

Kelime Anlamları

(Word Semantics)



Prof.Dr. Banu Diri

İçerik

- Kelime Anlamları ve WORDNET
 - Rakamlarla Wordnet
 - İlişki türleri
 - Wordnet'te anlamsal benzerlik ölçümleri
 - Bağ sayma
 - Leacock & Chodorow (1998)
 - Wu & Palmer (1994)
 - Ortak / Müşterek bilgi
 - Jiang-Conrath (1997)
 - Lin (1998)



Kelime Anlamları ve WORDNET

- WordNet sistemi kelimeleri, anlamlarını (tanımlarını), ilişkili olduğu kelimeleri ve ilişki türlerini, kelimelere erişim için çeşitli yazılımları içerir.
- Eş anlamlı kelimeler, eş küme (synset) ler içinde birlikte yer alırlar.
- İlişkiler eş kümeler arasında tanımlanmıştır.
- Elle oluşturulmuş (yüksek kalite)
- 10 yıldan fazla emek (yüksek iş gücü)
- **İNGİLİZCE** için.



Rakamlarla Wordnet

Category	Unique Forms	# of Senses
Noun	117,097	145,104
Verb	11,488	24,890
Adjective	22,141	31,302
Adverb	4,601	5,720



WordNet Noun Relations

Relation	Definition	Example
Hypernym	From concepts to superordinates	<i>breakfast</i> → <i>meal</i>
Hyponym	From concepts to subtypes	<i>meal</i> → <i>lunch</i>
Has-Member	From groups to their members	<i>faculty</i> → <i>professor</i>
Member-Of	From members to their groups	<i>copilot</i> → <i>crew</i>
Has-Part	From wholes to parts	<i>table</i> → <i>leg</i>
Part-Of	From parts to wholes	<i>course</i> → <i>meal</i>
Antonym	Opposites	<i>leader</i> → <i>follower</i>

Hyponym (alt kavram)/Hypernym (üst kavram) (Is-A relationships)

Bütün ilişkilerin %60'ı



Verb Relations in WordNet

Relation	Definition	Example
Hypernym	From events to superordinate events	<i>fly</i> → <i>travel</i>
Troponym	From events to their subtypes	<i>walk</i> → <i>stroll</i>
Entails	From events to the events they entail	<i>snore</i> → <i>sleep</i>
Antonym	Opposites	<i>increase</i> ↔ <i>decrease</i>

Adjective & Adverb Relations in WordNet

Relation	Definition	Example
Antonym	Opposite	<i>heavy</i> ↔ <i>light</i>
Adverb	Opposite	<i>quickly</i> ↔ <i>slowly</i>

Troponym: bir şeyin nasıl yapıldığı çeşidini gösterir

Entails: bir şeyin nasıl yapıldığını gösterir



Wordnet'le anlamsal benzerlik bulma

- İki temel yaklaşım
 - Bağ sayma
 - Taksonomi yeterli
 - Ortak / Müşterek bilgi (Mutual Information)
 - Taksonomi ve corpus kullanır



Leacock & Chodorow (1998)

$$\text{sim}_{LC}(c_1, c_2) = -\log \frac{\text{len}(c_1, c_2)}{2L}$$

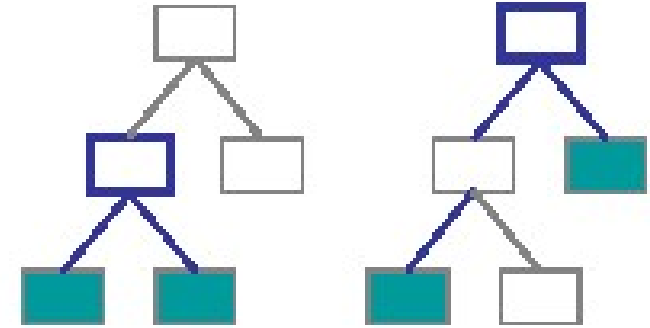
- $\text{len}(c_1, c_2)$ iki synset arasındaki en kısa yolun uzunluğu (*benzerlik değeriyle ters orantılı*)
- L , tüm taksonominin derinliği



Wu & Palmer (1994)

$$sim_{Wu\&Palmer}(c_1, c_2) = \frac{2 \times N_3}{N_1 + N_2 + 2 \times N_3}$$

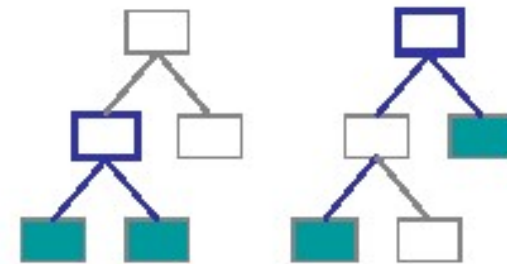
- N_1 ve N_2 , en yakın ortak üst synset'lerine iki synsetin IS-A bağlarıyla uzaklıkları (*benzerlik değeriyle ters orantılı*)
- N_3 , en yakın ortak üst synset'in kök synset'e IS-A bağlarıyla uzaklığı (*büyüklüğü ortak synset'in spesifikliğini gösterir*)



Jiang-Conrath (1997)- Lin (1998)

$$\begin{aligned} \text{sim}_{\mathcal{JC}}(c_1, c_2) &= \frac{2 \log(p(\text{lcs}(c_1, c_2)))}{\log(p(c_1)) + \log(p(c_2))} \\ \text{sim}_L(c_1, c_2) &= \frac{2 * \log(p(\text{lcs}(c_1, c_2)))}{\log(p(c_1)) + \log(p(c_2))} \end{aligned}$$

$\text{lcs}(c_1, c_2)$ en yakın ortak üst synset



- A, her iki kavramı da içeren en spesifik kavramı kullanır (İki kavramın beraber geçtiği doküman sayısına benzer)
- B, iki kavramdan herhangi birini içeren doküman sayısına benzer

Benzerlik Metotlarının Karşılaştırılması

- Bütün metotlar İngilizce 38 kelime çiftine uygulanmış.
- Bulunan benzerlik değerlerinin, insan yargılarıyla olan korelasyonları yandaki tabloda verilmiştir.

<i>Method</i>	<i>Type</i>	<i>Correlation</i>
Wu & Palmer 1994	Edge Counting	0.74
Li 2003	Edge Counting	0.82
Leacock & Chodorow 1998	Edge Counting	0.82
Resnik 1999	Info. Content	0.79
Lin 1998	Info. Content	0.82
Lord 2003	Info. Content	0.79
Jiang & Conrath 1998	Info. Content	0.83
Tversky 1977	Feature Based	0.73
Adapted Lesk 2002	Feature Based	0.37*
Rodriguez 2003	Hybrid	0.71



Kaynaklar

- Steve Vincent, Text Extraction, Similarity and WordNet
- http://www.ise.gmu.edu/~carlotta/teaching/INFS-795-s06/readings/Similarity_in_WordNet.pdf
- <http://www.cs.utah.edu/~sidd/documents/msthesis03ppt.pdf>



Son ☹️

