# Cross-checking WordNet and SUMO Using Meronymy

**Javier Álvez, Itziar Gonzalez-Dios, German Rigau**
LoRea Group, IXA Group, IXA Group
University of the Basque Country (UPV/EHU)
javier.alvez@ehu.eus, itziar.gonzalezd@ehu.eus, german.rigau@ehu.eus

**Abstract**

We report on the practical application of a black-box testing methodology for the validation of the knowledge encoded in WordNet, SUMO and their mapping by using automated theorem provers. Our proposal is based on the part-whole information provided by WordNet, out of which we automatically create a large set of tests. Our experimental results confirm that the proposed system enables the validation of some pieces of information and also the detection of missing information or inconsistencies among these resources.

**Keywords:** Meronymy, Knowledge validation, Automated Theorem Proving

## 1. Introduction

Despite being created manually, knowledge resources such as WordNet (Fellbaum, 1998) and SUMO (Niles and Pease, 2001) are not free of errors and inconsistencies. Unfortunately, improving, revising and correcting such large knowledge bases is a never-ending task that has been mainly carried out manually. A few automatic approaches have also been applied focusing on checking certain structural properties on WordNet e.g. Daudé et al. (2003) and Richens (2008) or using automated theorem provers on SUMO e.g. Horrocks and Voronkov (2006) and Álvez et al. (2012). Just a few more have studied automatic ways to validate the knowledge content encoded in these resources by cross-checking them. For instance, Álvez et al. (2008) exploit the EuroWordNet Top Ontology (Rodríguez et al., 1998) and its mapping to WordNet for detecting many ontological conflicts and inconsistencies in the WordNet nominal hierarchy.

In Álvez et al. (2017), we proposed a method for the automatic creation of *competency questions* (CQs) (Grüninger and Fox, 1995), which enabled to evaluate the competency of SUMO-based ontologies. Our proposal was based on several predefined question patterns (QPs) that were instantiated using information from WordNet and its mapping into SUMO (Niles and Pease, 2003). In addition, we described an application of automated theorem provers (ATPs) for the automatic evaluation of the proposed CQs. This proposal was used in Álvez and Rigau (2018) for a preliminary validation of WordNet, SUMO and their mapping.

The main contribution of this paper is a proposal for the automatic validation of WordNet, SUMO and their mapping. For this purpose, we create a new set of CQs that is obtained on the basis of the part-whole data of WordNet and which is an improved version of the set proposed in Álvez and Rigau (2018). By means of our proposal, we demonstrate the practical capabilities of the method introduced in Álvez et al. (2017) for the automatic detection of semantic agreements and inconsistencies in large-scale knowledge resources. For example, the knowledge encoded in the WordNet relations *part*($elementary\_particle^1_n$,$atom^1_n$), *member*($national^1_n$,$country^3_n$) and *substance*($cartilage^1_n$,$cartilaginous\_structure^3_n$) can also be inferred from SUMO. On the contrary, according to our interpretation of the meronymy relations of WordNet, the knowledge in the relations *part*($cell^2_n$,$cell\_nucleus^1_n$) and *substance*($grape^1_n$,$wine^1_n$) is incompatible with SUMO. In addition, our proposal enables the detection of missing knowledge: for instance, WordNet relates $waist^1_n$ and $torso^1_n$ by *part*, but the resulting conjectures are not proved to be entailed by SUMO in our experiments using ATPs. By a manual inspection of the ontology, we discover that this issue is due to the fact that the SUMO concepts connected to $waist^1_n$ and $torso^1_n$ according to the mapping between WordNet and SUMO —*Waist* and *Torso* respectively— are not related in SUMO.

*Outline of the paper*. In the following three sections, we introduce WordNet (Section 2), SUMO (Section 3), and their mapping (Section 4). Then, we describe the process for the creation of CQs in Section 5. Next, we report on and discuss our evaluation results in Sections 6 and 7 respectively. Finally, we conclude in Section 8.

## 2. Meronymy Relations in WordNet

In WordNet, meronymy —the part-whole relation— holds between synsets like $backrest^1_n$ and $seat^1_n$ (i.e. parts) and $chair^1_n$ (i.e. whole). Parts are inherited from their superordinates: if a chair has a seat, then an armchair has a seat as well. But parts are not inherited "upward" as they may be characteristic only of specific kinds of things rather than the class as a whole: chairs and kinds of chairs have a seat, but not all kinds of furniture have a seat.

There are 3 main meronymy relations in WordNet v3.0 that relate noun synsets: i) *part*, the general meronymy relation; ii) *member*, which relates particulars and groups; and iii) *substance*, which relates physical matters and things. In total, there are 22,187 (ordered) meronymy synset pairs (around a %10 of the synset pairs in WordNet): 9,097 pairs using *part*, 12,293 pairs using *member* and 797 pairs using *substance*. For example, the synsets $police\_officer^1_n$ and $police\_force^1_n$ are related by *member*, while $grape^1_n$ and $wine^1_n$ are related by *substance*.

## 3. SUMO

SUMO[1] has its origins in the nineties, when a group of engineers from the IEEE Standard Upper Ontology Working Group pushed for a formal ontology standard. Their goal was to develop a standard upper ontology to promote data interoperability, information search and retrieval, automated inference and natural language processing.

Currently, SUMO consists of about 20,000 terms and about 70,000 axioms organized in several levels. In the upper two levels —*Top* and *Middle* levels— the concepts, relations and axioms that are meta, generic or abstract can be found. From now on, we refer to the upper two levels of SUMO as its *core*. On the basis of these two levels, concepts that are specific to particular domains are in the so-called *domain* ontologies. Adimen-SUMO (Álvez et al., 2012) is obtained by means of a suitable transformation of the knowledge in the core of SUMO into First-Order Logic (FOL), which enables its use by FOL ATPs such as Vampire (Kovács and Voronkov, 2013) and E (Schulz, 2002).

The knowledge in SUMO is organized around the notions of *individuals* and *classes* —the main SUMO concepts. These concepts are respectively defined in Adimen-SUMO by means of the *meta*-predicates *$instance* and *$subclass*. SUMO objects and classes are not disjoint, since every SUMO class is defined to be instance of *Class*, and thus every SUMO class is also a SUMO object. Additionally, SUMO differentiates *relations* and *attributes*. In particular, SUMO distinguishes between *individual* relation and attributes —that is, instances of the SUMO classes *Relation* and *Attribute* respectively— and *classes* of relations and attributes —that is, subclasses of the SUMO classes *Relation* and *Attribute* respectively. SUMO provides specific predicates for dealing with relations and attributes. Some of the most important ones are *subrelation*, *subAttribute*, $holds^k$ and *attribute*. For example, in the next SUMO axiom the predicate *attribute* is used for the characterization of *subAttribute*:

```
(forall (?ATTR1 ?ATTR2)                         (1)
  (=>
    (subAttribute ?ATTR1 ?ATTR2)
    (forall (?OBJ)
      (=>
        (attribute ?OBJ ?ATTR1)
        (attribute ?OBJ ?ATTR2)))))
```

From now, on we denote the nature of SUMO concepts by adding as subscript the following symbols:

- $o$ SUMO individuals that are neither classes nor relations nor attributes
- $c$ SUMO classes that are neither classes of relations nor classes of attributes
- $r$ individual SUMO relations
- $a$ individual SUMO attributes
- $R$ classes of SUMO relations
- $A$ classes of SUMO attributes

For example, $Cell_c$ represents the *Cell* class, $member_r$ the *member* individual relation and $Larval_a$ the *Larval* attribute.

## 4. The Mapping Between WordNet and SUMO

WordNet is linked to SUMO by means of the mapping described in Niles and Pease (2003). This mapping connects synsets of WordNet to terms of SUMO using three relations: *equivalence*, *subsumption* and *instance*. The relation *equivalence* denotes that the related WordNet synset and the SUMO concept are equivalent in meaning, whereas the relations *subsumption* and *instance* indicate that the WordNet synset is subsumed by the SUMO concept or is an instance of the SUMO concept respectively. Furthermore, the mapping uses the complementaries of *equivalence* and *instance*. We denote mapping relations by concatenating the symbols '$=$' (*equivalence*), '$+$' (*subsumption*), '@' (*instance*), '$\hat{=}$' (complementary of *equivalence*) and '$\hat{+}$' (complementary of *subsumption*) to the corresponding SUMO concept. For example, the synsets $waist_n^1$ and $torso_n^1$ are connected to $Waist_o=$ and $Torso_c+$ respectively.

From the 82,115 noun synsets defined in WordNet v3.0, 73,472 noun synsets are directly connected to concepts that are defined in the core of SUMO —and, thus, in Adimen-SUMO—, while only 7,578 synsets are linked to SUMO concepts defined in domain ontologies. As described in Álvez et al. (2017), those synsets linked to concepts defined in domain ontologies can be connected to concepts from the core of SUMO by means of the SUMO structural relations *$subclass*, $subrelation_r$ and $subAttribute_r$. Finally, it is worth to remark that some noun synsets are connected to several SUMO concepts (concretely, 1,043 synsets).

The knowledge in the mapping between WordNet and SUMO can be formalized by means of Adimen-SUMO statements as described in Álvez et al. (2017). For this purpose, we take advantage from the fact that most of the SUMO knowledge is based on the notion of *objects* and that only a few of SUMO predicates provide information at the level of *classes*. Thus, the proposed Adimen-SUMO statements relate synsets to sets of SUMO objects in most cases. For the construction of those Adimen-SUMO statements, we choose the most suitable SUMO predicate according to nature of the SUMO concepts to which the synset is connected: $equal_r$ for SUMO objects, *$instance* for SUMO classes and $attribute_r$ for SUMO attributes.[2] In addition, we use *$subclass* for the construction of Adimen-SUMO statements in the few cases where synsets have to relate the knowledge of SUMO at the level of classes. Independently from the SUMO predicate that is selected, we introduce a new variable that is associated to the given synset in the proposed Adimen-SUMO statement. For example, the next Adimen-SUMO statements relate the synsets

[1] http://www.ontologyportal.org

[2] The mappings to SUMO relations are discarded for the moment.

| Meronymy relations | Pairs | | 1st QP | | 2nd QP | | 3rd QP | | 4th QP | | Total | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Total | Error | Pairs | CQs | Pairs | CQs | Pairs | CQs | Pairs | CQs | **Pairs** | **CQs** |
| *part* | 9,097 | 1,367 | 6,797 | 1,346 | 122 | 110 | 750 | 445 | 61 | 59 | **7,730** | **1,960** |
| *member* | 12,293 | 11,939 | 318 | 88 | 19 | 19 | 15 | 12 | 2 | 2 | **354** | **121** |
| *substance* | 797 | 632 | 147 | 47 | 6 | 6 | 10 | 9 | 2 | 2 | **165** | **64** |
| **Total** | **22,187** | **13,938** | **7,262** | **1,481** | **147** | **135** | **775** | **466** | **65** | **63** | **8,249** | **2,145** |

Table 1: CQs obtained from WordNet meronymy

*waist*$^1_n$, *atom*$^1_n$ and *police_officer*$^1_n$—respectively connected to *Waist*$_o$, *Atom*$_c$ and *PoliceOfficer*$_a$— to sets of SUMO objects by introducing the new variables *?W*, *?A* and *?O*:

(equal ?W Waist) (2)

($instance ?A Atom) (3)

(attribute ?O PoliceOfficer) (4)

Similarly, the next Adimen-SUMO statement relates the synsets *cartilage*$^1_n$, which is connected to *Tissue*$_c$, to a set of SUMO classes introducing the new variable *?T*:

($subclass ?T Tissue) (5)

Finally, the quantification of the introduced variables is decided according to the mapping relation that connects synsets and the SUMO concepts:

- If the given synset is connected using *equivalence* (resp. the complement of *equivalence*), then we can assume that the synset is related to all (resp. is not related to any of) the potential SUMO objects that are characterized by the Adimen-SUMO statement proposed above. For this purpose, the variable associated to the given synset is considered to be universally quantified.
- Otherwise —the synset is connected using *subsumption* (resp. the complement of *subsumption*) or *instance*—, we can only assume that the synset is related to (resp. is not related to) some of the potential SUMO objects that are characterized by the Adimen-SUMO statement proposed above. This means that the variable associated to the given synset is considered to be existentially quantified.

In the next section, we describe the use of question patterns for the combination of the Adimen-SUMO statements that are obtained for the synsets in a given WordNet pair.

## 5. Competency Questions Based on Meronymy

The automatic validation of WordNet and SUMO on the basis of CQs and ATPs requires to translate all the information into a formal language. By means of Adimen-SUMO, the core information of SUMO is already written in FOL. In addition, the mapping information between WordNet and SUMO can be translated into Adimen-SUMO statements as we describe in the above section. Thus, it suffices to translate the semantics of the WordNet meronymy relations in terms of Adimen-SUMO. For this purpose, we have inspected SUMO in order to find the relations that are synonym or semantically similar to them. In SUMO, the main meronymy relation is *part*$_r$ and we can find 30 different subrelations of *part*$_r$ in its core. Among them, we have selected the SUMO predicates *part*$_r$ and *member*$_r$ as counterpart of the WordNet relations *part* and *member* respectively. In addition, we have selected *material*$_r$, which is not subrelation of *part*$_r$, as counterpart of *substance*. As for every SUMO relation, SUMO provides *domain* axioms that restrict the set of SUMO objects that can be related by the above predicates as follows:

- *part*$_r$ relates pairs of *Object*$_c$ individuals.
- *member*$_r$ relates *SelfConnectedObject*$_c$ objects (first argument) to *Collection*$_c$ objects (second argument).
- *material*$_r$ relates subclasses of *Substance*$_c$ (first argument) to *CorpuscularObject*$_c$ objects (second argument).

Additionally, SUMO provides incompatibilities between SUMO objects. Among others, objects of *CorpuscularObject*$_c$ are incompatible with both *Collection*$_c$ and *Substance*$_c$.

On the basis of the above formalization of the meronymy information of WordNet in terms of Adimen-SUMO, we proceed to the creation of CQs. For this purpose, we propose 4 QPs depending on the mapping relation by which the synsets in the given WordNet pair are connected to SUMO. Those QPs are instantiated by using (a) the Adimen-SUMO statements that formalize the mapping information of synsets, and (b) the SUMO predicate that is selected depending on the given WordNet meronymy relation. During the process of instantiation, we can already detect some incompatibilities on the basis of individual SUMO incompatibilities.

We report on the number of WordNet pairs tested and the number of CQs resulting from each QP in Table 1 and in the next subsections we describe the proposed QPs.

```
(exists (?X, ?Y)
  (and
    < s_part, ?X >
    < s_whole, ?Y >
    (< SUMO predicate > ?X ?Y)))
```

Figure 1: First question pattern for $\langle s\_part, s\_whole \rangle$ meronymy pairs

### 5.1. First Question Pattern

The first question pattern is designed to be applied to meronymy pairs where both synsets are connected using (the negation of) *subsumption* and *instance*.
In Figure 1, we describe the combination of the selected SUMO predicate and the Adimen-SUMO statements that result from the mapping information of synsets by considering the introduced variables to be existentially quantified.

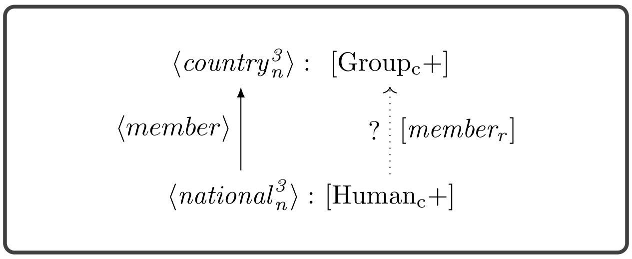


Figure 2: *national*$^1_n$ and *country*$^3_n$.

Next, we illustrate the instantiation of the resulting question pattern by considering the WordNet pair *member*(*national*$^1_n$,*country*$^3_n$). As described in Figure 2, the synsets in that pair are respectively connected to *Human*$_c$+ and *Group*$_c$+. Thus, the combination of the SUMO statements that result from their mapping information with the SUMO predicate *member*$_r$ yields the following CQ:

```
(exists (?X, ?Y)                                    (6)
  (and
    ($instance ?X Human)
    ($instance ?Y Group)
    (member ?X ?Y)))
```

In the same way, the synsets in the WordNet pair *substance*(*cartilage*$^1_n$,*cartilaginous_structure*$^1_n$) are respectively connected to *Tissue*$_c$ and *BodyPart*$_c$. In this case, we have to relate *cartilaginous_structure*$^1_n$ with a set of SUMO classes, since the selected SUMO predicate is *material*$_r$. Thus, the resulting CQ is:

```
(exists (?X, ?Y)                                    (7)
  (and
    ($subclass ?X Tissue)
    ($instance ?Y BodyPart)
    (material ?X ?Y)))
```

Using this first QP, we obtain 1,481 CQs from 7,262 WordNet pairs (see Table 1).

### 5.2. Second Question Pattern

The second question pattern is designed for the meronymy synset pairs $\langle s\_part, s\_whole \rangle$ where $s\_part$ is connected by (the negation of) *equivalence* and $s\_whole$ is connected by (the negation of) *subsumption* or *instance*.
In this case, the variable associated to $s\_whole$ is considered to be universally quantified, while the variable associated to $s\_part$ is considered to be existentially quantified. The resulting question pattern is described in Figure 3.

```
(forall (?X)
  (=>
    < s_part, ?X >
    (exists (?Y)
      (and
        < s_whole, ?Y >
        (< SUMO predicate > ?X ?Y)))))
```

Figure 3: Second question pattern for $\langle s\_part, s\_whole \rangle$ meronymy pairs

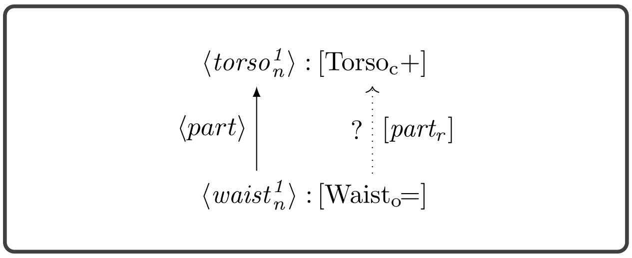


Figure 4: *waist*$^1_n$ and *torso*$^1_n$.

In order to illustrate the instantiation of this second question pattern, we consider the synset pair *part*(*waist*$^1_n$,*torso*$^1_n$), where the involved synsets are respectively connected to *Waist*$_o$= and *Torso*$_c$+ as described in Figure 4. On the basis of the above mapping information, the selected SUMO predicate is *part*$_r$ and we obtain the following CQ:

```
(forall (?X)                                        (8)
  (=>
    ($instance ?X Waist)
    (exists (?Y)
      (and
        ($instance ?Y Torso)
        (part ?X ?Y)))))
```

As reported in Table 1, from this QP we obtain 135 CQ on the basis of 147 WordNet pairs.

```
(forall (?Y)
  (=>
    < s_whole, ?Y >
    (exists (?X)
      (and
        < s_part, ?X >
        (< SUMO predicate > ?X ?Y)))))
```

Figure 5: Third question pattern for $\langle s\_part, s\_whole \rangle$ meronymy pairs

### 5.3. Third Question Pattern

The third question pattern is the dual of the second one because it is designed for meronymy synset pairs $\langle s\_part, s\_whole \rangle$ where $s\_part$ is connected by (the negation of) *subsumption* or *instance*, and $s\_whole$ is connected by (the negation of) *equivalence*. Consequently, the variables associated to $s\_whole$ and $s\_part$ are considered to be universally and existentially quantified respectively, as described in Figure 5.

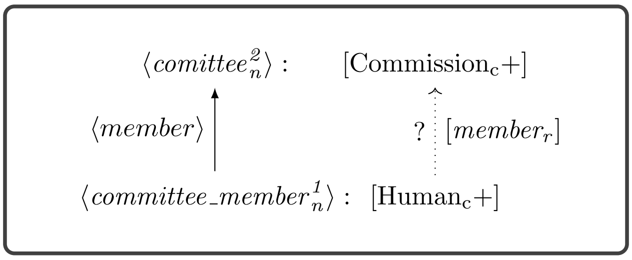


Figure 6: *committee_member*$^1_n$ and *committee*$^1_n$.

This third question pattern is applied to synset pairs like *member*(*committee*$^1_n$,*committee_member*$^1_n$), where synsets are respectively connected to *Human*$_c$+ and *Commission*$_c$=. By using the SUMO predicate *member*$_r$, the resulting CQ is:

```
(forall (?Y)                                   (9)
  (=>
    ($instance ?Y Commission)
    (exists (?X)
      (and
        ($instance ?X Human)
        (member ?X ?Y)))))
```

Using this third QP, we obtain 466 CQs from 775 WordNet pairs (see Table 1).

### 5.4. Fourth Question Pattern

The last question pattern is designed for its application to meronymy pairs where both synsets are connected using (the negation of) *equivalence*.

In this case, the question pattern is obtained by the conjunction of the second and the third question patterns (see Figure 7). In order to illustrate its application, we consider the synset pair *part*(*elementary_particle*$^1_n$,*atom*$^1_n$), where synsets are respectively connected to *SubatomicParticle*$_c$=

```
(and
  (forall (?X)
    (=>
      < s_part, ?X >
      (exists (?Y)
        (and
          < s_whole, ?Y >
          (< SUMO predicate > ?X ?Y)))))
  (forall (?Y)
    (=>
      < s_whole, ?Y >
      (exists (?X)
        (and
          < s_part, ?X >
          (< SUMO predicate > ?X ?Y))))))
```

Figure 7: Fourth question pattern for $\langle s\_part, s\_whole \rangle$ meronymy pairs

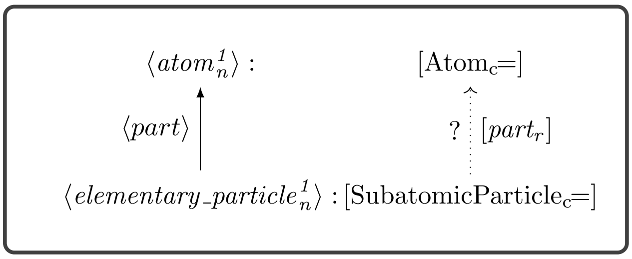


Figure 8: *elementary_particle*$^1_n$ and *atom*$^1_n$.

and *Atom*$_c$= as described in Figure 8. The resulting CQ is:

```
(and                                           (10)
  (forall (?X)
    (=>
      ($instance ?X SubatomicParticle)
      (exists (?Y)
        (and
          ($instance ?Y Atom)
          (part ?X ?Y)))))
  (forall (?Y)
    (=>
      ($instance ?Y Atom)
      (exists (?X)
        (and
          ($instance ?X SubatomicParticle)
          (part ?X ?Y))))))
```

From this last QP, we obtain 63 CQs on the basis of 65 WordNet pairs (see Table 1).

| Meronymy relations | QPs | | | | | | | | Total | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | QP #1 | | QP #2 | | QP #3 | | QP #4 | | | | | |
| | Pairs | CQs | Pairs | CQs | Pairs | CQs | Pairs | CQs | **Pairs** | **CQs** | **Pairs** | **CQs** |
| *part* | +3,272 | 599 | +122 | 56 | +288 | 162 | +8 | 8 | **+3,690** | **825** | **47.74%** | **42.09%** |
| | -21 | 6 | -0 | 0 | -1 | 1 | -5 | 5 | **-27** | **12** | **0.35%** | **0.61%** |
| *member* | +20 | 10 | +1 | 1 | +1 | 1 | +0 | 0 | **+22** | **12** | **6.21%** | **9.92%** |
| | -24 | 9 | -0 | 0 | -0 | 0 | -0 | 0 | **-24** | **9** | **6.78%** | **7.44%** |
| *substance* | +94 | 17 | +1 | 1 | +3 | 2 | +0 | 0 | **+98** | **20** | **59.39%** | **31.25%** |
| | -0 | 0 | -2 | 2 | -0 | 0 | -0 | 0 | **-2** | **2** | **1.21%** | **3.13%** |
| **Total** | **+3,386** | **626** | **+124** | **58** | **+292** | **165** | **+8** | **8** | **+3,810** | **857** | **46.19%** | **39.95%** |
| | **-45** | **15** | **-2** | **2** | **-1** | **1** | **-5** | **5** | **-53** | **23** | **0.64%** | **1.07%** |

Table 2: Evaluation results

## 6. Evaluation

In this section, we discuss the results obtained from the experimental evaluation of the set of CQs described in the above section. This experimentation has been performed in an Intel® Xeon® CPU E5-2640v3@2.60GHz with 2GB of RAM memory per processor and different. All the required knowledge resources—the ontology Adimen-SUMO v2.6,(Álvez et al., 2017) the set of CQs and conjectures, the mapping between Adimen-SUMO v2.6 and WordNetv3.0, WordNet v3.0 meronymy pairs— and the resulting execution reports are publicly available.[3]

For the evaluation of CQs, we have applied the framework proposed in Álvez et al. (2017). By following this proposal, we get two conjectures for each CQ: i) the conjecture that describes the CQ itself, which is expected to be entailed by the ontology (called *truth-test*), and ii) its negation, which is expected not to be entailed (called *falsity-test*). Each of the resulting 4,830 conjectures has been tested using several versions of Vampire[4] (Kovács and Voronkov, 2013) and E[5] (Schulz, 2002). In each test, we provide the conjecture together with Adimen-SUMO as input to the ATP system.

If a truth-test is proved to be entailed by the ontology, then we decide that the knowledge in the WordNet pairs that yield the corresponding CQ is *compatible* with SUMO and the mapping information. Thus, we conclude that those WordNet pairs are *well-aligned* —or simply *aligned*— with SUMO and the mapping information. On the contrary, if a falsity-test is proved to be entailed by the ontology, then the knowledge in the WordNet pairs that yield the corresponding CQ is *incompatible* with SUMO and the mapping information. Consequently, we say that those WordNet pairs are *misaligned* with SUMO and the mapping information. If either the truth-test or the falsity-test of a given CQ is proved to be entailed by the ontology, then we say than the corresponding CQ is *resolved*.

In Table 2 we summarize our experimental results as follows. For each meronymy relation (1st column) and each QP, we provide the following information from the 2nd to the 9th column:

- In Pairs columns, prefixed by '+' (resp. by '-') we give the number of WordNet pairs that are well-aligned (resp. misaligned) with SUMO and the mapping information.
- In CQs columns, we give the number of truth-tests/falsity-tests that have been proved for the classification of WordNet pairs as aligned/misaligned with SUMO and the mapping information.

In addition, in the last 4 columns we summarize the number/percentage of WordNet pairs that are well-aligned/misaligned with SUMO and the mapping information, and the number and percentage of truth-tests/falsity-tests that have been proved.

To sum up, the knowledge in 3,810 WordNet pairs (46.19 % of the tested WordNet pairs) is decided to be compatible with SUMO and the mapping information, while the knowledge in only 53 WordNet pairs (0.64 % of the tested pairs) is decided to be incompatible. Among them, it is easy to see that *part* and *substance* pairs are better aligned with the knowledge in SUMO than *member* pairs: on one hand, 3,690 *part* pairs (47.74 % of the tested *part* pairs) and 98 *substance* pairs (59.39 % of the tested *substance* pairs) are decided to be well-aligned with SUMO and the mapping information against 22 *member* pairs (6.21 % of the tested *member* pairs); on the other hand, 24 *member* pairs (6.78 % of the tested *member* pairs) are decided to be misaligned with SUMO and the mapping information, against 27 *part* pairs (0.35 % of the tested *part* pairs) and 2 *substance* pairs (0.64 % of the tested *substance* pairs). Further, if we consider the total number of meronymy pairs, the percentage of *part* pairs that are well-aligned with SUMO and the mapping information —40.56 % (3,690 from 9,097 *part* pairs)— is clearly larger than the percentage of *member* and *substance* pairs that are well-aligned —0.18 % (22 from 12,293 *member* pairs) and 12.3 % (98 from 797 *substance* pairs respectively). Finally, the percentage of *member* pairs that are tested —only 2.88 % (354 from 12,293 *member* pairs)— is the lowest one and, in addition, the number of *member* pairs that are decided to be misaligned with SUMO and the mapping information is larger than the number of *member* pairs that are decided to be well-aligned —24 against 22. These two facts, but especially the first

[3] `http://adimen.si.ehu.es/web/AdimenSUMO`

[4] Using the following parameters: `--proof tptp --output_axiom_names on --mode casc -t 600 -m 2048`.

[5] Using the following parameters: `--auto --proof-object -s --cpu-limit=600 --memory-limit=2048`.

one, lead us to conclude that the mapping information of the involved synsets can be substantially improved.

With respect to CQs, ATPs are able to successfully resolve 880 CQs (41.03 % of 2,145 CQs): 857 truth-tests (39.95 % of 2,145 truth-tests) plus 23 falsity-tests (1.07 % of 2,145 falsity-tests). It is worth noting that the percentage of meronymy-based CQs that are resolved by ATPs is quite similar to the percentage of CQs for the validation of the knowledge in the ontology that were resolved by ATPs in the experimentation reported in Álvez et al. (2017). Regarding QPs, the percentage of CQs obtained from the the first 3 QPs that are resolved —43.28 % (641 from 1,481 CQs), 44.44 % (60 from 158 CQs) and 35.62 % (166 from 466 CQs) respectively— is clearly larger than the percentage of CQs obtained from the last QP that are resolved — 20.63 % (13 from 63 CQs). This fact is not surprising since the truth-tests that result from the first 3 QPs are much weaker than the truth-tests that are obtained from the last QP.

## 7. Discussion

In this section, we proceed to discuss some of the particular results that we have extracted from our experiments. First of all, our proposal enables the detection of semantics agreements between WordNet, SUMO and their mapping. For example, the synset pairs *member*($national^1_n$,$country^3_n$), *substance*($cartilage^1_n$,$cartilaginous\_structure^3_n$) and *part*($elementary\_particle^1_n$,$atom^1_n$) are decided to be well-aligned with SUMO and the mapping information since the truth-tests obtained from CQs (6,7,10) are proved to be entailed by Adimen-SUMO v2.6. Similarly, the synsets in the pair *member*($police\_officer^1_n$,$police\_force^1_n$) are respectively connected to $PoliceOfficer_a$= and $PoliceOrganization_c$+. Thus, by applying the second QP we obtain the following CQ:

```
(forall (?Y)                                          (11)
  (=>
    (attribute ?Y PoliceOfficer)
    (exists (?X)
      (and
        ($instance ?X PoliceOrganization)
        (member ?X ?Y)))))
```

The above CQ is resolved by ATPS since its truth-test is proved to be entailed by Adimen-SUMO v2.6. Consequently, the pair *member*($police\_officer^1_n$,$police\_force^1_n$) is decided to be well-aligned with SUMO and the mapping information.

Second, our proposal enables the detection of inconsistencies among WordNet, SUMO and their mapping. Indeed, we do not always require the help of ATPs, since some inconsistencies can be discovered during the process of instantiation of QPs on the basis of incompatibilities among SUMO objects. For instance, the synsets $grape^1_n$ and $wine^1_n$ are related by *substance* and respectively connected $FruitOrVegetable_c$+ and $Wine_c$=. According to our interpretation of the semantics of *substance* and the mapping information, we have to use the third QP and the SUMO predicate $material_r$ in order to translate the knowledge in *substance*($grape^1_n$,$wine^1_n$) in terms of Adimen-SUMO. However, $FruitOrVegetable_c$ is defined to be subclass of $CorpuscularObject_c$ in SUMO. Therefore, $FruitOrVegetable_c$ is incompatible with $Substance_c$, which prevents the use of $material_r$ for the instantiation of the third QP. This fact leads us to discover that the pair *substance*($grape^1_n$,$wine^1_n$) should be better represented by *substance*($grape\_juice^1_n$,$wine^1_n$), where $grape\_juice^1_n$ is connected to $Substance_c$+. This also would require a new relation between $grape^1_n$ and $grape\_juice^1_n$. Additional examples can be taken from many synset pairs related by *member*, where both synsets are connected to the same SUMO concept although the first one denotes an individual organism and the second one the species, genus or family to which the organism belongs. For instance, $bear^1_n$ and $ursidae^1_n$ are both connected to $Mammal_c$+, which reveals the existence of an incompatibility according to our interpretation of the meronymy relations in WordNet and the mapping information. Anyway, inconsistencies are also detected by means of the use of ATPs. For example, the synsets in *part*($cell^2_n$,$cell\_nucleus^1_n$) are connected to $Cell_c$= and $CellNucleus_c$= respectively. Hence, we obtain the following QP by the instantiation of the fourth QP:

```
(and                                                  (12)
  (forall (?X)
    (=>
      ($instance ?X CellNucleus)
      (exists (?Y)
        (and
          ($instance ?Y Cell)
          (part ?X ?Y)))))
  (forall (?Y)
    (=>
      ($instance ?Y Cell)
      (exists (?X)
        (and
          ($instance ?X CellNucleus)
          (part ?X ?Y))))))
```

ATPs are able to prove the falsity-test that results from the above CQ, which enables the detection of an inconsistency according to our interpretation of the knowledge in WordNet and the mapping information. Concretely, the falsity-test that results from our interpretation is incompatible with the fact that some cells lack a nucleus, as stated by the following SUMO axiom:

```
(=>                                                   (13)
  ($instance ?C RedBloodCell)
  (not
    (exists (?N)
      (and
        ($instance ?N CellNucleus)
        (part ?N ?C)))))
```

Similarly, the knowledge in the pair *member*(*malacosoma_americana*$^1_n$,*genus_malacosoma*$^1_n$) is detected to be incompatible with SUMO and the mapping information. As the involved synsets are respectively connected to $Insect_c+$ and $Larval_a+$, we obtain the following CQ by the instantiation of the first QP:

$$\begin{aligned}&(\textbf{exists}\ (?\texttt{X}, ?\texttt{Y})\\&\quad(\textbf{and}\\&\qquad(\$\text{instance } ?\texttt{X} \text{ Insect})\\&\qquad(\text{attribute } ?\texttt{Y} \text{ Larval})\\&\qquad(\text{member } ?\texttt{X}\ ?\texttt{Y})))\end{aligned} \tag{14}$$

By inspecting the proof of the falsity-tests that results from the above CQ, we discover that the problem is due to the mapping of *genus_malacosoma*$^1_n$ to $Larval_a+$, since the second argument of $member_r$ is restricted to be instance of $Collection_c$ and $Larval_a+$ cannot be applied to instances of $Collection_c$.

Finally, our proposal also enables the detection of missing knowledge. For example, ATPs cannot prove the truth- and falsity-tests inherited from (8) because $Waist_o$ and $Torso_c$ are not properly related in SUMO, as discussed in Section 1. In the same fashion, we discover that SUMO lacks the appropriate knowledge relating the concepts of $Human_c$ and $Commission_c$ by $member_r$ since the truth- and falsity-tests inherited from (9) cannot be proved by ATPs.

## 8. Conclusions and Future Work

By analysing our experimentation results, we can conclude that our proposal enables a sophisticated cross-checking of the information in WordNet, SUMO and their mapping. In particular, by means of practical examples, we have illustrated that the proposed system enables (a) the validation of some pieces of knowledge and (b) the detection of inconsistencies and missing knowledge. Further, our results also demonstrate the suitability of the involved resources for their application to practical tasks related to natural language processing. In future work, we plan to correct some of the issues detected with the mapping information and to augment the knowledge in SUMO in order to increase the level of alignment among WordNet, SUMO and their mapping.

## 9. Acknowledgements

This work has been partially funded by the Spanish Projects TUNER (TIN2015-65308-C5-1-R) and GRAMM (TIN2017-86727-C2-2-R), the Basque Project LoRea (GIU15/30) and the UPV/EHU project OEBU (EHUA16/33).

## 10. Bibliographical References


Álvez, J. and Rigau, G. (2018). Towards cross-checking WordNet and SUMO using meronymy. In P. Vossen, et al., editors, *Proceedings of the 9$^{th}$ Global WordNet Conference (GWC 2018)*.

Álvez, J., Atserias, J., Carrera, J., Climent, S., Laparra, E., Oliver, A., and Rigau, G. (2008). Complete and consistent annotation of WordNet using the Top Concept Ontology. In Nicoletta Calzolari, et al., editors, *Proc. of the 6$^{th}$ Int. Conf. on Language Resources and Evaluation (LREC 2008)*, pages 1529–1534. European Language Resources Association (ELRA), may.

Álvez, J., Lucio, P., and Rigau, G. (2012). Adimen-SUMO: Reengineering an ontology for first-order reasoning. *Int. J. Semantic Web Inf. Syst.*, 8(4):80–116.

Álvez, J., Lucio, P., and Rigau, G. (2017). Black-box testing of first-order logic ontologies using WordNet. *CoRR*, abs/1705.10217.

Daudé, J., Padró, L., and Rigau, G. (2003). Making WordNet mapping robust. *Procesamiento del lenguaje natural*, 31.

C. Fellbaum, editor. (1998). *WordNet: An Electronic Lexical Database*. MIT Press.

Grüninger, M. and Fox, M. S. (1995). Methodology for the design and evaluation of ontologies. In *Proc. of the Workshop on Basic Ontological Issues in Knowledge Sharing (IJCAI 1995)*.

Horrocks, I. and Voronkov, A. (2006). Reasoning support for expressive ontology languages using a theorem prover. In J. Dix et al., editor, *Foundations of Information and Knowledge Systems*, LNCS 3861, pages 201–218. Springer.

Kovács, L. and Voronkov, A. (2013). First-order theorem proving and Vampire. In N. Sharygina et al., editors, *Computer Aided Verification*, LNCS 8044, pages 1–35. Springer.

Niles, I. and Pease, A. (2001). Towards a standard upper ontology. In Guarino N. et al., editor, *Proc. of the 2$^{nd}$ Int. Conf. on Formal Ontology in Information Systems (FOIS 2001)*, pages 2–9. ACM.

Niles, I. and Pease, A. (2003). Linking lexicons and ontologies: Mapping WordNet to the Suggested Upper Merged Ontology. In H. R. Arabnia, editor, *Proc. of the IEEE Int. Conf. on Inf. and Knowledge Engin. (IKE 2003)*, volume 2, pages 412–416. CSREA Press.

Richens, T. (2008). Anomalies in the WordNet verb hierarchy. In *Proc. of the 22$^{nd}$ Int. Conf. on Computational Linguistics-Volume 1*, pages 729–736. Association for Computational Linguistics.

Rodríguez, H., Climent, S., Vossen, P., Bloksma, L., Peters, W., Alonge, A., Bertagna, F., and Roventini, A. (1998). The top-down strategy for building EuroWordNet: Vocabulary coverage, base concepts and top ontology. In *EuroWordNet: A multilingual database with lexical semantic networks*, pages 45–80. Springer.

Schulz, S. (2002). E - A brainiac theorem prover. *AI Communications*, 15(2-3):111–126.


## 11. Language Resource References


J. Álvez and P. Lucio and G. Rigau. (2017). *Adimen-SUMO*. 2.6.