Semantic Selectional Restrictions for Disambiguating Meronymy Relations

Ashwin Ittoo[1], Gosse Bouma[2]

[1]Faculty of Economics and Business,
[2] Faculty of Arts,
University of Gronigen,
Groningen, The Netherlands
{r.a.ittoo,g.bouma}@rug.nl

Abstract. In this paper, we present an unsupervised approach to automatically learn lexico-syntactic patterns encoding meronymy relations from texts. Our major contribution lies in alleviating the challenge of disambiguating polysemous patterns that encode meronymy only in some contexts. We rely on the linking theory to posit that semantic features of the Part and Whole instances participating in a meronymy relation facilitate the identification of meronymy-encoding patterns. We abstract the instances to their hypernyms, enforcing semantic selectional restrictions to constrain the contexts within which patterns participate in meronymy. We disambiguate polysemous patterns using their contexts based on a modified version of Harris' distributional hypothesis by postulating that similar patterns share similar contexts. Our experiments revealed that enforcing semantic selectional restrictions to constrain the contexts within which patterns participate in meronymy leads to the identification of high-quality patterns. Furthermore, our method does not require annotated data, and has a broader coverage compared to previous studies.

Key words: Meronymy, Part-Whole Relations, Unsupervised Learning

1 Introduction

Meronymy is a semantic relation between an object corresponding to a "part" and to its corresponding "whole"[1]. If an entity X is the meronym of another entity Y, then sentences of the form *"Xs are parts of Y"* or *"Y has Xs"* are valid when noun phrases X and Y are interpreted generically; for example, *"an engine is a part of a car"*. The inverse of meronymy is holonymy. Meronymy relations between part and whole objects are crucial for various Natural Language Processing tasks, such as question-answering, information extraction and text summarization [2].

Recently, several research efforts geared towards automatically identifying meronymy patterns from texts have been proposed [3, 1, 2]. However, the long-standing challenge of resolving pattern ambiguity has not yet been adequately addressed. Pattern ambiguity arises when an expression encodes meronymy only when it occurs within specific contexts. For example, the genitive pattern *"of"* is polysemous since it encodes different semantic relations depending on its contexts. It conveys meronymy in *"engine of the car"*, but not in *"book of the student"*, in which it indicates a *possession* relation.

In this paper, we present a novel, unsupervised approach to automatically learn high-quality lexico-syntactic patterns encoding meronymy relations from unstructured, natural language texts. We address the pattern ambiguity issue by considering part-whole relations as analogous to semantic frames, introduced by Fillmore in his work on Natural Language Analysis [4], and commonly used in Semantic Role Labeling [5]. A semantic frame schematically depicts an action or relation, together with the participating concepts that are labeled depending on their roles [5]. Based on this analogy, meronymy is a specialized semantic frame, illustrating a part-whole relation between an instance playing the role of a Part, and another one with the corresponding Whole role. The instances occurring in the Part and Whole roles of a pattern constitute its contexts, and determine, for ambiguous patterns, whether they encode meronymy. The novelty in our approach to disambiguate polysemous meronymy expressions lies in enforcing semantic selectional restrictions on the instances in the Part and Whole roles of patterns to constrain the possible contexts within which these patterns encode part-whole relations. These meronymy-indicative contexts form the basis of our pattern disambiguation procedure. Our selectional restrictions are obtained by abstracting the Part and Whole instances to their hypernyms.

We disambiguate polysemous patterns based on a modified version of Harris' distributional hypothesis [6], in which we claim that "patterns are similar to the extent to which they share similar contexts". Hence, to disambiguate a pattern (i.e. to determine whether it encodes meronymy given its contexts), we find out if its contexts are similar to any of the meronymy-indicative contexts within which the pattern can participate in a part-whole relation.

Besides acting as selectional restrictions to facilitate the disambiguation of polysemous expressions, converting Part and Whole instances to their conceptual classes also helps in broadening the scope of our approach, and in alleviating issues related to data sparsity. Furthermore, our strategy detects meronym noun pairs that are not documented in the WordNet semantic database [7]. To enhance the quality of the patterns acquired by our unsupervised procedure, we rely on certain mathematical and logical properties of meronymy relations.

Compared to previous related studies, our methodology has broader coverage, and its unsupervised learning procedure reduces the reliance on large amount of annotated training data. We also innovate in our choice of the English Wikipedia corpus for mining and evaluating meronymy relations. Wikipedia has till now been exploited to extract very general relations, as in [8, 9].

The remainder of this paper is structured as follows. We present an overview of the theories of parthood in Section 2. Section 3 discusses some recent studies related to the automatic discovery of part-whole relations. We present our unsupervised approach for automatically identifying meronymy relations in Section

4. Our experiment results follow in Section 5, and we conclude and highlight potential areas for future investigation in Section 6.

2 Meronymy Relations

In linguistics, the semantic relation between part entities and their corresponding wholes is known as meronymy [10]. An entity X is the meronym of another entity Y if *"Xs are parts of Y"*, or *"Y has Xs"*. Conversely, Y is the holonym of X.

Meronymy relations can be established both at the conceptual (class) level and at the individual instances (objects) level [11]. Class-level part-whole relations indicate that every instance of the Whole concept includes one or more instances of the Part concepts. They are generic statements at the class level. Instance-level meronymy relations imply that the (specific) Whole object includes one or more of the Part objects. For example, the part-whole relation between Engine and Car is at the conceptual level; while that between 3000CC Engine and Lamborghini is an instance-level relation.

2.1 Formal and Mathematical Properties of Parthood

Simons [12, 13] and Varzi [14] defined meronymy as a strict partial ordering, with the following axioms: existence, asymmetry, supplementarity, transitivity, extensionality, irreflexivity, existence of mereological sum, existence of mereological atoms, and decomposition into atoms. Of particular interest to us are asymmetry, transitivity and irreflexivity, which we rely upon to validate the part-whole relations acquired by our approach. They are illustrated in Equations (1) - (3). $PWR(X, Y)$ indicates a part-whole relation between X and Y.

The irreflexivity property states that a concept X cannot be a part or a whole of itself [15], and is formalized in equation (1).

$$\forall X; \neg PWR(X, X) \tag{1}$$

Equation (2) illustrates the asymmetry property, according to which, if X is part of Y, then Y is not part of X [15].

$$\forall X, Y; PWR(X, Y) \rightarrow \neg PWR(Y, X) \tag{2}$$

The transitivity property, in equation (3), states that if X is part of Y, and Y is part of Z, then X is also part of Z [15].

$$\forall (X, Y, Z); PWR(X, Y) \wedge PWR(Y, Z) \rightarrow PWR(X, Z) \tag{3}$$

Although transitivity is a minimal requirement for part-whole relations [16], it applies only across relations of the same types according Winston's taxonomy [10].

2.2 Taxonomy of Meronymy Relations

Winston et al. [10] developed a taxonomy of part-whole relations based on psycho-linguistics experiments, which illustrates the kinds of semantic relations expressed by the ordinary English speakers' use of the phrase *"part of"* and its cognates. Six major types of meronymy relations were classified by Winston et al.(Winston's). They are: 1) Component-Integral, depicting the relation between a component(part) and its integral(whole), as in ENGINE-CAR, 2)Member-Collection, representing membership in a collection, such as SHIP-FLEET, 3)Portion-Mass, capturing the relations between portions and masses, and physical dimensions, for example, METRE-KILOMETRE or SLICE-PIE, 4)Material-Object, encoding the relation between an object and its constituent material, such as GRAPE-WINE, 5)Place-Area, indicating the relation between areas and locations within them, as in NETHERLANDS-EUROPE, and 6)Feature-Activity, capturing the relation between phases of an activitiy, for example, SCORING-PLAYING.

2.3 Ambiguity of Meronymy Encoding Patterns

The many different ways in which an entity can be expressed as a part of another entity give rise to a wide variety of lexico-syntactic structures that can encode meronymy [17]. Previous studies [1] recognized four types of meronymy-indicating lexico-syntactic patterns. They are: 1)genitives and the verb *"to have"*, such as ENGINE *"of the"* CAR, CAR's ENGINE and CAR *"has an"* ENGINE, 2)noun compounds, as in CAR DOOR, 3) prepositional phrases, like MAN *"with"* two HEADS, and 4) other rare patterns, for example BANK *"is a branch of"* COMPANY.

Some lexico-syntactic patterns are unambiguous and always depict meronymy. Other polysemous patterns are ambiguous, and convey meronymy depending on the contexts in which they occur. Tables 1-3 list examples of ambiguous meronymy patterns for genitives, noun compounds and prepositional phrases respectively.

Table 1. Ambiguous Genitives.

ENGINE "of the" CAR	True
BOOK "of the" STUDENT	False

Table 2. Ambiguous Noun Compounds.

CAR ENGINE	True
STUDENT BOOK	False

Table 3. Ambiguous Prepositional Constructs.

CAR "with" ENGINE	True
STUDENT "with" BOOK	False

Disambiguating polysemous lexico-syntactic structures, to determine whether their occurrences are indicative of meronymy, is a challenge that has not been adequately addressed in past research.

3 Related Work

Hearst [18] proposed an approach for identifying meronymy relations from text based on the occurrence of a restricted set of lexico-syntactic patterns.Although successful employed for detecting hypernymy [19], this technique was unsuccessful for learning part-whole relations. The poor performance could be attributed to the wider variety of meronymy-invoking constructs (compared to a smaller set of hypernymy-expressing patterns).Berland and Charniak [3] also relied on specific lexico-syntactic patterns to discover part-whole relations. They focused on a small set of genitive patterns and on six seed nouns representing WHOLE concepts. Statistical measures over a large corpus were then employed to locate meronymy relations.

Neither Hearst nor Berland and Charniak discriminated between ambiguous and unambiguous meronymy constructs, which resulted in the low performance of their approaches. They also suffered from low coverage, compounded by their small seed sets, and their discarding of terms with suffixes "ing", "ness", or "ity". Girju et al. [2][1] attempted to alleviate the pattern ambiguity challenge by adopting a machine-learning approach based on decision tree classifiers. Despite being more accurate than [3] and [18], they required substantial manual intervention, for example, to identify an initial set of meronymy-encoding patterns, to annotate positive and negative examples, and to disambiguate word senses. Furthermore, annotating training data for supervised learning algorithms is tedious and expensive.

Our methodology to automatically acquire high-quality meronymy lexico-syntactic patterns addresses the above-mentioned issues, as well as the crucial challenge of resolving meronymy pattern ambiguity.

4 Methodology

The overall framework in which we embed our unsupervised algorithms to automatically discover meronymy relations consists of a knowledge acquisition and an evaluation phase. Similar to other data-driven methodologies, we start by acquiring seed PART-WHOLE noun pairs from a semantic database. These *noun* pairs are abstracted to their corresponding PART-WHOLE *concept* pairs. Lexico-syntactic patterns relating (noun) instances of the PART-WHOLE concept pairs are then searched for, validated and extracted from a corpus. Evaluation involves measuring the precision with which the acquired patterns encode meronymy over a completely different test corpus.

4.1 Acquiring Seeds

The aim of this phase is to automatically acquire meronym or holonym noun pairs, $<n_1,n_2>$, participating in part-whole relations. This means either n_1 is a meronym of n_2 (n_2 is the holonym of n_1) or vice-versa. The pair n_1 and n_2 constitute a Part-Whole noun pair. We implemented a meronym/holonym noun pair harvesting procedure that takes as input a seed noun, s_i, and traverses WordNet to collect its meronyms and holonyms for all its different senses. The collected meronyms and holonyms are then treated as seeds, and fed to our harvesting algorithm so that their meronyms and holonyms are acquired recursively. Unlike past studies $[3, 2, 1, 18]$, our seed harvesting and expansion strategy enables us to broaden our coverage since the holonym of a seed s_i could have other meronyms s_j (besides s_i). For example, s_i = "$engine\#3$", has as holonym "$train\#1$", which in turn has another meronym synset $s_j = < car\#2, railcar\#1, railway\_car\#1, railroad\_car\#1 >$. Our approach also efficiently handles seeds that are both meronyms and holonyms, such as s_i = "$engine\#3$" which is a holonym of "$footplate\#1$", but a meronym of "$train\#1$". Furthermore, our seed collection procedure requires the specification of either a part or a whole noun while previous studies needed both the part and whole seeds.

Thus, from a minimal list of input nouns, this phase outputs an augmented seed set of Part-Whole noun pairs. Since they were harvested by traversing the meronymy and holonymy semantic links of WordNet, which we consider a reliable reference, the co-occurences of these noun pairs in any lexico-syntactic patterns always indicate meronymy.

4.2 Conceptual Semantic Abstraction

This stage takes as input the previously harvested Part-Whole *noun* pairs, and generalizes them to their super-ordinate concepts (classes), resulting into Part-Whole *concept* pairs. We achieve this abstraction by replacing the Part and Whole noun instances with their respective direct ancestors, one level up the WordNet hypernymy chain. For example, the Part-Whole *noun* pair $< grape, wine >$ will be abstracted to the Part-Whole *concept* pair $< edible\_fruit\#1, alcohol\#1 >$, where $\#x$ indicates WordNet sense x.

Our rationale for incorporating semantic features such as hypernyms stems from the linking theory $[4, 5]$, which is at the core of frame-based semantic role labeling [5]. According to this theory, lexico-syntactic realizations of patterns that relate arguments to their predicates can be accurately predicted from the arguments' semantics. Our reformulation of meronymy relation discovery as a special case of semantic role labeling, with meronymy relations corresponding to semantic frames, enables us to postulate that semantic information about

Part and Whole arguments could facilitate identifying lexico-syntactic realizations of patterns that encode meronymy (Section 4.3). Abstracting Part and Whole noun pairs to their hypernyms also enables the implicit capture of certain mathematical properties (irreflexivity, asymmetry, transitivity) as well as other characteristics (homogeneous and homeomerous) of the relations in which they participate. Furthermore, replacing individual noun instances with their conceptual classes broadens the coverage of lexical items that are characterized by large vocabularies, and hence, reduces problems of data sparsity.

The output of this phase is a set of Part-Whole concept pairs of the form $< Hypernym(n_1\#x), Hypernym(n_2\#y) >$, where n_1 and n_2 are Part-Whole noun pairs with WordNet sense numbers x and y respectively. The function $Hypernym(w\#s)$ returns the immediate hypernym of noun w with sense s based on WordNet's hierarchy. When they appear as arguments of lexico-syntactic patterns, these concept pairs impose semantic selectional restrictions that constrain the possible contexts within which these pattern encode meronymy, facilitating the disambiguation of polysemous patterns (Section 4.3). We index such Part-Whole concept pair in a Part-Whole concept lexicon.

One issue with generalizing noun instances to their immediate WordNet hypernyms is that the most informative hypernym could reside a few level up the WordNet hierarchy. For example, abstracting the noun pair $< robin, wing >$ to its corresponding concept pair yields $< thrush, organ >$, while the pair $< bird, organ >$ would have been more desirable. We discuss how to handle this problem in our future work (Section 6).

4.3 Acquiring and Disambiguating Meronymy-Indicating Patterns

The objective in this step is to acquire lexico-syntactic expressions that convey meronymy from a corpus. This process entails the proper disambiguation of polysemous patterns to determine whether their occurrences indicate meronymy. For each identified Part-Whole concept pairs $<$ *Part_Concept*, *Whole_Concept* $>$ from the previous phase, we search a portion of the English Wikipedia corpus for sentences containing noun pairs $< n_1, n_2 >$ such that n_1 is an instance of the Part concept *Part_Concept*, and n_2 is an instance of the corresponding Whole concept *Whole_Concept*.

Since *Part_Concept* and *Whole_Concept* participate in meronymy at the class-level, according to Cruse [11], every instance of type *Whole_Concept*(e.g.: n_2) should contain one or more instances of type *Part_Concept* (e.g.:n_1). Based on Cruse's idea and on the extensionality property of part-whole relations, which states that objects with the same parts are identical, we infer that n_1 and n_2 are Part-Whole noun pairs that also participate in meronymy relations. As example, consider the Part-Whole concept pair $< edible\_fruit\#1, alcohol\#1 >$ from the Part-Whole concept lexicon (Section 4.2). Searching for instances of this meronymy concept pair in the Wikipedia corpus could lead to the identification of the Part-Whole noun pair $< grape\#1, wine\#1 >$ from Wikipedia sentences.(We do not discuss determining the correct senses of nouns in this paper). Since these identified noun pairs are instances of known Part-Whole concept

pairs, and hence also participate in part-whole relations, the lexico-patterns that relate them in their sentences encode meronymy. Similar to [20], we formalize the representational space of lexico-syntactic patterns by dependency paths. We syntactically parse sentences containing occurrences of the Part and Whole noun pairs, and define a meronymy lexico-syntactic construct as the shortest path in the parsed dependency structure that relates the Part and the Whole nouns. For example, given sentence
S1= "Plan Bordeaux calls for a simplification of French wine labels , including the name of the grapes ...".
S1 was identified from our corpus since it contains instances of the Part-Whole concept pair $< edible\_fruit\#1, alcohol\#1 >$, namely the Part and the Whole noun pairs *"grape"* and *"wine"*. Parsing *S1* results in
P'=[Alcohol#1: wine#1]+nn+labels+pobj+of <-prep->
including+pobj+name+prep+of+pobj+[Edible_Fruit#1:grape#1]..
The shortest dependency path between the Part and the Whole noun pairs extracted from *P'* is
P= "+nn+labels+pobj+of <- prep ->including+pobj+name+prep+of+pobj+".
In the above example, the lexico-syntactic pattern P encodes meronymy since its arguments are instances of the known Part-Whole concept pair *"edible_fruit"* and *"alcohol"*. In this way, we acquire a set of promising meronymy lexico-syntactic patterns by searching our corpus for sentences containing instances of previously identified Part-Whole concept pairs, parsing these sentences and extracting the shortest dependency paths between the noun instances. We index the learnt patterns in a meronymy-pattern lexicon.

However, our approach could prove to be too general, and identify instances that do not partcipate in meronymy relations, such as
$< kiwi\#1, wine\#1 >$ which are valid instances of $< edible\_fruit\#1, alcohol\#1 >$. Subsequently extracted patterns relating such false meronymy noun pairs do not encode part-whole relations. To keep only those constructs that are likelier to express meronymy, we maintain a frequency count, and discard patterns with frequencies below an experimentally set threshold. We assume that true meronymy patterns such as *"made of"* will not relate instances that are not parts and wholes, as in *"wine is made of kiwi"*. Furthermore, invalid meronym noun pairs are in most cases related by conjunctions or negations, as in *"he likes wine and kiwi"* or *"wine is not made of kiwi*, which we filter in our approach to improve the quality of the acquired part-whole patterns.

To disambiguate polysemous patterns, such as genitives (*"of"*) that encode meronymy only in certain contexts, we rely on the previously acquired Part-Whole concept pairs indexed in the Part-Whole concept lexicon, and on the unambiguous patterns recorded in the meronymy-pattern lexicon. The key idea underlying our disambiguation procedure is that Part-Whole concept pairs enforce semantic selectional restrictions on the lexico-syntactic patterns they co-occur with, constraining the possible contexts in which these patterns express meronymy, thereby facilitating their disambiguation. Contextual information has been traditionally employed in word sense disambiguation, following Harris' distributional hypothesis, which states that words are similar to the extent to which they share similar contexts. We reformulate this hypothesis as "lexico-syntactic patterns are similar to the extent to which they share similar contexts", and use this modified hypothesis in disambiguating polysemous patterns. In our case, the contexts of patterns are defined by the arguments that they sub-categorize. For example, given an unambiguous meronymy lexico-syntactic pattern $P =$*" contains"*, acquired from our corpus and indexed in the meronymy-pattern lexicon, and its meronymy-indicating contexts consisting of the PART-WHOLE concept pair $< vehicle\#1, motor\#1 >$. P expresses a conceptual level part-whole relation *"vehicle contains motor"*. To determine whether the occurence of an ambiguous pattern *P'* (e.g.: *"of"*) with arguments *"args1"* (e.g.: *"engine"*) and *"args2"* (e.g.: *"car"*) encodes meronymy, we compute its contextual similarity with the meronymy-pattern *P*. We consider *P'* to be similar to *P*, and hence, to also encode a part-whole relation, if the arguments defining its contexts are instances of the known PART-WHOLE concept pairs that are sub-categorized by the meronymy pattern *P*. Thus, in this example, since *"engine"* and *"car"* are instances of *"motor#1"* and *"vehicle#1"* respectively, we infer that the ambiguous meronymy pattern *P'= "of"* indicates a part-whole relation. The semantic classes (e.g.: concepts *"motor#1"* and *"vehicle#1"*) act as selectional modifiers on ambiguous patterns, and restrict the allowable contexts in which these patterns participate in meronymy. Had the contexts of *P'* been characterized by the arguments *"book"* and *"student"*, as in *"book of student"*, our technique would infer that *P'* does not encode meronymy in these contexts. It reaches this conclusion on two bases. First, the corpus from which the patterns are extracted does not contain invalid English constructs involving meronymy patterns as in *"student is made of book"*. Such invalid expressions will cause our algorithm to identify *"student"* and *"book"* as PART-WHOLE noun pairs, and to infer that the pattern *"made of"* encodes meronymy when it sub-categorizes instances of type *"enrollee#1"* and *"publication#1"* (the conceptual classes of *"student"* and *"book"* respectively). The second basis is that meronymy and holonymy relations documented in WordNet are correct. Otherwise, we will count pairs such as $<$*"enrollee#1"*, *"publication#1"*$>$ as valid PART-WHOLE concepts, and will wrongly consider the patterns in which their instances occur, such as *"student buys book"* or *"student has book"*, as encoding meronymy.

Our approach, although simple and intuitive, enables the discovery of meronymy (and holonymy) relations between pairs that are not mentioned in WordNet. For example, consider the PART-WHOLE noun pair $< base\#2, construction\#4 >$, harvested during our seed acquisition procedure (Section 4.1), and its corresponding PART-WHOLE concept $< support\#7, artifact\#1 >$ (Section 4.2). Our algorithm mines for meronymy lexico-syntactic patterns by searching our corpus for sentences containing co-occuring instances of concepts *"support#7"* and *"artifact#1"*. One such sentence could be *"buttress of excavation"* since the noun pair $< buttress, excavation >$ is an instance of the PART-WHOLE concept pair $< support\#7, artifact\#1 >$. Besides concluding that the genitive *"of"* expresses meronymy in the context of $< support\#7, artifact\#1 >$, we also deduce that *buttress* is a meronym of *excavation* - a valid fact that is not mentioned in WordNet. Hence, we augment WordNet with new relations to improve its completeness. To further enhance the performance of our unsupervised learning process, we discard relations that do not satisfy the asymmetry and irreflexivity properties of meronymy; for example, *"engine is part of engine"*, which violates irreflexivity or *"car is part of engine"*, which violates asymmetry.

The output of this stage is a set of (both ambiguous and unambiguous) lexico-syntactic patterns, together with the contexts in which they encode meronymy. These meronymy-indicating contexts are PART-WHOLE concept pairs. They enforce semantic selectional constraints on the patterns, and restrict the set of possible contexts in which the patterns participate in meronymy, thereby enabling their disambiguation. This phase also augments the PART-WHOLE concept lexicon (Section 4.2) with new PART-WHOLE concept pairs.

4.4 Evaluating the Acquired Patterns

The objective in this phase is to evaluate the quality of the patterns acquired. We want to determine whether enforcing semantic selectional restrictions on the PART and WHOLE instances of the patterns contributed in disambiguating polysemous meronymy patterns. Evaluation involves searching a test corpus for occurences of the acquired patterns, and measuring the precision with which they are indicative of part-whole relations. As will be shown in the next section, our unsupervised methodology is able to identify highly-precise meronymy lexico-syntactic patterns.

5 Experimental Evaluations

In our experimental setup, we employed WordNet 3 to acquire our seed set of meronym pairs (Section 4.1) and hypernyms (Section 4.2), and the English Wikipedia corpus [21] to extract candidate meronymy patterns and to evaluate them. Syntactic parsing was achieved by the Stanford parser [22], and programmatic access to WordNet was enabled via the WordNet-QueryData interface [23].

Our meronym noun pair acquisition starts with five seeds that participate in the different types of Winston's part-whole relations (we did not consider Feature-Activity relations, which are realized by verb entailments). These seeds are fed to our harvesting procedure which traverses WordNet, collects their meronyms and holonyms, and recursively expands the seed set. This strategy led to the discovery of some interesting PART-WHOLE noun pairs for a given input seed as shown in Table 4. A total of 12389 PART-WHOLE noun pairs were thus collected, many of them duplicates.

The collected PART-WHOLE noun pairs were abstracted to their corresponding PART-WHOLE concept pairs based on the WordNet hypernymy hierarchy. Duplicate pairs and those violating the irreflexivity and asymmetry properties of meronymy were discarded, resulting in a concise and informative set

Table 4. Some harvested Part-Whole noun pairs (sense numbers omitted).

Winston's Relation	Seed	PART-WHOLE noun pairs collected
Component-Integral	engine	$< camshaft, engine >, < arrester, attackaircraftcarrier >,$ $< odometer, automotivevehicle >, ...$
Member-Collection	ship	$< bay, ship >, < commode, lavatory >,$ $< boilerplate, steamboiler >, < davit, ship >, ...$
Portion-Mass	metre	$< metre, decameter >, < angstrom, nanometer >,$ $< adenine, dna >, < mebibyte, gb >, ...$
Material-Object	grape	$< grape, grapevine >,$ $< dimocarpus\_longan, genus\_dimocarpus >, ...$
Place-Area	USA	$< last\_frontier, united\_states >,$ $< empire\_state, united\_states >, < st.\_lawrence, united\_states >, ...$

of 580 PART-WHOLE concept pairs, such as $< military\_academy, agency >, < power\_brake, selfpropelled\_vehicle >, < letter, bicameral\_script >, < edible\_fruit, alcohol >, < metric\_linear\_unit, metric\_linear\_unit >, < partition, vessel >, < stroke, table\_game >, < propeller, internal - combustion\_engine >$,....These PART-WHOLE concept pairs enforce semantic selectional restrictions on lexico-syntactic patterns by constraining the possible contexts within which these patterns encode meronymy.

Our candidate meronymy pattern acquisition procedure begins by syntactically parsing around 25% of the English Wikipedia corpus ($\approx$ 140M words and 10M sentences), provided by the University of Amsterdam [21], using the Stanford parser [22]. We saved both the phrase structure tree and the results of the dependency analysis, from which we extracted meronymy-indicating patterns as the shortest dependency paths between instances of our previously acquired PART-WHOLE concepts. This procedure yielded 11351 individual lexico-syntactic expressions together with their co-occuring PART-WHOLE pairs that restrict the contexts in which these expression convey meronymy.

The most frequent meronymy-indicating pattern was *"+pobj+with+prep+shoulders+nsubj$< -led - >$prep+in+pobj+parts+prep+of+pobj+"* which co-occurred with instances of types *"homo#n#2"* (such as *"world#n#8"*) and instances of type *"group#n#1"* (such as *"people#n#1"*) 355 times [1]. The next most frequent pattern was the ambiguous genitive *"of"*. It occurred 91 times within different meronymy-indicating concept pairs, such as $< feature\#n\#2, external\_body\_part\#n\#1 >$, for example in *"temple of the head"*, which indeed conveys meronymy. This corroborates with the findings of [1], who also observed that genitives were the most frequent (and most ambiguous) meronymy patterns. Around 95% of all the acquired patterns had a frequency of one. These typically were domain-specific meronymy patterns, such as *"+poss+capital+nsubj$< -fell- >$ prep+After+pobj+days+prep+of+pobj+raids+nn+"*

[1] We later found that the high count for this pattern was due to its repetitive occurrence within a single segment of the corpus, illustrating a case of Wikipedia spamming.

that occured in the contexts of the PART-WHOLE concept pair $< chemical\_element\#n\#1, gas\#n\#2 >$.

We evaluated the patterns, P, thus acquired by measuring their precision in expressing meronymy when their contexts are constrained by (i.e.: they subcategorize) certain PART-WHOLE concept pairs. We parsed an additional 130K sentences from the Wikipedia corpus and extracted 88219 lexico-syntactic patterns together with their co-occurring noun pairs as our test set. For each pattern P and its PART-WHOLE concept pairs, $< Part\_Concept, Whole\_Concept >$, we searched for its occurrences, P_{test}, in the test set. Only those P_{test} co-occurring with nouns m_1 and m_2 that are respectively instances of PART_CONCEPT and WHOLE_CONCEPT were considered. If m_1 and m_2 are meronyms or holonyms pairs, we incremented a count *P.true_positive* to reflect that the lexico-syntactic pattern P indicates meronymy when it occurs within restricted contexts consisting of instances of PART_CONCEPT and WHOLE_CONCEPT. If m_1 and m_2 are not meronyms or holonyms, we incremented a count *P.false_positive*, indicating that P may not always indicate meronymy within the restricted contexts of instances of PART_CONCEPT and WHOLE_CONCEPT. We manually determined whether m_1 and m_2 are meronyms or holonyms instead of relying (automatically) on WordNet since, as shown before, WordNet is sparse and does not document many valid meronymy or holonymy relations that our approach discovers. The precision of a pattern P occurring within the context c (defined by a PART-WHOLE concept pair) is

$$Precision(P_c) = \frac{P.true\_positive}{P.true\_positive + P.false\_positive} \tag{4}$$

We do not compute recall as we are unaware beforehand of the actual number of meronymy patterns in the corpus. Table 5 illustrates evaluating the precision with which the genitive pattern *"of"* encodes meronymy when its contexts are constrained by instances of specific PART-WHOLE conceptual classes. Noun pairs marked with "*" are meronyms that our technique identified but that are not documented in WordNet, while those marked with "**" are false positives, i.e. we wrongly identified them as possible contexts within which *"of"* conveys meronymy. The precision of all the evaluated patterns is given in Table 6.

Table 5. Evaluating the precision of the genitive pattern *"of"* in the test set.

PART: *Person#n#1* WHOLE: *People#n#1* Example: *(warrior#n#1, nation#n#2)*; (inhabitant#n#1,world#n#5)*; (leader#n#1, business#n#8)*; (owner#n#2, land#n#8)*; (creditor#n#1 land#n#8)*** Total Count:17, $Precision = \frac{16}{17} = 0.94$
PART: *Point#n#6* WHOLE: *Time_Interval#n#1* Example: *(beginning#n#2, period#n#2)*; (start#n#2,round#n#2)*; (start#n#2, period#n#2)*; (end#n#2, period#n#2)** Total Count:4, $Precision = \frac{4}{4} = 1$

Table 6. Precision of test patterns in encoding meronymy when constrained by respective contexts.

Lexico-Syntactic Pattern	*Part Whole* Contexts	Precision
"pobj+of+prep" ("of")	*Part#n#2 Object#n#1*	1
	Room#n#1 Housing#n#1	1
	Administrative_District#n#1 Administrative_District#n#1	1
	Object#n#1 Object#n#1	1
	Point#n#6 Time_Interval#n#1	1
	Concept#n#1 Concept#n#1	0.94
	Tract#n#1 Structure#n#1	1
	Room#n#1 Dwelling#n#1	1
	Person#n#1 People#n#1	0.94
	Room#n#1 Structure#n#1	1
	Area#n#6 Document#n#1	1
	Facility#n#1 Store#n#2	1
"pobj+in+prep" ("in")	*Structure#n#1 Structure#n#1*	1
	Room#n#1 Structure#n#1	0.5
	Administrative_District#n#1 Administrative_District#n#1	1
	Time_Unit#n#1 Time_Unit#n#1	0.67
	Region#n#3 Administrative_District#n#1	1
	Area#n#5 Structure#n#1	0.75
	Body#n#1 Organism#n#1	1
	Collection#n#1 Natural_Object#n#1	1
	Tract#n#1 Tract#n#1	0.5
"pobj+in+prep+located+ "partmod" ("located in")	*Administrative_District#n#1 Administrative_District#n#1*	1
	Structure#n#1 Structure#n#1	1
"pobj+as+prep" ("as")	*Room#n#1 Structure#n#1*	1
	Object#n#1 Object#n#1	0.5
"pobj+at+prep" ("at")	*Structure#n#1 Structure#n#1*	1
	Area#n#5 Structure#n#1	1
"pobj+by+prep" ("by")	*Region#n#3 Administrative_District#n#1*	1
"pobj+for+prep" ("for")	*Administrative_District#n#1 Administrative_District#n#1*	1
"pobj+on+prep" ("on")	*Structure#n#1 Structure#n#1*	1
	Area#n#6 Document#n#1	1
	Geological_Formation#n#1 Geological_Formation#n#1	1
"pobj+with+prep" ("with")	*Body#n#2 Educational_Institution#n#1*	1

Our results indicate that in 88% of cases, the patterns learnt together with their respective PART-WHOLE concept pairs, which restrict the contexts within which they participate in part-whole relations, expressed meronymy with 100% precision. As example, the pattern *"of"*, which as indicated in Table6, encodes meronymy when its arguments are instances of the semantic classes *"Point#n#6"* and *"Time_Interval#n#1"*, expresses a part-whole relation in *"start of period"*. The high precision obtained validates our underlying hypothesis that semantic features of the PART and WHOLE instances (in the form of their respective PART-WHOLE conceptual classes) participating in meronymy relations facilitate the identification and disambiguation of these relations by constraining the set of allowable contexts within which they encode meronymy. To confirm these findings, we repeated our experiments, but without abstracting the PART and WHOLE noun pairs (Section 4.1) to their semantic classes. We search for the exact occurences of these noun pairs in the Wikipedia corpus, and extracted potential meronymy-encoding lexico-patterns that relate these noun pairs (Section 4.3). As could be expected, precision was higher, around 90%, but our coverage was much less. The smaller relative gain in precision compared to the much substantial reduction in coverage, did not outweigh the benefits gained by exploiting the semantic features of PART and WHOLE noun pairs. Furthermore, defining meronymy at the class (concept)-level allowed the discovery of meronym noun pairs that are not defined in WordNet, such as *(warrior#n#1, nation#n#2),(room#n#1,level#n#8), (center#n#2,field#n#2), (section#n#4,film#n#4),(beginning#n#2,period#n#2) ,(section#n#6,law#n#3), (division#n#2,property#n#4), (center#n#2,platform#n#4), (site#n#1,area#n#5), (site#n#1,erection#n#2)* . We also observed that semantic selectional restrictions do not suffice to constrain some contexts within which patterns participate in meronymy. One such context is defined by the PART-WHOLE concept pair *Person#1* and *People#1* . Ambigous patterns occurring with instances of these concepts might not always express meronymy, such as in *"creditor of land"*.

6 Conclusion and Future Work

We presented an unsupervised approach to automatically acquire high-quality meronymy patterns from texts. Our approach considers meronymy as a specialized semantic frame, depicting a part-whole relation between instances playing a PART and a WHOLE role. To disambiguate polysemous patterns that encode meronymy only within certain contexts, we rely on the linking theory to posit that semantic features of the PART-WHOLE instances could facilitate the identification of lexico-syntactic patterns encoding meronymy. We abstract instances to their hypernyms, enforcing semantic selectional restrictions to constrain the contexts within which the patterns participate in meronymy. These meronymy-encoding contexts are the crux of our disambiguation procedure, in which we extend Harris' hypothesis to postulate that similar patterns share similar contexts. To validate our approach, we measured the precision of the acquired patterns in expressing meronymy over a test corpus. Our evaluation results indicate that 88% of the learnt patterns expressed meronymy with 100% precision when constraints were applied to restrict the contexts within which they participate in part-whole relations. This highlights the major contribution of our methodology in the identification and disambiguation of meronymy-indicating patterns. Furthermore, our unsupervised approach circumvents the need for annotated training data and manual intervention, as opposed to those based on supervised learning. Also, our definition of meronymy at the conceptual level enabled the identification of part-whole noun pairs that are not documented in WordNet as participating in meronymy. By expanding an initial seed set, the technique we present does not suffer from low coverage, as do previous related studies.

As future work, we will address the issue of obtaining the *basic level category* [24], i.e. most informative hypernym (semantic class), for a given Part or Whole noun, instead of merely abstracting them to the immediate WordNet parent in the hypernymy hierarchy. As example, the basic level category of *"robin"* is *"bird"*, which is more informative than the hypernym *"thrush"*. Another improvement could be reducing the reliance on WordNet (or basic level categories) in abstracting nouns to their concepts by using Wikipedia instead. We also intend to validate our findings by calculating the inter-annotator agreement score between judges who have to decide whethe a given pattern encodes meronymy when appearing within a particular context.

Acknowledgments. Ashwin Ittoo is a PhD candidate under the IOP-IPCR program of the Dutch Ministry of Economic Affairs.

References

1. Girju, R., Badulescu, A., Moldovan, D.: Automatic discovery of part-whole relations. Comput. Linguist. **32**(1) (2006) 83–135
2. Girju, R., Badulescu, A., Moldovan, D.: Learning semantic constraints for the automatic discovery of part-whole relations. In: NAACL '03: Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology, Morristown, NJ, USA, Association for Computational Linguistics (2003) 1–8
3. Berland, M., Charniak, E.: Finding parts in very large corpora. In: Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics, Morristown, NJ, USA, Association for Computational Linguistics (1999) 57–64
4. Fillmore, C.: Frame semantics and the nature of language. Annals of the New York Academy of Sciences **280** (1976) 20–32
5. Gildea, D., Jurafsky, D., Gildea, D., Jurafsky, D.: Automatic labeling of semantic roles. Computational Linguistics **28** (2002) 245–288
6. Harris, Z.S.: Mathematical Structures of Language. Wiley, New York, NY, USA (1968)
7. Fellbaum, C., ed.: WordNet An Electronic Lexical Database. The MIT Press, Cambridge, MA ; London (May 1998)

8. Nguyen, D.P.T., Matsuo, Y., Ishizuka, M.: Exploiting Syntactic and Semantic Information for Relation Extraction from Wikipedia. IJCAI Workshop on Text-Mining & Link-Analysis (TextLink 2007) (2007)
9. Suchanek, F., Kasneci, G., Weikum, G.: YAGO–a core of semantic knowledge. In: Proc. Int. Conf. on World Wide Web. (2007) 697–706
10. Winston, M.E., Chaffin, R., Herrmann, D.: A taxonomy of part-whole relations. Cognitive Science **11**(4) (1987) 417–444
11. Cruse, D.: Lexical Semantics. Cambridge University Press, Cambridge, UK (1986)
12. Simons, P.: Parts: A Study in Ontology. Routledge (1987)
13. Simons, P.: Part/whole ii: Mereology since 1900. In Burkhardt H., S.B., ed.: Handbook of Metaphysics and Ontology. (1991) 261–288
14. Varzi, A.: Mereology. In Zalta, E., ed.: The Stanford Encyclopedia of Philosophy. (2004)
15. Keet, M.: Introduction to part-whole relations: mereology, conceptual modelling and mathematical aspects. Technical report, Free University of Bozen-Bolzano, Italy (2006)
16. Artale, A., Franconi, E., Guarino, N., Pazzi, L.: Part-whole relations in object-centered systems: An overview (1996)
17. Iris, M.A., Litowitz, B.E., Evens, M.: Problems of the part-whole relation. In Evens, M.W., ed.: Relational Models of the Lexicon: Representing Knowledge in Semantic Networks. Cambridge University Press, Cambridge (1988) 261–288
18. Hearst, M.: Automated discovery of wordnet relations. In Fellbaum, C., ed.: WordNet: An Electronic Lexical Database and Some of its Applications. MIT Press (1998)
19. Hearst, M.: Automatic acquisition of hyponyms from large text corpora. In: Proceedings of the 14th International Conference on Computational Linguistics, Nantes, France (1992)
20. Snow, R., Jurafsky, D., Ng, A.: Learning syntactic patterns for automatic hypernym discovery. In: Proceedings of the 17th Conference on Advances in Neural. (2005)
21. http://ilps.science.uva.nl/WikiXML/: English wikipedia corpus
22. http://nlp.stanford.edu/software/lex parser.shtml: Stanford statistical parser
23. http://search.cpan.org/dist/WordNet QueryData/: Wordnet-querydata
24. Izquierdo, R., Surez, A., Rigau, G.: Exploring the automatic selection of basic level concepts. In: Proceedings of the International Conference on Recent Advances on Natural Language Processing (RANLP'07). (2007)