Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Using Wordnets in Teaching Virtual Courses of Computational Linguistics

Claudia Kunze, Lothar Lemnitzer

{kunze,lothar}@sfs.uni-tuebingen.de.

The Second Global Wordnet Conference 2004, Brno, 23 January 2004

Topics of the talk

- The context
- Student projects
  - Linguistic aspects of wordnets
  - Structural aspects of wordnets
  - Tools
  - Evaluation
  - Use of GermaNet as a lexical resource
- Concluding remarks

The context

We use wordnets as subjects of teaching and learning in two ’virtual seminars’. The scenario:

- Students prepare a topic in advance, with a textbook
- They discuss the topic, in plenary and group sessions, with the help of synchronous computer mediated communication (text based chat)
- They solve exercises
- They work on their own or in groups on a larger project to earn a certificate.

Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

The virtual courses

GermaNet plays a key role in two courses

- Applied Computational Linguistics
  - Topic: Language Engineering for intelligent dictionary access in text comprehension
  - GermaNet plays the role of an underlying lexical resource
- Computational Lexicography
  - Topics (a.o.): Lexical Semantics, modelling of lexical data, wordnets
  - GermaNet plays the role of a reference resource and test case

Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Student projects

Student projects which have been assigned so far, in relation to GermaNet or wordnets in general, encompass:

- Linguistic issues
- Data model and standardization issues
- Implementation of tools
- Evaluation issues
- Use of GermaNet as lexical resource in applications

Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Linguistic aspects

- Project: Analysis of the meronymy / holonymy relation
- Question: given the German lexical data, is it justified to subdivide this relation into a set of (sub)relations?
- Question: given that, is the subdivision implemented in the Princeton WordNet adequate for GermaNet?
- Status: not yet finished

Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Linguistic aspects ctnd

- Project: Analysis of the antonymy relation
- Question: given the German lexical data, is it justified to subdivide this relation into a set of (sub)relations?
- If so, develop an empirically justified subdivision of this relation.
- Status: not yet finished

Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Linguistic aspects ctnd

- Project: Applicability of regular polysemy in wordnets
- Question: which are the advantages and disadvantages of implementing regular polysemy ?
- Question: is there an appropriate hierarchical level for the implementation of such rules?
- Status: not yet finished

Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Linguistic aspects cntd

- Idea: to introduce aspect and Aktionsart as features of verbal semantics into the verb hierarchy
- Task: for a given verb field, test the possbility of creating an inheritance hierarchy for aspectual features in interaction with argument structure
- This task builds upon the valency patterns which are already available in GermaNet
- Status: not yet finished

Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Structural aspects

- Task: Convert the lexicographer's files of GermaNet to an XML compliant format
- The conversion has been based on an Entity Relationship model of the GermaNet data structure
- Extended Links (XLink) have been used to model the relations between synsets and literals ('lexical units')
- The XML version of GermaNet is the basis for data exchange, visualisation and further conversions.

Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Structural aspects ctnd

- Task: Explore the possibility of integrating GermaNet data into the framework of an ontology language (OWL)
- An experiment with a data sample showed the feasibility of this approach
- The use of OWL and DL as modelling language seems to be superior to RDF(S)
- Only a well-known class-as-instance problem remains to be solved

Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Tools

- Task: Develop a tool for the extraction of relational neighbours of any synset or lexical unit
- two approaches have been tested. One based on GermaNet as a relational database, one on the XML encoded version
- Both approaches turned out to be reliable, with the RDMBS approach being slightly faster

Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

A screenshot

Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Gerna Net Browser

System Buffer SQL Help

Synsets LexUnits Words

WordClass: nomen LexGroup: Artefakt SynsetId: 222

Search hyperonymy Chains

Result#1 Result#2 Result#3

```
<results>
  <result>
      <synset id="nArtefakt.222" wordClass="nomen" lexGroup="Artefakt">
        <lexUnit id="nArtefakt.222.Stütze" stilMarkierung="nein" sense="1" orthVar="ne
          <orthForm>Stütze</orthForm>
        </lexUnit>
      </synset>
      <con_rel name="hyperonymy" dir="->">
          <synset id="nArtefakt.221" wordClass="nomen" lexGroup="Artefakt">
            <lexUnit id="nArtefakt.221.Stützkonstruktion" stilMarkierung="nein" sense
              <orthForm>Stützkonstruktion</orthForm>
            </lexUnit>
          </synset>
          <con_rel name="hyperonymy" dir="->">
              <synset id="nArtefakt.27" wordClass="nomen" lexGroup="Artefakt">
                <lexUnit id="nArtefakt.27.Konstruktion" stilMarkierung="nein" sense=
                  <orthForm>Konstruktion</orthForm>
                </lexUnit>
                <lexUnit id="nArtefakt.27.Vorrichtung" stilMarkierung="nein" sense='
                  <orthForm>Vorrichtung</orthForm>
                </lexUnit>
              </synset>
              <con_rel name="hyperonymy" dir="->">
                  <synset id="nArtefakt.1" wordClass="nomen" lexGroup="Artefakt">
                    <lexUnit id="nArtefakt.1.Artefakt" stilMarkierung="nein" sense='
                      <orthForm>Artefakt</orthForm>
                    </lexUnit>
```

Conceptual chains(hyperonymy) for synset ("nomen", "Artefakt", "222")

Tools ctnd

- Task: Build an wordnet visualization tool built on the XML version of the data
- Result: the implementation (using Java) does not work very well
- Consequence: we will try to convert GermaNet structures into SVG and look for ’off-the-shelve’ visualization tools

Evaluation

- Task: Evaluate the reliability of synset-to-synset links via the Interlingual index for the purpose of machine translation
- The intuition is that synset-to-synset links are not adequate for translation purposes because interlingual equivalence is a relation between lexical units
- A bilingual dictionary for a language pair under examination will be used as reference
- Status: not yet assigned

Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Evaluation cntd

- Task: based on a corpus of citation, compare manual / semi-automatic sense division with the senses of that lexeme which are established in GN
- Goal: Find a criterion for sense clustering which is based on observations of language use
- Remark: The project can only be a case study as long as we do not have reliable sense clustering software
- Status: not yet assigned

Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Application

- Task: Use GermaNet relations for vocabulary training
- Remark: the vocabulary trainer has been a project in another virtual seminar
- GermaNet data have successfully been used for the training module

Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN

Concluding remarks

Concerning the use of wordnets for educational purposes, we want to propose

- to build a collection of small to medium-sized projects for advanced undergraduate and graduate students
- to make the results of successful projects known to the WordNet community
- to establish a forum for the discussion of didactical issues in conjunction with wordnets

Seminar für Sprachwissenschaft
EBERHARD KARLS UNIVERSITÄT TÜBINGEN