Part-Whole Relations

Meronymy refers to part-whole relations, which are ubiquitous.

Example from [Girju et al., 2003]:

> The car's mail messenger is busy at work in the **mail car** as the train moves along. Through the open **side door** of the car, moving scenery can be seen. The worker is alarmed when he hears an unusual sound. He peeks through the door's **keyhole** leading to the tender and locomotive **cab** and sees the two bandits trying to break through the express **car door**.

Part-of (mail car, train)	*Part-of* (cab, locomotive)
Part-of (side door, car)	*Part-of* (door, express car)
Part-of (keyhole, door)	*Part-of* (car, express [train])

Types of Part-Whole Relations

WordNet includes three types of part-whole relations:

Member Of	*MemberOf* (UK, NATO)
Stuff Of	*StuffOf* (carbon, coal)
General Part Of	*PartOf* (leg, table)

Researchers have also identified additional types, such as:

Portion-Mass	meter, kilometer
Feature-Activity	paying, shopping
Place-Area	oasis, desert

Finding Parts in Very Large Corpora

- [Berland & Charniak, ACL 1999] published an early study using statistical methods to produce word lists that represent "parts" of a object.
- Motivated by Hearst's hyponym patterns, they identified meronym patterns and instantiated them with a target object (as the "whole").
- Probabilistic measures were then used to rank the words that occurred in these patterns.
- The top 50 proposed words for 6 objects (*book, building, car, hospital, plant, school*) were judged to be 55% accurate.

Top Learned Words

Book: author, subtitle, co-author, foreword, publication, epigraph, co-editor, cover, copy, page, title, authorship, manuscript, chapter, epilogue, publisher, jacket, subject, double-page, sale, excerpt

Building: rubble, floor, façade, basement, roof, atrium, exterior, tenant, foortop, wreckage, stairwell, shell, demolition, balcony, hallway, renovation, janitor, rotunda, entrance, hulk, wall, ruin

Car: trunk, windshield, dashboard, headlight, wheel, ignition, hood, driver, radiator, shifter, occupant, brake, vent, fender, tailpipe, bumper, pipe, airbag, seat, speedometer, converter, backseat

Hospital: ward, radiologist, trograncic, mortuary, hopewell, clinic, aneasthetist, ground, patient, floor, unit, room, entrance, doctor, administrator, corridor, staff, department, bed, pharmacist, director

Learning Semantic Constraints for Meronomy

[Girju et al., HLT-NAACL 2003]

- [Girju et al., 2003] developed a technique to identify instances of part-whole relations in text.
- They use lexico-syntactic meronym patterns to locate potential instances of part-whole relations.
- They use supervised machine learning to automatically learn semantic constraints that can be applied to the nouns in the patterns.
- The learned semantic constraints produce good accuracy for identifying part-whole instances that occur in three specific patterns.

Lexico-Syntactic Meronymy Patterns

- Explicit, unambiguous part-whole constructions

  *The substance **consists of** two ingredients.*

  *The cloud **was made of** dust.*

  *Iceland **is a member of** NATO.*

- Explicit, ambiguous part-whole constructions

  *The horn **is part of** the car.*

  *The car **has** a horn.*

- Implicit part-whole constructions (ambiguous)

  *girl**'s** mouth, eyes **of** the baby, door knob*

Identifying Lexico-Syntactic Patterns

Lexico-syntactic patterns were extracted from 20,000 sentences to see what expressions are most common.

1) Selected 100 part-whole pairs from WordNet and extracted sentences that contain both concepts.
2) Manually inspected the sentences and retained only those where the pair refers to meronymy.
3) Extracted lexico-syntactic expressions that link the two concepts.

The Extracted Part-Whole Patterns

- 535 part-whole occurrences were found.
- 493 (92%) were phrase-level patterns; 36 distinct patterns. The most frequent:

  NP1 of NP2 : 173 times (35%)

  NP1 ‘s NP2 : 71 times (14%)

- 42 (8%) were sentence-level patterns; 18 distinct. The most frequent:

  NP1 verb NP2 : 18 times (43%)

Learning Semantic Constraints

- The general approach is to automatically learn semantic constraints for the nouns in meronymy patterns.
- WordNet's semantic hierarchy is used as the source for semantic categories.
- The C4.5 decision tree machine learning algorithm learns rules to decide whether a pair of semantic classes is likely to be in a part-whole relation.
- Training data consisted of:

  34,609 positive NP pairs (from manual annotations + WordNet)

  46,971 negative NP pairs (from manual annotations)

Generalizing the Training Examples

- For each NP pair, generalize the words to their semantic classes and represent as triples with the class label (*yes* if meronymy, *no* if not meronymy). For example:

  <aria#1 ; opera#1; yes> → *<entity#1 ; abstraction#1; yes>*

- Group the semantic class pairs based on whether they are all **Positive** examples, **Negative** examples, or **Ambiguous**. For example:

  <apartment#1 ; woman#1; no> → *<entity#1 ; entity#1; no>*

  <hand#1 ; woman#1; yes> → *<entity#1 ; entity#1; yes>*

  The unambiguous examples are used for learning.

Learning from Unambiguous Examples

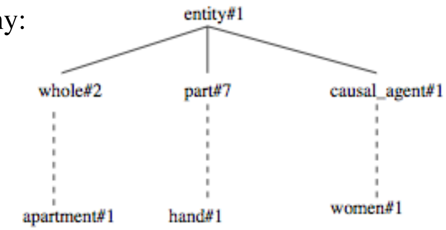
- The C4.5 decision tree algorithm is applied to the unambiguous examples and rules are extracted from the learned decision tree.
- There are two features, *part* and *whole*, and the values are the semantic classes.
- Using 10-fold cross-validation, 10 sets of rules were learned.
- All of the learned rules were ranked based on frequency and average accuracy. Rules that occurred in at least 7 of the 10 sets with accuracy > 50% were selected.

Specializing the Ambiguous Examples

- The ambiguous examples are gradually specialized until they are unambiguous.
- Initially, each semantic class is a root node in WordNet. Each class is specialized by replacing it with its first hyponym.
- If the first attempt at specialization still produces ambiguous results, then it is further specialized.
- The specialization process stops when there is no ambiguity or when it can't be specialized any further.

Specialization Example

Section of WordNet hierarchy:



<apartment#1, women#1, no> → *<entity#1, #entity#1, no>*
<hand#1, women#1, yes> → *<entity#1, #entity#1, yes>*

→ *<whole#1, #causal_agent#1, no>*
→ *<part#1, #causal_agent#1, yes>*

Constraint Learning Flowchart

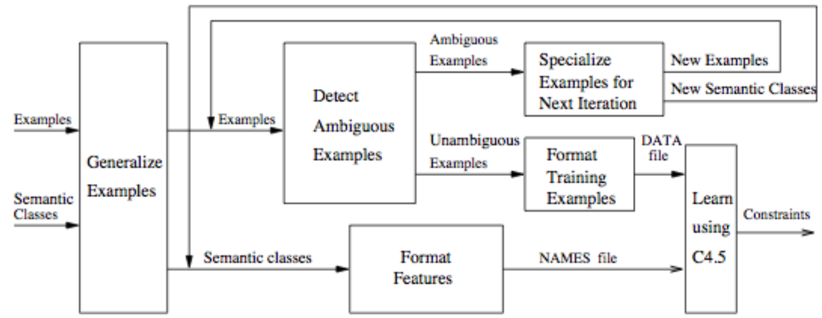


Figure 2: Architecture of the constraint learning procedure.

Evaluation Results

- Accuracy was measured based on manual validation of a test set containing pairs from the 3 meronym patterns.

  Recall = 117/119 (98.3%)

  Precision = 117/140 (83.6%)

- An additional 43 manner relations occurred in different contexts. Including them would give 72% recall.
- Errors were primarily attributed to the verb "have" and genitives ('s), which are very ambiguous in the relations that they can convey.

Summary of Results

Number of Relations	Y verb X	Y's X	X of Y	All patterns
Number of patterns	280	225	962	1467
Number of correct relations	18	23	78	119
Number of relations retrieved	25	24	91	140
Number of correctly retrieved relations	18	22	77	117
Precision	72%	91.16%	84.61%	83.57%
Recall	100%	95.65%	98.71%	98.31%

Examples of Learned Constraints

object#1	social_event#1	1	69.84	9	*scene#4 - movie#1*
whole#2	social_event#1	1	63.00	7	*sequence#3 - movie#1*
entity#1	**group#1**	**1**			*academician#1 - academy#2* *king#1 - royalty#1*
location#1	people#1	0	50.00	8	*section#3 - nation#1*
organism#1	system#1	0	50.00	8	*archbishop#1 - York#1*
group#1	**group#1**	**1**			*military_reserve#1 - military_unit#1* *amoebida#1 - genus_amoeba#1*
collection#1	arrangement#2	0	92.60	10	*data#1 - table#1*
arrangement#2	social_group#1	0	85.70	10	*hierarchy#1 - church#1*
system#1	collection#1	0	85.70	10	*economy#1 - selection#2*
entity#1	**entity#1**	**1**			*door#4 - car#4* *point#15 - knife#1*
point#2	object#1	0	89.55	10	*place#1 - wall#2*
location#1	object#1	1			*base#16 - box#1*
geographic_area#1	instrumentality#3	0	80.75	8	*graveyard#1 - ship#1*
person#1	person#1	0	89.55	10	*child#1 - woman#1*
object#1 !substance#1 !natural_object#1	organism#1 !plant#2 !animal#1	0			*desk#1 - man#1* - No *feather#1 - bird#1* - Yes *blade#1 - grass#1* - Yes *body#1 - man#1* - Yes

Summary

- Meronymic relations are common and difficult to recognize accurately based solely on contextual patterns.
- Applying semantic constraints to noun pairs can yield good accuracy, and semantic constraints can be learned.
- Both positive constraints (meronymic pair) and negative constraints (not meronymic pair) were useful.
- However, this work only showed the benefit of semantic constraints applied to a few meronymic patterns. Finding all part-whole relations is still an open problem!