NAME

wngloss − glossary of terms used in WordNet system

DESCRIPTION

The *WordNet Reference Manual* consists of Unix-style manual pages divided into sections as follows:

Section	Description
1	WordNet User Commands
3	WordNet Library Functions
5	WordNet File Formats
7	Miscellaneous Information about WordNet

System Description

The WordNet system consists of lexicographer files, code to convert these files into a database, and search routines and interfaces that display information from the database. The lexicographer files organize nouns, verbs, adjectives and adverbs into groups of synonyms, and describe relations between synonym groups. **grind**(1WN) converts the lexicographer files into a database that encodes the relations between the synonym groups. The different interfaces to the WordNet database utilize a common library of search routines to display these relations. Note that the lexicographer files and **grind**(1WN) program are not generally distributed.

Database Organization

Information in WordNet is organized around logical groupings called synsets. Each synset consists of a list of synonymous words or collocations (eg. "**fountain pen**", "**take in**"), and pointers that describe the relations between this synset and other synsets. A word or collocation may appear in more than one synset, and in more than one part of speech. The words in a synset are grouped such that they are interchangeable in some context.

Two kinds of relations are represented by pointers: lexical and semantic. Lexical relations hold between semantically related word forms; semantic relations hold between word meanings. These relations include (but are not limited to) hypernymy/hyponymy (superordinate/subordinate), antonymy, entailment, and meronymy/holonymy.

Nouns and verbs are organized into hierarchies based on the hypernymy/hyponymy relation between synsets. Additional pointers are be used to indicate other relations.

Adjectives are arranged in clusters containing head synsets and satellite synsets. Each cluster is organized around antonymous pairs (and occasionally antonymous triplets). The antonymous pairs (or triplets) are indicated in the head synsets of a cluster. Most head synsets have one or more satellite synsets, each of which represents a concept that is similar in meaning to the concept represented by the head synset. One way to think of the adjective cluster organization is to visualize a wheel, with a head synset as the hub and satellite synsets as the spokes. Two or more wheels are logically connected via antonymy, which can be thought of as an axle between the wheels.

Pertainyms are relational adjectives and do not follow the structure just described. Pertainyms do not have antonyms; the synset for a pertainym most often contains only one word or collocation and a lexical pointer to the noun that the adjective is "pertaining to". Participial adjectives have lexical pointers to the verbs that they are derived from.

Adverbs are often derived from adjectives, and sometimes have antonyms; therefore the synset for an adverb usually contains a lexical pointer to the adjective from which it is derived.

See **wndb**(5WN) for a detailed description of the database files and how the data are represented.

GLOSSARY OF TERMS

Many terms used in the *WordNet Reference Manual* are unique to the WordNet system. Other general terms have specific meanings when used in the WordNet documentation. Definitions for many of these terms are given to help with the interpretation and understanding of the reference manual, and in the use of the WordNet system.

In following definitions **word** is used in place of **word or collocation**.

adjective cluster A group of adjective synsets that are organized around antonymous pairs or triplets. An adjective cluster contains two or more **head synsets** which represent antonymous concepts. Each head synset has one or more **satellite synsets**.

attribute A noun for which adjectives express values. The noun **weight** is an attribute, for which the adjectives **light** and **heavy** express values.

base form The base form of a word or collocation is the form to which inflections are added.

basic synset Syntactically, same as **synset**. Term is used in **wninput**(5WN) to help explain differences in entering synsets in lexicographer files.

collocation A collocation in WordNet is a string of two or more words, connected by spaces or hyphens. Examples are: **man-eating shark**, **blue-collar**, **depend on**, **line of products**. In the database files spaces are represented as underscore (_) characters.

coordinate Coordinate terms are nouns or verbs that have the same **hypernym**.

cross-cluster pointer A **semantic pointer** from one adjective cluster to another.

derivationally related forms Terms in different syntactic categories that have the same root form and are semantically related.

direct antonyms A pair of words between which there is an associative bond resulting from their frequent co-occurrence. In **adjective clusters**, direct antonyms appears only in **head synsets**.

domain A topical classification to which a synset has been linked with a CATEGORY, REGION or USAGE pointer.

domain term A synset belonging to a topical class. A domain term is further identified as being a CATEGORY_TERM, REGION_TERM or USAGE_TERM.

entailment A verb **X** entails **Y** if **X** cannot be done unless **Y** is, or has been, done.

exception list Morphological transformations for words that are not regular and therefore cannot be processed in an algorithmic manner.

group Verb senses that similar in meaning and have been manually grouped together.

gloss Each synset contains **gloss** consisting of a definition and optionally example sentences.

head synset Synset in an adjective **cluster** containing at least one word that has a **direct antonym**.

holonym The name of the whole of which the meronym names a part. **Y** is a holonym of **X** if **X** is a part of **Y**.

hypernym The generic term used to designate a whole class of specific instances. **Y** is a hypernym of **X** if **X** is a (kind of) **Y**.

hyponym The specific term used to designate a member of a class. **X** is a hyponym of **Y** if **X** is a (kind of) **Y**.

indirect antonym — An adjective in a **satellite synset** that does not have a **direct antonym** has an indirect antonyms via the direct antonym of the **head synset**.

instance — A proper noun that refers to a particular, unique referent (as distinguished from nouns that refer to classes). This is a specific form of hyponym.

lemma — Lower case ASCII text of word as found in the WordNet database index files. Usually the **base form** for a word or collocation.

lexical pointer — A lexical pointer indicates a relation between words in synsets (word forms).

lexicographer file — Files containing the raw data for WordNet synsets, edited by lexicographers, that are input to the **grind** program to generate a WordNet database.

lexicographer id (lex id) — A decimal integer that, when appended onto **lemma**, uniquely identifies a sense within a lexicographer file.

monosemous — Having only one sense in a syntactic category.

meronym — The name of a constituent part of, the substance of, or a member of something. **X** is a meronym of **Y** if **X** is a part of **Y**.

part of speech — WordNet defines "part of speech" as either noun, verb, adjective, or adverb. Same as **syntactic category**.

participial adjective — An adjective that is derived from a verb.

pertainym — A relational adjective. Adjectives that are pertainyms are usually defined by such phrases as "of or pertaining to" and do not have antonyms. A pertainym can point to a noun or another pertainym.

polysemous — Having more than one sense in a syntactic category.

polysemy count — Number of senses of a word in a syntactic category, in WordNet.

postnominal — A postnominal adjective occurs only immediately following the noun that it modifies.

predicative — An adjective that can be used only in predicate positions. If **X** is a predicate adjective, it can only be used in such phrases as "it is **X**" and never prenominally.

prenominal — An adjective that can occur only before the noun that it modifies: it cannot be used predicatively.

satellite synset — Synset in an adjective **cluster** representing a concept that is similar in meaning to the concept represented by its **head synset**.

semantic concordance — A textual corpus (e.g. the Brown Corpus) and a lexicon (e.g. WordNet) so combined that every substantive word in the text is linked to its appropriate sense in the lexicon via a **semantic tag**.

semantic tag — A pointer from a word in a text file to a specific sense of that word in the WordNet database. A semantic tag in a semantic concordance is represented by a **sense key**.

semantic pointer — A semantic pointer indicates a relation between synsets (concepts).

sense — A meaning of a word in WordNet. Each sense of a word is in a different **synset**.

sense key — Information necessary to find a sense in the WordNet database. A sense key combines a **lemma** field and codes for the synset type, lexicographer id, lexicographer file number, and information about a satellite's **head synset**, if required. See **senseidx**(5WN) for a description of the format of a sense key.

subordinate Same as **hyponym**.

superordinate Same as **hypernym**.

synset A synonym set; a set of words that are interchangeable in some context without changing the truth value of the preposition in which they are embedded.

troponym A verb expressing a specific manner elaboration of another verb. **X** is a troponym of **Y** if **to X** is **to Y** in some manner.

unique beginner A noun synset with no **superordinate**.