FROM DICTIONARIES TO KNOWLEDGE REPRESENTATION FORMALISMS

VERGINICA BARBU MITITELU, EDUARD BARBU

Abstract. In this paper we lay out various ways of representing knowledge. They start from the older and simpler way: the printed dictionary. However useful for linguists and for the general public, the printed dictionary is not able to answer the nowadays needs in linguistics and in interdisciplinary domains that deal with various semantic aspects of language. The advent of computers imposed the development of knowledge representation formalisms such as semantic networks, description logics, conceptual graphs and frame systems. Given the fact that many linguistic resources are used in Natural Language Processing tasks we think that Romanian lexicography and lexical semantics has to catch up with the formalization systems. In this way new competitive Romanian language resources could be developed and used in Artificial Intelligence tasks.

INTRODUCTION

From early age we make use of dictionaries as great repositories of knowledge. We start using them in order to enrich our vocabulary and continue using them either to learn new words, new terms, or to acquire new meanings of the already learned words. The time when people were the only ones using dictionaries has passed. Nowadays we need to teach computers how to use such linguistic resources in order to be able to solve various tasks. The aim of this article is to make a presentation of the evolution of dictionaries[1] from printed books aimed for humans' use to the machine readable version of such dictionaries and further on to knowledge representation formalisms. The paper has two parts: in the first one we discuss the concept of dictionary and how it evolved. We begin with a definition of the dictionary and we try to identify different kinds of dictionaries starting from this definition. After motivating the centrality of lexicon for a linguistic theory we present the inner organization of dictionaries and thesauri. Then we discuss the ideal dictionary of Mel'cuk *et al.* (1995).

[1] We are not concerned here with making a historic presentation of lexicography (for the history of the Romanian lexicography see Seche 1966, 1969). We are not concerned with a critical evaluation of dictionaries either (for a presentation of the deficiencies that dictionaries usually have see Hanks 2003).

RRL, **LII**, *1–2*, p. 135–159, Bucureşti, 2007

Nowadays the knowledge encoded in dictionaries needs to be accessible to computers. That is why in the second part of our paper we introduce the main Knowledge Representation formalisms. We start by defining knowledge representation, then we present the main Knowledge Representation formalisms: semantic networks, description logics, conceptual graphs and frame systems.

1. DICTIONARIES

1.1. Defining the term

A rule of the thumb for any author writing on whatever topic is to make sure from the very beginning that s/he will be understood by her/his readers by defining or specifying the meaning of the words s/he makes use of, especially when those words are used to cover more (various) meanings. This is our case at the moment. The word *dictionary* has more meanings, thus covering various concepts[2]. If one looks the word up in a dictionary s/he gets explanations similar to the following[3]:

1. a reference book containing words usually alphabetically arranged along with information about their forms, pronunciations, functions, etymologies, meanings, and syntactical and idiomatic uses;
2. a reference book listing alphabetically terms or names important to a particular subject or activity along with discussion of their meanings and applications;
3. a reference book giving for words of one language equivalents in another;
4. a list (as of items of data or words) stored in a computer for reference (as for information retrieval or word processing).

(Merriam-Webster Online Dictionary)

From these definitions it is clear that a dictionary is either a "flesh-and-blood" book or the resource that is the core part of a software product. Until some decades ago the fourth definition was inexistent. But the first three ones cover concepts that have a long existence in the history of mankind. Their long existence over the time can be explained by the fact that people really need dictionaries.

If we draw a comparison between dictionary and another language resource, namely grammar, we notice that we hardly make use of grammars of a language, especially of the mother tongue. The explanation for this resides in the basic difference between the two components of a language: the lexicon and the grammar. If the latter consists of a fixed and (relatively) stable inventory of rules

[2] A concept is an abstract, universal idea or notion that serves to designate a category or class of entities, events or relations. Concepts are bearers of meanings.

[3] The definitions provided here are the most comprehensive ones we could find by checking various online English dictionaries.

(and some unavoidable exceptions to these rules) and once learned they need no further improvement and refinement, the former is open-ended in the sense that during our lifetime we keep on learning both new words (we acquire vocabulary) and new meanings of the already known words (we acquire new uses of vocabulary). It is a process to which we are all "condemned", in a higher or lower degree, depending on our field of activity, on our education, on our age, on the social conditions in which we live, etc. Dictionaries are needed to guide us through the thick forest of meanings, even in our native language.

1.2. Types of dictionaries – meeting various needs

The definitions above already suggest the existence of a classification of dictionaries: the first definition[4] (a reference book containing words usually alphabetically arranged along with information about their forms, pronunciations, functions, etymologies, meanings, and syntactical and idiomatic uses) applies to *explanatory dictionaries*, the second one (a reference book listing alphabetically terms or names important to a particular subject or activity along with discussion of their meanings and applications) to *specialised dictionaries* (useful for various domains of activity), the third (a reference book giving for words of one language equivalents in another) to *bilingual dictionaries* (and can be extended to account for *multilingual dictionaries*, as well), and the last one (a list (as of items of data or words) stored in a computer for reference (as for information retrieval or word processing)) to a dictionary in electronic form (a machine readable dictionary).

Lexicographers used different criteria for classifying dictionaries[5]. Under one criterion, Jackson (2002: 23-25) distinguishes *monolingual* dictionaries from *bilingual* (or even multilingual) ones. Among the former, a distinction should be made between *historical* and *synchronic* dictionaries. Dictionaries can also be classified according to their intended audience: for students, for the target language as a second or foreign language, general-purpose dictionaries (these are the most widely used), specialist dictionaries (dictionaries of pronunciation, of spelling, of etymology, of synonyms, of a certain domain of activity, etc.). Landau (2001) classifies dictionaries also according to their form of presentation, thus distinguishing between dictionaries in which words are arranged alphabetically or thematically, and, allied to this, between dictionaries in the form of a book and those in electronic form. This correlation will become more explicit further on in this article.

[4] With this meaning, the word *dictionary* is synonym with the word *lexicon*, the latter being preferred in modern linguistic theories, as well as in Natural Language Processing literature. *Lexicon* is also used to refer to a language user's knowledge of words, thus being synonym with *mental dictionary*, *mental lexicon*.

[5] For an overview of the history and of the different types of dictionaries see also Canarache (1970).

Dictionaries are "living" social institutions (see Leech 1974: 203): there is a social need for them to change in order to better serve the continuously changing society. The lexicographic history of a language reflects the needs of the society. Consider the case of Romanian lexicography: the first dictionaries appeared to ease the work of those translating religious texts from old Slavonic into Romanian in the 16th century, thus the oldest Romanian dictionaries are bilingual, namely Slavonic-Romanian dictionaries (cf. Seche 1966). English lexicography has also bilingual beginnings: glossaries were collected to facilitate monks' and priests' understanding of the theological texts written in Latin (the language of the Roman form of Christianity that spread in the old English territories in the first century AD) (see Jackson 2002).

Nowadays, when computers have entered our lives and become a more and more important part of it, the computer user cannot but be satisfied with the existence of dictionaries in electronic form. At home, at the office, in the plane, in the street, wherever one may be, if s/he has a computer, then s/he may also have a dictionary of any of the types enumerated above. They are available either as (purchasable) software products (for English: *Cambridge Advance Learner's Dictionary*, *Oxford Spellchecker and Dictionary*, and others; for Romanian: *NODEX – Noul Dicționar Explicativ al Limbii Române*, *Dicționar morfologic al limbii române contemporane în format electronic*, and others) or as (free) online dictionaries (requiring Internet connection in order to be accessed) (for English: *Merriam-Webster Online Dictionary*: http://www.m-w.com/dictionary, *Compact Oxford English Dictionary*: http://www.askoxford.com/dictionaries, and others; for Romanian: *DEX online*: www.dexonline.ro). Some online dictionaries offer a nice or user-friendly way of visualising the senses of a word. *Visual Thesaurus* (http://www.visualthesaurus.com/), for instance, is a dictionary and a thesaurus[6] with an intuitive interface that allows for the exploration of the word senses, altogether with the lexico-semantic relations they establish with other words. In Fig. 1 we present the image the thesaurus returns for the searched word *happiness*.
This image is interactive. Clicking on each of the nodes will display a different definition. *Happiness* is linked to two nodes, so it has got two meanings. Words linked to a certain node via a continuous line are synonyms. Thus, in one of its meanings (the one associated with the node for which the definition is expanded in the image), *happiness* lacks a synonym. But in the other meaning, *happiness* is synonymous with *felicity*. The dotted lines linked to *happiness* represent the antonymy relation[7]. So, *happiness* has *unhappiness* and *sadness* as antonyms. The dotted lines leading to the nodes to which *unhappy* and *happy* are associated represent the *is_an_attribute_of* relation. So, *happiness* is an attribute having as

[6] A thesaurus is a listing of words with similar, related, or opposite meanings. See below for a presentation of thesauri.

[7] In the Internet browser going with the mouse on each line the name of the relation for which the line stands is displayed.

possible values *happy* and *unhappy*. The dotted line leading to the node to which *feeling* is associated represents the hyponymy relation. So, *happiness* is a type of *feeling*.

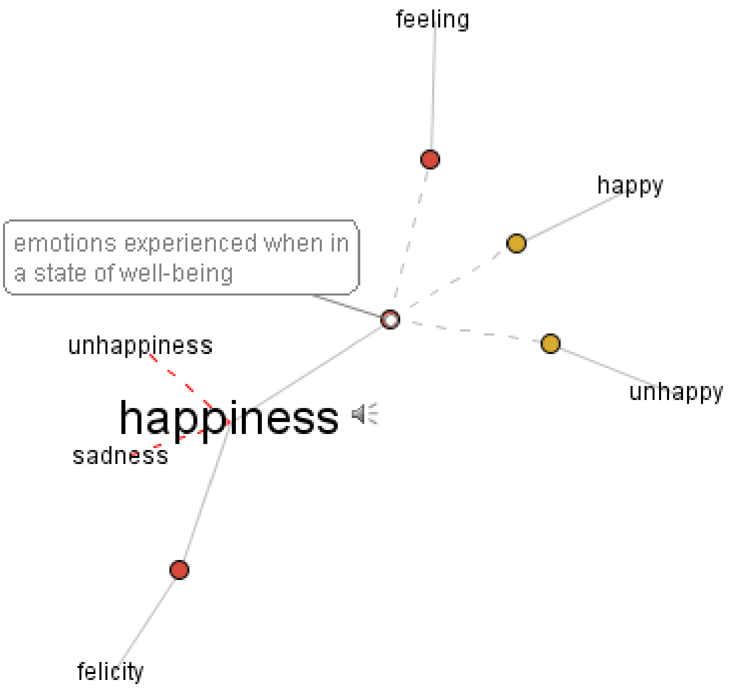

Fig. 1. Information displayed by the Visual Thesaurus for the word *happiness*.

In the last century the emergence of Natural Language Processing (NLP) as a new field of activity aiming at processing and manipulating natural language created the need for linguistic resources in electronic form. Thus, the lexical entries in the dictionary are processed not by human beings, but by the machines (computers), so information needs be presented in a way that computers can "understand" and use for further applications or reasoning.

1.3. The importance of the lexicon

The Chomskian revolution in linguistics (started in 1950s) placed syntax at the core of a linguistic theory. However, late 20th century brought about a change of emphasis on the language components: grammatical theories such as Head-driven Phrase Structure Grammar (Pollard and Sag 1987, 1994) lay stress on the lexical component of the theory. Lexical entries in the lexicon are attached rich phonologic, syntactic, semantic and even pragmatic information relevant for the possible uses of words. Another important characteristic of a grammar as this one is the organization of its lexicon: it contains types of objects organized in an inheritance hierarchy that eliminates a lot of the redundancy: once the subtype of a word specified, all properties of its supertypes are inherited.

In the early NLP the lexicon was a peripheral component (Gazdar and Mellish 1989: 257). The evolution from the traditional task of constructing computational components for the automatic generation and analysis of natural language sentences in isolation to the more recent task of constructing components suitable for the treatment of large amount of texts increased the lexicon importance (Ooi 1998: 26) to such an extent that "natural language systems cannot understand text for which they do not possess the lexicon" (Amsler 1989:12). It was natural then that researchers oriented their attention towards *machine readable dictionaries* (MRD), i.e. dictionaries in electronic form. MRDs are *Longman Dictionary of Contemporary English* (LDOCE), *Oxford Advanced Learner's Dictionary* (OALD), and some others.

1.4. The Inner organization of dictionaries – dictionaries and thesauri

There are two ways of arranging lexical entries in a dictionary: alphabetically and conceptually. In the case of printed dictionaries, the alphabetical organization of lexical entries is of great help for readers looking for a word's meaning. However, such an organization allows limited access to words related to the target one: for instance, from *uncle* one cannot get to *aunt* (although there is clearly a relation between the two in our mind). Conceptual dictionaries, also called thesauri, do not have this disadvantage. On the one hand, the user may start from the index (where words are arranged alphabetically) and then go to the thematic area/s to which the target word belongs and find its meanings there, altogether with the conceptually related words (synonyms, antonyms, hyponyms). On the other hand, the user may start from a particular thematic area and find there the words used to express the concepts belonging to it[8]. To put it differently, (alphabetically organized) dictionaries only allow access from word to its meanings, while conceptually organized dictionaries, i.e. thesauri, allow access both from word to its meanings and from meaning to the appropriate word expressing it.

Adopting the natural language generation perspective, Zock (2005) speaks about two views concerning lexicalization: the process is conceptually-driven (meaning precedes the linguistic form) or lexically-driven (access the target word via a source word). A hybrid input is also possible: in Mel'cuk *et al.* (1995) (see

[8] It seems that a thesaurus is preferred to a dictionary especially in the case of specialized vocabularies, where the hierarchical organization of material is of greater help for the user. Moreover, due to the reduced dimensions of the vocabulary of a certain domain, a thesaurus for a special domain is easier to create than one for general language. There are thesauri for domains such as medicine, astronomy, art and architecture, cooking, etc.

1.5. below) the notion of intensity, for instance, is expressed by various words, depending on the form of the argument: we say in Romanian *grav bolnav*, *profund mâhnit*, *adânc întristat*, but not **adânc bolnav*, **grav întristat*, etc.

For dictionaries in electronic form the alphabetical organization is of less importance, as the user introduces in the searching interface his/her target word, and the machine finds its entry in less than no time, irrespective of the alphabetical or conceptual organization of the material.

1.4.1. The *Organization of lexical entries*

Here is a snapshot of the online version of LDOCE[9]:

goat /gəut $ gout/ [countable]
1 an animal that has horns on top of its head and long hair under its chin, and can climb steep hills and rocks. Goats live wild in the mountains or are kept as farm animals.
2 get somebody's goat *spoken informal* to make someone extremely annoyed.
3 old goat *informal* an unpleasant old man, especially one who annoys women in a sexual way.
4 act/play the goat *British English informal* to behave in a silly way.

Just like in the paper dictionaries, the orthographic form of the word is followed by its pronunciation, by morphologic information, then by the sense definitions[10] (alongside with stylistic information about the usage of the respective sense).

Starting from this snapshot, we can present the characteristics of a MRD as being the following (Ooi 1998):

- reliance on the user's background linguistic and common-sense knowledge to retrieve and comprehend the information it contains;
- use the human's reaction to visual attributes (bold face, italics, etc.) implicitly to convey semantic information and help in the interpretation of sub-sections of an entry;
- tendency to be organized as a list of lexical entries sorted alphabetically.

What makes a good dictionary for humans does not necessarily make a good one for a computer. The latter needs an explicit dictionary which contains formal descriptions of lexical data, and a systematic and flexible one, i.e. it can be able to compare or generalise across lexical entries. Boguraev and Briscoe (1989: 4-5) identify five broad types of knowledge that are relevant for various NLP tasks:

[9] Due to publishing reasons we decided not to include here the picture (representing a goat) that appears in the online LDOCE at this entry and also the icon whose clicking gives the user the pronunciation of the word in .wav form.

[10] For a study of dictionary lexical entries (with special emphasis on Romanian DEX) from both a lexicographic and a semantic perspective, see Bidu Vrănceanu (1993).

- *phonological knowledge* – concerns the sound system and structure of words and utterances;
- *morphological knowledge* – concerns the internal structure of words;
- *syntactic knowledge* – concerns the organization of words into phrases and sentences;
- *semantic knowledge* – concerns the meaning of words and the way they combine to form the meaning of sentences;
- *pragmatic / 'encyclopaedic' knowledge* – concerns the circumstances of an utterance.

An ideal lexicon would contain such information for all lexical entries. However, on the one hand such a lexical resource is very difficult to create: it requires a lot of time, effort and money. On the other hand, it is unlikely that an NLP application would require access to all these types of information.

Just like printed dictionaries, MRDs organize their material under lexical entries, where each lexical entry corresponds to a homograph. That is, if a word form belongs to different parts of speech, then it appears in the dictionary as many times as many parts of speech it belongs to. However, polysemy of a word form belonging to a certain part of speech is treated under the same lexical entry.

Unlike dictionaries addressed to humans, those for machines need ensure access to the lexical entries not only in a word-by-word manner, but also via other information in the entry (imagine a speech recognition system that requires access to entries via the phonological field), or even access to classes of entries sharing some properties (for instance access needed to all transitive verbs that take a human subject and a human direct object).

1.4.2. *Beyond traditional dictionaries – foreseeing the needs in NLP*

We mentioned above the existence of dictionaries in electronic form. In this section we focus on four such dictionaries that present characteristics that qualify them as representative of a new type of dictionary, going beyond the limitations of the traditional ones, from an NLP perspective: *Longman Dictionary of Contemporary English* (LDOCE), *Longman Lexicon of Contemporary English* (LLOCE), *Cambridge International Dictionary of English* (CIDE), and *Göteborg Lexical DataBase* (GLDB).

LDOCE is still a traditional MRD. However, the hierarchical organization of the semantic codes (simple ones such as Animal, Concrete, Gas, Human, or compound ones such as Female Animal, Inanimate Concrete, Male Animal or Human, etc.) associated to words represents already a step forward towards the MRDs satisfying some needs in NLP.

One further step is represented by LLOCE that combines the entry format of LDOCE (which provides detailed syntactic information) with the semantic structure of a thesaurus. Such an organization offers two advantages: on the one hand, syntactic and semantic properties are related, on the other, dependencies between semantic predicates classes and subcategorization frames are individuated.

In CIDE senses are classified into 900 categories that are structured in a hierarchy with four levels. Moreover, selectional preferences (syntactic and semantic restrictions) are also coded.

Not only English beneficiates of such a resource. A lexical database (GLDB) has been created for Swedish. Besides the information commonly encountered in a dictionary, GLDB also contains selectional restrictions for arguments of verbs and adjectives, semantic relations (such as hyponymy, co-hyponymy, synonymy, semantic opposition), and terminological domains (such as *medicine*, *sport*, *music*, *gardening*, etc.).

1.5. Mel'cuk *et al.*'s explanatory combinatorial dictionary

Aiming at creating an ideal dictionary, Mel'cuk *et al.* (1995) design some elaboration principles that ensure the dictionary the following characteristics:

- *pre-established metalanguage* (used consistently). It consists of words having the simplest sense units which are used to decompose the defined word sense. The main characteristic of the simplest sense units is their lack of ambiguity: they have only one meaning in the pre-established metalanguage;
- *coherence*: the different linguistic information layers need to be interconnected;
- *uniform treatment* of the semantically similar words senses (those belonging to the same semantic field, like nationalities for example, beneficiate of similarly structured definitions);
- *exhaustivity*: include all the necessary information for establishing the exact place a word sense has in the vocabulary, its relations with other words senses in the vocabulary, thus ensuring its correct understanding and usage by the speakers.

These are the principles proposed by Mel'cuk *et al.* (1995) for creating an ideal explanatory combinatorial dictionary in electronic form for French.

The basic unit in lexicology is called *lexeme* (a word or phrase with a specific meaning) by Mel'cuk *et al*. For each lexeme the following information is specified:

- phonological – obligatory information
- semantic – obligatory information containing the definition and the connotation of the respective lexeme
- combinatorial – obligatory information about the combinatorial possibilities of the respective lexeme from stylistic, morphologic, syntactic and lexical perspectives
- illustration – optional information
- phraseology – optional information.

Unlike in ordinary dictionaries, the lexicographic definition here has formal characteristics. The definition has two parts: the *defined* term and the *defining* phrase (the semantic description): A= “B”. The latter presents in a formal way the sense of the former. It is a semantic representation written in the form of a semantic network, but presented as a definition in French (for ease of understanding, we present the definition in Romanian below). It takes the form of a proposition in which the variables are semantic arguments:

(1) exila = “X exilează pe Y din L_1 în L_2 pentru Z pentru o perioadă de T”

The defined and the defining terms should be reciprocally substitutable in the same context, without altering its meaning (the defining condition of synonymy). Let’s consider the following definition for the Romanian equivalent of *teacher*:

(2) profesor de X/ de Y la Z = persoană a cărei profesie este aceea de a preda X la nivel de Y (fiind angajată de o instituție de învățământ Z)

Consider now the following sentences, which different from each other only in respect to the fact that in (3)a we use the word *profesoară* and in (3)b we use its defining phrase:

(3) a. Avem o nouă *profesoară* de chimie.
b. Avem o nouă *persoană a cărei profesie este aceea de a preda* chimie.

There is no difference in meaning between the two sentences.

With respect to the content of the lexicographic definition, there are four criteria that need to be obeyed:

- linguistic pertinence (this criterion applies only to the lexemes that designate objects or substances): the definition of a lexeme L must include a semantic component if and only if there is another lexeme L’ in the same language that is formally linked with L and that has the respective semantic component in its definition. Two lexemes (L and L’) are formally linked if any of the conditions below holds:
  - L and L’ are in polysemy relation
  - L is derived from L’
  - L is a phrase including L’ (for instance L’ is *zăpadă* and L is *alb ca zăpada*)
- the definition should reflect the co-occurrence with qualifying modifiers;
- the definition should reflect the co-occurrence with quantifiers;
- the definition should reflect the co-occurrence with negation.

The syntactic combinatorial area of the lexicographic definition contains the deep syntactic arguments a lexeme may take and the restriction on their lexicalization (case, preposition, function) (see above the definition for *exila*).

The lexical combinatorial area contains syntagmatic information about the way the target lexeme combines with other lexemes in a sentence. This is the area where lexical functions are represented. Mel'cuk *et al*. make use of 56 lexical functions that cover syntagmatic (collocations) and paradigmatic (synonymy, antonymy, etc.) functions. These are functions in the mathematical sense, representing certain extremely general ideas, such as 'very', 'begin' or 'implement', or else certain semantico-syntactical roles (Mel'cuk and Zholkovsky 1984). The lexical functions serve two uses within the Meaning-Text Theory (Mel'cuk *et al.* 1981, Mel'cuk *et al*. 1984, Mel'cuk, Polguère 1987). One is to control the proper choice of lexical items, whereas the other is to describe sentence synonymy. The latter is done by describing a number of equivalences in terms of lexical functions. Such paraphrasing rules account for the equivalence of the sentences below:

(4) Am primit un răspuns satisfăcător.
(5) Am primit un răspuns care m-a satisfăcut.

Mel'cuk *et al.*'s *Explanatory Combinatorial Dictionary* can be viewed as a semantic network (see below a more detailed presentation of semantic networks) where the nodes are lexemes and the links are either lexical functions giving related collocations or argument positions for predicates in the meaning decomposition for a lexeme.

2. KNOWLEDGE REPRESENTATION FORMALISMS

This section of the paper has two parts. In the first one we will try to answer the question: "What is Knowledge Representation (KR)?" In the second part we will briefly present the main KR paradigms.

Generally, in the literature, much effort was dedicated to describing and comparing different KR systems. The pluses and minuses of each system were analyzed: the adequacy and expressivity of the chosen KR language, the decidability of the language, the kind of problems that a knowledge based system has, etc. Much less effort was put in answering the foundational question: "What is Knowledge Representation?" We will address this problem in the first part of our discussion of KR systems, in section **2.1**.

In the second part of section **2** we will shortly introduce the main formalisms for knowledge representation: Semantic Networks (SN), Description Logics (DL), Conceptual Graphs (CG) and Frame Systems (FS). The first formalism presented is the older one: SN.

2.1. What is knowledge representation?

For responding this question we will follow Davis *et al.*'s (1993) influential paper that extensively addressed this problem. The thing to start with is the obvious and fundamental observation that what we want to represent is the world in the broad sense. We notice that the representation is different from the object that it should represent. Therefore, the object is a substitute, a replacement of the thing itself.

The second observation is that we are committed to a certain vocabulary for representation. It is up to us to define and select the relevant concepts and attributes for representing the world; technically speaking we make an *ontological commitment.* Without entering deep philosophical discussions, the ontological commitment can be seen by analogy with a pair of glasses with which we choose to see the world. Some things will be brought in focus, while others will be in the background or not even considered at all.

The third thing we notice is that KR has always been seen as related to intelligence. The representation is not arbitrary, but a sign of intelligent behaviour.

The fourth observation is that KR systems were designed primarily for computer understanding. Consequently, they should be adequate to this purpose.

Last but not least, we notice that KR language is not only the language that facilitates computer understanding, but also a language that is meant for human understanding. It is the language in which the programmers encode the knowledge and, in many situations, it is the language in which they communicate with each other.

If we detail the above summary presentation we can say that a KR is:

1. **a surrogate,** a substitute for the thing to be represented
2. **a set of ontological commitments**
3. **a fragmentary theory of intelligent reasoning**
4. **a medium for computation**
5. **a medium for human expression.**

Below we will discuss each of these in turn.

2.1.1. *KR is a surrogate*

The underlying fundamental philosophical assumption here is that reality exists independently of the agent who perceives it, be it a human or a computer. The representation of reality, the surrogate, is in the mind of the cognitive agent. The relation of correspondence between the surrogate and the things in the world constitutes the meaning of the representation. Moreover, the surrogate, by means of inference[11] operations, enables us to draw conclusions about the world by reasoning and not by acting. If we take, for example, the planning process of building a house we see that we have internal representations for things like bricks, walls, etc., and that the planning process itself is a reasoning process that takes place inside the cognitive agent's mind. Even if this process is internal, in the sense that we do not act against the represented reality, we reach conclusions about the

[11] By inference we understand the operation of derivation of new knowledge from existent one with the aid of rules. We do not imply that the only inference rules are logical rules.

external world. In a way, we can say that we reach those conclusions by mentally manipulating the representations of things as opposed to directly acting against the things themselves. Because it is clear that we cannot represent everything, one very important question is the adequacy or fidelity of the representation. We can ask how faithful the surrogate is to the original, that is what attributes of the real thing are represented and what attributes are not represented. Generally, in the KR community this question receives a pragmatic answer. Roughly speaking, what we represent depends on our purpose, which in turn depends on the application for which we use the KR. However, if we go beyond application and we do cognitive modelling, then the pragmatic principle will be replaced by particular scientific criteria.

2.1.2. *KR is a set of ontological commitments*

In discussing this point we can make a simplifying assumption. We will consider that the substitute is a language. The studies in cognitive science will ultimately tell us if our internal representation of the world has or has not a language-like structure. This is an interesting problem, but it is beyond the scope of this paper. We can ask many important questions about this language: what are its syntactic rules, how it gets semantics, etc. The answer to the second question was sketched above. We saw that the meaning of the language is the correspondence relation between the language and the world. But a more primary question about the language is: to what kind of entities is the language committed? For example, the language of first order logic (FOL) is committed to the existence of individuals, concepts and properties, the language of physics is committed to the existence of matter, of atoms, of waves, of energy, etc., the language of genetics is committed to the existence of genes, chromosomes, DNA, etc. By employing a language and not another we make certain choices about what is out there in the world. Technically speaking, we make an *ontological commitment,* we represent the world in a certain way.

2.1.3. *KR is a fragmentary theory of intelligent reasoning*

One of the main properties of a representation is that it is not seen just as a mere sign but as a mark of intelligent behaviour. But what constitutes an intelligent behaviour? In Artificial Intelligence (AI) there are two main paradigms that answer this question. The first one, called *logical paradigm*, is very old and can be traced back to Aristotle and especially to Leibniz's *mathesis universalis.* According to it the mark of intelligence is rationality, where rationality is meant to mean logical soundness. Therefore, whatever behaves according to logical rules is intelligent. The other approach, called *psychologist paradigm*, takes inspiration from the work of Miller (1951), Newell, Simon (1972). It opposes logical paradigm in that it states that logical rules are inadequate as an explanation of intelligent behaviour. In these paradigms the logical rules are replaced with a collection of cognitive inspired mechanisms.

These two paradigms have a clear influence on KR representation systems. The KR systems inspired by logical paradigms, like description logics, put emphasis on logical deduction and on completeness and consistency of the system. The offsprings of psychological paradigm, like frame systems for example, employ cognitive inspired mechanisms like analogical reasoning or reasoning based on partial matching.

2.1.4. *KR is a medium for computation*

Because KR languages are designed especially for computer manipulation they should be computationally efficient. Therefore, from this point of view we are concerned with the computational properties of different languages for KR. For example, the whole FOL language is not decidable; therefore decidable fragments of it were identified and used as KR languages. Moreover it is known that there is trade-off between expressivity of the language and its computational efficiency: the more powerful the language, the harder to compute. For a brief presentation of a computationally tractable language see the Description Logics section below (**2.3.**).

2.1.5. *KR is a medium for human expression*

This is perhaps the weakest characteristic of KR, but, nevertheless, one that cannot be neglected. As stated in **2.1.4.** above, an artificial KR language is mainly designed for computers. From the human beings' point of view the most efficient KR language is by far the natural language. However, we can evaluate different artificial KR language with respect to the easiness with which they are learned and understood by human beings. Obviously, a very difficult to learn knowledge representation language will be more difficult to accept by the community than an easy one.

2.2. Semantic networks

Semantic Networks (SNs) were introduced by Quillian (1967). Because Quillian's SNs were inspired by cognitive studies on human memory, they were named semantic memory models. With his paper, Quillian inaugurated an important line of research in AI. However, SNs were not Quillian's invention. Even if they were not known as such, they have a long history that started with Aristotle. The first to construct a semantic network was Porphyry, a student of the celebrated philosopher Plotinus. He arranged Aristotle's categories in a tree known as *the tree of Porphyry*. Bellow we show this famous tree, the first Semantic Network:

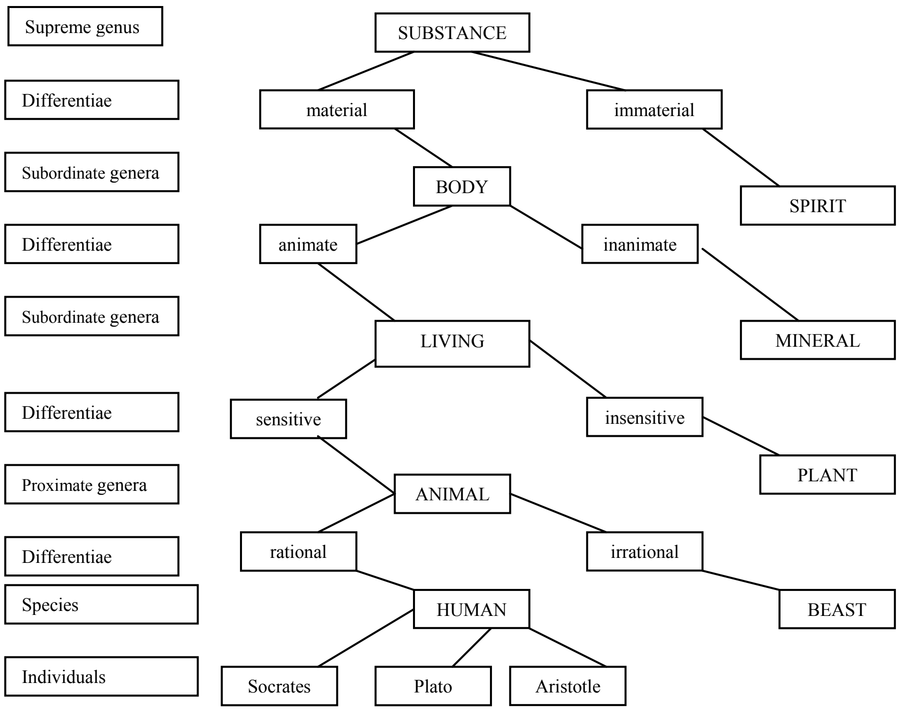

Fig. 2 – The tree of Porphyry.

The top of the tree is the most general category, SUBSTANCE. As we descend we find the differentia that distinguishes the subcategories. For example, a SUBSTANCE with differentia *material* gives us the subcategory BODY and a substance with differentia *immaterial* gives us the category SPIRIT. A HUMAN will be a *rational* ANIMAL or, if we want to define it using the top category SUBSTANCE, we will say that a HUMAN is a *rational sensitive animate material* SUBSTANCE. This is achieved by successively replacing each definable term with its definies: the subcategory ANIMAL will be replaced by *sensitive* LIVING, LIVING will be replaced by *animate* BODY, and BODY by *material* SUBSTANCE. As one can easily notice, everything could be defined using only the category SUBSTANCE and the differentiae.

It is interesting to notice that Quillian's semantic memory models and many of other semantic networks that followed him (for a nice overview see Brachman 1979) aimed at representing concepts similarly to an encyclopaedic or dictionary definition. Because there did not exist an acknowledged terminology, the first semantic networks had some terminological differences. Therefore, we will not try to present semantic memory models or other specific network. Instead, we will introduce a *general semantic network* by abstracting the main characteristics from a family of semantic networks. Bellow a small semantic network is depicted. It will be used for introducing the Semantic Networks terminology.

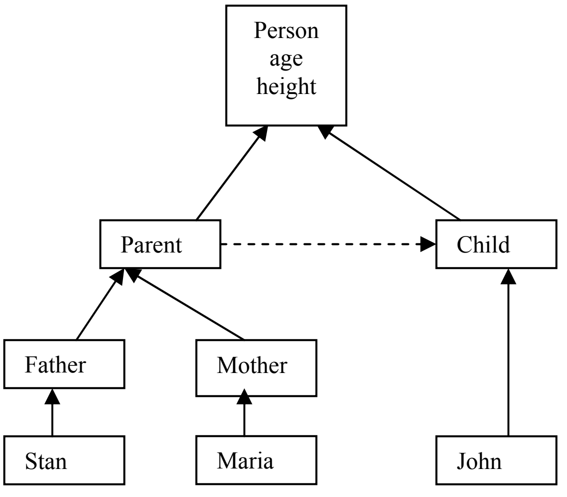

Fig. 3 – A general semantic network.

As one can see in the image above, a semantic network is a graph whose nodes represent concepts and individuals, and whose arcs represent relations between nodes. Besides the relations, attached to each node there is a set of attributes. One can also notice that a semantic network can be extracted from a dictionary or from a body of knowledge like a corpus. The difference between a traditional dictionary and a semantic network is the more efficient way in which the semantic network stores the information. A more detailed definition of all the terms used in the definition of a semantic network is given bellow.

Individuals or instances represent the objects of the domain of discourse. Because each Semantic Network aims at representing the knowledge associated with a particular domain we consider that the domain of discourse is fixed. The objects of the domain of discourse are tokens of specific kind. In Semantic Networks they are represented by individual nodes. In our example Maria, Stan and John are individuals.

Classes or concepts are interpreted as sets of individuals. They are universals that exist in their respective instances. The individuals are grouped in classes on the basis of the common features they have, therefore classes can be thought of as a way of partitioning the domain of discourse. In our network the concepts are represented by concept nodes. The Parent node represents the set of all parents in our domain of discourse, the node Mother represent all the mothers, etc.

There are two kinds of **properties**: relations and attributes. The relations link two or more individuals. In modern semantic networks the arity of relations is two: that is, the relations link two individuals or two concepts. The attributes are simple properties attached to the concept nodes. In the above example an attribute is age and a relation is hasChild (represented by means of the dashed arrow in Figure 3).

We introduce now the most important relations that organize the semantic networks. It is essential to understand that these are distinct relations, as they are often confused in practice. The relations are called: *instance-of* and *IS-A*. The former relation holds between an instance and a class, in this order. For example, Maria is an instance of the concept Mother. The IS-A (called also hyponymy in lexical semantics) relation defines the hierarchy of concepts, called taxonomy. For example, Mother IS-A Parent (in Figure 3 all IS-A relations are represented as continuous arrows). Mother is called the subclass or the subconcept (also called hyponym in lexical semantics) and Parent is called the superclass or superconcept (in lexical semantics also called hyper(o)nym). The IS-A relation has a very clear meaning that can be easily specified formally:

(6) A concept A IS-A a concept B if every instance of A is also an instance of B.

The IS-A relation induces the inheritance of properties, meaning that every subclass inherits all the properties of its superclasses. For example, all subclasses of Persons inherit from this class the attributes *age* and *height*. The advantage of inheritance is that this leads to economy of representation. It is not necessary to copy all the shared properties between a certain concept and its subconcepts.

At this point it is important to introduce the two components of a knowledge base:

1. The Terminology or the T-BOX. This is meant to represent the universal part of a knowledge base. It includes the concepts, their relationships and all the additional axioms that constrain the meaning of the concepts. Nowadays the term **ontology** imposed as a proper name for this intensional structure. In the Semantic Network above the subgraph that includes the taxonomy plus the relation hasChild constitute the ontology.
2. The assertional part or A-BOX. This includes all the knowledge about instances. For example, the fact that Stan is an instance-of Father or that the pair of instances (Maria, John) stays in hasChild relation are part of A_BOX.

One of the problems of early semantic networks formalism was the lack of precise semantics for the elements of the net. The most typical instance for an ad-hoc interpretation of the elements of a network is that of inheritance by default. Let’s take, for example, the concept *penguin*. We all know that a penguin is a bird and therefore it will inherit from the concept *bird* properties like being warm-blooded, laying eggs, etc. However, birds have by default the property of flying, which penguins lack. For modelling this in the early Semantic Networks the penguin's property *no-fly* was let to override the superconcept's property *fly*. This inheritance by default is not a desirable aspect of a formalism, because it makes the IS-A relation meaningless.

To overcome this problem one solution was the proposal of non-monotonic formalisms. A well known example of non-monotonic formalism can be found in the work of Touretzky *et al.* (1987). Another solution is to study the monotonic part of a Semantic Network. In this direction one of the main descendents of semantic networks were Description Logics, the topic of the next section.

2.3. Description logics

Description Logics is regarded as one of the direct successors of Semantic Networks and nowadays this is the most popular formalism for representing knowledge. Because it is a solid formalism tested in many applications the future Semantic Web will extensively use it. Not accidentally the OWL-DL language (http://www.w3.org/TR/owl-guide/), the language that describes the decidable and complete fragment of OWL language, has an acknowledged correspondence with DL.

A DL language is usually a fragment of FOL, but there are DL languages that are more powerful than FOL, being species of modal logics. In what follows we will introduce this formalism by presenting the main elements of the language. The vocabulary of a DL language has the following components:

1. **Concepts** that correspond to unary predicates in FOL;
2. **Roles** that correspond to binary relations in FOL;
3. **Constructors** that are operators for building new concepts from existing ones.

For example, if we consider that we have in T-BOX the concepts Herbivore and Mammal we can use the constructor $\cap$ (intersection) for building a new concept: *Mammal*$\cap$*Herbivore*. This concept will denote those individuals that are both mammals and herbivores. The most interesting construct in DL is the construct which allows us to establish relations between concepts.

$\exists$*bodypart. Trunk* will denote the set of individuals that have at least a body part relation with a concept Trunk. In plain English this is read as: the things that have a trunk.

$\forall$*colour. Grey* will denote the set of individuals that have a colour relations only with the concept Grey. In natural language this is equivalent to: the set of individuals that have only grey colour.

Using the introduced concepts we can build the concept Elephant in a definitional manner: Elephant = *Mammal*$\cap\exists$*bodypart.Trunk*$\cap\forall$*colour.Grey*. According to this definition an elephant is a Mammal who has a trunk as a body part and has only grey colour. One can notice again that, in this way, we can give a logic translation to any dictionary definition. A name can be assigned to the newly created concept in T-BOX.

In addition to the possibility of building new concepts with the aid of constructors and to the possibility of assigning them a name, a DL system offers reasoning capabilities. One very important and much studied reasoning service is the automatic computation of taxonomy with the aid of a reasoner. The usual procedure for building a knowledge base in DL formalism is to start with a set of atomic concepts and then use the constructors to built new ones. Afterwards we can make use of the reasoner for automatic computation of the taxonomy. The reasoner not only allows us to automatically compute the taxonomy, but it also finds inconsistent assertions. This is a very important function that allows us to maintain the internal consistency of large knowledge bases.

As stated above, there are a variety of DL languages but here we will limit at presenting the most basic language. The Attributive Language (AL) was introduced by Schmidt-Schauß and Smolka (1991). In this language we have:

A | an atomic concept
$\top$ | the universal concept
$\bot$ | the bottom concept

The language constructors are:

~A | atomic negation
C∩D | intersection
$\forall R.C$ | (value restriction)
$\exists R.T$ | Limited existential qualification

A formal semantics for AL language is easily given by considering an interpretation function which maps every atomic concept to a subset of the domain of discourse and every role to a pair of individuals in the domain of discourse. Without introducing any supplementary notation we give an informal semantics:

1. The universal concept is interpreted as the entire domain;
2. The bottom concept is the void set;
3. The C ∩ D is interpreted as the set of C intersected with the set of D;
4. $\exists R.T$ is the set of individuals that have at least a relation R with an individual from the domain;
5. $\forall R.C$ is the set of individuals in the domain that have all R relations only with individuals that are instances of concept C.

Considering Person and Male as basic concepts in this sample language and hasChild a role in this language, we can build the following expressions:

Person∩Male .The set of persons that are males.

Person∩$\exists hasChild.T$ The set of persons that have a child.

Person∩$\forall hasChild.Male$. The set of persons that have only male children.

The DL can be used for giving a precise semantic to Semantic Networks. For example, the nodes of the network are mapped to the concepts of DL language the relations to roles, etc. One alternative formalism for representing knowledge that can be regarded as a competitor for DL is represented by Conceptual Graphs. They make the topic of the next section.

2.4. Conceptual graphs

Conceptual Graphs (CG) are a very popular formalism for representing knowledge and were introduced by John Sowa (1984). They are largely used in the NLP community mainly for two reasons: they allow a direct mapping to language and they have a supposed cognitive significance. It is claimed that in the process of understanding a sentence there is a stage which involves the translation of the respective sentence in a conceptual graph-like format.

Sowa acknowledges two sources of inspiration for his formalism: Charles Sanders Peirce's existential graphs and semantic networks.

Formally a conceptual graph is a bipartite graph that has two kinds of nodes: concepts and relations. Nodes are related by directed arcs. In graphic notation, which is only one possible notation for CG, the direction is indicated by an arrow head. The simplest conceptual graph is drawn bellow. It represents the phrase *conceptual graph*.

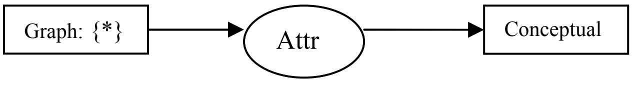

Fig. 4 – The conceptual graph for the phrase *conceptual graph*.

As opposed to SNs or DL, in CG a concept is considered to be a tuple [Type: Referent]. The type of a concept is a name that we give to a group of entities that have similar traits and the referent of the concept is the object that is designed in the actual situation. There are two kinds of concepts:

- Individual concepts like [Dog: Grivei] where the referent is present.
- Generic concepts, i.e. concepts that lack the referent, as, for example, [Dog]. However, in this case the presence of a general referent is implied. [Dog] is just shorthand for [Dog: *] which means that there exists a dog.

Types are organized in hierarchies which are represented as lattices with an upper element Entity called *minimal common supertype* and the lowest element Absurdity called the *minimum common subtype*. The figure bellow depicts such a lattice hierarchy:

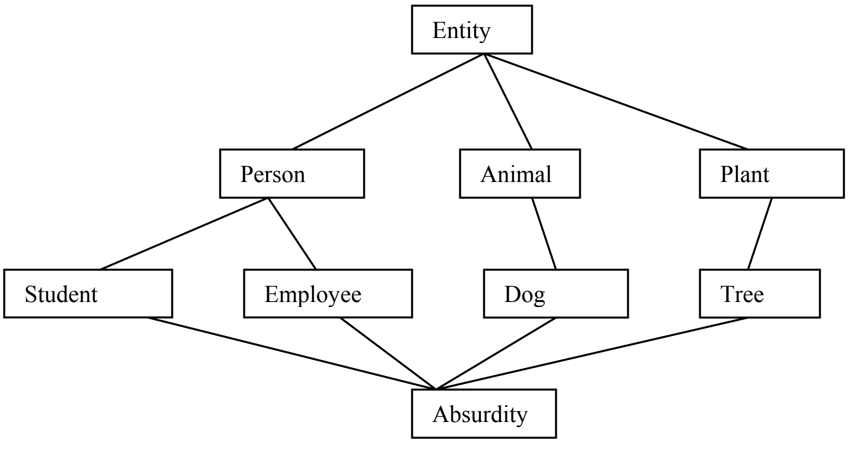

Fig. 5 – A lattice hierarchy.

In CG a relation links two or more concepts. The most important characteristics of a relation are the following ones:

- **type**. If the type of a concept is the name we give to a set of individuals that have common traits, the type of a relation is the name that we give to a set of relations that have common characteristics.
- **valence –** represents the number of arcs that belong to the relation
- **signature –** represents the concepts involved in the relation.

Next we will present two examples of conceptual graphs. The first graph represents the sentence *A cat is on the mat*.

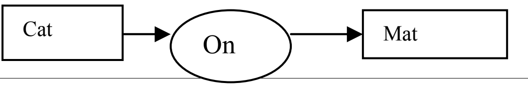

Fig. 6 – The graph representation for the sentence *A cat is on the mat*.

The concept [cat] represents an instance of cat and the concept [mat] an instance of mat and the relation *on* links the two concepts.

An example a little bit more complicated is the following: *John hit the table with a hammer*.

The relation types AGNT, INST and PTNT represent the semantic roles Agent, Instrument and Patient, respectively. The conceptual graph can be paraphrased as: John is the agent of hit, hammer is the instrument of hit, and table is the patient of hit.

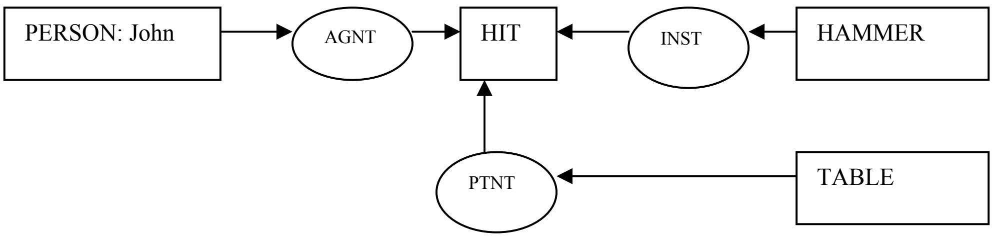

Fig. 7 – The graph representation for the sentence *John hit the table with a hammer*.

From the point of view of CG a knowledge base will have the following structure:

1. A type hierarchy which represents the categories we want to talk about. As stated above, the type hierarchies are arranged in lattice structures.
2. A type relation hierarchy.
3. A set of individuals in the KB.
4. A collection of conceptual graphs which represents the A-BOX.

According to Sowa, CG receive semantics by being translated in FOL. This being the case, they are too powerful for the reasoners. Therefore, small decidable fragments of CG were identified, the best known being the simple conceptual graphs.

2.5. Frame systems

The third main formalism for representing knowledge, Frame Systems, was introduced by Marvin Minsky in his seminal paper *A framework for representing knowledge* (1975). Frame Systems are a reaction to the logic approach for representing knowledge.

One consequence of this fact is the vague definition that frames receive in the above mentioned paper. The original definition of frames is given by showing examples and making analogies with other similar approaches in science or philosophy. According to Minsky a frame is a pattern for representing knowledge:

> A *frame* is a data-structure for representing a stereotyped situation, like being in a certain kind of living room, or going to a child's birthday party. Attached to each frame there are several kinds of information. Some of this information is about how to use the frame. Some is about what one can expect to happen next. Some is about what to do if these expectations are not confirmed. (Minsky 1975)

From this definition it can be seen that the main goal of an approach based on frames was to gather all the relevant knowledge about a situation in one data structure instead of distributing it over different axioms. Another important characteristic of frames is that they are basic building blocks of frame systems: "We can think of a frame as a network of nodes and relations. The "top levels" of a frame are fixed, and represent things that are always true about the supposed situation. The lower levels have many *terminals* – "slots" that must be filled by specific instances or data. Each terminal can specify conditions its assignments must meet. (The assignments themselves are usually smaller "sub-frames.")" (Minsky 1975)

As Minsky himself acknowledges, the concept of frames is neither original nor new. Structures that resemble frames have a respectable tradition in philosophy and psychology. For example, writing an important essay in the philosophy of science, Kuhn (1970) defined the so-called *paradigms*. They were meant as an explanation for the scientific community practice. According to Kuhn the scientific practice takes place inside a set of received beliefs called paradigms. The beginner scientist should assimilate these beliefs without questioning them. S/He will see the world through this paradigm and her/his work will be that of discovering new things inside this paradigm. In the same way when we confront a new situation, like entering a room, or going to someone's party, we have a kind of pattern in our mind, a general schema of the respective situation. We have prior expectations about what we encounter and also possible responses if the situation does not meet these expectations. In assessing the situation we can say that in a way we "retrieve" this structure, from memory and instantiate it by filling it with particular aspects of the situation. Even if Minsky was deliberately vague and did not give a more precise definition for frames, when it comes to building applications we should be very specific. Formalizing frames is the best way to better understand what the advantages of frame-based systems are and how frame systems are different from the other KR systems. Based on Minsky's insights frame systems were built immediately after the publication of his foundational paper.

The experience the scientific community gathered after implementing frame systems was synthesized by Patrick Hayes (1980): "most of the frames are just a new syntax for parts of first order logic". This means that the monotonic part of this formalism can be captured by FOL.

Despite this disconcerting result we see two advantages of using frame-based systems. The first is the compactness of representation. Instead of being distributed across many axioms the knowledge is encapsulated in an object-like structure. The second one, which is still under-explored, is the fact that frames employ not standard reasoning procedures which like reasoning by analogy or reasoning based on partial matching. As frame-based formalisms are much more difficult than semantic networks, we prefer not to give examples of frame structures. Moreover, in contrast with Semantic Networks, the instantiation of frames is a very complicated procedure.

After presenting all these formalisms we can ask a legitimate question: which is the best formalism for representing knowledge? The answer we give is that there is no such thing as the best formalism for representing knowledge. Every knowledge representation formalism has its strengths and its weaknesses. If in the 70's Minsky believed that the systems based on frames will be much better than other systems, now he changed his opinion: "To solve really hard problems, we'll have to use several different representations. This is because each particular kind of data structure has its own virtues and deficiencies, and none by itself would seem adequate for all the different functions involved with what we call common sense." (see http://www.aaai.org/AITopics/html/repr.html)

But we did not learn only this lesson from the past. We also learnt that formalizing the common sense is a problem much harder than it was thought to be. Impossible would one say. Why? Perhaps because the human brain is not a computer and the notion of semantics employed in the interpretation of KB is too weak to account for human semantic capacity.

CONCLUSIONS

In this paper we showed how the representation of knowledge evolved from simple dictionaries to modern knowledge representation formalisms. Our presentation is far from being exhaustive or even complete. Many important types of dictionaries and many relevant formalisms for representing knowledge were not presented. We strongly believe that the KR field did not say its last word. We hope that in the future, with the progress of cognitive science, we will have a better understanding of the nature of meaning and, implicitly, we will build better knowledge representation systems.

REFERENCES

Amsler, R. A., 1989, "Words and Worlds", in: Y. Wilks (ed.), *Theoretical Issues in Natural Language Processing*, New Jersey, Lawrence Erlbaum Associates.

Bidu Vrănceanu, A., 1993, *Lectura dicționarelor*, București, Editura şi Atelierele Tipografice „Metropol".

Boguraev, B., E. Briscoe, 1989, "Introduction", in: B. Boguraev, T. Briscoe (eds) *Computational Lexicography for Natural language Processing*, London, New York, Longman, 1–40.

Brachman, R. J., 1979, "On the epistemological status of semantic networks", in N.V. Findler (ed.): *Associative Networks*, Academic Press, 3–50.

Canarache, A., 1970, *Lexicografia de-a lungul veacurilor. De când există dicționare?*, București, Editura Științifică.

Coteanu, I., L. Seche, M. Seche, (eds), 1996, *Dicționarul Explicativ al Limbii Române* (DEX), București, Editura Univers Enciclopedic.

Davis, R., H. Shrobe, P. Szolovits, 1993, "What is A Knowledge Representation?", *AI Magazine*, 14 (1), 17–33.

Gazdar, G., C. Mellish, 1989, *Natural Language Processing in PROLOG: An Introduction to Computational Linguistics*, Wokingham, Addison-Wesley.

Hanks, P., 2003, "Lexicography", in: R. Mitkov: (ed.), *The Oxford Handbook of Computational Linguistics*, Oxford University Press.

Hayes, P. J., 1980, "The logic of frames", in: D. Metzing (ed.), *Frame Conceptions and Text Understanding*, Berlin, deGruyter, 46–61.

Jackson, H., 2002, *Lexicography. An Introduction*, London and New York, Routledge.

Kuhn, T., 1970, *The Structure of Scientific Revolutions,* University of Chicago Press, 2nd Ed.

Landau, S.I., 2001, *Dictionaries. The Art and Craft of Lexicography*, Cambridge University Press, Second Edition.

Leech, G., 1974, *Semantics*, Harmondsworth, Penguin.

Merriam Webster Online Dictionary, http://www.m-w.com/

Mel'cuk, I.A., N. Arbatschevsky-Jumarie, L. Elnitsky, L. Iordanskaya, A. Lessard, 1984, *Dictionnaire explicatif et combinatoire du français contemporain*, Montréal, Les Presses de l'Université de Montréal.

Mel'cuk, I., A. Clas, A. Polguère, 1995, *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve, Duculot.

Mel'cuk, I., L.N. Iordanskaya, N., Arbatchewsky-Jumarie, 1981, "Un nouveau type de dictionnaire", *Cahiers de lexicologie*, 38, 1, 3–34.

Mel'cuk, I., A. Polguère, 1987, "A Formal Lexicon in the Meaning-Text Theory (Or How to Do Lexica with Words)", *Computational Linguistics*, 13, 3/4, 261–275.

Mel'cuk, I., A. Zholkovsky, 1984, *Explanatory Combinatorial Dictionary of Modern Russian*, Vienna, Wiener Slawistischer Almanach.

Miller, G.A., 1951, *Language and Communication*, New York, McGraw-Hill.

Minsky, M., 1975, "A framework for representing knowledge", in: P.H. Winston (ed.), *The Psychology of Computer Vision*, New York, McGraw-Hill, 211–277.

Newell, A., H. A. Simon, 1972, *Human Problem Solving*, Englewood Cliffs, NJ, Prentice-Hall.

Ooi, V.B.Y., 1998, *Computer Corpus Lexicography*, Edinburgh University Press.

Pollard, C., I.A. Sag, 1987, *Information-based Syntax and Semantics. Volume 1. Fundamentals*, Stanford, University of Chicago Press.

Pollard, C., I.A. Sag, 1994, *Head-driven Phrase Structure Grammar*, Chicago, University of Chicago Press.

Quillian, M.R., 1967, "Word concepts: A theory and Simulation of Some Basic Semantic Capabilities", *Behavioral Sciences*, 12, 410–430.

Schmidt-Schauß, M., G. Smolka, 1991, "Attributive concept descriptions with complements", *Artificial Intelligence*, 48, 1, 1–26.

Seche, M., *Schiță de istorie a lexicografiei române*, București, Editura Științifică, I, 1966, II, 1969.

Sowa, J.F., 1984, *Conceptual Structures: Information Processing in Mind and Machine*, Addison-Wesley, Reading, MA.

Touretzky, D.S., J.F. Horty, R. H. Thomason, 1987, "A Clash of Intuitions: the Current State of Non-monotonic Multiple Inheritance Systems", in *Proceedings of the 10th International Joint Conference on Artificial Intelligence* (IJCAI'87), 476–482.

Visual Thesaurus, http://www.visualthesaurus.com/

Zock, M., 2005, *What does the word 'dictionary' evoke in the minds of a linguist, a computer scientist, or a psychologist?*, Chiang Rai, Thailand, Papillon meeting.