Semantic and ontological coherence of lexical resources
Towards practical methods

Laure Vieu
IRIT-CNRS & LOA-ISTC-CNR
with the contribution of Nervo Verdezoto, Alessandro Oltramari & Ekaterina Ovchinnikova

Workshop Senso Comune, 23 marzo 2011, Roma

The issue

- Lexical resources such as WordNet and FrameNet are widely used nowadays
- Not only local information like synonymy or argument structure, but also the overall structure, especially WN's noun taxonomy
  - Increasingly used for standard NLP tasks (e.g., similarity measures)
  - Reasoning tasks: recognizing textual entailment, question-answering...
  - Even as ontologies in AI applications
- Yet it is well-known that the structural information in lexical resources is far from perfect

The issue

- For WN's noun taxonomy, all repair proposals beyond the top-level are based on the manual inspection of the 82,155 synsets, and require cross-validation among experts
  - alignment with an ontology (e.g., SUMO)
  - annotation with features
  - OntoClean methodology

The issue

- For WN's noun taxonomy, all repair proposals beyond the top-level are based on the manual inspection of the 82,155 synsets, and require cross-validation among experts
  - alignment with an ontology (e.g., SUMO)
  - annotation with features
  - OntoClean methodology
- Couldn't some more focused and less time-consuming and expertise-requiring improvement methods be conceived?

The idea: semi-automatic repair procedure

- Automatically search for incoherences to focus the attention of the experts on problems

The idea: semi-automatic repair procedure

- Automatically search for incoherences to focus the attention of the experts on problems
- Two experiences: one on FN, one on WN
  - FN: Check on annotated corpus for textual entailment if FN enables inferences or not (17%-83%)
  - WN: Contrast two sources of knowledge and use semantic and ontological constraints to look for incoherences

The idea: semi-automatic repair procedure

- Automatically search for incoherences to focus the attention of the experts on problems
- Two experiences: one on FN, one on WN
  - FN: Check on annotated corpus for textual entailment if FN enables inferences or not (17%-83%)
  - WN: Contrast two sources of knowledge and use semantic and ontological constraints to look for incoherences
- Example developed here: contrast WN's noun taxonomy and meronymy data

Outline

- The method
  Five tests based on semantic and ontological constraints
- Analysis of taxonomy errors and types of repairs
  Classical ontological confusions yield standard solutions
- Conclusion

The method: data set constitution

- Need for aligned data sources. Here, data annotated with WN's synset id (sense key)
- Data source1: extract pairs of synsets linked by a parthood relation from
  - WordNet's meronymy relations (22,187 pairs)
  - corpora annotated with parthood relations *and* WN sense keys of its arguments
    SemEval 2007 Task 4 training and test datasets (89 pairs)
- Data source2: enrich these pairs with, for each synset, an instance or class tag and its hypernymy chain up to the top-category in WN taxonomy

Example pair

```
<pair relationOrder="(e1, e2)" comment="meronym_part" source="WordNet-3.0">
    <e1 synset="head%1:06:04" isInstance="No">
        <hypernym>
            {obverse%1:06:00}...{surface%1:06:00}...{artifact%1:03:00 }...
            {physical_object%1:03:00}{entity%1:03:00}
        </hypernym>
    </e1>
    <e2 synset="coin%1:21:02" isInstance="No">
        <hypernym>
            ...{metal_money%1:21:00}{currency%1:21:00}...{quantity%1:03:00}
            {abstract_entity%1:03:00}{entity%1:03:00}
        </hypernym>
    </e2>
</pair>
```

The method: searching for incoherences exploiting semantic and ontological constraints-1

- Standard parthood
  - √ Individual-individual $\langle a, b \rangle$: $P(a, b)$
    $\langle balthazar\%1:18:00, magi\%1:14:00 \rangle$
- Semantics of the *derived* meronymy relation
  - √ Class-class $\langle A, B \rangle$:
    $\forall x(A(x) \rightarrow \exists y(B(y) \wedge P(x, y))) \vee \forall y(B(y) \rightarrow \exists x(A(x) \wedge P(x, y)))$
    $\langle roof\%1:06:00, building\%1:06:00 \rangle$
    $\langle handle\%1:06:00, umbrella\%1:06:00 \rangle$
  - √ Class-individual $\langle A, b \rangle$: $\forall x\ A(x) \rightarrow P(x, b)$
    $\langle sura\%1:10:00, koran\%1:10:00 \rangle$
  - × Individual-Class $\langle a, B \rangle$: $\forall y\ B(y) \rightarrow P(a, y)$
    $\langle gospel\text{-}according\text{-}to\text{-}mark\%1:10:00, new\text{-}testament\%1:10:00 \rangle$
- Test0: no individual can be a meronym of a class

The method: searching for incoherences exploiting semantic and ontological constraints-2

Semantics of the parthood relation imposes an ontological homogeneity between part and whole

- Three top-level categories, based on a (very rough) mapping between DOLCE top-level and WN 3.0 top-level:
  - endurants (ED) or physical entities: a dog, a table, some smoke
  - perdurants (PD) or eventualities: a talk, a sleep, a downpour
  - abstract entities (AB): a number, a text's contents
- Three incoherence tests spotting ontological heterogeneity:
  - Test1: $\langle$ED,AB$\rangle$ or $\langle$AB,ED$\rangle$
  - Test2: $\langle$ED,PD$\rangle$ or $\langle$PD,ED$\rangle$
  - Test3: $\langle$PD,AB$\rangle$ or $\langle$AB,PD$\rangle$

Example of ontological heterogeneity extracted with Test1

⟨ED - physical object, AB - abstract entity⟩

```
<pair relationOrder="(e1, e2)" comment="meronym_part" source="WordNet-3.0">
    <e1 synset="head%1:06:04" isInstance="No">
        <hypernym>
            {obverse%1:06:00}...{surface%1:06:00}...{artifact%1:03:00 }...
            {physical_object%1:03:00}{entity%1:03:00}
        </hypernym>
    </e1>
    <e2 synset="coin%1:21:02" isInstance="No">
        <hypernym>
            ...{metal_money%1:21:00}{currency%1:21:00}...{quantity%1:03:00}
            {abstract_entity%1:03:00}{entity%1:03:00}
        </hypernym>
    </e2>
</pair>
```

The method: searching for incoherences exploiting semantic and ontological constraints-3

Semantics of "Member" parthood implies that the whole has to be a collection, i.e., an hyponym or an instance of *group%1:03:00*.

- Test4

  examples

  - ⟨*altair%1:17:00*, *aquila%1:17:00*⟩
    aquila%1:17:00 is an instance of *constellation%1:17:00* which is not a group
    (NB: *galaxy%1:14:00* is a group)
  - ⟨*ethiopian%1:18:00*, *ethiopia%1:15:00*⟩
    ethiopia%1:15:00 is an instance of *country%1:15:00*, the "location" sense of country
    (NB: *country* also has a "people" sense, but *Ethiopia* has only one sense)

Quantitative results

Number of pairs extracted by the tests from the 22,187 / 89 pairs

Test	WordNet		SemEval	
0	349	1.57%	0	0%
1	163	1.62%	2	2.78%
2	45	0.45%	2	2.78%
3	108	1.07%	0	0%
4	550	4.47%	7	7.87%

(% calculated over the Member pairs (12,293 in WN) in Test 4)

After manual inspection, all pairs indeed point at an error of some sort, or even several errors

The method: guided manual error analysis

Incoherence between two data sources (here meronymy and taxonomy hierarchies) may be caused by an error in *either source*

- Taxonomy errors: all previous examples
- Meronymy errors
  - confusion with other relations
    - "is located in"
      ⟨*balkan-wars%1:04:00*, *balkan-peninsula%1:15:00*⟩ (Test 2)
    - "participates in"
      ⟨*feminist%1:18:00*, *feminist-movement%1:04:00*⟩ (Test 2)
    - "is a quality of"
      ⟨*personality%1:07:00*, *person%1:03:00*⟩ (Test 1),
      ⟨*regulation-time%1:28:00*, *athletic-game%1:04:00*⟩ (Test 3)
  - wrong synsets selected (small slips?)
    ⟨*seat%1:06:01*, *seating_area%1:06:00*⟩: chair sense of *seat* instead of its area sense, *seat%1:15:01* (Test 1)

Analysis of taxonomy errors

Three types of taxonomy errors occur:

- Instance-class confusion (Tests 0 and 4)
- Wrong hypernym for a synset (Tests 1-4)
- Missing sense of a word (all tests)

They reveal confusions that are often systematic and correspond to classical ontological confusions, among which:

- Confusion between class and group
- Confusion between an entity and a property of that entity
- Confusion between two senses of a polysemic word

Such regularities can be turned into guidelines for the manual analysis and repair of errors extracted by the tests

Confusion between class and group

- ⟨*gospel-according-to-mark%1:10:00*, *new-testament%1:10:00*⟩: instance, hyponym of *written-communication%1:10:00*
- ⟨*air-force-research-laboratory%1:06:00*, *us-air-force%1:14:00*⟩: instance, hyponym of group

$\Rightarrow$ General cure: transform these classes in instances

Confusion between class and group

- ⟨*gospel-according-to-mark%1:10:00*, *new-testament%1:10:00*⟩: instance, hyponym of *written-communication%1:10:00*
- ⟨*air-force-research-laboratory%1:06:00*, *us-air-force%1:14:00*⟩: instance, hyponym of group

⇒ General cure: transform these classes in instances

Issue with genera like *genus-australopithecus%1:05:00*: ⟨*lucy%1:05:00*, *genus-australopithecus%1:05:00*⟩, instance of *australopithecus-afarensis%1:05:00*, hyponym of group

- *genus-australopithecus%1:05:00* should be an instance of *genus%1:14:00*, so not a class, and not instance of group either
- but which link with class *australopithecus-afarensis%1:05:00* (hyponym of *organism%1:03:00*)?
- and which link with instances of this class (e.g. *lucy%1:05:00*)?

Confusion between an entity and a property of that entity or a relation involving it

Mostly Tests 1 and 3, confusion between `ED` or `PD` and `AB`

- Quantity or measure (`AB`)
  coin%1:21:02, *haymow%1:23:00*, *tear%1:08:01*, *helping%1:13:00* (and hyponyms *drumstick*, *fillet*, *sangria*...)
- Shape (`AB`)
  corolla%1:20:00, *mothball%1:06:00*
- Relation (`AB`)
  possession%1:03:00 (and hyponyms *credit-card*, *hacienda*...)

$\Rightarrow$ General cure: move the synset to another branch of the taxonomy

Confusion between two senses of a polysemic word

- "Dot objects" and systematic polysemy, not coherently handled with multiplication of senses or multiple inheritance
  - *book%1:10:00* only `ED`: 1 sense
  - *document%1:10:00* `AB` and *document%1:06:00* `ED`: 2 senses
  - *letter%1:10:00* both `AB` (hypernym *text%1:10:00*) and `ED` (hypernym *document%1:06:00*): multiple inheritance

Confusion between two senses of a polysemic word

- "Dot objects" and systematic polysemy, not coherently handled with multiplication of senses or multiple inheritance
  - *book%1:10:00* only `ED`: 1 sense
  - *document%1:10:00* `AB` and *document%1:06:00* `ED`: 2 senses
  - *letter%1:10:00* both `AB` (hypernym *text%1:10:00*) and `ED` (hypernym *document%1:06:00*): multiple inheritance
- Multiple senses not inherited by hyponyms or instances
  - *ethiopia%1:15:00* is an instance of *country%1:15:00*, the "location" sense of country, "people" sense missing
  - *rain%1:19:00* hyponym of *precipitation%1:19:00* (process, `PD`)
    no *rain* hyponym of *precipitation%1:23:00* (quantity, `AB`)
    no *precipitation* hypernym of *rain%1:27:00* (water fallen `ED`)

Confusion between two senses of a polysemic word

- "Dot objects" and systematic polysemy, not coherently handled with multiplication of senses or multiple inheritance
  - *book%1:10:00* only ED: 1 sense
  - *document%1:10:00* AB and *document%1:06:00* ED: 2 senses
  - *letter%1:10:00* both AB (hypernym *text%1:10:00*) and ED (hypernym *document%1:06:00*): multiple inheritance
- Multiple senses not inherited by hyponyms or instances
  - *ethiopia%1:15:00* is an instance of *country%1:15:00*, the "location" sense of country, "people" sense missing
  - *rain%1:19:00* hyponym of *precipitation%1:19:00* (process, PD)
    no *rain* hyponym of *precipitation%1:23:00* (quantity, AB)
    no *precipitation* hypernym of *rain%1:27:00* (water fallen ED)
- Issue with fictional entities (*psychological_feature%1:03:00*), abstract and yet similar to concrete entities: angels have "wings", Hell has "rivers" (Acheron)

Summing up

- Automatic extraction of errors possible
- Efficient method: tests are accurate
- Classical regular errors make it quite simple to produce repair guidelines
- Difficult ontological issues also show up

Looking forward

Refine and extend the method

- Extend the coverage to larger annotated corpus
- Refine the tests into subtests, exploiting observed regularities: ideally, one error type per test
- New tests to search for further incoherences moving down the top-level: e.g., location vs. physical objects in ED
- New test exploiting semantics of Substance type of meronomy (like Test 4 for Member)
- New tests and other corpora for other semantic relations

Exploit the lessons learned

- Tool for cleaning up Princeton WN for a next release
- Tool for assisting WN development for other languages
- Tool for assisting the development of any lexical resource?