Towards a Logical Foundation of Semantic Networks – A Typology of Descriptive Means for Semantic Inference

Hermann Helbig, Ingo Glöckner
Intelligent Information and Communication Systems
FernUniversität in Hagen, Germany

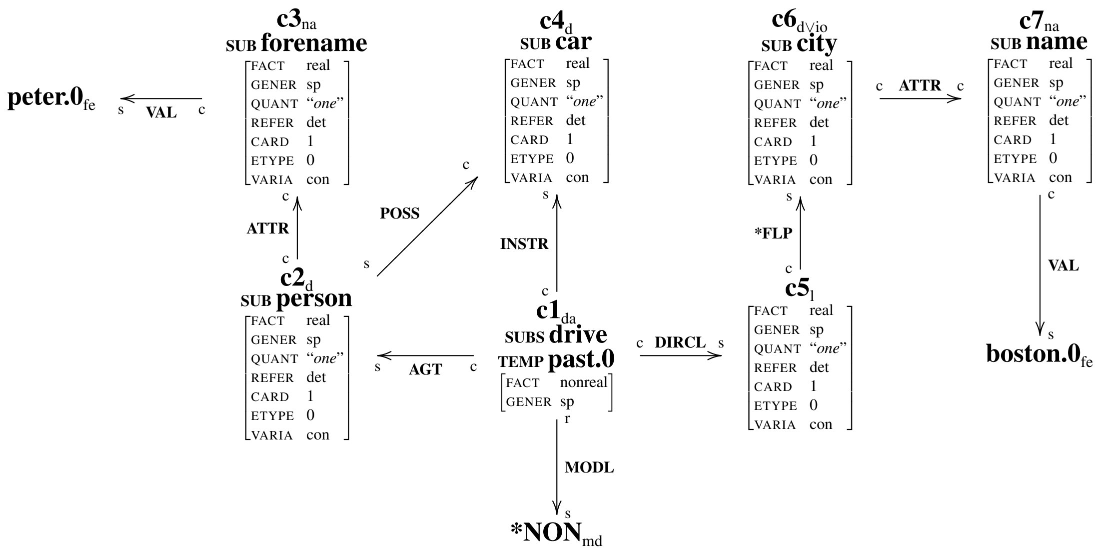
Abstract

Semantic Networks (SN) are a knowledge representation paradigm especially suited for the meaning representation of natural language expressions. In order to clearly define their basic constructs, the relations and functions used in a semantic network must be given a logical characterization. The paper exemplifies this strategy for Multilayered Extended Semantic Networks (the so-called MultiNet paradigm). In particular, it is shown that the axioms characterizing the logical properties of the expressional means of an SN have to be classified according to different criteria which are connected with specific types of inference.

1 Introduction

Semantic Networks have a long tradition as a paradigm for representing cognitive structures, starting with Quillian [9]. As to the 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 [10], and Brachmann's KL-ONE [1], 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 [5], however not on a formal logical level. Moreover the proposed relations and roles are not contrasted with each other to form a balanced system of expressional means. In SNePS [10], for example, the guideline for choosing the appropriate relations is deliberately given the status of a recommendation only. To meet the requirements on a knowledge representation formalism useful for NLP (especially the universality, homogeneity, and interoperability criteria [4, Chap. 1]), one needs a commitment on a clearly defined repertory of expressional means. Multilayered Extended Semantic Networks (the MultiNet paradigm [4]) were designed to fulfill these requirements, by fixing a set of expressional means and formalizing axioms which describe these functions and relations. The formalism is comprehensively documented and successfully used in NLP applications, e.g. for describing large semantically based computational lexica [3], and as the backbone of NL interfaces [7]. Its application in the

Figure 1: *"It is not true that Peter didn't drive to Boston with his car"*



InSicht question-answering system [2] has been successfully evaluated in CLEF 2004, with a MultiNet knowledge base generated from 4,9 Mio sentences.

MultiNet extends simple semantic networks by the following features: (1) Every node is labeled by a sort from a predefined ontology of sorts and by bundles of layer attributes. (2) MultiNet admits functions and relations of arbitrary arity. (3) The arcs (relations) are formally characterized by associated axioms. (4) Subnetworks can be encapsulated to form concepts of higher order which can be connected to other concepts by relations and functions. (5) The relationships in the network are assigned a knowledge type and thus marked as categorically valid (c), prototypically valid (p), modally restricted (r), or situationally bounded (s) with regard to each argument, as shown in Fig. 1.

2 The Expressional Means of MultiNet

To characterize an SN, we need a precise specification of the relations corresponding to the arcs (links). Our formalism provides about 140 relations and functions characterized on the basis of a uniform schema and by logical axioms. Table 1 sketches the relations and functions needed in this paper. MultiNet distinguishes 45 sorts of conceptual entities used to define the signatures of relations and functions [4, Sect. 17.1]. Some of these sorts are explained in Table 1. The sorts are also needed to constrain the applicability of logical rules. For example, the law of double negation holds for semantically total properties (sort [tq]) like *dead* and its negation *alive*, where *not alive* means the same as *dead*. Another sort [gq] is used for gradable properties like *friendly* and *unfriendly*. Although *unfriendly* means *not friendly*, the law of double negation does not hold in this case, i.e. if someone is *not unfriendly* this does not mean that the person is *friendly*.

MultiNet not only classifies conceptual entities by their sorts but also by the values of

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	$si^{\prime} \times si^{\prime}$	Relation between cause and effect (Causality)
CIRC	$si \times si$	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
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

Table 1: Strongly abbreviated description of relations used in this paper. Explanation of sorts: objects o include concrete objects co (*house*) and attributes at (*height*); situations si (*write*); locations l (*here*), times t (*now*), modal descriptors md (*impossible*); properties p (*dead*), quantificators and measurements qn (*many*, *two litres*), formal entities fe (figures or names). The notation $si^{\prime}$ demands [FACT = *real*], and the notation $\overline{si}$ demands [GENER = *ge*].

six so-called 'layer attributes'. Here we are concerned only with two of these attributes:
Facticity. We discern three kinds of facticity: [FACT=*real*] for existing entities (*Eiffel tower*), [FACT=*non*] for non-existing entities (*the light ether*), and [FACT=*hypo*] for hypothetical entities (*quarks*). Apart from the *extensional negation* expressed by a non-existing situation with [FACT=*non*], MultiNet supports the *intensional negation* of a situation s, expressed by the relation $(s\,\mathrm{MODL}\ {*}\mathrm{NON})$. Both types occur in the example shown in Fig. 1. Facticity must be anchored in the logical language since special inference rules apply to hypothetical and non-existing objects.
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 "*Lions feed on antelopes*". A modeling by a universal quantifier ranging over all lions would be inadequate because the sentence expresses only default knowledge.

3 A Typology of Axioms for Inferences over an SN

While a logical expression is either true or false in first-order logic (FOL), a semantic formalism dealing with NL must support different degrees of reliability. Moreover, logical calculi normally do not give a clue how to use the axioms in an effective inference strategy. These considerations suggest the following cross-classification of axioms.

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 effects never take place before the cause: $(x\,\mathrm{CAUS}\,y) \rightarrow \neg(y\,\mathrm{ANTE}\,x)$ (1)
Other examples are given by axioms (4), (5), (7) below. 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). Inter alia, R-axioms serve to express the symmetry or transitivity of relations, i.e. properties which are difficult to handle efficiently.
B-axioms. Axioms containing the representative of at least one concept are called *conceptually bound* or B-axioms. Thus, with every selling act s there is a buying act b entailed by s. The corresponding relationship is given by the following axiom:

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

Another example of a B-axiom is (6) which contains only one concept. Such B-axioms have only a local effect, i.e. they are applied only in those cases where one concept has to be connected to another during the inference process. 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 $s_2 = \mathrm{s}k(s_1)$. For the latter class we have:

$$(s_1\,\mathrm{SUBS}\,\langle \mathit{send\text{-}act}\rangle) \wedge (s_1\,\mathrm{TEMP}\,t_1) \wedge (\mathrm{s}k(s_1)\,\mathrm{SUBS}\,\langle \mathit{receive\text{-}act}\rangle) \wedge (\mathrm{s}k(s_1)\,\mathrm{TEMP}\,t_2) \rightarrow (t_1\,\mathrm{ANTE}\,t_2) \quad (3)$$

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 axiom expresses knowledge which is categorically valid:

$$(p_1\, \mathrm{COMPL}\, p_2) \rightarrow (o\, \mathrm{PROP}\, p_1) \vee (o\, \mathrm{PROP}\, p_2) \tag{4}$$

Axiom (4) states that one from two complementary properties (if applicable at all) must hold. It is obvious that there is no exception from this rule.

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

$$(d_1\, \mathrm{SUB}\, d_2) \wedge (d_3\, \mathrm{PARS}\, d_2) \rightarrow \exists d_4[(d_4\, \mathrm{SUB}\, d_3) \wedge (d_4\, \mathrm{PARS}\, d_1)] \tag{5}$$

It is a good assumption that a conceptual object subordinated to a generic object inherit known parts from the latter. However, there are exceptions. While ships normally have a keel, there are also ships which have not.

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, though. In MultiNet, we warrant that every deduction step involving a default again produces only default knowledge. The newly generated default knowledge has to be checked for *local* contradictions in a neighborhood of the concepts involved. Semantic networks can help defining such neighborhoods as their link structure gives a natural notion of vicinity for concepts.

3.3 Deductive Axioms vs. Destructive Axioms

Deductive Axioms. Many axioms, like (1) through (5), 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.

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:

$$(e\, \mathrm{SUBS}\, \mathit{end}) \wedge (e\, \mathrm{AFF}\, s) \wedge (e\, \mathrm{TEMP}\, t) \rightarrow (s\, \mathrm{FIN}\, t) \,|\, \mathrm{DEL}\, (s\, \mathrm{TEMP}\, \_) \tag{6}$$

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. While the first type of axioms can be treated by symbolic derivations, the latter type requires actions on the knowledge base like deleting arcs.

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:

$$(k_1\, \mathrm{CAUS}\, k_2) \wedge (k_2\, \mathrm{CAUS}\, k_3) \rightarrow (k_1\, \mathrm{CAUS}\, k_3) \tag{7}$$

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 [8], i.e. humans asserting a causal relation emphasize

a certain cause and neglect other necessary conditions for this relationship.
Epistemically Non-restricted Axioms. For most axioms no epistemically motivated restriction can be observed. In particular, the transitivity of conceptual subordination (8) and of spatial inclusion (9) hold unconditionally:

$$(o_1\,\mathrm{SUB}\,o_2) \wedge (o_2\,\mathrm{SUB}\,o_3) \rightarrow (o_1\,\mathrm{SUB}\,o_3) \tag{8}$$

$$(o\,\mathrm{LOC}\,(\text{*IN}\,m)) \wedge (m\,\mathrm{LOC}\,(\text{*IN}\,n)) \rightarrow (o\,\mathrm{LOC}\,(\text{*IN}\,n)) \tag{9}$$

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.

4 Conclusion

The MultiNet formalism is intended for the semantic representation of unrestricted language and thus supposed to represent the facticity status, degree of generality, modal embedding, and other characteristics of NL concepts. The meaning of the relations and functions on which the formalism is based, can be made precise by axioms which capture their expected behaviour. These axioms differ with respect to the classificatory dimensions of categoricity, conceptual boundedness, and epistemic restriction. We have shown that these dimensions also affect the validity and efficiency of inference.

References

[1] R. J. Brachman and J. G. Schmolze. An overview of the KL-ONE knowledge representation system. *Cognitive Science*, 9(2):171–216, 1985.

[2] S. Hartrumpf. Question answering using sentence parsing and semantic network matching. In *Proc. of CLEF 2004*, pages 385–392, Bath, England, Sept. 2004.

[3] S. Hartrumpf, H. Helbig, and R. Osswald. The semantically based computer lexicon HaGenLex. *Traitement automatique des langues*, 44(2):81–105, 2003.

[4] H. Helbig. *Knowledge Representation and the Semantics of Natural Language*. Springer, Berlin, 2006.

[5] R. Jackendoff. *Semantic Structures*. MIT Press, Cambridge, Massachusetts, 1990.

[6] F. Lehmann, editor. *Semantic Networks in Artificial Intelligence*. Pergamon Press, 1992.

[7] J. Leveling and H. Helbig. A robust natural language interface for access to bibliographic databases. In *Proceedings of SCI 2002*, pages 133–138, Orlando, Florida, July 2002.

[8] J. L. Mackie. *The Cement of Universe*. Oxford University Press, Oxford, 1974.

[9] M. R. Quillian. Semantic memory. In M. Minsky, editor, *Semantic Information Processing*, pages 227–270. MIT Press, Cambridge, Massachusetts, 1968.

[10] S. C. Shapiro and W. J. Rapaport. The SNePS family. In Lehmann [6], pages 243–275.