Laura Gilda Paccagnella
University of Padua
Dept. Pure and Applied Mathematics
Via Trieste 63, 35121 Padova, Italy
Tel.: (39–049) 82 71 417; (39–049) 82 71 406
E-mail: laurap@math.unipd.it
Research interests: natural language processing, mathematical data bases.

Carlo Minnaja
University of Padua
Dept. Pure and Applied Mathematics
Via Trieste 63, 35121 Padova, Italy
Tel.: (39–049) 82 71 417; (39–049) 82 71 406
E-mail: minnaja@math.unipd.it
Research interests: history of mathematics, etymology of mathematical words, planned languages.

A MATHEMATICAL DICTIONARY FOR ESPERANTO, ACCORDING TO WORDNET*

WordNet is a database started in 1985 at Princeton, where the search of a word is based on psycholinguistic theories. In WordNet, information is stored according to syntactic categories and is linked through different types of relations. Both synonyms and polysemous words appear; synonyms are collected in sets called synsets. The database is organized by hierarchies between concepts. Later, EuroWordnet was born, a multilingual WordNet with nine different European languages, and some specific domains have been added to get different WordNets, one for each language and for each subject. We arranged a mathematical WordNet for Esperanto, which is at present limited to nouns, where connections among words are classed as follows: 1) synonym – when two words are synonyms; 2) antonym – when two words have opposite meanings, e. g. adicio (addition) and subtraho (subtraction); 3) hypernym/hyponym – when a concept is more or less general of another, e.g. formulo de Taylor (Taylor's formula) is a hyponym of formulo; 4) holonym/meronym – when an object has another as its part, or, on the contrary, it is a part of another one, e.g. adiciato (addend) is a part of adicio, and adicio has adiciato as one of its parts; 5) relation – when two words are related, either because they occur in the same situation, or they have the same holonym or hypernym, like adicio and divido (division), both having operacio as a hypernym.

The database is stored at the URL: http://www.math.unipd.it/~minnaja/ WORDNETESP/vortaro.html where a slot is offered. Typing in the word to be sought one gets its antonyms, hypernyms, hyponyms, etc. Hypernyms and hyponyms can be nested: funkcio (function) has funkcio kontinua (continuous function) as its hyponym, which has in turn funkcio derivebla (function with derivative) as its hyponym. We added a link to definitions, formulae and more extended comments; when it deals with a name of a mathematician, the link points to his/ her biography (The Mac Tutor History of Mathematics Archive, in English).

KEY WORDS: WordNet, mathematical dictionary, Esperanto, word taxonomies.

* The authors thank Prof. F. Sullivan for helping with the English version.

1. Dictionaries in Esperanto: a short introduction

Esperanto was launched in 1887 through the first book in Russian, with title and author as follows:

Дᵖ ЭСПЕРАНТО
МЕЖДУНАРОДЫЙ
ЯЗЫКЪ

and in the same year the versions in Polish, French and German came out; the edition in English was issued in the following year. The name of the author, Dr. Esperanto, was the pseudonym of Lejzer Ludwik Zamenhof, a Polish ophthalmologist.

So, the very first Esperanto dictionary was attached to the first grammar in Russian: it was printed on a unique large page and had 917 roots with their translations. We speak about roots, because many words can be regularly formed from just one root, and the regularity of making words was, at the beginning, the main argument for the easiness of an "artificial" language, as it was called at that time, or a "planned" language, as we would say today.

At the end of 1888 the *Plena vortaro Rusa-Internacia* (Full Russian-International Dictionary) came out: it is the first dictionary *from* an ethnic language *into* the international one. The name "Esperanto" was not yet generally used: its first appearance was in some Russian announcements in the review Ķčāą in 1888 written by Zamenhof himself, where he calls his language "Всемирный Языкъ Эсперанто", in order to distinguish it from other projects also having the epithet "international". This dictionary included nearly 500 new roots, mainly of international origin; but precisely because it was from Russian into Esperanto, it remained not very widely known, and subsequent dictionaries did not incorporate all the words it listed.

Finally, in 1893 (with date of issue 1894) the so called *Universala vortaro* (Universal Dictionary) in five languages (German, French, English, Russian, Polish) proposed 1710 more roots, so at that time the whole corpus was of 2627 roots. Of course, roots in Esperanto were translated by complete words into the different ethnic languages, according to the morphological meaning of each root. This dictionary became the standard, as the first General Congress of Esperantists, which took place in 1905 at Boulogne-sur-Mer, considered it as an essential part of the fundamental texts and stated that it would not be changed. The language was to develop like an ethnic language, by way of neologisms and archaisms, but these words will never change their meanings or their spelling.

Some dictionaries devoted to specific subjects began to appear in the first years of the 20th century; the first papers on mathematics appeared in France (Bardellé 1901) and in Italy. In particular, Dr. Ceretti (1903) wrote a small Italian-Esperanto mathematical dictionary: just a few pages, but with some interesting insights. The first sufficiently rich text of mathematics fully in Esperanto appeared a couple of years later (Bricard 1905): a 60-page booklet with sections concerning arithmetic, algebra, analysis, geometry and mechanics, involving 472 specifically mathematical words.

Then, further dictionaries were issued, especially for mathematics or for general science. Some specific dictionaries for different subjects used to appear as the second part of the Yearbook of the Universal Esperanto-Association, at that time, and even now, the greatest association of the Esperanto-speaking community. In such a series, a 1031-word mathematical dictionary appeared in 1954 (Bean 1954).

A very interesting effort was made in 1980, when a mathematical lexicon in eight languages, with Esperanto as the leading one, was issued in Germany (Hilgers and Yashovardan 1980). It was a significant initiative, not so much for the number of words, just 460, but for the fact that the other seven languages were the languages of the European Community of that time: Danish, Dutch, English, French, German, Italian, Portuguese. Some words had their equivalents also in Irish, Greek, Spanish and Turkish, announcing in advance the coming of other countries into the European Community. Concise, but precise, definitions were in Esperanto.

Further dictionaries became bigger and bigger: the *Matematika Vortaro Esperanta-Ceha-Germana* (Esperanto-Czech-German Mathematical Dictionary), by Jan Werner (1990), issued by the International Academy of Sciences San Marino, has 3722 words, and it is up to now the most extensive mathematical dictionary in Esperanto. Its second edition can be downloaded for free at: rasmus.uib.no/~st01490/esperanto/JanWerner.pdf.

At present the most modern and really the best such dictionary is the *Matematika vortaro kaj oklingva leksikono* (Mathematical Dictionary and 8-languages Lexicon) by Marc Bavant (2003), a French engineer, published in the Czech Republic. It is not so rich in entries, only 1328, but many words can be found through the examples. In addition, there are several plates with the names of symbols, curves and other specific situations. The eight languages are Esperanto as the leading one, and then Czech, German, English, French, Hungarian, Polish and Russian. There are also some very interesting historical comments about previous dictionaries and the evolution of the meanings of the words.

This brief look at the mathematical dictionaries in Esperanto has considered both traditional alphabetically-ordered dictionaries and some types of small encyclopedias, were words are explained by their use in texts. Two of these dictionaries, namely those of Bricard and Bavant, also arrange a sort of hierarchy among terms.

2. WordNet and its Extensions

2.1 General principles

Let us now have a look at the way WordNet is arranged. A traditional dictionary is based on historical procedures for the classical ways to organize information: words are listed alphabetically, in order to facilitate searches. Their meanings are usually listed in order of their frequencies of use, or, less often, in order of the historical use, or, sometimes, in a nested order: the most general is registered first, and more and more particular meanings come afterwards.

Another way to organize information is to collect words in different groups according to their meanings: synonyms and similar meanings are put together; the list is no longer alphabetical, so the search must be done a different way.

An elementary way to classify words is to divide them according to their morphological categories. This looks to be easy, but it involves strong differences depending on the different languages. In English, for instance, the difference between nouns and adjectives is very weak: a lot of nouns behave like adjectives when they are put in some positions, or at least they should be translated into other languages by adjectives. In similar cases, German or Esperanto make a frequent use of compound words; French and Italian prefer the use of adjectives or development of the couple of neighboring nouns into a noun and a genitive phrase.

WordNet is a system of lexical search for English, based on up-to-date psycholinguistic theories. It has been developed at the Cognitive Science Laboratory of Princeton University (Miller *et al.* 1990). This is not the first attempt in this field, of course, several dictionaries having been constructed with this goal. The Duden Wörterbuch, for instance, is of this type, and it has the Esperanto version by Rüdiger Eichholz (1988); another good example just for Esperanto is Mariano (1980). There are also dictionaries listing both synonyms and antonyms.

Ontology is usually thought of as a philosophical discipline, far away from the world of Information Technology. In philosophy it is a branch of metaphysics and it concerns the study of being or existence and their fundamental categories.

In Information Technology an ontology is the attempt to formulate an exhaustive and conceptual scheme in a specific domain. The net and the communication through it have made the ontological aspects of information strategic. In order to organize information, the contents become fundamental. The most widely accepted definition of ontology is that of Tom Gruber (1993): "ontology is an explicit specification of a conceptualization".

We will deal now with lexical ontologies, which mark a language or more than one language, or one of its parts. A lexical ontology is independent of the domain and expresses the semantics of linguistic constructions. We would like to express two features:

- a lexical knowledge, constituted by a set of words (strings of characters);
- a semantic knowledge, collecting the word meanings and the relations between them.

WordNet fulfils these features, as it allows one not only to take in the meaning of a word, as in a dictionary, but, above all, it gives relations with other words based on their meanings.

At first glance, WordNet may look like an online dictionary, but it is really used as a lexical ontology (WordNet, without date[1]). As we said before, in a traditional dictionary words are listed alphabetically, and their meanings are listed in accordance with their frequencies. The search can get tedious. In WordNet, information is stored in accordance with both the syntactic categories and the meaning, and there is a connection through different types of hierarchical relations.

WordNet classifies words via two concepts: the Word Form, i. e., the written form, and the Word Meaning, i. e., the concept expressed. So the initial point of the WordNet classification of the words is the relations between the headword and the meaning. For instance, the word *function* can have several meanings, as *mathematical relation*, *subroutine*, *religious ceremony*, and it can be a verb, too. On the other hand, in mathematics it has some synonyms, like *application*, *map* and so on.

A Lexical Matrix can be constructed this way: the meanings are put in the rows and the headwords in the columns. When a word matches a meaning, the square of the matrix is filled; when there is no match, it remains empty. Squares filled in the same column indicate that a word has some different meanings; squares filled in the same row indicate that a concept can be expressed by various different words. In this last case, the words are collected in a unique family, called "synonym set", or, briefly, *synset*.

2.2 Extensions for WordNet: Multilingualism and Specific Domain

A linguistic ontology should have all meanings arising from the different fields of knowledge, and this is very complex. Extensions are of two types: multilingualism and domain-specific.

2.2.1 Multilingualism

The project "EuroWordNet" (Vossen 2001), financed by the European Community (1996-1999), aimed at constructing a consistent and reliable multilingual lexical database and to maintain the diversity and richness of the different languages.

There were nine languages involved in this project: Dutch, English, Italian and Spanish from its beginning, with Czech, Estonian, French, German and Swedish being added in 1998. Big databases were established, nearly 30.000 concepts and 50.000 word senses. Such databases have been very useful for subsequent dictionaries made outside the program itself; e. g., the biggest Lithuanian-English dictionary is based on the WordNet database.

[1] The most recent Windows version is 2.1, released in March, 2005. Version 3.0 for Unix/Linux/Solaris was released in December, 2006.

Each database is structured following the same guidelines as Wordnet, i. e., synonyms are gathered in synsets, which are associated to the synsets of the Princeton WordNet, in order to create a multilingual database.

The equivalence relations among synsets of different languages are realized by a structure called an Inter-Lingual-Index (ILI). An ILI uses essentially the English WordNet with the addition of specific concepts from other languages. In Diagram 1 we have an example of the architecture, dealing with the word "drive". This word in English is compared with Spanish, Dutch and Italian words having the same meanings. As the diagram shows, the word is linked to the general ontology about traffic, which is independent of the language. The ethnic languages refer, in turn, to the ILI.

ILI is in English, but its function is just for the connection among the different databases: further extensions of the database use ILI as a set of concepts.

The Top-Concept Ontology is a common structure for the most important concepts of all wordnets in the different languages. It consists of 63 basic semantic groups and 1610 fundamental concepts common to all languages.

The Domain Ontology is important in order to separate specific words from generic ones; this is important for solving problems of ambiguity.

Sometimes a concept does not exist in a specific language, while it exists in another. For instance the Italian word *dito* and the Spanish word *dedo* have both the meaning of *fingers* and *toes*, having a more general meaning than the corresponding English word. The inverse happens for

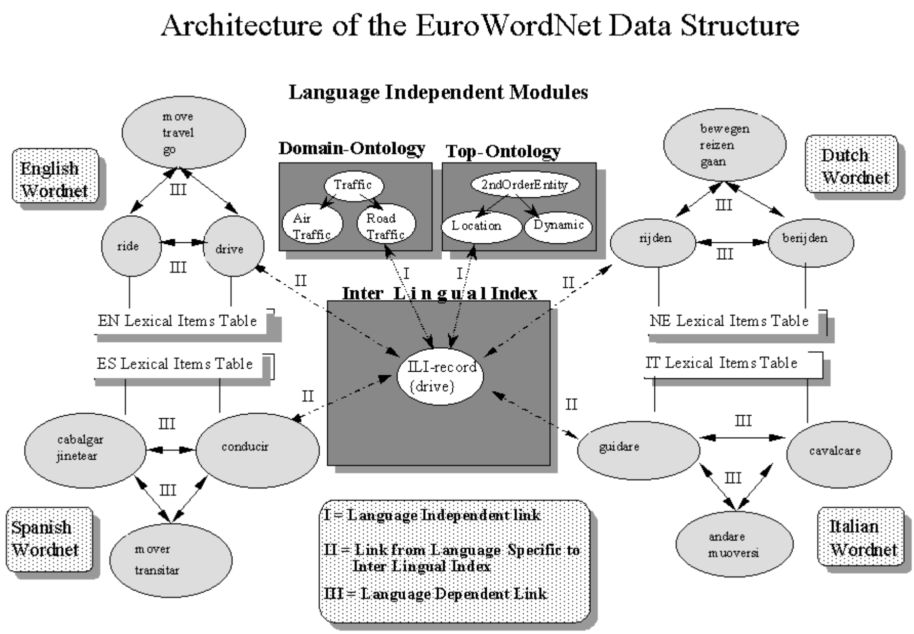

Diagram 1. Scheme of EuroWordNet for word "drive"[2]

[2] The diagram is taken from http://www-ksl.stanford.edu/onto-std/eurowordnet.pdf

Inter-Lingual-Index
Unstructured Superset of Concepts

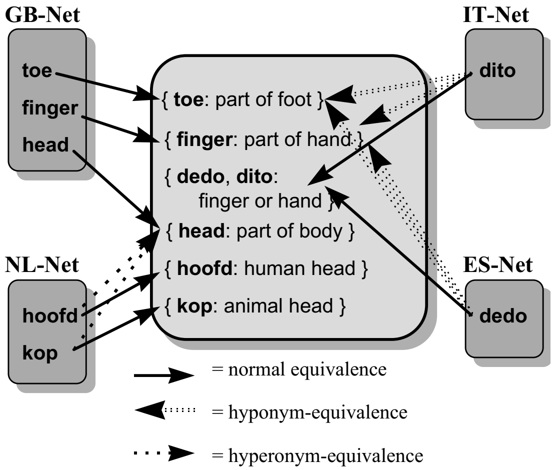

Diagram 2. Record for a new word in ILI[3]

the Dutch words *kop* and *hoofd*, which are more specific than the English *head*. In both cases the relation will be realized by adding a new (non-English) record to ILI (look at Diagram 2).

Multilingualism plays an important role in this case, as different languages do not behave in the same way; feelings are not interpreted with the same sense, and often a word in one language needs several terms to be expressed in another language. Although for mathematics this phenomenon is reduced to the minimum, it is, nevertheless, present.

2.2.2 Specific Domains

Generic lexical ontologies have lexical databases containing a knowledge level without specific coverage. Specific databases gather specific concepts, and provide lexical hierarchies. The synsets collected this way have the form of compound words and the co-operation of an expert is crucial in order to establish the right connections among them.

Several specific lexical dictionaries have been developed, mostly for English which now has dictionaries for medicine, art, architecture and geography. Some specific databases also exist for other languages. For Italian, e. g., specific databases have been implemented for the maritime domain and for economics, architecture and law; in particular, a mathematical database has been developed at the Department of Pure and Applied Mathematics of the University of Padova (Giunta, Minnaja and Paccagnella 2005).

3. Architetcture of Wordnet

3.1 Hierarchical and semantic relations

WordNet was born in 1985. The most evident difference between WordNet and an ordinary dictionary is that the former divides the lexicon into five syntactic categories: nouns, verbs, adjectives,

[3] The diagram is taken from http://www.ercim.org/publication/ws-proceedings/DELOS3/Vossen.pdf

adverbs and function words. At present WordNet does not deal with the last category. Nouns are arranged in a lexical memory as hierarchies (hypernyms/hyponyms); verbs are structured in a hierarchy of troponymy (a verb is a troponym of another when it expresses a specific action in comparison of a more general one, e.g. to walk/to move himself/herself). Adjectives, in turn, are arranged as *n*-dimensional hyperspaces.

WordNet is based on the semantic relations among concepts; among these relations, synonymy plays a fundamental role, but it is not the only one employed. We will deal with the following main classification:

– Synonymy,
– Antonymy,
– Hyponymy/hypernymy,
– Meronymy/holonymy,
– Morphological relations.

Synonymy has a very rigid definition (attributed to Leibniz):

> Two concepts are synonymic if the truth value of the sentence does not change when one concept is substituted by the other.

According to this definition, synonyms would be very rare. So a weaker definition is used, not connected to the sentence, but to the relevant context:

> Two concepts are synonymic in a linguistic context C if the substitution of one concept by the other in the linguistic context C does not change its truth value.

With this definition, *function* and *map* are synonyms in mathematics, but not in cartography.

Antonymy is a type of relation, which cannot be defined easily and in the same way for all situations, but it is quite common. Usually the antonym of a word *x* is defined as *not-x*. For instance *rich* and *poor* are usually classed as antonyms, although a person can be not rich, but also not poor.

Hyponymy/Hypernymy is a relation defined as follows:

> a concept represented by the synset {x1, x2, x3, …} is called hyponym of the concept represented by the synset {y1, y2, y3, …} if we can accept a sentence constructed this way: An x is a type of y.

Hyponymy is transitive and antisymmetric; it generates a hierarchical semantic structure where hyponyms (concept: *son*) are under their hypernyms (concept: *father*).

Meronymy/Holonymy is a semantic relation, and expresses the concept of *part of*:

> {x1, x2, x3, …} is a meronym of a concept expressed by {y1, y2, y3, …} if we can accept sentences like x is a part of y

The meronymy/holonymy relation is transitive (with the reservations we will explain later) and antisymmetric, and it can be used to construct hierarchical relations. E. g. *beak* and *wing* are meronyms of *bird*, *eagle* is a hyponym of *bird*. So, because of the hierarchical system, *beak* and *wing* are themselves meronyms of *eagle*.

3.2 Developing WordNet

WordNet uses some symbols in its coding; they are pointers which are recognized by the program, which will organize the words in a chart. We list some of them, not all of which will be useful in mathematics.

The symbols for nouns are:

!	Antonym
@	Hypernym
@i	Instance hypernym
~	Hyponym
i	Instance hyponym
#m	Member holonym
#s	Substance holonym
#p	Part holonym
%m	Member meronym
%s	Substance meronym
%p	Part meronym
=	Attribute
+	Derivationally related form

;c Domain of synset – TOPIC
-c Member of this domain – TOPIC
;r Domain of synset – REGION
-r Member of this domain – REGION
;u Domain of synset – USAGE
-u Member of this domain – USAGE

For instance, let us look how the word "Lithuania" is treated:

Lithuania

Noun

1. a republic in northeastern Europe on the Baltic Sea

(synonym) Republic of Lithuania, Lietuva
(hypernym) Baltic State, Baltic Republic
(member-meronym) Lithuanian
(part-meronym) Klaipeda, Memel

As is easy to guess, the category "substance meronym" is useless in mathematics, as well as "substance holonym". Also the different types of domain are not meaningful in our field.

We do not provide the complete list of the pointers for other morphological categories; here we mention just a few of them: for verbs, an interesting relation is the one indicating cause, for adjectives, pertaining to a noun or being a participle have a specific pointer, as well as the relation of similarity; the derivation from an adjective is pointed out for an adverb.

Many pointers have the reflexivity property, so in the following list we find the relation at left and its reflected one at right, and vice versa.

Antonym – Antonym,
Hyponym – Hypernym,
Hypernym – Hyponym,
Holonym – Meronym,
Instance Hyponym – Instance Hypernym,
Instance Hypernym – Instance Hyponym,
Meronym – Holonym,
Similar to – Similar to,
Attribute – Attribute,
Verb Group – Verb Group,
Derivationally Related – Derivationally Related,
Domain of synset – Member of Domain.

The *source files* have to contain the lexical database, which is the kernel of WordNet, and can be increased with time. WordNet works associating four characteristics to a single word:

- its orthographic form,
- the syntactic category (noun, adjective, verb, adverb),
- the semantic field,
- the sense number, i. e. the number identifying the meaning.

WordNet arranges the words following a deep semantic and morphological analysis, dividing nouns from adjectives and verbs; nouns and verbs are, in turn, divided into semantic fields. Adjectives are divided into two categories: descriptive adjectives and relational adjectives. "Glossa" is a commentary, usually the definition or an example sentence, and it must be included in any synset. In the source file every synset occupies one line. The general syntax is as follows:

{word pointer (glossa)}

Every synset must have at least one headword and the *glossa*; our mathematical WordNet has for every headword at least the semantic field, too.

4. A mathematical database for Esperanto

The aim to create a lexical resource for mathematics which does not give only definitions led us to choose the principles of WordNet as a base for the construction of the dictionary. In addition, such a scheme allows interfacing it with other databases. The mathematical WordNet deals with nouns and has the semantic fields.

A Specific Domain for mathematics has never been established and usually mathematical dictionaries have no taxonomies.

In the present work we show both the most recent changes to the structure of the database and the Esperanto version of the mathematical WordNet. In order to make it easy to enrich the database we have introduced some new structures and organized a new part, which deals with biographies of mathematicians and formulae, and which has no analogue in other specific dictionaries. In particular, nothing for Esperanto was done in WordNet.

4.1 The source files

Our mathematical WordNet deals only with nouns and organizes them by semantic fields; for instance *bazo* (base) has different meanings in geometry, arithmetic and linear algebra. It is founded on a source file, which is a text file containing a set of strings consisting of four fields:

- identifier of the entry (the headword itself),
- identifier of the subject (e.g. geometry, functional analysis, probability, …),
- pointer to semantic relations,
- identifier of the definition.

The identifier of the entry deals only with singular nouns, or compound nouns joined by an underscore character. The subject indicates which branch of mathematics is involved; if the word is involved in more than one branch, the main branch is indicated, and the other fields can be found in the definition. The pointer to semantic relations is essential in the construction of the structure; the symbols and the taxonomy are those of WordNet. We made just a couple of changes: in our mathematical WordNet the synonymy relation is marked by symbol S; actually, there are few synonyms in mathematics, but they are not completely absent. In addition, we introduced a new relation R with the meaning "in relation with". Other relations have the same pointer symbols as WordNet. Finally, the identifier of the definition is the *glossa*.

The subject identifies the concept uniquely. The classification of the mathematical subjects follows the Dewey classification used in libraries. The subjects are:

#1 Mathematics (in general) and Biographies,
#2 Algebra,
#3 Arithmetic,
#4 Topology,
#5 Analysis,
#6 Geometry,
#9 Applied mathematics.

The scheme is hierarchical, so, for instance, we indicated algebra (general) by #2, rings, groups and modules by #21, algebraic equations, logarithms, exponentials by #23; general analysis is indicated by 5, real analysis by #51, theory of functions by #55; applied mathematics is indicated by #9, operational research by #96 and so on.

This dictionary is for students and for novices of mathematics, so the *glossa* plays a crucial role. Therefore we created a file devoted just to the *glossa*, so it can be enriched without burdening the other source files.

Every line of the noun file is structured this way:

{headword identifier_subject [f1,pointer] [f2,pointer],…, [fn, pointer] (identifier_glossa)}

where the square brackets indicate that the fields included are optional.

The variables in the line can take the following values:

- *headword* is the form of the word to which the synset represented by the line is related, and it is its main key;
- *identifier_subject* is given by the key of the subject to which the headword refers; it has the numerical value of the mathematical subject according to the classification mentioned above;
- the pairs *fi,pointer* are constituted as follows: *fi* is another headword, *pointer* is a symbol which connects *fi* to the main headword and is one of the symbols mentioned in Section 3;
- *identifier_glossa* is the headword itself and points to the corresponding definition, which can be enriched by examples, formulae and comments.

In mathematics some relations do not happen; e. g., distinctions between part meronyms, substance meronyms and member meronyms are meaningless; so pointers have been restricted to the following types:

! Antonym
@ Hypernym
~ Hyponym
#p Part holonym
%p Part meronym
S Synonym
R Relation

Our *identifier_glossa* is merely a string included between round brackets. It points to another file, apart from the main source file of the nouns, so it can be managed autonomously. This *glossa* file describes the headword itself: it is a definition, or an explanation, or a formula with a comment or an example, or a complete biography when it deals with a mathematician. Everything is in Esperanto, except the last case, where the biography is obtained by a link to a specific URL in English.

4.2 Examples

Let us show, as an example, how the source file with headword *operacio* (operation) appears:

{operacio 2 operacio_binara,~ operacio_ekstera,~ operacio_interna,~ inversigo,~ adicio,~ subtraho,~ multipliko,~ divido,~ konjugo,~ radikado,~ potencigo,~ transpono,~ malvastigo,~ kunligo,~ kunfaldo,~ logaritmo,~ derivado,~ integrado,~ operatoro,R elemento_neutra,R operaciato,%p rezulto,%p strukturo_algebra,R (operacio)}

In the *glossa* the definition, or a useful comment, or a biography are registered.

The programs operating on these files have been created and improved over the years (Croin 2004, Giunta, Minnaja and Paccagnella 2005, Niero 2007).

The word search must be done at the following URL: http://www.math.unipd.it/~minnaja/WORDNETESP/vortaro.html

A screen opens with directions (in Esperanto) for the word search, and with a slot to insert the word. If a full expression is searched, the single words must be joined into only one word by use of the underscore character, and the adjective follows the noun. The spelling used for the word being sought in Esperanto is the x-system, and the breve over "u" must be omitted. Clicking on "sercxu" (search), one gets the result.

For instance, let us apply the procedure to the example above. We type "*operacio*" in the slot and we get the screen for "*vorto: operacio*" (word: operation) with the following result (*"rezulto"*):

Vorto: operacio
Rezulto:

TEMO: Algebro (gxenerala)
DIFINO
regulo au aro da reguloj asociantaj al n elementoj donitaj en difinita ordo alian elementon (au aliajn elementojn) nomata(j)n rezultato de la operacio
HIPONIMOJ DE operacio

- operacio_binara
- operacio_ekstera
- operacio_interna
- inversigo
- adicio
- subtraho
- multipliko
- divido
- konjugo
- radikado
- potencigo
- transpono
- malvastigo
- kunligo
- kunfaldo
- logaritmo
    - logaritmo_dekuma
    - logaritmo_de_Briggs
    - ogaritmo_natura
    - logaritmo_kompleksa
    - ogaritmo_nepera
    - modulo_de_logaritmo
- derivado
- integrado

HIPERNIMOJ DE operacio
Neniu elemento
MERONIMOJ DE operacio
    operaciato
    rezulto
HOLONIMOJ DE operacio
Neniu elemento
SINONIMOJ DE operacio
Neniu elemento
ANTONIMOJ DE operacio
Neniu elemento
EN RILATO KUN operacio
operatoro
elemento_neutra
strukturo_algebra

As is easily seen, the definition comes at the beginning; then the various relations to the headword (hyponyms, hypernyms, meronyms, holonyms, synonyms, antonyms and relation in general) are represented by nests, and every entry of the nest has, in turn, its synset. When no entry matches a specific relation, the result offers: *neniu elemento* (no elements). We remark that a hyponym (son) of *operacio* is *adicio* (addition), and if we go back to the URL and type in *adicio* we find that *operacio* is its hypernym (father):

Vorto: adicio
Rezulto:

TEMO: Aritmetiko
DIFINO
operacio kiu aldonas kvantojn unu al la alia kaj kunigas ilin en unu solan, nomatan sumo
HIPONIMOJ DE adicio
Neniu elemento
HIPERNIMOJ DE adicio
    operacio
MERONIMOJ DE adicio
    adiciato
    sumo
HOLONIMOJ DE adicio
Neniu elemento
SINONIMOJ DE adicio
Neniu elemento
ANTONIMOJ DE adicio
subtraho
EN RILATO KUN adicio
multipliko
divido

Another example is *funkcio* (function), which has several hyponyms and hypernyms:

Vorto: funkcio
Rezulto:

TEMO: Analitiko (gxenerala)
DIFINO
bildigo, precipe kiam la fonta kaj cela aroj konsistas el nombroj;
vd. ankau korespondo
HIPONIMOJ DE funkcio
    funkcio_abela
    funkcio_barita
    funkcio_kontinua
    funkcio_derivebla
    funkcio_de_klaso_Cn
    funkcio_holomorfa
    funkcio_elvolvebla_en_serion_de_Taylor
    funkcio_analitika
    funkcio_de_Bessel
    funkcio_de_Diricxlet
    funkcio_fucxsia
    funkcio_integralebla
    funkcio_integrala
    funkcio_kunligita
    funkcio_meromorfa
    funkcio_monotona
    funkcio_kreskanta
        funkcio_kreskanta_strikte
    funkcio_malkreskanta
        funkcio_malkreskanta_strikte
    funkcio_perioda
    funkcio_trigonometria
        sinuso
        kosinuso
        tangento
    funkcio_inversa
    funkcio_plurvalora
    funkcio_unuvalora
HIPERNIMOJ DE funkcio
    rilato_binara
    rilato
MERONIMOJ DE funkcio
    argumento
    celo-aro
    fonto-aro
    valoro

HOLONIMOJ DE funkcio
Neniu elemento
SINONIMOJ DE funkcio
apliko
korespondo
bildigo
transformo
ANTONIMOJ DE funkcio
Neniu elemento
EN RILATO KUN funkcio
vastigo
malvastigo

As seen in the previous example, in the noun

"funkcio_elvolvebla_en_serion_de_Taylor"
(function which can be expanded
in Taylor's series)

the name of a mathematician appears. At the School of Mathematics and Statistics of the University of St. Andrews (Scotland) an URL collecting biographies of mathematicians (in English) is available on line: The MacTutor History of Mathematics Archive (http://www-gap.dcs.st-and.ac.uk/~history/BiogIndex.html). The Mathematical WordNet has been extended to these mathematical biographies. When looking at the word "Taylor" we have the following result, where the link points to the biography of Taylor:

Vorto: Taylor
Rezulto:

TEMO: Gxenerala matematiko kaj Biografioj
DIFINO
Brook Taylor 1685-1731, anglo
HIPONIMOJ DE Taylor
funkcio_elvolvebla_en_serion_de_Taylor
funkcio_analitika
formulo_de_Taylor
HIPERNIMOJ DE Taylor
Neniu elemento
MERONIMOJ DE Taylor
Neniu elemento
HOLONIMOJ DE Taylor
Neniu elemento
SINONIMOJ DE Taylor
Neniu elemento
ANTONIMOJ DE Taylor
Neniu elemento
EN RILATO KUN Taylor
Neniu elemento

If we go back to the URL and search for "formulo_de_Taylor" we get the following result:

Vorto: formulo_de_Taylor
Rezulto:

TEMO: Kompleksa analitiko
DIFINO
formulo; gxi estas utila por aproksimi funkcion havantan derivajxojn en cxirkauajxo de punkto pere de polinomo
HIPONIMOJ DE formulo_de_Taylor
Neniu elemento
HIPERNIMOJ DE formulo_de_Taylor
formulo
Taylor
MERONIMOJ DE formulo_de_Taylor
resto_de_formulo
HOLONIMOJ DE formulo_de_Taylor
Neniu elemento
SINONIMOJ DE formulo_de_Taylor
Neniu elemento
ANTONIMOJ DE formulo_de_Taylor
Neniu elemento
EN RILATO KUN formulo_de_Taylor
formulo_de_MacLaurin
serio_de_Taylor

where "formulo" is a link for a definition of the formula. A click on it gives the following text, which is an extended definition with a short comment:

This mathematical WordNet can be easily extended with new terms and new morphological categories.

Ĉe funkcio $f(x)$ havanta derivaĵojn sinsekvajn oni nomas *formulo de Taylor* la egalaĵon

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \frac{h^3}{3!}f'''(x) + \ldots + \frac{h^{n-1}}{(n-1)!}f^{n-1}(x) + R_n(h)$$

kiun oni povas skribi pli kompakte

$$f(x+h) = \sum_{i=0}^{n-1} \frac{h^i}{i!} f^{(i)}(x) + R_n(h)$$

kie la 0-orda derivaĵo koincidas kun la funkcio mem, kaj laŭ konvencio oni metas $0! = 1$. La lasta termo nomiĝas *resto de Taylor* kaj ĝi estas infinitezimo kun h de ordo pli granda ol $n-1$.

References

BARDELLÉ, CH., 1901. L'Esperanto et les Mathématiciens. *L'Enseignement mathématique*, 444-445.

BAVANT, M., 2003. *Matematika vortaro kaj oklingva leksikono*. Dobųichovice: Kava-Pech.

BEAN, C. M., 1954. Matematika terminaro. *In: Jarlibro – Dua parto*. Rotterdam: UEA, 39-77.

BRICARD, R., 1905. *Matematika terminaro kaj krestomatio*. Paris: Hachette.

CERETTI, U., 1903. Matematica ed Esperanto. *Rivista di Fisica, Matematica e Scienze Naturali*, IV, 45, Sett. 1903, 1-16.

CROIN, S., 2004. *Sviluppo di un dizionario matematico nella struttura di WordNet* (Development of a mathematical dictionary in the WordNet structure). Thesis (MA). Univ. Padova, Engineering, Academic Year 2002-2003.

EICHHOLZ, R., 1988. *Esperanta Bildvortaro*. Mannheim: Bibliographisches Institut.

HILGERS, R.; YASHOVARDHAN, Hrsg., 1980. *EG-Wörterbuch mathematischer Begriffe*. Alsbach: Leuchtturm.

GIUNTA, A.; MINNAJA, C.; PACCAGNELLA L. G., 2005. Extending the Italian WordNet with the Specialized Language of the Mathematical Domain. *Grundlagenstudien aus Kibernetik und Geisteswissenschaft*, 46, 3: 3-12.

GRUBER, T. R., 1993. A Translation Approach to Portable Ontologies. *Knowledge Acquisition*, 5(2): 199-220.

MARIANO, M., 1980. *Esperanto-Vortaro per temo-tabeloj* (Esperanto-dictionary by subject tables). Ascoli Piceno: Gabrielli.

MILLER, G. A.; BECKWITH, R.; FELLBAUM, C.; GROSS, D.; MILLER, K., 1990. WordNet: an On-line Lexical Database. *International Journal of Lexicography (special issue)*, 3 (4): 235-312.

NIERO, F., 2007. *WordNet e sue applicazioni. Revisione e implementazione di un database di termini matematici* (WordNet and its applications. Revision and implementation of a database of mathematical terms). Thesis (MA). Univ. Padova, Engineering, Academic Year 2005-2006.

VOSSEN, P. ed., 2001. *EuroWordNet* [online]. Available from: http://www.illc.uva.nl/Euro WordNet/ [Accessed 27 May 2007].

WERNER, J., 1990. *Matematika Vortaro Esperanta-Ceha-Germana*, Poprad: Akademio Internacia de la Sciencoj (AIS) San Marino.

Wordnet. Available from: http://www.cogsci.princeton.edu/~wn/ [Accessed 27 May 2007].

Laura Gilda Paccagnella
Padujos universitetas, Italija
Moksliniai interesai: natūralių kalbų apdorojimas, matematinės duomenų bazės.
Carlo Minnaja
Padujos universitetas, Italija
Moksliniai interesai: matematikos istorija, matematinių žodžių etimologija, dirbtinės kalbos.

MATEMATIKOS TERMINŲ ŽODYNAS ESPERANTO KALBA PAGAL WORDNET

Santrauka

WordNet yra duomenų bazė, veikianti nuo 1985 metų Princetono universitete. Joje žodžio paieška grindžiama psicholingvistinėmis teorijomis. Bazėje WordNet informacija yra saugoma pagal sintaksines kategorijas ir yra sujungta įvairiais santykių tipais. Yra ir sinonimų, ir poliseminių žodžių; sinonimai yra skirstomi į grupes, vadinamas sinonimų eilėmis (*synsets*). Visi konceptai duomenų bazėje yra hierarchiškai struktūruoti. Vėliau buvo sukurta duomenų bazė EuroWordNet. Tai daugiakalbis WordNet, parengtas devyniomis Europos kalbomis. Be to, jis buvo papildytas tam tikromis sritimis, kad atsirastų skirtingi WordNet, po vieną kiekvienai kalbai ir dalykui. Mes sukūrėme matematikos terminų duomenų bazę WorNet esperanto kalba. Bazėje šiuo metu yra tik daiktavardžiai, o santykiai tarp žodžių yra klasifikuojami taip: 1) sinonimas – kai du žodžiai yra sinonimai; 2) antonimas – kai du žodžiai turi priešingas reikšmes, pvz., *adicio* (*sudėtis*) ir *sutraho* (*atimtis*); 3) hiperonimas/ hiponimas – kai konceptas yra daugiau ar mažiau bendresnis nei kitas, pvz., *formulo de Taylor* (*Tayloro formulė*) yra *formulo* hiponimas; 4) holonimas/ meronimas – kai vienas objektas turi savyje kito objekto dalį ar, atvirkščiai, jis yra kito objekto dalis, pvz., *adiciato*

(*dėmuo*) yra dalis *adicio,* o *adicio* turi *adiciato* kaip vieną iš savo dalių; 5) santykis – kai du žodžiai yra susiję tuo, kad jie yra toje pačioje situacijoje, ar tuo, kad jie turi tą patį holonimą ar hiperonimą, pvz., *adicio* ir *divido* (*dalinimas*) hiperonimas yra *operacio*. Šios duomenų bazės prieiga: http://www.math.unipd.it/~minnaja/ WORDNETESP/vortaro.html. Įrašius norimą žodį, jums bus pateikti antonimai, hiperonimai, hiponimai ir kt. Hiperonimai ir hiponimai gali būti įterpti – *funkcio* (*funkcija*) hiponimas yra *funkcio kontinua* (*tęstinė funkcija*), o šio hiponimo hiponimas yra *funkcio derivebla* (*išvestinė funkcija*). Visa tai mes papildėme nuoroda į definicijas, formules ir išsamius komentarus. O tais atvejais, kai užklausa siejama su matematikų vardais, pateikiama nuoroda į jų biografiją (The Mac Tutor History of Mathematics Archive, anglų kalba).

REIKŠMINIAI ŽODŽIAI: WordNet, matematikos terminų žodynas, esperanto, žodžio taksonomijos.

Gauta 2007 11 12
Priimta publikuoti 2008 04 09