NON}) \wedge [\mathrm{FACT}(s) = \mathit{real}] \tag{9}$$

This axiom says in particular that the right side of the formula can not be monotonously derived from the left side and additionally inserted into the knowledge base. Given one of the expressions (left side or right side of the axiom) is true, one can take the expression on the other side of $\stackrel{\cdot}{\longleftrightarrow}$ instead of the former (but not both of them at a time, as would be the case with a normal equivalence). It is not possible to regard this a deductive logical inference because simply adding the derived relationships would result in an analysis involving two negating constructions, which corresponds to a double negation (as shown in Fig. 2).

3.4 Epistemically Restricted vs. Non-restricted Axioms

Epistemically Restricted Axioms. There are axioms which are epistemically restricted in the sense that their validity is only warranted within a certain epistemic or cognitive context. A typical example is the restricted transitivity of CAUS:

$$(k_1\,\mathrm{CAUS}\,k_2) \wedge (k_2\,\mathrm{CAUS}\,k_3) \rightarrow (k_1\,\mathrm{CAUS}\,k_3) \tag{10}$$

This axiom is connected with a fading effect preventing infinite prolongation of causality chains by a presumed (but not strongly valid) transitivity of CAUS. This effect is due to the so-called INUS-conditions [18], i.e. humans asserting a causal relation emphasize a certain cause and neglect other necessary conditions for this relationship. A long causal chain thus weakens the connection between the original cause and final effect. A similar behavior is shown by the relation PARS [8, Sect. 4.2]. It is not at all obvious how such epistemic level and the observed fading effect could be expressed by purely logical means.

Epistemically Non-restricted Axioms. For most axioms no epistemically motivated restriction can be observed. In particular, the transitivity of conceptual subordination (11) and of spatial inclusion (12) hold unconditionally:

$$(o_1\,\mathrm{SUB}\,o_2) \wedge (o_2\,\mathrm{SUB}\,o_3) \rightarrow (o_1\,\mathrm{SUB}\,o_3) \tag{11}$$

$$(o\,\mathrm{LOC}\,({*}\mathrm{IN}\,m)) \wedge (m\,\mathrm{LOC}\,({*}\mathrm{IN}\,n)) \rightarrow (o\,\mathrm{LOC}\,({*}\mathrm{IN}\,n)) \tag{12}$$

Axiom (11) concludes from the fact that pidgeons are birds and that birds are animals that indeed pidgeons are animals. Axiom (12) concludes from the fact that Marc's sun glasses are in his car and the fact that the car is in the parking garage that Marc's sun glasses are in the parking garage.

4 Effects on Inference and Logical Answer-Finding

4.1 The Role of Different Axiom Types in Inferences

Axioms which are *conceptually not bound* have to be treated with care by the reasoner, since an R-axiom for a relation R can be applied in inferences over the SN wherever R is involved (global effect). In addition, symmetry axioms may cause oscillations in the set of literals to be proved, while transitivity axioms blow up searches and repeatedly create literals containing only variables as arguments (see e.g. [19]). By contrast, *conceptually bound* axioms have only local effects. They establish a kind of hyperrelation between two or more concepts in a meta-level SN. Here, we meet the **Frame Problem** in Artificial Intelligence: In a B-axiom like (2), only the change of participant roles (like AGT, AFF, AVRT, and OBJ) is specified, but nothing is said about the local, temporal and circumstantial embedding of the main situation (mainly represented by LOC, TEMP, and CIRC, resp.) The transfer of these specifications must be handled by axiom schemata for classes of concepts: While the temporal specification of a selling act like s in (2) transfers unchanged to b, there is no such transfer of the specification $(s_1\,\mathrm{TEMP}\,t_1)$ of a sending act s_1 to the corresponding receiving act $\mathrm{sk}(s_1)$. For the latter class we have:

$$(s_1\,\mathrm{SUBS}\,\langle \textit{send-act}\rangle) \wedge (s_1\,\mathrm{TEMP}\,t_1) \wedge (\mathrm{sk}(s_1)\,\mathrm{SUBS}\,\langle \textit{receive-act}\rangle) \wedge (\mathrm{sk}(s_1)\,\mathrm{TEMP}\,t_2) \rightarrow (t_1\,\mathrm{ANTE}\,t_2) \quad (13)$$

In other words, the act of receiving always takes place after the corresponding sending act.

Categorically valid axioms lead to monotonic reasoning, while *prototypically valid* axioms call for nonmonotonic reasoning. The standard approach to default reasoning based on a truth-maintenance system does not scale up to knowledge bases of realistic size, though. The basic technique proposed in MultiNet is to warrant that every deduction step involving default knowledge again produces only default knowledge. The newly generated default knowledge has to be checked for *local* contradictions in a well-defined neighborhood of the concepts involved. Semantic networks can help defining such neighborhoods as their link structure gives a natural notion of vicinity for conceptual entities.

For *epistemically restricted* axioms, we propose the use of built-in procedures which treat borderlines of epistemic or functional levels by special parameters for controlling the inference process. For transitivity and symmetry axioms, we propose a procedural treatment in any case (be they epistemically restricted or not).

4.2 The Effect of Sorts and Layers on Inferences

Sorts. In the context of NL semantics, not every rule of standard FOL is valid without restriction. The proposed sorts are useful to constrain the applicability of logical rules. For example, the distinction of gradable properties (sort [gq]) and total properties (sort [tq]) prevents a wrong application of the law of double negation. In FOL, the meanings of adjectives are usually formalized as unary predicates, e.g. "*friendly*" as $\mathrm{FRIENDLY}(x)$ and "*unfriendly*" as $\mathrm{UNFRIENDLY}(x)$. The normal way to express the dependency of the predicates would be UNFRIENDLY$(x) \leftrightarrow \neg$FRIENDLY(x). We then obtain $\neg$UNFRIENDLY$(x) \leftrightarrow$ FRIENDLY(x), which is clearly wrong. This indicates that the law of double negation does not hold unconditionally. Unlike total properties (*dead*), the gradable properties *friendly* and *unfriendly* are not stable under double negation.

Genericity. The classification into generic concepts vs. specific instances is important for inferences, because it distinguishes between different mechanisms of knowledge processing and inheritance. For example a specific individual [GENER=*sp*] can inherit from a generic concept [GENER=*ge*]. However, it is impossible that a generic concept or an individual will ever inherit from a specific instance.

Facticity. The rules of inference must be different for existing and non-existing objects. If there exists no object with property B, there is also no object with a stronger property B'. Existing objects [FACT=*real*] show the opposite pattern: If there is an object with property B, then it also has any weaker property B''. In general, the support for non-denoting terms will affect the calculus rules [20]. Facticity must be anchored in the logical language since special inference rules apply to hypothetical and non-existing objects.

5 Conclusion

The described semantic network formalism targets at the formal representation of unregimented natural language. To achieve this, we introduced representational means which capture the facticity status, degree of generality, modal embedding, and other characteristics of NL concepts in terms of a multidimensional assignment of layer attributes and knowledge types. Our final goal is that of developing a ‘Logic of MultiNet’ which fully captures the combinatorics of these dimensions. This will demand, among other things, an integration of reasoning with modalities, prototypes, pluralities and generalized quantifiers. In the paper, we have demonstrated that the relations and functions on which the formalism is based can be made precise by axioms which formally describe their expected behaviour. These axioms differ with respect to a number of characteristics and we have shown that the proposed classificatory dimensions of categoricity, conceptual boundedness and epistemic restriction are especially important in this respect because they affect the validity and efficiency of inference. The ‘knowledge types’ mentioned in the introduction also have a significant influence on the type of inference and on the answer-generation strategy (this topic has been dealt with in [21]). MultiNet as described in this paper and the NLP technology connected to it are the cornerstone of the semantically based search engine SEMPRIA [22].

References

1. Quillian, M.R.: Semantic memory. In Minsky, M., ed.: Semantic Information Processing. MIT Press, Cambridge, Massachusetts (1968) 227–270
2. Lehmann, F., ed.: Semantic Networks in Artificial Intelligence. Pergamon Press (1992)
3. Woods, W.A.: What’s in a link: Foundations for semantic networks. In Bobrow, D.G., Collins, A., eds.: Representation and Understanding. Academic Press (1975)
4. Baader, F., Calvanese, D., McGuinness, D., Nardi, D., Patel-Schneider, P., eds.: The Description Logics Handbook. Cambridge University Press, Cambridge (2004)

5. Shapiro, S.C., Rapaport, W.J.: The SNePS family. In Lehmann, F., ed.: Semantic Networks in Artificial Intelligence. Pergamon Press (1992) 243–275
6. Sowa, J.F.: Conceptual Structures. Information Processing in Mind and Machine. The System Programming Series. Addison-Wesley, Reading, Massachusetts (1984)
7. Brachman, R.J., Schmolze, J.G.: An overview of the KL-ONE knowledge representation system. Cognitive Science **9**(2) (1985) 171–216
8. Helbig, H.: Knowledge Representation and the Semantics of Natural Language. Springer, Berlin (2006)
9. Hartrumpf, S., Helbig, H., Osswald, R.: The semantically based computer lexicon HaGenLex. Traitement automatique des langues **44**(2) (2003) 81–105
10. Hartrumpf, S.: Hybrid Disambiguation in Natural Language Analysis. Der Andere Verlag, Osnabrück, Germany (2003)
11. Leveling, J., Helbig, H.: A robust natural language interface for access to bibliographic databases. In: Proceedings of SCI 2002, Orlando, Florida (2002) 133–138
12. Hartrumpf, S.: Question answering using sentence parsing and semantic network matching. In: Results of the CLEF 2004 Cross-Language System Evaluation Campaign, Bath, England (2004) 385–392
13. Leveling, J., Hartrumpf, S.: University of Hagen at CLEF 2004: Indexing and translating concepts for the GIRT task. In Peters, C., Borri, F., eds.: Results of the CLEF 2004 Cross-Language System Evaluation Campaign, Working Notes for the CLEF 2004 Workshop, Bath, England (2004) 199–208
14. Helbig, H., Hartrumpf, S., Phoenix, I.: Automatic generation of large knowledge bases using deep semantic and linguistically founded methods. In: 6th International Conference of Agents and Artificial Intelligence (ICAART-2014), Angers, France (2014) 297–304
15. Evens, M.W.: Relational Models of the Lexicon. Cambridge University Press, Cambridge, England (1988)
16. Jackendoff, R.: Semantic Structures. MIT Press, Cambridge, Massachusetts (1990)
17. Hartrumpf, S., Helbig, H.: The generation and use of layer information in multilayered extended semantic networks. In: Proc. of the 5th Int. Conf. on Text, Speech and Dialogue (TSD 2002). Volume 2448 of LNAI., Springer (2002) 89–98
18. Mackie, J.L.: The Cement of Universe. Oxford University Press, Oxford (1974)
19. Ng, K.W., Ma, W.Y.: Pitfalls in PROLOG programming. SIGPLAN Notices **21**(4) (1986) 75–79
20. Lambert, K., ed.: Philosophical Applications of Free Logic. Oxford University Press, Oxford (1991)
21. Glöckner, I., Helbig, H.: Wissensrepräsentation und Antwortfindung mit Mehrschichtigen Erweiterten Semantischen Netzen. KI **2** (2005) 49–55
22. Hartrumpf, S., Helbig, H.: Semantische Suche und Stichwortsuche: ein Vergleich von Ansatz und Ergebnissen auf einem deutschsprachigen Dokumentenarchiv. info 7: Medien, Archive, Information **29**(2) (2014) 62–66