

# Using the WordNet ontology for interpreting Medical Records

Jon Espen Ingvaldsen and Csaba Veres

Norwegian University of Science and Technology  
N-7491 Trondheim-NTNU  
jonespi@stud.ntnu.no, Csaba.Veres@idi.ntnu.no

As hospitals throughout Europe are striving exploit advantages of IT and network technologies, electronic medical records systems are starting to replace paper based archives. This paper suggests and describes an add-on service to electronic medical record systems that will help regular patients in getting insight to their diagnoses and medical record. The add-on service is based annotating polysemous and foreign terms with WordNet synsets. By exploiting the way that relationships between synsets are structured and described in WordNet, it is shown how patients can get interactive opportunities to generalize and understand their personal records.

## Introduction

At Norwegian Hospitals, patients have the right to access all information that is stored in their medical record and add comments if written information are incorrect. However, such a right and opportunity vanishes if the medical record is written in such a language that it is not understandable for patients without medical background.

The add-on service that we will present in this paper exploits how relationships between words and word-meanings are structured and described in WordNet. In difference from just looking up in a thesaurus and give patients a formal definition of foreign words, we will show how the relationships in WordNet can be used to give patients interactive opportunities to generalize and understand foreign words.

## WordNet

WordNet is an electronic lexical database that has been developed and maintained at Princeton University since 1985 [3]. Unlike standard alphabetical dictionaries, which organize vocabularies using morphological logical similarities, WordNet structures lexical information in terms of word meanings. Words of the same syntactic category that can be used to express the same meaning or concept are grouped into a single synonym set, called synset. Words with multiple meanings (polysemous words) belong to multiple synsets. In addition to a set of words of the same syntactic category, each synset has a unique identifier and a gloss that defines the synset.

Several types of semantic relations between synsets are recorded in WordNet. These include [2][3]:

- **Hypernymy** - Hypernymy (specific-generic) is the most dominant semantic relation and it structures noun concepts into 11 hierarchies. **A** is a hyponym of **B** if **A** is a (kind of) **B**.
- **Meronymy / Holonymy** - A part-whole inversible relation between nouns. *If A is a meronym of B, then B one part and a holonym of A.*
- **Entailment**- A verb **A** entails **B** if **A** cannot be done unless **B** is, or has been, done. I.e., *snore* lexically entails *sleep*. Another property of entailment is that negation reverses the direction of entailment, i.e., *Not sleeping* entails *not snoring*.
- **Troponymy** - A concept relation between two verbs **A** and **B** that can be expressed by the formula: "*To A is to B in some particular manner*". I.e., *to limp* is also *to walk* in a certain manner; *limp* is a troponym of *walk*. Activities referred to by a troponym and its more general superordinate are always temporally co-extensive, in that one must necessarily be walking every instant that one is limping. Troponymy therefore represents a special case of Entailment [5].

## Medical Records

The primary purpose of medical records is [4]:

1. To provide documentation on the cause of an individual's health care.
2. To provide a means of communication amongst health care professionals for current and future patient care.

A lot of research has been done in order to replace much of the paper based medical record archives with electronic systems. Some hospitals have recently entered the era of such systems; introducing electronic medical records. Medical records are characterized by short and precise sentences. There are seldom use of hard grammars, but they are written with use of professional language with medical words and abbreviations. An example of such characteristics and information that can be found in a medical record is given in table 1.

**Table 1.** Examples of information in a medical record.

<b>History</b>	Date December 10, 2001
<b>ID</b>	Ms. XY is a 23-year-old single female college student
<b>CC</b>	"Headache"
<b>HPI</b>	Ms. XY presents with a one-month history of headaches. The headaches are located in a band like distribution around her forehead and occipital area.
<b>Physical Examination</b>	
<b>Vital Signs</b>	BP 110/68 P 72 RR 14 T 98.4F Wt 160lb Ht 68 in
<b>CV</b>	No jugular venous distension. No carotid bruits. Apical impulse mid-clavicular line, 5th intercostal space. Regular rhythm. Normal S1, S2 with physiological split. No S3 or S4. No murmur or rub. No femoral bruit. Pulses: dorsalis pedis, posterior tibialis, femoral and radial 2+. Abdomen Flat,

	without scars. Bowel sounds present. No aortic, renal or iliac bruits. Liver span 8cm in mid-clavicular line; edge palpable 1 cm below costal margin, smooth. Spleen and kidneys not palpable.
--	--

Advantages of giving patients access to their medical records includes [4] (a) making them involved in their own health care, and able to understand treatment and follow medication programs in an informed manner, (b) to be sure that the record is accurate and relevant, and (c) to be able to make a complaint. Just having a look at the medical record in table 1, or other medical records, it is easy to see that such advantages vanishes if the respective patient is unable to read medical profession language and understand the content.

### Suggested Solution

Our solution for giving patients interpretative support is to annotate foreign word-meanings with the respective WordNet synset. Having such annotations available in an electronic medical record system, we are able to exploit the gloss and conceptual relationships in WordNet and interactively generalize sentences in the medical record such that patients can grasp their meanings. The conceptual relationships that we want to make use of are the *synonym terms* in the *synsets*, *hypernymy* and *troponymy*. We will also show the use of *meronymy* and *holonymy* in this context.

Uncertainty about word meaning might disappear by a look at the gloss. If not, the patient can try finding familiar and known words in the synonym terms of the synset<sup>1</sup>. If the words are annotated correctly, it is possible to change a word with a synonymous term in the respective synset, and still preserve the semantical meaning of the sentence. If none of the synonym terms in the synset are familiar for the patient, the patient can use hypernymy or troponymy relationships (if existing) to generalize the word to a level that is found familiar<sup>2</sup>.

We will illustrate the approach with an example. Assume that you are the patient with the medical record in table 1, and you are struggling to understand CV field.

Say, you are uncertain about the meaning of the first annotated word “*jugular*”. Given that this word is annotated with the synset “*jugular -- (relating to or located in the region of the neck or throat; "jugular vein")*”, most patients are able to grasp the meaning of the word by reading the gloss. Knowing the meaning of “*jugular*”, we may still be unfamiliar with the meaning of “*distension*”. -Making us unable to grasp the meaning of the whole sentence. Having a look at the annotated synset “*distention, distension -- (the act of expanding by pressure from within)*” we may not get enough information by looking at the gloss or the synonym terms. Choosing to look at the next level the hypernymous hierarchy, we find the synset “*expansion, enlargement -- (the act of increasing (something) in size or volume or quantity or scope)*”. With such information we might be able to grasp the meaning of the word “*distension*” and the

<sup>1</sup> Note that not all synsets have a set of synonym terms.

<sup>2</sup> Our approach to generalize word-meanings is similar for both hypernymy and troponymy relationships. If the foreign word is a noun, we will use hypernyms (if existing) to generalize. Similarly, we would use troponyms if the foreign word is a verb.

whole sentence. If the patient is still uncertain about the word, he or she can continue to iterate to more general levels in the hypernym-hierarchy. At each level of generality the patient has the opportunity to get information about the gloss and synonym terms.

Meronymy and holonymy are also relationships that can be used to clarify the intended meaning of foreign words. Assume that the patient is unfamiliar with the word “*abdomen*” which is annotated with the synset “*abdomen, venter, stomach, belly...*”. By having a look at the synonym terms in the synset most patients are able to grasp the meaning of the word, but we can also find additional “*part of*” information in holonyms that is useful for narrowing the meaning. In our case of “*abdomen*”, WordNet provides the holonym “*torso, trunk, body -- (the body excluding the head and neck and limbs; "they moved their arms and legs and bodies")*.”

## Concluding Remarks

Presented functionality that is provided on top of WordNet annotations will give patients abilities to interactively generalize and grasp the meaning of their medical record. The most important advantages of annotating words to WordNet synsets rather than definitions in a dictionary of foreign words are:

1. Patients are able to iterate to more and more general synsets.
2. The rich set of relationships in WordNet allows patients to explore the meaning of words from different directions.
3. The word-meaning for polysemous words are disambiguated preventing patients from interpreting faulty word meanings.

It is not only the patients at the hospital that gain advantages of such WordNet annotations. As ambiguity about word-meanings disappears, employees at the hospital become protected from doing erroneous decisions and actions based on faulty interpretations of word-meanings in medical records.

## References

- [1] Vider, and Orav, 1990, WordNet: An Online Lexical Database, University of Tartu, Estonia
- [2] Miller, Beckwith, Fellbaum, Gross, Miller, 1993, Introduction to WordNet: An On-Line Lexical Database, [www.cogsci.princeton.edu/~wn/5papers.pdf](http://www.cogsci.princeton.edu/~wn/5papers.pdf), last accessed 11th of January, 2004
- [3] Montague, 1995, Consumer access to medical records, Victorian Office of the Public Advocate and the Office of the Privacy Commissioner