A Taxonomic Classification of WordNet Polysemy Types

Abed Alhakim Freihat
Qatar Computing Research Institute
Doha, Qatar
afrihat@qf.org.qa

Fausto Giunchiglia
University of Trento
Trento, Italy
fausto@disi.unitn.it

Biswanath Dutta
Indian Statistical Institute (ISI)
Bangalore, India
bisu@drtc.isibang.ac.in

Abstract

WordNet represents polysemous terms by capturing the different meanings of these terms at the lexical level, but without giving emphasis on the polysemy types such terms belong to. The state of the art polysemy approaches identify several polysemy types in WordNet but they do not explain how to classify and organize them. In this paper, we present a novel approach for classifying the polysemy types which exploits taxonomic principles which in turn, allow us to discover a set of polysemy structural patterns.

1 Introduction

Polysemy in WordNet (Miller, 1995) corresponds to various kinds of linguistic phenomena and can be grouped into various polysemy types (Falkum, 2011). Although WordNet was inspired by psycholinguistic and semantic principles (Miller et al., 1990), its conceptual dictionary puts greater emphasis on the lexical level rather than on the semantic one (Dolan, 1994). Lexicalizing polysemous terms without any further information about their polysemy type affects the usability of WordNet as a knowledge resource for semantic applications (Mandala et al., 1999).

In general, the state of the art approaches suggests different solutions to the polysemy problem. The most prosperous among these approaches are the regular/systematic polysemy approaches such as (Buitelaar, 1998) (Barque and Chaumartin, 2009) (Veale, 2004) (Peters, 2004). These approaches propose the semantic regularity as a basis for classification of the polysemy classes and offer different solutions that commensurate the nature of the discovered polysemy types.

Despite the diversity and depth of the state of the art solutions, no or very little attention has been given, so far, to the principles or rules used to identify polysemy types. In fact, none of these approaches can explain how to identify the polysemy types of the discovered polysemy structural patterns or how to differentiate for example, between homonymy and metaphoric structural patterns. Although Apersejan's semantic similarity criterion (Apresjan, 1974) can be used to account for regularity in polysemy, it can not predict the polysemy type of the regular polysemy types in WordNet. Our hypotheses in this paper is that identifying and differentiating between the polysemy types of the regular polysemy structural patterns requires understanding the hierarchical structure of WordNet and, thus, the criteria related to the taxonomic principles that the hierarchical structure of WordNet comply with or violates. In this paper, we show how to use two taxonomic principles as criteria for identifying the polysemy types in WordNet. Based on these principles, we introduce a semi automatic method for discovering and identifying three polysemy types in WordNet.

The paper is organized as follows. In Section two, we discuss the problem. In Section three, we introduce the formal definitions we use. In Section four, we discuss the taxonomic principles that we use to discover three of the polysemy types in WordNet. In Section five, we give an overview of our approach. In Section six, we show how to use the taxonomic principles to identify metaphoric structural patterns. In Section seven, we demonstrate how to determine specialization polysemy structural patterns. In Section eight, we describe how to discover homonymy structural patterns. In Section nine, we explain how to handle false positives in the structural patterns. In Section ten, we present the results of our approach. In Section eleven, we conclude the paper and depict our future work.

2 Problem Statement

WordNet is a machine readable online lexical database for the English language. Based on psycholinguistic principles, WordNet has been developing since 1985, by linguists and psycholinguists as a conceptual dictionary rather than an alphabetic one (Miller et al., 1990). Since that time, several versions of WordNet have been developed. In this paper, we are concerned with WordNet 2.1. WordNet 2.1. contains 147,257 words, 117,597 synsets and 207,019 word-sense pairs. The number of polysemous words in WordNet is 27,006, where 15776 are nouns.

In this paper, we deal with polysemous nouns at the concept level only. We do not consider polysemy at the instance level. After removing the polysemous nouns that refer to proper names, the remaining polysemous nouns are 14530 nouns.

WordNet does not differentiate between the types of the polysemous terms and it does not contain any information in terms of polysemy relations that can be conducted to determine the polysemy type between the synsets of a polysemous term. The researchers who attached the polysemy problem in WordNet gave different descriptions for the polysemy types in WordNet. For example, polysemy reduction approaches (Edmonds and Agirre, 2008) (Mihalcea R., 2001) (Gonzalo J., 2000) differentiate between contrastive polysemy and complementary polysemy. Regular polysemy approaches such as (Barque and Chaumartin, 2009) (Veale, 2004) (Peters, 2004) (Freihat et al., 2013) (Lohk et al., 2014) give more refined classification of the polysemy types into metonymy, metaphoric, specialization polysemy, and homonymy. In one of our recent papers, *compound noun polysemy* is introduced as a new polysemy type beside the former four polysemy types in WordNet (Freihat et al., 2015).

So far, no polysemy reduction approaches have introduced a mechanism for classifying the polysemy types into contrastive and complementary. Instead, these approaches adopt semantic and probabilistic rules to discover redundant and/or very fine grained senses. On the other hand, the regular polysemy approaches embrace a clear definition for classifying polysemous terms into regular and non regular polysemy (Apresjan, 1974). Although, the definition of regular polysemy in these approaches is useful to distinguish between regular and non regular polysemy, these approaches do not reveal the principles or the criteria used to classify polysemous terms into polysemy types.

In this paper, we explain how to use the exclusiveness property and the collectively exhaustiveness property (Bailey, 1994) (Marradi, 1990) for identifying the following polysemy types.

1 **Metaphoric polysemy:** Refers to the polysemy instances in which a term has literal and figurative meanings (Evans and Zinken, 2006). In the following example, the first meaning of the term `fox` is the literal meaning and the second meaning is the figurative.
#1 `fox: alert carnivorous mammal.`
#2 `dodger, fox, slyboots: a shifty deceptive person.`

2 **Specialization polysemy:** A type of related polysemy which denotes a hierarchical relation between the meanings of a polysemous term. In the case of abstract meanings, we say that a meaning A is a more general meaning of a meaning B. We may also use the taxonomic notations type and subtype instead of more general meaning and more specific meaning respectively. For example, we say that the first meaning of `turtledove` is a subtype of the second meaning.
#1 `australian turtledove, turtledove: small Australian dove.`
#2 `turtledove: any of several Old World wild doves.`

3 **Homonymy:** Refers to the contrastive polysemy instances, where meanings are not related. Consider for example the following polysemy instance of the term `bank`.
#1 `depository financial institution, bank: a financial institution.`
#2 `bank: sloping land (especially the slope beside a body of water).`

3 Approach Notations

We begin with the basic notations. Lemma is the basic lexical unit in WordNet that refers to the base form of a word or a collocation. Based on this definition, we define a natural language term or simply a term as a lemma that belongs to a grammatical category; i.e., noun, verb, adjective or adverb.

Definition 1 *(Term).*

A term T is a quadruple ⟨Lemma, Cat⟩, where
a) Lemma is the term lemma;
b) Cat is the grammatical category of the term.

Synset is the fundamental structure in WordNet that we define as follow.

Definition 2 *(WordNet synset).*

A synset S is defined as ⟨Cat, Terms, Gloss, Relations⟩, where
a) Cat is the grammatical category of the synset;
b) Terms is an ordered list of synonymous terms that have the same grammatical category Cat;
c) Gloss is a text that describes the synset;
d) Relations is a set of semantic relations that hold between synsets.

Now, we move to the hierarchical structure of WordNet. WordNet uses the relation direct hypernym to organize the hierarchical relations between the synsets. This relation denotes the superordinate relationship between synsets. For example, the relation direct hypernym holds between `vehicle` and `wheeled vehicle` where `vehicle` is hypernym of `wheeled vehicle`. The direct hypernym relation is transitive. In the following, we generalize the direct hypernym relation to reflect the transitivity property, where we use the notion hypernym instead of a direct hypernym.

Definition 3 *(hypernym relation).*

For two synsets s and $s^{'}$, s is a hypernym of $s^{'}$, if the following holds: s is a direct hypernym of $s^{'}$, or there exists a synsets $s^{''}$ such that s is a direct hypernym of $s^{''}$ and $s^{''}$ is a hypernym of $s^{'}$.
For example, `vehicle` is a hypernym of `car`, because `vehicle` is direct hypernym of `wheeled vehicle` and `wheeled vehicle` is a direct hypernym of `car`.
We use the following symbols to denote direct hypernym/hypernym relations:
a) $s \prec s^{'}$ if s is a direct hypernym of $s^{'}$
c) $s \prec^{*} s^{'}$ if s is a hypernym of $s^{'}$

Using the direct hypernym relation, wordNet organizes noun-synsets in a hierarchy that we define as follows.

Definition 4 *(wordNet hierarchy).*

Let $S = \{s_1, s_2, ..., s_n\}$ be the set of noun-synsets in WordNet. WordNet hierarchy is defined as a connected and rooted digraph $\langle S, E \rangle$, where
a) `entity` $\in S$ is the single root of the hierarchy;
b) $E \subseteq S \times S$;
c) $(s_1, s_2) \in E$ if $s_1 \prec s_2$;
d) For any synset s $\neq$ `entity`, there exists at least one synset $s^{'}$ such that $s^{'} \prec s$.

In this definition, point (a) defines the single root of the hierarchy and point (d) defines the connectivity property in the hierarchy.
We move now to the semantics of WordNet. We define the subset of the semantics of WordNet hierarchy that is relevant for our approach. A full definition of the WordNet semantics is described in approaches such as (Alvarez, 2000) (Rudolph, 2011) (Breaux et al., 2009).
We define the semantics of WordNet using an Interpretation $I = \langle \Delta^I, f \rangle$, where Δ^I is an non empty set (the domain of interpretation) and f is an interpretation function.

Definition 5 *(Semantics of WordNet Hierarchy).*

Let $WH = \langle S, E \rangle$ be wordNet hierarchy. We define an Interpretation of WH, $I = \langle \Delta^I, f \rangle$ as follows:
a) $entity^I = \Delta^I$;
b) $\perp^I = \varnothing$;
c) $\forall s \in S$: $s^I \subset \Delta^I$;
d) $(s_1 \sqcap s_2)^I = s_1^I \cap s_2^I$;
e) $(s_1 \sqcup s_2)^I = s_1^I \cup s_2^I$;
f) $s_1 \sqsubseteq s_2$ if $s_1^I \subseteq s_2^I$.

In points a) and b), we define the empty and universal concepts. Point c) states that Δ^I is closed under the interpretation function f. In and d) and e), we define the conjunction and disjunction operations. In f), we define the subsumption relation.

We present now the polysemy notations. A term is polysemous if it is found in the terms of more than one synset. A synset is polysemous if it contains at least one polysemous term. In the following, we define polysemous terms.

Definition 6 *(polysemous term).*

A term t = ⟨Lemma, Cat, T-Rank⟩ is polysemous if there is a term $t^{'}$ and two synsets s and $s^{'}$, $s \neq s^{'}$ such that
a) t $\in$ s.Terms and $t^{'} \in s^{'}$.Terms;
b) t.Lemma = $t^{'}$.Lemma;
c) t.Cat = $t^{'}$.Cat.

In the following, we define polysemous synsets.

Definition 7 *(polysemous synset).*

A synset s is polysemous if any of its terms is a polysemous term.
It is possible for two polysemous synsets to share more than one term. Two polysemous synsets and their shared terms constitute a polysemy instance. In the following, we define polysemy instances.

Definition 8 *(polysemy instance).*

A polysemy instance is a triple $[\{T\}, s_1, s_2]$, where s_1, s_2 are two polysemous synsets that have the terms {T} in common.

For example, the term `bazaar` belongs to the following polysemy instances: $[\{bazaar, bazar\}, \#1, \#2]$, $[\{bazaar\}, \#1, \#3]$, and $[\{bazaar\}, \#2, \#3]$.
#1 `bazaar, bazar: a shop where a variety of goods are sold.`
#2 `bazaar, bazar: a street of small shops.`
#3 `bazaar, fair: a sale of miscellany; often for charity.`
We move now to the last part of our definitions. We exploit the structural properties in WordNet hierarchy to identify the polysemy types of the polysemy instances in WordNet. According to the connectivity property of WordNet hierarchy in definition 4, any two synsets in wordNet have at least one common subsumer that we define as follows.

Definition 9 *(common subsumer).*

Let s_1, s_2, and s be synsets in wordNet. The synset s is a common subsumer of s_1 and s_2 if $s \prec^* s_1$ and $s \prec^* s_2$.

The WordNet hierarchy is a DAG (directed acyclic graph). This implies that it is possible for two synsets to have more than one common subsumer. We define the least common subsumer as the subsumer with the least height.
In the following, we define structural patterns.

Definition 10 *(structural pattern).*

A structural pattern of polysemy instance I = $[\,\{T\}, s_1, s_2]$ is a triple $P = \langle r, p_1, p_2 \rangle$, where
a) r is the least common subsumer of s_1 and s_2;
b) $r \prec p_1$ and $r \prec p_2$;
c) $p_1 \prec^* s_1$ and $p_2 \prec^* s_2$.

We call r *the pattern root* and p_1, p_2 *the pattern hyponyms.* For example, the structural pattern of the polysemy instance $[\,\{bazaar, bazar\}, s_1, s_2]$ is $\langle mercantile\ \ establishment, marketplace, shop \rangle$ as shown in Figure 1, where `mercantile establishment` is the pattern root and `marketplace` and `shop` are the pattern hyponyms. A special structural pattern is the

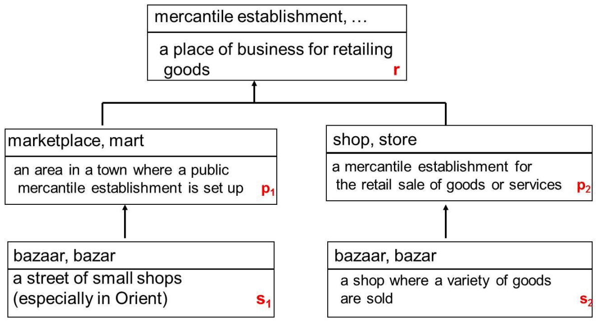

Figure 1: Example of a structural pattern

common parent structural pattern as illustrated in Figure 2. A strcutural pattern $P = \langle r, p_1, p_2 \rangle$ of a polysemy instance I = $[\,\{T\}, s_1, s_2]$ is a common parent structural pattern if $p_1 = s_1$ or $p_2 = s_2$.

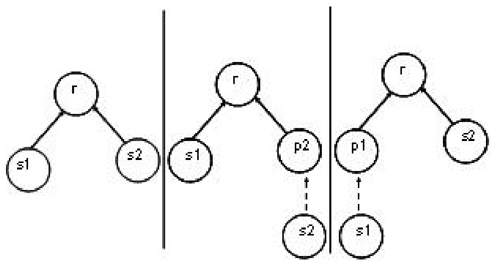

Figure 2: Common parent structural pattern

4 Taxonomic principles in WordNet

WordNet hierarchy represents a classification hierarchy where synsets are the nodes. Classification hierarchies should fulfill among other requirements the exclusiveness property and the exhaustiveness property.
We begin with the exclusiveness property.

Definition 11 *(Exclusiveness property).*

Two synsets $s_1, s_2 \in S$ fulfill the exclusiveness property if $s_1^I \sqcap s_2^I = \bot^I$. For example, `abstract entity` and `physical entity` fulfill the exclusiveness property. On the other hand `expert` and `scientist` do not fulfill this property because $expert^I \sqcap scientist^I \neq \bot^I$.
The exclusiveness property means that any two

sibling nodes n_i, nj in the hierarchy are disjoint, i.e., $n_i^I \not\sqsubseteq n_j^I$ and $n_j^I \not\sqsubseteq n_i^I$. Analyzing the structural patterns in WordNet shows that the exclusiveness property is not always guaranteed in WordNet. For example, the pattern $\langle person, expert, scientist \rangle$ shown in Figure 3 does not fulfill this property because forcing this property would result in preventing a scientist to be an expert or an expert to be a scientist. We

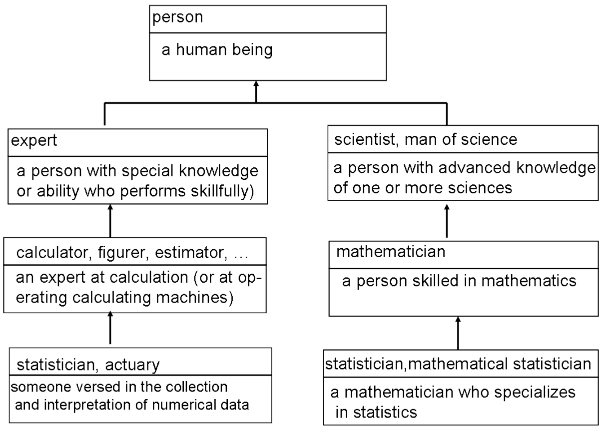

Figure 3: An example of exclusiveness property violation

are concerned with the cases, where the synsets s_1 and s_2 are not disjoint and each of them subsumes a synset of the same polysemous term such as the term *statistician* in Figure 3. The fact that the two synsets of the polysemous terms are not disjoint implies that the polysemy type of these two synsets can not be homonymy, metonymy, or metaphoric. This can be explained as follow. The polysemy type homonymy implies that the two synsets are unrelated and that the disjointness between the two synsets indicates a relation between the two synsets. Metonymy on the other hand means that one synset is a part of the other synset. Now, we explain the exhaustiveness property.

Definition 12 *(Collective Exhaustiveness).*

Two synsets $s_1, s_2 \in S$ are collectively exhaustive if it is possible to find a synset s such that $s^I = s_1^I \sqcup s_2^I$ and s_1, s_2 fulfill the exclusiveness property.

For example, `abstract entity` and `physical entity` fulfill the collectively exhaustiveness property because $entity^I = abstract\ \ entity^I \sqcup physical\ \ entity^I$. On the other hand `worker` and `female` in the pattern $\langle peron, worker, female \rangle$ do not fulfill this property because `worker` corresponds to a role and `female` to a concept. This is because `person` is a direct hypernym of the concept `organism` and the role `causal agent`.

5 Approach Overview

We exclude the structural patterns whose pattern root resides in the first and second level in WordNet hierarchy. Accordingly, any structural pattern whose root belongs to the synsets {`entity, abstract entity, abstraction, physical entity, physical object` } was automatically excluded. Our hypothesis is that the pattern hyponyms in these structural patterns in general fulfill the exclusiveness and the exhaustiveness property. These patterns are subject to our current research in discovering metonymy structural patterns. On the other hand, exclusiveness and exhaustiveness property are not guaranteed for all structural patterns whose roots reside in the third level and beyond. The input of the algorithm is the taxonomic structure of WordNet, starting from level 3, after removing lexical redundancy in compound nouns (Freihat et al., 2015). The output consists of three lists that contain specialization polysemy, metaphoric polysemy and homonymy instances. The first step of our algorithm is automatic, while the other two are manual.

S1. **Structural pattern discovery**: The input of this step is the current structure of WordNet after removing lexical redundancy. The algorithm returns structural patterns associated with their corresponding polysemy instances.

S2. **Structural pattern classification**: In this step, we manually classify the structural patterns returned in the previous step. The output consists of four lists of patterns associated with their polysemy instances. These four lists are:
Specialization polysemy patterns: This list contains the patterns whose corresponding instances are specialization polysemy candidates.
Metaphoric patterns: This list contains the patterns whose corresponding instances are metaphoric candidates.
Homographs patterns: This list contains the patterns whose corresponding instances are homonymy candidates.
Singleton patterns: The patterns in this group are those patterns that have one polysemy in-

stance only and thus cannot be considered to be regular.

S3 **Identifying false positives**: In this step, we manually process the polysemy instances in the four lists from the previous step. Our task is to decide the polysemy type for the instances in the singleton patterns list and remove false positives form the other three lists.

6 Metaphoric Structural Patterns

Identifying metaphoric patterns is based on the distinction between the literal meaning and the figurative meaning. Our idea is that it is not possible for a literal and the figurative meaning to be collectively exhaustive. Violating the exhaustiveness property in a structural pattern $\langle r, p_1, p_2 \rangle$ may be a result of the following:
a) p_1 and p_2 belong to different types and can not be subsumed by the pattern root r, or
b) $p_1 \sqsubset p_2$ or $p_2 \sqsubset p_1$.
For example $female$ and $worker$ can not be subsumed by $person$ in the pattern $\langle person, female, worker \rangle$ as shown in Figure 4. On the other hand, it is correct that

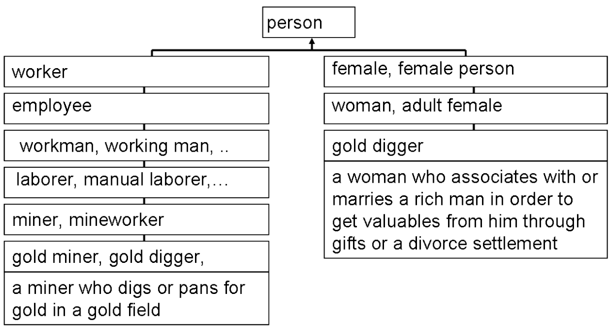

Figure 4: Example of a metaphoric polysemy instance

person and animal are organisms in the structural $\langle organism, animal, person \rangle$ but it is clear that $person^I \sqsubset animal^I$
In the following, we define metaphoric patterns structural pattern as follows.

Definition 13 *(Metaphoric structural pattern).*

A pattern $p = \langle r, p_1, p_2 \rangle$ is metaphoric if p_1 and p_2 do not fulfill the collectively exhaustiveness property.
In the following we give examples for identified metaphoric patterns. The pattern $\langle organism, animal, person \rangle$ is metaphoric. Although both synsets share the same hypernym `organism`, they are not collectively exhaustive as explained. The polysemy instances that belong to this pattern are 326 instances. Consider for example the following instance.
#1 `snake, serpent, ophidian: limbless scaly elongate reptile.`
#2 `snake, snake in the grass: a deceitful or treacherous person.`
Another example is the pattern $\langle attribute, property, trait \rangle$. Although, both synsets share the same hypernym `attribute`, they are not collectively exhaustive because $trait^I$ is a special case of $property^I$ ($trait^I = property^I \sqcap person^I$). The polysemy instances that belong to this pattern are 111 instances. Consider for example the following instance.
#1 `softness:the property of giving little resistance to pressure and being easily cut or molded.`
#2 `gentleness, softness, mildness: acting in a manner that is gentle and mild and even-tempered.`

7 Specialization Polysemy Structural Patterns

We use the exclusiveness property and the pattern root in a structural pattern to discover specialization polysemy candidates indirectly. The relation between the synsets in specialization polysemy is hierarchical. The hierarchical relation between the synsets in a specialization polysemy instance indicates that the exclusiveness property does not hold between synsets and thus between the structural pattern hyponyms.
We define specialization polysemy patterns as follows.

Definition 14 *(specialization polysemy structural pattern).*

A pattern $p = \langle r, p_1, p_2 \rangle$ is a specialization polysemy pattern if a) and b) hold
a) p_1 and p_2 do not fulfill the exclusiveness property.
b) p_1 and p_2 fulfill the exhaustiveness property.
In the following we give examples for identified specialization polysemy patterns. All instances that belong to the common parent structural patterns are classified as specialization polysemy instances. The polysemy instances that belong to this pattern are 2879 instances. Consider for example the following instance.

#1 `capital, working capital: assets available for use in the production of further assets.`

#2 `capital: wealth in the form of money or property owned by a person or business and human resources of economic value.`

Another example is the pattern $\langle act, action, activity\rangle$. The polysemy instances that belong to this pattern are 406 instances. Consider for example the following.

#1 `employment, work: the occupation for which you are paid.`

#2 `employment, engagement: the act of giving someone a job.`

Another example, is the pattern $\langle animal, invertebrate, larva\rangle$. The polysemy instances that belong to this pattern are 17 instances. Consider for example the following.

#1 `ailanthus silkworm, Samia cynthia: large green silkworm of the cynthia moth.`

#2 `cynthia moth, Samia cynthia, Samia walkeri: large Asiatic moth introduced into the United States; larvae feed on the ailanthus.`

8 Homonymy Structural Patterns

We define homonymy patterns as follows.

Definition 15 *(Homonymy structural pattern).*

A pattern $p = \langle r, p_1, p_2\rangle$ is homonymy pattern if the following condition hold.
a) p_1 and p_2 fulfill the exclusiveness property;
b) p_1 and p_2 fulfill the exhaustiveness property;
c) There is no relation between p_1 and p_2.

In the following we give examples for identified homonymy patterns. The pattern $\langle organism, person, plant\rangle$. The polysemy instances that belong to this pattern are 40 instances. Consider for example the following instance.

#1 `spinster, old maid: an elderly unmarried woman.`

#2 `zinnia, old maid, old maid flower: any of various plants of the genus Zinnia.`

Another example is the pattern $\langle organism, animal, plant\rangle$. The polysemy instances that belong to this pattern are 41 instances. Consider for example the following.

#1 `red fox, Celosia argentea: weedy annual with spikes of silver-white flowers.`

#2 `red fox, Vulpes fulva: New World fox; often considered the same species as the Old World fox.`

Another example is the pattern $\langle vertebrate, bird, mammal\rangle$. The polysemy instances that belong to this pattern are 13 instances. Consider for example the following.

#3 `griffon, wire-haired pointing griffon: breed of medium-sized long-headed dogs.`

#4 `griffon vulture, griffon, Gyps fulvus: large vulture of southern Europe and northern Africa.`

9 False Positives Identification

In this section, we describe the third step of our approach. Our task here is to process the four lists returned at the end of the pattern classification and remove false positives. These lists are the metaphoric polysemy list, the specialization polysemy list, the homonymy list, and a list of non regular (singleton patterns) list. This task can only be performed manually due to the implicit and missing information in synset glosses. Our procedure for determining the polysemy class of a polysemy instance is based on the three definitions in the previous section, where we process the polysemy instances instance by instance to determine the the relation between the synsets of the polysemy instances.

If a polysemy instance does not belong to the polysemy type it was assigned to (false positive instance), we assign it to its corresponding polysemy type.

In the following, we give examples for false positives. The common parent structural pattern which was automatically assigned to the specialization polysemy type (step 1 in Section 5) contains 180 false positive polysemy instances, 98 of them were identified as homonymy instances. One example is:

#1 `cardholder: a person who holds a credit card or debit card.`

#2 `cardholder: a player who holds a card or cards in a card game.`

Metaphoric false positives (82 instances) were also identified in the common parent class. Consider for example the following instance.

#1 `game plan: (figurative) a carefully thought out strategy for achieving an objective in war.`

#2 `game plan: (sports) a plan for achieving an objective in some sport.`

Another example is the pattern $\langle organism, animal, person \rangle$ which was assigned to the metaphoric polysemy type contains 326 polysemy instances, 74 of them were identified as homonyms such as the following instance.

#2 `Minnesotan, Gopher: a native or resident of Minnesota.`

#3 `ground squirrel, gopher, spermophile: any of various terrestrial burrowing rodents of Old and New Worlds.`

10 Results and Evaluation

The number of polysemy instances computed by the polysemy instances discovery algorithm is 41306. We excluded 28318 instances because the pattern roots of these instances reside in the first and the second level of the hierarchy as per the approach discussed in Section 5.The remaining number of polysemy instances is 12988. These instances are divided in two groups as follow. 12988 of these instances belong to 1028 regular type compatible patterns and 1569 instances belong to single tone patterns. The classification of the pasterns and the result of the false positive removing is shown in the following tables.

#Type	#patterns	#instances
Specialization	823	9902
Metaphoric	134	1697
Homonymy	71	1389
Total	1028	12988

Table 1: Classification of the regular structural patterns

In Table 2, we show the results removing false positive instances, where we see that the average false positives is about 17%.

#Poly Type	#Instances	#False Positives
Specialization	9902	1740
Metaphoric	1697	175
Homonymy	1389	295
Total	12988	2210

Table 2: False Positives in Pattern Classification

To evaluate our approach, 3797 polysemy instances were evaluated by two evaluators. The agreement of the evaluators with our approach was on 96.5% of the instances. In the following Table 3, a refers to our approach, e_1, e_2 refer to evaluator1 and evaluator 2 respectively.

$e_1 = a \vee e_2 = a$	3665 (96.5%)
$a = e_1$	3621 (95.3%)
$a = e_2$	3600 (94.8%)

Table 3: Evaluation of the polysemy classification

11 Conclusion and future Work

In this paper, we have presented how to use two taxonomic principles for classifying the polysemy types in WordNet. We have demonstrated the usefulness of our approach on classifying three polysemy types, namely, specialization, metaphoric and homonymy. In this approach, we were able to discover all specialization polysemy structural patterns and subsets of the metaphoric and metonymy structural patterns. We aim to continue our work to study the metonymy patterns in the upper level of WordNet hierarchy, where we generalize our structural pattern definition as follows.

Definition 16 *(generalized structural pattern).*

A structural pattern of polysemy instance I = $[\ \{T\}, s_1, s_2]$ is a triple $P = \langle r, p_1, p_2 \rangle$, where
a) r is the least common subsumer of s_1 and s_2;
b) $r \prec^* p_1$ and $r \prec^* p_2$;
c) $p_1 \prec^* s_1$ and $p_2 \prec^* s_2$.

Our hypothesis is that in case of metonymy structural patterns: the nodes p_1 and p_2 fulfill the exclusiveness and the exhaustiveness properties and there is a part of relation between p_1 and p_2. The conditions for metaphoric and homonymy structural patterns obtained by adapting the new structural definition remain the same as explained in this paper.

Acknowledgment

The research leading to these results has received partially funding from the European Community's Seventh Framework Program under grant agreement n. 600854, Smart Society (http://www.smart-society-project.eu/).

References

Jordi Alvarez. 2000. Integrating the wordnet ontology into a description logic system.

JU. Apresjan. 1974. Regular polysemy. *Linguistics*, pages 5–32.

Kenneth D. Bailey. 1994. *Typologies and Taxonomies: An Introduction to Classification Technique*. Sage Publications, Thousand Oaks, CA, 1994///.

Lucie Barque and Franois-Rgis Chaumartin. 2009. Regular polysemy in wordnet. *JLCL*, 24(2):5–18.

Travis D. Breaux, Annie I. Anton, and Jon Doyle. 2009. Semantic parameterization: A process for modeling domain descriptions. *ACM Transactions on Software Engineering Methodology*, 18(2).

Paul Buitelaar. 1998. Corelex: Systematic polysemy and underspecification. *PhD thesis,Brandeis University, Department of Computer Science*.

W. B. Dolan. 1994. Word sense ambiguation: clustering related senses. In *Proceedings of COLING94*, pages 712–716.

Philip Edmonds and Eneko Agirre. 2008. Word sense disambiguation. *Scholarpedia*, 3.

Vyvyan Evans and Jrg Zinken. 2006. Figurative language in a modern theory of meaning construction: A lexical concepts and cognitive models approach.

Ingrid Lossius Falkum. 2011. The semantics and pragmatics of polysemy: A relevance-theoretic account. *PhD thesis, University College London*.

Abed Alhkaim Freihat, Fausto Giunchiglia, and Biswanath Dutta. 2013. Regular polysemy in wordnet and pattern based approach. *International Journal On Advances in Intelligent Systems*, 6(3&4), jan.

Abed Alhkaim Freihat, Biswanath Dutta, and Fausto Giunchiglia. 2015. Compound noun polysemy and sense enumeration in wordnet. *In Proceedings of the 7th International Conference on Information, Process, and Knowledge Management (eKNOW)*, pages 166–171.

Verdejo F. Gonzalo J., Chugur I. 2000. Sense clusters for information retrieval: Evidence from semcor and the eurowordnet interlingual index. *ACL-2000 Workshop on Word Senses and Multi-linguality, Association for Computational Linguistics*, pages 10–18.

Ahti Lohk, Kaarel Allik, Heili Orav, and Leo Vhandu. 2014. Dense components in the structure of wordnet. In Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis, editors, *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Reykjavik, Iceland, may. European Language Resources Association (ELRA).

Rila Mandala, Takenobu Tokunaga, and Hozumi Tanaka. 1999. Complementing wordnet with roget's and corpus-based thesauri for information retrieval. In *EACL*, pages 94–101. The Association for Computer Linguistics.

Alberto Marradi. 1990. Classification, typology, taxonomy. *Quality & Quantity: International Journal of Methodology*, 24(2):129–157.

MoldovanD. I. Mihalcea R. 2001. Ez.wordnet: Principles for automatic generation of a coarse grained wordnet. *FLAIRS Conference*, pages 454–458.

G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, and K. J. Miller. 1990. Introduction to WordNet: an on-line lexical database. *International Journal of Lexicography*, 3(4):235–244.

George A. Miller. 1995. Wordnet: A lexical database for english. *Commun. ACM*, 38(11):39–41, November.

Wim Peters. 2004. Detection and characterization of figurative language use in wordnet. *PhD thesis, Natural Language Processing Group, Department of Computer Science, University of Sheffield*.

Sebastian Rudolph. 2011. Foundations of description logics. In Arenas C. Polleres A., dÁmato C., editor, *Reasoning Web. Semantic Technologies for the Web of Data - 7th International Summer School 2011*, volume 6848 of *LNCS*, pages 76–136. Springer.

Tony Veale. 2004. Pathways to creativity in lexical ontologies. In *In Proceedings of the 2nd Global WordNet Conference*.