

Constructing Fuzzy Relations from WordNet for Word Sense Disambiguation

Christos Diou, Giorgos Katsikatsos and Anastasios Delopoulos

Multimedia Understanding Group

Electrical and Computer Engineering Department

Aristotle University of Thessaloniki

Word Sense Ambiguity

- Play (n):
 1. “He wrote several plays, but only one was produced on Broadway
 2. “The play lasted two hours”
 3. “The coach drew up the plays for her team”
 4. “The runner was out in a play by the shortstop”
 5. “The ball was still in the play”
 6. etc. etc. ...

Automatic Word Sense Disambiguation:

Determine the sense of a word, as used in a text.

Very difficult problem. Often human annotators disagree when senses are fine grained!

Previous work and applications

- An old problem.
 - Knowledge based approaches
 - Statistical WSD (supervised and unsupervised)
 - “Deep” and “shallow” WSD.
- Applications
 - Information Retrieval
 - Personalization
 - Content management and annotations
- In this work:
 - *WSD as a problem of imprecision and uncertainty*

Assumptions

- Context of a word: $C = [w_{-N} \dots w_0 \dots w_N]$
- Assumptions
 1. The sense of a word depends on the words appearing in its context. The syntax of the context and senses of its words can be ignored without significant loss of information.
 2. Although the exact nature of the association between a word w in C and a sense S of w_0 is unknown, it can be quantified as a degree of membership of w in the fuzzy set of words that are associated with S .
- Goal: construct $S_{w_0C} = d_1 / S_1 + \dots + d_n / S_n$
- Select the sense that corresponds to $\max_i(d_i)$

Senses as fuzzy sets of words

- Each sense is characterized by a fuzzy set

$$W_{S_i} = d_{i1} / w_{i1} + \dots + d_{ik} / w_{ik}$$

- Represents the degree of word-sense association in a context. Can also be interpreted as the degree of truth of the proposition $w_{ij} \Rightarrow S_i$
- By using conjunctive inference $d_i = U_{w_{ij}=w_l} (d_{ij})$
- Where d_i is the degree of S_i in the set $S_{w_0 C}$ and w_l are words in C . U is a fuzzy union operator, usually *not* the standard union.

Algorithm 1: Sense ranking

Algorithm 1 Sense ranking.

Input: n possible senses S_i associated with word w_0 to be disambiguated, a fuzzy set W_{S_i} for each sense and the context C .

Output: The membership degrees d_i of S_i in the fuzzy set $S_{w_0 C}$.

```
1: for  $i := 1$  to  $n$  do
2:    $d_i := 0$ 
3:   for  $l := -N$  to  $N$  do
4:      $j := \text{search\_word}(W_{S_i}, w_l)$ 
5:     if  $j > 0$  then
6:        $d_i = \mathcal{U}(d_i, d_{ij})$ 
7:     end if
8:   end for
9: end for
```

Notes: `search_word` is an arbitrary search algorithm that returns the positive index j of the word w_{ij} if $w_l = w_{ij}$ (i.e., if w_l is found in the set W_{S_i}) or negative otherwise.

Need to define:

The senses S_i

A fuzzy set W_{S_i} for each sense

WordNet

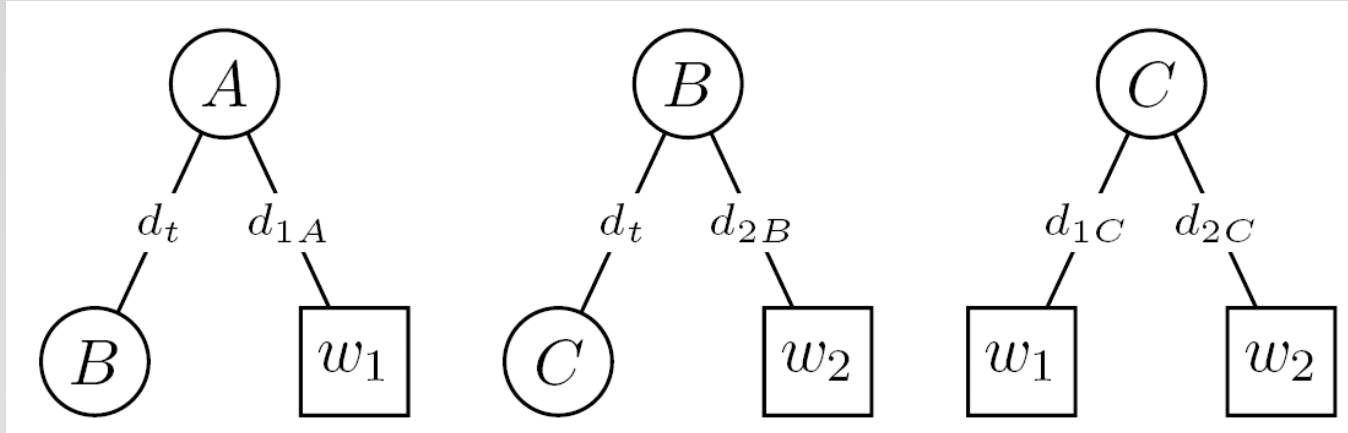
- Senses
 - Correspond to conceptual entities
 - Defined by a set of synonymous words (synsets)
- Relations between senses
 - E.g., hyponymy (is a kind of), meronymy (is part of).
- Additional information
 - A small description of each sense (gloss)
 - Estimated frequency of appearance of a sense w.r.t. a word.

WordNet, a lexical database for the English language,
homepage: <http://wordnet.princeton.edu>

WordNet as a fuzzy Knowledge Base

- Crisp relations in WordNet: “*allergology#1 is a kind of medicine#1*”.
- Additional assumption:
 - A sense S_j has a degree of membership to the fuzzy set of senses that are associated with S_i in a context if the two senses are related in WordNet with $R_t(S_j, S_i)$
- In our experiments:
 - Hyponymy, hyperonymy, holonymy, meronymy and domain relations with constant weights (0.9, 0.4, 0.4, 0.9 and 0.9 resp.)
- Hence for each $R_t(S_j, S_i)$ in WordNet, a fuzzy relation $\mathbf{R}_t(S_j, S_i) = d_t$ is constructed.

Quantification of word – sense relations

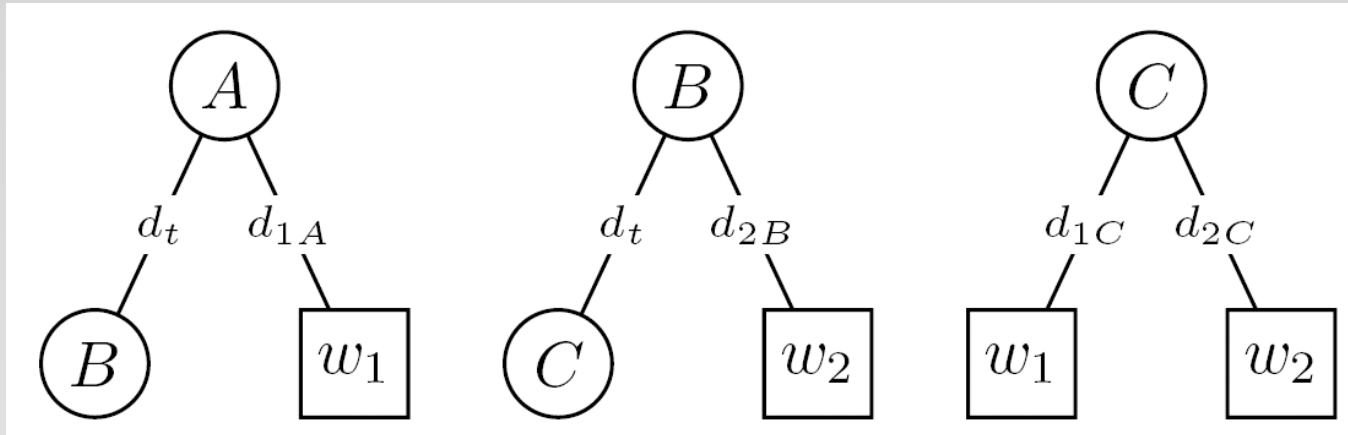


Extend \mathbf{R}_t to include word-sense associations and calculate its transitive closure \mathbf{R}_t^+

If the properties of sup- t composition are not desirable, then generalized composition can be used. In the Figure

$$\mathbf{R}_t^+(w_2, A) = I(U(I(d_{1C}, d_t), d_{2B}), d_t)$$

Quantification of word – sense relations (2)



The direct word – sense relation is given by

$$d_{wS_i} = 0.3 + 0.7 \frac{1}{1 + e^{-2(x-c)}}$$

Where x is the number of occurrences of the sense based on WordNet and $c = fmax - fmin$ is the difference of occurrence between the most and least frequent senses of WordNet.

Algorithm 2: traverse tree

Algorithm 2 Depth first tree traversal of relation R_t and retrieval of the elements of W_{S_i} (or equivalently calculate $R_t^+(w, S_i)$ for all words with significant membership degree).

Input: WordNet information, W_{S_i} , t , the current sense S the current depth of the traversal and a depth threshold T .

Output: Updated W_{S_i} .

```

1: if depth := 0 then
2:    $S :=$  next sense in depth first traversal
3:   depth := 1
4:    $W_{S_i} :=$ traverse_tree( $W_{S_i}, t, S, \text{depth}, T$ )
5:   return
6: end if
7: if depth >  $T$  then
8:   return
9: end if
10: for all elements  $w$  in the synset of  $S$  do
11:    $d_f :=$  (given by Equation (4) if frequency is taken
             into account or 1 otherwise)
12:    $d_w := \mathcal{I}(\underbrace{d_t, \dots, d_t}_{\text{depth times}}, d_f)$ 

```

```

13:   if  $w$  already in  $W_{S_i}$  with  $d'_w$  then
14:      $d_w := \mathcal{U}(d'_w, d_w)$ 
15:     Change membership degree to  $d_w$ 
16:   else
17:     Add  $w$  to  $W_{S_i}$  with degree  $d_w$ 
18:   end if
19: end for
20:  $S :=$  next sense in depth first traversal
21: depth := depth + 1
22:  $W_{S_i} :=$ traverse_tree( $W_{S_i}, t, S, \text{depth}, T$ )

```

Algorithm 3: Construct all W_{S_i}

Algorithm 3 Calculate all W_{S_i}

Input: WordNet information, the n senses of the word w_0 to be disambiguated and the traversal depth threshold T .

Output: $W_{S_i}, i = 1, \dots, n$.

```
1: for  $i := 1$  to  $n$  do
2:   /* Initialize and get synset of  $S_i$ . */
3:    $W_{S_i} = \emptyset$ 
4:   for all elements  $w$  in the synset of  $S_i$  do
5:     Add  $w$  to  $W_{S_i}$  with degree given by (4) or 1 if
       frequency is not taken into account
6:   end for
7:   for all types of relations  $t$  do
8:     /* traverse_tree is given in Algorithm 2 */
9:     depth := 0
10:     $W_{S_i} := \text{traverse\_tree}(W_{S_i}, t, S_i, \text{depth}, T)$ 
11:  end for
12: end for
```

Experimental Setup

- 103 tagged texts (brown1 collection of the brown corpus).
- Automatic sense tagging on all tagged nouns.
- Fine – grained senses are used. An upper limit of 80% precision is assumed (72.5% interannotator agreement in Senseval – 3 for all words).
- Parameters:

Parameter	Value
$\mathcal{U}(a, b)$	$a + b - ab$
$\mathcal{I}(a, b)$	ab
N	36
d_t	(see Section 3.1)
T	4

(or slide 9 of the presentation)

Results

Without frequency information

<i>Measure</i>	<i>Precision (%)</i> (polysemous)	<i>Precision (%)</i> (all words)
Mean	58.5	67.2
σ	6.6	6.4
max	80	83.6
min	43.9	49.7

With frequency information

<i>Measure</i>	<i>Precision (%)</i> (polysemous)	<i>Precision (%)</i> (all words)
Mean	62.7	70.5
σ	6.6	6.0
max	81.4	84.8
min	44.3	56.6

Conclusions and Future Work

- Modeling WSD as a problem of imprecision seems promising.
- No training or online statistical computations are performed. The fuzzy KB can be constructed from WordNet and is then used directly.
- Computationally efficient.
- Open issues:
 - Model WSD as a problem of uncertainty, e.g., using possibility theory.
 - Apply the proposed approach to practical application scenarios to further identify its weaknesses.

Thank You!

Multimedia Understanding Group

Department of Electrical & Computer Engineering

Aristotle University of Thessaloniki