WordNet 2 – A Morphologically and Semantically Enhanced Resource

Boris Fersing

Reference:
S. Harabagiu, G. A. Miller, D. Moldovan (1999): WordNet 2 - A Morphologically and Semantically Enhanced Resource. Proc. of SIGLEX 1999

Why Extended WordNet?

WordNet is a machine readable dictionary which is already a valuable resource for the language technology and knowledge processing communities.
In order to understand why enhancing WordNet may be useful, lets have a look at its current features.

Why Extended WordNet?

Feature 1: Synonymy

In WordNet, the synonymy relation is represented through *synsets*. The words which express the same meaning are grouped into a *synset*.

Why Extended WordNet?

Feature 1: Synonymy

In WordNet, the synonymy relation is represented through *synsets*. The words which express the same meaning are grouped into a *synset*.

Noun

- S: (n) **computer**, computing machine, computing device, data processor, electronic computer, information processing system (a machine for performing calculations automatically)

Why Extended WordNet?

Feature 1: Synonymy

In WordNet, the synonymy relation is represented through *synsets*. The words which express the same meaning are grouped into a *synset*.

Noun

- S: (n) **computer**, computing machine, computing device, data processor, electronic computer, information processing system (a machine for performing calculations automatically)

Why Extended WordNet?

Feature 2: Hypernymy

Hypernymy is the dominant semantic relation in WordNet.

Why Extended WordNet?

Feature 2: Hypernymy

Hypernymy is the dominant semantic relation in WordNet.

Noun

- S: (n) **computer**, computing machine, computing device, data processor, electronic computer, information processing system (a machine for performing calculations automatically)
  - ***direct hyponym*** / *full hyponym*
    - S: (n) analog computer, analogue computer (a computer that represents information by variable quantities (e.g., positions or voltages))
    - S: (n) digital computer (a computer that represents information by numerical (binary) digits)
    - S: (n) home computer (a computer intended for use in the home)
    - S: (n) node, client, guest ((computer science) any computer that is hooked up to a computer network)
    - S: (n) number cruncher (a computer capable of performing a large number of mathematical operations per second)
    - S: (n) pari-mutuel machine, totalizer, totaliser, totalizator, totalisator (computer that registers bets and divides the total amount bet among those who won)
    - S: (n) predictor (a computer for controlling antiaircraft fire that computes the position of an aircraft at the instant of a shell's arrival)
    - S: (n) server, host ((computer science) a computer that provides client stations with access to files and printers as shared resources to a computer network)
    - S: (n) Turing machine (a hypothetical computer with an infinitely long memory tape)
    - S: (n) web site, website, internet site, site (a computer connected to the internet that maintains a series of web pages on the World Wide Web) "*the Israeli web site was damaged by hostile hackers*"

Why Extended WordNet?

Feature 2: Hypernymy

Hypernymy is the dominant semantic relation in WordNet.

Noun

- S: (n) **computer**, computing machine, computing device, data processor, electronic computer, information processing system (a machine for performing calculations automatically)
  - ***direct hyponym*** / *full hyponym*
    - S: (n) analog computer, analogue computer (a computer that represents information by variable quantities (e.g., positions or voltages))
    - S: (n) digital computer (a computer that represents information by numerical (binary) digits)
    - S: (n) home computer (a computer intended for use in the home)
    - S: (n) node, client, guest ((computer science) any computer that is hooked up to a computer network)
    - S: (n) number cruncher (a computer capable of performing a large number of mathematical operations per second)
    - S: (n) pari-mutuel machine, totalizer, totaliser, totalizator, totalisator (computer that registers bets and divides the total amount bet among those who won)
    - S: (n) predictor (a computer for controlling antiaircraft fire that computes the position of an aircraft at the instant of a shell's arrival)
    - S: (n) server, host ((computer science) a computer that provides client stations with access to files and printers as shared resources to a computer network)
    - S: (n) Turing machine (a hypothetical computer with an infinitely long memory tape)
    - S: (n) web site, website, internet site, site (a computer connected to the internet that maintains a series of web pages on the World Wide Web) "*the Israeli web site was damaged by hostile hackers*"

Is-a

Why Extended WordNet?

And more...

WordNet has some other features, for example:

Why Extended WordNet?

And more...

WordNet has some other features, for example:

- Meronymy for nouns.

Why Extended WordNet?

And more...

WordNet has some other features, for example:

- Meronymy for nouns.
- Entailment and cause for verbs.

Why Extended WordNet?

And more...

WordNet has some other features, for example:

- Meronymy for nouns.
- Entailment and cause for verbs.
- Similarity for adjectives.

Why Extended WordNet?

And more...

WordNet has some other features, for example:

- Meronymy for nouns.
- Entailment and cause for verbs.
- Similarity for adjectives.
- ...

Why Extended WordNet?

A feature WordNet doesn't have

Each synset in WordNet has a related gloss which explains the sense of the words contained in the synset.

Why Extended WordNet?

A feature WordNet doesn't have

Each synset in WordNet has a related gloss which explains the sense of the words contained in the synset.

Noun

- S: (n) **computer**, computing machine, computing device, data processor, electronic computer, information processing system (a machine for performing calculations automatically)

Why Extended WordNet?

A feature WordNet doesn't have

Each synset in WordNet has a related gloss which explains the sense of the words contained in the synset.

Noun

- S: (n) **computer**, computing machine, computing device, data processor, electronic computer, information processing system (a machine for performing calculations automatically)

The glosses were written manually and are not machine readable. Furthermore, the words of the glosses are not connected with the concepts contained in the WordNet database.

Why Extended WordNet?

A feature WordNet doesn't have

Each synset in WordNet has a related gloss which explains the sense of the words contained in the synset.

Noun

- S: (n) **computer**, computing machine, computing device, data processor, electronic computer, information processing system (a machine for performing calculations automatically)

The glosses were written manually and are not machine readable.
Furthermore, the words of the glosses are not connected with the concepts contained in the WordNet database.
Extended WordNet tries to provide machine readable glosses whose words are connected with concepts of the WordNet database.

How to do that?

4 main steps

In order to provide machine readable glosses 4 steps are necessary:

How to do that?

4 main steps

In order to provide machine readable glosses 4 steps are necessary:

- POS tagging: The words of the glosses are tagged using a POS tagger.

How to do that?

4 main steps

In order to provide machine readable glosses 4 steps are necessary:

- POS tagging: The words of the glosses are tagged using a POS tagger.
- Word sense disambiguation: For each polysemous word in the gloss, find the corresponding sense and link the sense with the word.

How to do that?

4 main steps

In order to provide machine readable glosses 4 steps are necessary:

- POS tagging: The words of the glosses are tagged using a POS tagger.
- Word sense disambiguation: For each polysemous word in the gloss, find the corresponding sense and link the sense with the word.
- Logical form transformation: Transform the words of the gloss into predicates with arguments.

How to do that?

4 main steps

In order to provide machine readable glosses 4 steps are necessary:

- POS tagging: The words of the glosses are tagged using a POS tagger.
- Word sense disambiguation: For each polysemous word in the gloss, find the corresponding sense and link the sense with the word.
- Logical form transformation: Transform the words of the gloss into predicates with arguments.
- Semantic form transformation: Identify the thematic and semantic relations.

How to do that?
POS tagging

This step is not explained in the paper but there is a lot of Part-of-Speech taggers available: TnT, TATOO, SVMTool, TreeTagger, ...

How to do that?
POS tagging

This step is not explained in the paper but there is a lot of Part-of-Speech taggers available: TnT, TATOO, SVMTool, TreeTagger, ...

```
<gloss pos="NOUN" synsetID="00012748">
  <synonymSet>animal, animate_being, beast, brute, creature, fauna</synonymSet>
 <text>
   a living organism characterized by voluntary movement
 </text>
  <wsd>
      <wf pos="DT" >a</wf>
      <wf pos="VBG" >living</wf>
      <wf pos="NN"  >organism</wf>
      <wf pos="VBN" >characterized</wf>
      <wf pos="IN" >by</wf>
      <wf pos="JJ" >voluntary</wf>
      <wf pos="NN" >movement</wf>
  </wsd>

(...)

</gloss>
```

How to do that?
Word sense disambiguation

Word sense disambiguation is the biggest part of this paper. To disambiguate word senses, the authors used 3 complementary methods:

How to do that?
Word sense disambiguation

Word sense disambiguation is the biggest part of this paper. To disambiguate word senses, the authors used 3 complementary methods:

- Heuristics: 6 Classes of heuristics which use the information contained in WordNet.

How to do that?
Word sense disambiguation

Word sense disambiguation is the biggest part of this paper. To disambiguate word senses, the authors used 3 complementary methods:

- Heuristics: 6 Classes of heuristics which use the information contained in WordNet.
- Conceptual density: Based on the information contained in WordNet.

How to do that?
Word sense disambiguation

Word sense disambiguation is the biggest part of this paper. To disambiguate word senses, the authors used 3 complementary methods:

- Heuristics: 6 Classes of heuristics which use the information contained in WordNet.
- Conceptual density: Based on the information contained in WordNet.
- Statistics on large corpora: For example, the output of Internet search engines.

Word sense disambiguation

Heuristics: 1. Class – Hypernyms

The gloss may begin with a hypernym of the described word:

- S: (n) **intrusion** (entrance by force or without permission or welcome)

Word sense disambiguation

Heuristics: 1. Class – Hypernyms

The gloss may begin with a hypernym of the described word:

- S: (n) **intrusion** (entrance by force or without permission or welcome)

Word sense disambiguation

Heuristics: 1. Class – Hypernyms

The gloss may begin with a hypernym of the described word:

- S: (n) **intrusion** (entrance by force or without permission or welcome)

- S: (n) **entrance**, entranceway, entryway, entry, entree (something that provides access (to get in or get out)) "*they waited at the entrance to the garden*"; "*beggars waited just outside the entryway to the cathedral*"
- S: (n) **entrance**, entering (a movement into or inward)
- S: (n) **entrance**, entering, entry, ingress, incoming (the act of entering) "*she made a grand entrance*"

Word sense disambiguation

Heuristics: 1. Class – Hypernyms

The gloss may begin with a hypernym of the described word:

- S: (n) **intrusion** (entrance by force or without permission or welcome)

- S: (n) **entrance**, entranceway, entryway, entry, entree (something that provides access (to get in or get out)) "*they waited at the entrance to the garden*"; "*beggars waited just outside the entryway to the cathedral*"
- S: (n) **entrance**, entering (a movement into or inward)
- S: (n) **entrance**, entering, entry, ingress, incoming (the act of entering) "*she made a grand entrance*"

- S: (n) **intrusion** (entrance by force or without permission or welcome)
  - ***direct hypernym*** / *inherited hypernym* / *sister term*
    - S: (n) entrance, entering, entry, ingress, incoming (the act of entering) "*she made a grand entrance*"

Word sense disambiguation

Heuristics: 1. Class – Hypernyms

The gloss may begin with a hypernym of the described word:

- S: (n) **intrusion** (entrance by force or without permission or welcome)

- S: (n) **entrance**, entranceway, entryway, entry, entree (something that provides access (to get in or get out)) "*they waited at the entrance to the garden*"; "*beggars waited just outside the entryway to the cathedral*"
- S: (n) **entrance**, entering (a movement into or inward)
- S: (n) **entrance**, entering, entry, ingress, incoming (the act of entering) "*she made a grand entrance*"

- S: (n) **intrusion** (entrance by force or without permission or welcome)
  - ***direct hypernym*** / *inherited hypernym* / *sister term*
    - S: (n) entrance, entering, entry, ingress, incoming (the act of entering) "*she made a grand entrance*"

Word sense disambiguation

Heuristics: 2. Class – Linguistic Parallelism

Words that are syntactic parallel may be semantic parallel:

- S: (n) **interaction** (a mutual or reciprocal action; interacting)

Word sense disambiguation

Heuristics: 2. Class – Linguistic Parallelism

Words that are syntactic parallel may be semantic parallel:

- S: (n) **interaction** (a mutual or reciprocal action; interacting)

Word sense disambiguation

Heuristics: 2. Class – Linguistic Parallelism

Words that are syntactic parallel may be semantic parallel:

- S: (n) **interaction** (a mutual or reciprocal action; interacting)

- S: (adj) common, **mutual** (common to or shared by two or more parties) "*a common friend*"; "*the mutual interests of management and labor*"
- S: (adj) reciprocal, **mutual** (concerning each of two or more persons or things; especially given or done in return) "*reciprocal aid*"; "*reciprocal trade*"; "*mutual respect*"; "*reciprocal privileges at other clubs*"

Word sense disambiguation

Heuristics: 2. Class – Linguistic Parallelism

Words that are syntactic parallel may be semantic parallel:

- S: (n) **interaction** (a mutual or reciprocal action; interacting)

- S: (adj) common, **mutual** (common to or shared by two or more parties) "*a common friend*"; "*the mutual interests of management and labor*"
- S: (adj) reciprocal, **mutual** (concerning each of two or more persons or things; especially given or done in return) "*reciprocal aid*"; "*reciprocal trade*"; "*mutual respect*"; "*reciprocal privileges at other clubs*"

Word sense disambiguation

Heuristics: 3. Class – Gloss Comments

A specialization or typical relation may exist between the elements of the comment and the elements of the gloss:

- S: (n) **scuff** (the act of scuffing (scraping or dragging the feet))

Word sense disambiguation

Heuristics: 3. Class – Gloss Comments

A specialization or typical relation may exist between the elements of the comment and the elements of the gloss:

- S: (n) **scuff** (the act of scuffing (scraping or dragging the feet))

Word sense disambiguation

Heuristics: 3. Class – Gloss Comments

A specialization or typical relation may exist between the elements of the comment and the elements of the gloss:

- S: (n) **scuff** (the act of scuffing (scraping or dragging the feet))

- S: (v) **scuff**, drag (walk without lifting the feet)
- S: (v) **scuff** (get or become scuffed) "*These patent leather shoes scuffed*"
- S: (v) **scuff** (mar by scuffing) "*scuffed shoes*"
- S: (v) **scuff** (poke at with the foot or toe)

Word sense disambiguation

Heuristics: 3. Class – Gloss Comments

A specialization or typical relation may exist between the elements of the comment and the elements of the gloss:

- S: (n) **scuff** (the act of scuffing (scraping or dragging the feet))

- S: (v) **scuff**, drag (walk without lifting the feet)
- S: (v) **scuff** (get or become scuffed) "*These patent leather shoes scuffed*"
- S: (v) **scuff** (mar by scuffing) "*scuffed shoes*"
- S: (v) **scuff** (poke at with the foot or toe)

Word sense disambiguation

Heuristics: 4. Class – Gloss Examples

Like comments, examples also provide useful information.

```
3. penetration -- (the act of forcing a way into or through something;
"the penetration of upper management by women")
```

Word sense disambiguation

Heuristics: 4. Class – Gloss Examples

Like comments, examples also provide useful information.

3. penetration -- (the act of forcing a way into or through something; "the penetration of upper management by women")

Word sense disambiguation

Heuristics: 4. Class – Gloss Examples

Like comments, examples also provide useful information.

3. penetration -- (the act of forcing a way into or through something; "the penetration of upper management by women")

9. way -- (doing as one pleases or chooses: "if I had my way")

Word sense disambiguation

Heuristics: 5. Class – Collocations

It's possible to use verbal collocations to identify the sense.

- S: (n) **escape**, flight (the act of escaping physically) "*he made his escape from the mental hospital*"; "*the canary escaped from its cage*"; "*his flight was an indication of his guilt*"

Word sense disambiguation

Heuristics: 5. Class – Collocations

It's possible to use verbal collocations to identify the sense.

- S: (n) **escape**, flight (the act of escaping physically) "*he made his escape from the mental hospital*"; "*the canary escaped from its cage*"; "*his flight was an indication of his guilt*"

Word sense disambiguation

Heuristics: 5. Class – Collocations

It's possible to use verbal collocations to identify the sense.

- S: (n) **escape**, flight (the act of escaping physically) "*he made his escape from the mental hospital*"; "*the canary escaped from its cage*"; "*his flight was an indication of his guilt*"

- S: (v) **escape**, get away, break loose (run away from confinement) "*The convicted murderer escaped from a high security prison*"
  - *direct troponym* / *full troponym*
  - ***direct hypernym*** / *inherited hypernym* / *sister term*
    - S: (v) flee, fly, take flight (run away quickly) "*He threw down his gun and fled*"

Word sense disambiguation

Heuristics: 6. Class – Lexical Relations

If there are two relations $[w_1 r w_2]$ and $[w_1 r w_3]$ where some senses of w_2 and w_3 have a common hypernym, then the sense of w_1 in $[w_1 r w_2]$ is likely to be the same than the sense of w_1 in $[w_1 r w_3]$:

- S: (n) **Underground Railroad**, Underground Railway (secret aid to escaping slaves that was provided by abolitionists in the years before the American Civil War)

Word sense disambiguation

Heuristics: 6. Class – Lexical Relations

If there are two relations $[w_1 rw_2]$ and $[w_1 rw_3]$ where some senses of w_2 and w_3 have a common hypernym, then the sense of w_1 in $[w_1 rw_2]$ is likely to be the same than the sense of w_1 in $[w_1 rw_3]$:

- S: (n) **Underground Railroad**, Underground Railway (secret aid to escaping slaves that was provided by abolitionists in the years before the American Civil War)

Word sense disambiguation

Heuristics: 6. Class – Lexical Relations

If there are two relations $[w_1 r w_2]$ and $[w_1 r w_3]$ where some senses of w_2 and w_3 have a common hypernym, then the sense of w_1 in $[w_1 r w_2]$ is likely to be the same than the sense of w_1 in $[w_1 r w_3]$:

- S: (n) **Underground Railroad**, Underground Railway (secret aid to escaping slaves that was provided by abolitionists in the years before the American Civil War)

$$[w_1 r w_2] = [\text{aid to slave}].$$

Word sense disambiguation

Heuristics: 6. Class – Lexical Relations

If there are two relations $[w_1 r w_2]$ and $[w_1 r w_3]$ where some senses of w_2 and w_3 have a common hypernym, then the sense of w_1 in $[w_1 r w_2]$ is likely to be the same than the sense of w_1 in $[w_1 r w_3]$:

- S: (n) **Underground Railroad**, Underground Railway (secret aid to escaping slaves that was provided by abolitionists in the years before the American Civil War)

$$[w_1 r w_2] = [\text{aid to slave}].$$

- S: (n) care, attention, **aid**, tending (the work of providing treatment for or attending to someone or something) "*no medical care was required*"; "*the old car needs constant attention*"

$[w_1 r w_3] = [\text{aid to someone}]$. A slave is someone $\Rightarrow$ Pick this sense of aid.

Word sense disambiguation

Conceptual density method

This method has already been explained last week.

Word sense disambiguation

Statistics on large corpora

This method doesn't use the information available in the WordNet database and can only be applied if we have collocating pairs of words where one word has already been disambiguated.
Lets consider the word pair w_1–w_2 where w_1 has already been disambiguated. To find out which sense w_2 has, we follow these steps:

- For each sense of w_2, create a similarity list which includes w_2 and its synonyms. $\{w_2^i, w_2^{i(1)}, w_2^{i(2)}, \dots\}$.
- Form pairs with w_1 and all the words in each similarity list.
- For each set, search a large corpus (Internet) for the occurences of any of the pairs.
- Compare the hits. The sense whose set got the most results is assigned to the word w_2.

Word sense disambiguation

The procedure

We combine the 3 methods to disambiguate as many words as possible:

Word sense disambiguation

The procedure

We combine the 3 methods to disambiguate as many words as possible:

- Identify and exclude the monosemous words. $\Rightarrow$ S0

Word sense disambiguation

The procedure

We combine the 3 methods to disambiguate as many words as possible:

- Identify and exclude the monosemous words. $\Rightarrow$ S0
- Apply Method 1 (Heuristics) to the remanining words (S0). $\Rightarrow$ S1

Word sense disambiguation

The procedure

We combine the 3 methods to disambiguate as many words as possible:

- Identify and exclude the monosemous words. $\Rightarrow$ S0
- Apply Method 1 (Heuristics) to the remanining words (S0). $\Rightarrow$ S1
- Apply Method 2 (Density) to S0. $\Rightarrow$ S2

Word sense disambiguation

The procedure

We combine the 3 methods to disambiguate as many words as possible:

- Identify and exclude the monosemous words. $\Rightarrow$ S0
- Apply Method 1 (Heuristics) to the remanining words (S0). $\Rightarrow$ S1
- Apply Method 2 (Density) to S0. $\Rightarrow$ S2
- Apply Method 3 (Statistics) to S0. $\Rightarrow$ S3

Word sense disambiguation

The procedure

We combine the 3 methods to disambiguate as many words as possible:

- Identify and exclude the monosemous words. $\Rightarrow$ S0
- Apply Method 1 (Heuristics) to the remanining words (S0). $\Rightarrow$ S1
- Apply Method 2 (Density) to S0. $\Rightarrow$ S2
- Apply Method 3 (Statistics) to S0. $\Rightarrow$ S3
- Combine S1 and S2. In the case of conflict give priority to S1. S1 $\cup$ S2 $\Rightarrow$ S4

Word sense disambiguation

The procedure

We combine the 3 methods to disambiguate as many words as possible:

- Identify and exclude the monosemous words. $\Rightarrow$ S0
- Apply Method 1 (Heuristics) to the remanining words (S0). $\Rightarrow$ S1
- Apply Method 2 (Density) to S0. $\Rightarrow$ S2
- Apply Method 3 (Statistics) to S0. $\Rightarrow$ S3
- Combine S1 and S2. In the case of conflict give priority to S1. S1 $\cup$ S2 $\Rightarrow$ S4
- Combine S4 and S3. In the case of conflict give priority to S4.

Word sense disambiguation

The procedure

We combine the 3 methods to disambiguate as many words as possible:

- Identify and exclude the monosemous words. $\Rightarrow$ S0
- Apply Method 1 (Heuristics) to the remanining words (S0). $\Rightarrow$ S1
- Apply Method 2 (Density) to S0. $\Rightarrow$ S2
- Apply Method 3 (Statistics) to S0. $\Rightarrow$ S3
- Combine S1 and S2. In the case of conflict give priority to S1. S1 $\cup$ S2 $\Rightarrow$ S4
- Combine S4 and S3. In the case of conflict give priority to S4.

(94% accuracy on a 1000 glosses set).

Logical form transformation

How does the LFT look like?

Once the words disambiguated we have the possibility to create the Logical Form Transformations (LFTs). The properties of the LFT are:

Logical form transformation

How does the LFT look like?

Once the words disambiguated we have the possibility to create the Logical Form Transformations (LFTs). The properties of the LFT are:

- The *Genus* predicate is at the first position and the rest of the LFT is the definition *differentia*.

Logical form transformation

How does the LFT look like?

Once the words disambiguated we have the possibility to create the Logical Form Transformations (LFTs). The properties of the LFT are:

- The *Genus* predicate is at the first position and the rest of the LFT is the definition *differentia*.
- A predicate is generated for every noun, verb, adjective or adverb of the gloss. The predicate is a concatenation of the morpheme's base form, the POS and the WordNet semantic sense. For example: `student:n#1`

Logical form transformation

How does the LFT look like?

Once the words disambiguated we have the possibility to create the Logical Form Transformations (LFTs). The properties of the LFT are:

- The *Genus* predicate is at the first position and the rest of the LFT is the definition *differentia*.
- A predicate is generated for every noun, verb, adjective or adverb of the gloss. The predicate is a concatenation of the morpheme's base form, the POS and the WordNet semantic sense. For example: `student:n#1`
- Verb predicates (as well as nominalizations of actions, events, states) have 3 arguments: `predicate`(e_1, x_1, x_2). e_1 is the *eventuality* of the action, x_1 the subject and x_2 the object.

Logical form transformation

How does the LFT look like? (2)

- Subject and object share the corresponding arguments with the verb predicate: "a person who backs a politician" = [person:n#1(x_1) & back:v#1(e_1, x_1, x_2) & politician:n#2(x_2)]

Logical form transformation

How does the LFT look like? (2)

- Subject and object share the corresponding arguments with the verb predicate: "a person who backs a politician" = [person:n#1(x_1) & back:v#1(e_1, x_1, x_2) & politician:n#2(x_2)]
- Modifiers share the same arguments with the predicate they modify: "a man-made object" = [object:n#1(x_1) & man-made:a#1(x_1)]

Logical form transformation

How does the LFT look like? (2)

- Subject and object share the corresponding arguments with the verb predicate: "a person who backs a politician" = [person:n#1(x_1) & back:v#1(e_1, x_1, x_2) & politician:n#2(x_2)]
- Modifiers share the same arguments with the predicate they modify: "a man-made object" = [object:n#1(x_1) & man-made:a#1(x_1)]
- Conjunction are transformed in predicates with variable number of arguments. The first argument being the result of the conjunction: "x_2 or x_3 or x_4" = [or(x_1, x_2, x_3, x_4)]

Logical form transformation

How does the LFT look like? (2)

- Subject and object share the corresponding arguments with the verb predicate: "a person who backs a politician" = [person:n#1(x_1) & back:v#1(e_1, x_1, x_2) & politician:n#2(x_2)]
- Modifiers share the same arguments with the predicate they modify: "a man-made object" = [object:n#1(x_1) & man-made:a#1(x_1)]
- Conjunction are transformed in predicates with variable number of arguments. The first argument being the result of the conjunction: "x_2 or x_3 or x_4" = [or(x_1, x_2, x_3, x_4)]
- Preposition predicates have 2 arguments: The first is the predicate of the head of the phrase to which the prepositional phrase is attached and the second is the prepositional object.

Logical form transformation

Sources of information

To implement LFTs, the following information are used:

Logical form transformation

Sources of information

To implement LFTs, the following information are used:

- Lexical and semantic disambiguation: contributes to the creation of predicate names.

Logical form transformation

Sources of information

To implement LFTs, the following information are used:

- Lexical and semantic disambiguation: contributes to the creation of predicate names.
- Phrasal parsing: detect phrase head and complements.

Logical form transformation

Sources of information

To implement LFTs, the following information are used:

- Lexical and semantic disambiguation: contributes to the creation of predicate names.
- Phrasal parsing: detect phrase head and complements.
- Syntactic transformation rules: detect subject and object.

Logical form transformation

Sources of information

To implement LFTs, the following information are used:

- Lexical and semantic disambiguation: contributes to the creation of predicate names.
- Phrasal parsing: detect phrase head and complements.
- Syntactic transformation rules: detect subject and object.
- Prepositional attachment resolution: determine the arguments of the preposition predicates.

Semantic form transformation

The purpose

The purpose of the SFTs is to provide information about the thematic (Who is the agent, experiencer, ...) and semantic relations (is something a part of something else, ...) of a gloss.

Semantic form transformation

Description

The SFTs are based on the LFT and the information provided by the WordNet database. The elements of the SFT are:

Semantic form transformation

Description

The SFTs are based on the LFT and the information provided by the WordNet database. The elements of the SFT are:

- Syntactic subject: recognized in the LFT by the formula $\texttt{noun}(x_1)\&\texttt{verb}(e_1, x_1, x_2)$. The definition of the thematic relation is based on information contained in WordNet: verb is hyponym of `cause` $\Rightarrow$ *agent*.

Semantic form transformation

Description

The SFTs are based on the LFT and the information provided by the WordNet database. The elements of the SFT are:

- Syntactic subject: recognized in the LFT by the formula $\texttt{noun}(x_1)\&\texttt{verb}(e_1, x_1, x_2)$. The definition of the thematic relation is based on information contained in WordNet: verb is hyponym of `cause` $\Rightarrow$ *agent*.
- Syntactic object: LFT form $\texttt{verb}(e_1, x_1, x_2)\&\texttt{noun}(x_2)$.

Semantic form transformation

Description

The SFTs are based on the LFT and the information provided by the WordNet database. The elements of the SFT are:

- Syntactic subject: recognized in the LFT by the formula $\texttt{noun}(x_1)\&\texttt{verb}(e_1, x_1, x_2)$. The definition of the thematic relation is based on information contained in WordNet: verb is hyponym of `cause` $\Rightarrow$ *agent*.
- Syntactic object: LFT form $\texttt{verb}(e_1, x_1, x_2)\&\texttt{noun}(x_2)$.
- Prepositional predicates: transformed into semantic or thematic relations. Example: ‘‘`sacrament of penance`’’ $\Rightarrow$ `kind-of(penance,sacrament)`.

Semantic form transformation

Description

The SFTs are based on the LFT and the information provided by the WordNet database. The elements of the SFT are:

- Syntactic subject: recognized in the LFT by the formula $\texttt{noun}(x_1)\&\texttt{verb}(e_1, x_1, x_2)$. The definition of the thematic relation is based on information contained in WordNet: verb is hyponym of `cause` $\Rightarrow$ *agent*.
- Syntactic object: LFT form $\texttt{verb}(e_1, x_1, x_2)\&\texttt{noun}(x_2)$.
- Prepositional predicates: transformed into semantic or thematic relations. Example: ‘‘`sacrament of penance`’’ $\Rightarrow$ `kind-of(penance,sacrament)`.
- Complex nominal predicates: If there is a semantic relation between the components of the predicate, then select the relation. Otherwise transform the predicate into thematic relation.

Semantic form transformation

Description

The SFTs are based on the LFT and the information provided by the WordNet database. The elements of the SFT are:

- Syntactic subject: recognized in the LFT by the formula $\texttt{noun}(x_1)\&\texttt{verb}(e_1, x_1, x_2)$. The definition of the thematic relation is based on information contained in WordNet: verb is hyponym of `cause` $\Rightarrow$ *agent*.
- Syntactic object: LFT form $\texttt{verb}(e_1, x_1, x_2)\&\texttt{noun}(x_2)$.
- Prepositional predicates: transformed into semantic or thematic relations. Example: ‘‘`sacrament of penance`’’ $\Rightarrow$ `kind-of(penance,sacrament)`.
- Complex nominal predicates: If there is a semantic relation between the components of the predicate, then select the relation. Otherwise transform the predicate into thematic relation.
- Adverbs and adjectives are connected to the predicate they modify through *attribute* relations.

Conclusion
Examples

Gloss	(a game played with rackets by two or four players who hit a ball back and forth over a net that divides a tennis court)
LFT	game:n#2(e_1,x_1,x_2) & play:v#2(e_1,x_1,x_2) & with(e_1,x_3) & racket:n#4(x_3) & by(e_1,x_1) & or(x_1,x_3,x_4) & two:n#1(x_3) & four:n#1(x_4) & player:n#1(x_1) & hit:v#1(e_2,x_1,x_5) & ball:n#1(x_5) & back_and_forth:r#1(e_2) & over(e_2,x_6) & net:n#5(x_6) & divide:v#5(e_3,x_6,x_7) & tennis_court:n#1(x_7)
SFT	*gloss*(tennis:n#1,game:n#2) *object*(game:n#2,play:v#2) *agent*(player:n#1,play:v#2) *attribute*(or(two:n#1,four:n#1),player:n#1) *agent*(player:n#1,hit:v#1) *object*(ball:n#1,hit:v#1) *location*(net:n#5,hit:v#1) *agent*(net:n#5,divide:v#5) *object*(tennis_court:n#1,divide:v#5)

Table 2: Transformations associated with the gloss of synset {tennis, lawn tennis}: (a game played with rackets by two or four players who hit a ball back and forth over a net that divides a tennis court)

```
<gloss pos="NOUN" synsetID="00012748">
  <synonymSet>animal, animate_being, beast, brute, creature, fauna</synonymSet>
 <text>
   a living organism characterized by voluntary movement
 </text>
  <wsd>
      <wf pos="DT" >a</wf>
      <wf pos="VBG" lemma="live" quality="gold" wnsn="3" >living</wf>
      <wf pos="NN" lemma="organism" quality="gold" wnsn="1" >organism</wf>
      <wf pos="VBN" lemma="characterize" quality="gold" wnsn="2" >characterized</wf>
      <wf pos="IN" >by</wf>
      <wf pos="JJ" lemma="voluntary" quality="gold" wnsn="2" >voluntary</wf>
      <wf pos="NN" lemma="movement" quality="gold" wnsn="1" >movement</wf>
  </wsd>
<parse quality="SILVER">
(TOP (S (NP (NN animal) )
        (VP (VBZ is)
            (NP (NP (DT a) (VBG living) (NN organism) )
                (VP (VBN characterized)
                    (PP (IN by)
                        (NP (JJ voluntary) (NN movement) ) ) ) ) )
        (. .) ) )
</parse>
 <lft quality="GOLD">
  animal:NN(x1) -> living:JJ(x1) organism:NN(x1) be:VB(e1, x1, x2) characterized:JJ(x2) by:IN(e1, x3) voluntary:JJ(x3) move
ment:NN(x3)
 </lft>
</gloss>
```

Conclusion
Possible uses

Informations provided by Extended WordNet could be used to prove or explain answers in *Question Answering* systems:

- D. Moldovan, C. Clark, S. Harabagiu, S. Maiorano (2003): COGEX: A Logic Prover for Question Answering HLT. Proceedings of HLT-NAACL 2003.
- D. Moldovan and V. Rus (2001): Explaining Answers with Extended WordNet, Proc. Of ACL 2001.