

NAME

2.0to2.1.noun.poly – mapping from polysemous noun senses in WordNet 2.0 to corresponding 2.1 senses

2.0to2.1.noun.mono – mapping from monosemous noun senses in WordNet 2.0 to corresponding 2.1 senses

2.0to2.1.verb.poly – mapping from polysemous verb senses in WordNet 2.0 to corresponding 2.1 senses

2.0to2.1.verb.mono – mapping from monosemous verb senses in WordNet 2.0 to corresponding 2.1 senses

DESCRIPTION

WordNet users who have semantically tagged text to senses in version 2.0, or who have statically assigned or used 2.0 senses in other applications should upgrade to WordNet 2.1 if possible. To help users automatically convert 2.0 noun and verb senses to their corresponding 2.1 senses, we provide sense mapping information with version 2.1.

The sense mapping was done as follows:

- Nouns and verbs unique to either database were ignored.
- Nouns and verbs that are monosemous in both databases were found and their *sense_keys* and *synset_offsets* were mapped. These sense mappings are in the files **2.0to2.1.{noun,verb}.mono**.
- All senses of polysemous nouns and verb in version 2.0 were mapped to senses in version 2.1. Various heuristics were used to evaluate the similarity of 2.0 and 2.1 senses, and a score was assigned to each comparison. For each word, each 2.0 sense was compared to all of the 2.1 senses for the same word, and the 2.1 sense (or senses) with the highest score was deemed the best mapping. These sense mappings are in the file **2.0to2.1.{noun,verb}.poly**.

Heuristics include comparison of sense keys, similarity of synset terms, and relative tree location (comparison of hypernyms). Glosses are not used for comparisons, as they are often significantly modified.

File Format

A sense mapping is generally represented by two *sense_key* *synset_offset* pairs, one for the 2.0 sense and one for its corresponding 2.1 sense. For the polysemous sense mappings, *sense_number* is also in each pair. This field is not needed in the monosemous mappings since all monosemous words are assigned *sense_number* 1. See **senseidx(5WN)** for a detailed description of these fields.

In all the mapping files, a space is the field delimiter unless otherwise noted, and each line is terminated with a newline character.

2.0to2.1.{noun,verb}.mono

These files contain the mapping of sense keys for nouns and verbs that are monosemous in both WordNet 2.0 and 2.1. Although the actual words and sense numbers are the same in both databases, not all *sense_keys* are the same, and the *synset_offsets* are different. This file is an alphabetized list of one mapping per line. Each line is of the form:

2.0_sense_key;2.0_synset_offset 2.1_sense_key;2.1_synset_offset

2.0to2.1.{noun,verb}.poly

These files contain the mapping of sense keys for nouns and verbs that are polysemous in WordNet 2.0 and are also found in 2.1. This file is sorted by score from highest score (100) to lowest (0), and then alphabetically within each score. Each line lists all 2.1 sense(s) that the corresponding 2.0 sense maps

to with that score. Each line is of the form:

score 2.0_sense_info 2.1_sense_info [2.1_sense_info...]

where *sense_info* consists of the following three fields:

sense_key;synset_offset;sense_number

SCORES AND STATISTICS

Scores range from 0 to 100, and are an indication of how confident the mapping heuristics are that the senses are the same – a higher score indicates greater reliability in the mapping. The vast majority of senses mapped with a score of 90 or 100. Mappings with a score of 90 or greater make up 98% of the total nouns senses mapped, and 97% of the total verb senses mapped.

Noun Statistics

There are 141,690 noun senses in WordNet 2.0. A total of 141,364 senses have been mapped to senses in version 2.1. The remaining 326 senses represent noun senses unique to version 2.0.

A total of 42,726 senses for polysemous nouns are mapped in the file **2.0to2.1.noun.poly**, with the following scores:

Score	Count
100	36992
90	5336
80	133
70	110
60	36
50	11
40	11
30	26
20	46
0	25

98,638 monosemous nouns are mapped in the file **2.0to2.1.noun.mono**.

Verb Statistics

There are 24,632 verb senses in WordNet 2.0. A total of 24,617 senses have been mapped to senses in version 2.1. The remaining 15 senses represent verb senses unique to version 2.0.

A total of 18,546 senses for polysemous verbs are mapped in the file **2.0to2.1.verb.poly**, with the following scores:

Score	Count
100	17851
90	554
80	12
70	86
60	8
50	6
40	4
30	10
20	12
0	3

6,071 monosemous verbs are mapped in the file **2.0to2.1.verb.mono**.

NOTES

The number of senses of a polysemous word in version 2.0 often differs from the number of senses for the same word in version 2.1. While there will always be a mapping for each 2.0 sense to one or more 2.1 senses, there may be 2.1 senses to which no 2.0 sense is mapped.

WordNet 2.0 words not found in either of the monosemous maps are unique to version 2.0, and therefore cannot be mapped to version 2.1.

ENVIRONMENT VARIABLES

WNHOME Base directory for WordNet. Unix default is **/usr/local/WordNet-2.1**, Windows default is **C:\Program Files\WordNet\2.1**.

FILES

All files are in **WNHOME/sensemap** on Unix platforms, **WNHOME\sensemap** on Windows.

FILES

2.0to2.1.noun.poly mapping of polysemous 2.0 noun senses to 2.1 senses
2.0to2.1.verb.poly mapping of polysemous 2.0 verb senses to 2.1 senses
2.0to2.1.noun.mono mapping of monosemous 2.0 noun senses to 2.1 senses
2.0to2.1.verb.mono mapping of monosemous 2.0 verb senses to 2.1 senses

SEE ALSO

senseidx(5WN), **wndb(5WN)**, **wnpkgs(7WN)**.