
Softwareprojekte

Anette Frank

Seminar für Computerlinguistik

Universität Heidelberg

SS 2009

Projekt 3

WSD für VerbNet Klassen

WSD für VerbNet Klassen

- Ressourcen: SemLink: VNet/PBank/WNet Mapping
- Training eines WSD Systems für VerbNet-Klassen
- Reimplementierung & Erweiterung von Abend et al. 2008
 - Verbesserung und Erweiterung von Merkmalen
 - Integration von WordNet senses/disambiguierung

Literatur

Literatur zu VerbNet und SemLink: siehe Projekt 2

Literatur zu VerbNet Class disambiguation

- Abend, Reichart and Rappoport (2008): A supervised algorithm for verb disambiguation into VerbNet classes. in: Coling 2008.
<http://www.cs.huji.ac.il/~omria01/papers/VN.pdf>

Literatur zu WSD

- D. Yarowsky and R. Florian (2002): Evaluating Sense Disambiguation Performance Across Diverse Parameter Spaces. In Journal of Natural Language Engineering, Natural Language Engineering 8 (4): 293-310. <http://www.cl.uni-heidelberg.de/courses/archiv/ws07/corpus/Semantics/yarowsky-florian-NLE.pdf>
- Gerard Escudero and Llus M Arquez and German Rigau: 2000: Naive Bayes and exemplar-based approaches to Word Sense Disambiguation revisited, In Proceedings of the 14th European Conference on Artificial Intelligence, 421--425. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.35.229>
- Claire Cardie (to appear): Word Sense Disambiguation. In: Language and Linguistics Compass.

VerbNet: A large Verb Lexicon for English

- **Event structure:** aspektuelle Klassen
 - *Activities and Processes*: only predicates referring to *during(E)* stage
 - *Activities w/ end point*: predicates holding *during(E)* & at *end(E)* stage
 - *Accomplishments*: also have predicates holding of result stage *result(E)*
 - *States*: predicates are stated over entire event (E)
- **Thematische Rollen** (23 Rollen)
- **Selektionsrestriktionen** (EWN top level entries)
 - Region: *from under the rug*
 - Place: *under the rug*
 - Object: *on the table*
- **Syntaktische Rahmen**
 - hit: Agent V Patient (*John hit the ball*)
Agent V at Patient (*John hit at the window*)
- **Semantische Prädikate**: Resultatzustände: Possession, Motion, ..

Beispiel

Hit class

Class	hit-18.1		
Parent	—		
Members	bang (1,3), bash(1), batter(1,2,3), beat(2,5), ..., hit(2,4,7,10), kick(3), ...		
Themroles	Agent Patient Instrument		
Selrestr	Agent[+int_control] Patient[+concrete] Instrument[+concrete]		
Frames	Name	Syntax	Semantic Predicates
	Transitive	Agent V Patient “Paula hit the ball”	cause(Agent, E) $\wedge$ manner(during(E),directedmotion,Agent) $\wedge$!contact(during(E), Agent, Patient) $\wedge$ manner(end(E),forceful, Agent) $\wedge$ contact(end(E), Agent, Patient)
	Transitive with Instrument	Agent V Patient Prep(with) Instrument “Paula hit the ball with a stick”	cause(Agent, E) $\wedge$ manner(during(E),directedmotion,Agent) $\wedge$!contact(during(E),Instrument,Patient) $\wedge$ manner(end(E),forceful, Agent) $\wedge$ contact(end(E), Instrument,Patient)

Levin's Hypothese und Levin Klassen

Beth Levin (1993): *English Verb Classes and Alternations*.

Hypothese: Syntaktische Alternationseigenschaften der Argumentstruktur sind (partiell) semantisch bedingt

Levin Classes / VerbNet:

- Umfangreiche empirische Analyse („Levin classes“)
 - *break Verben*: crack, rip (reißen), shatter (zerbrechen), snap (brechen), ...
 - *cut Verben*: cut, hack (zerhacken), saw (schneiden), scratch, slash (aufschlitzen), ...
 - *touch Verben*: pat (tätscheln), stroke (streicheln), tickle (kitzeln), ...
 - *hit Verben*: bash (verprügeln), kick, pound (hämmern, klopfen), tap, whack (schlagen, durchprügeln).
- VerbNet: Kipper-Schuler, K. (2005). *VerbNet: A broad-coverage, comprehensive verb lexicon*. PhD thesis, University of Pennsylvania.
- <http://verbs.colorado.edu/~mpalmer/projects/verbnet.html>

Verbalternationen: „Levin Classes“

	middle	conative	Body part possessor
cut	+	+	+
break	+	-	-
touch	-	-	+
hit	-	+	+

- *break* und *cut*-Verben:
Zustandswechsel (durch Schneiden, Hacken, Sägen, ...).
- *touch* und *hit* Verben:
kein Zustandswechsel

- *cut*-Verben:
bezeichnen die Art und Weise des Zustandswechsels
- *break*-Verben:
keine Anzeige der Art u. Weise des Zustandswechsels

Example

Class	Transfer_mesg-37.1		
Parent	—		
Members	cite(1,3,4), demonstrate(1), ...		
Themroles	Agent Topic Recipient		
Selrestr	Agent[+animate] Topic[+message] Recipient[+animate]		
Frames	Name	Syntax	Semantic Predicates
	Transitive	Agent V Topic	transfer_info(during(E),Ag,?,Top)∧ cause(Ag,E)
	Dative (to-PP variant)	Agent V Topic Prep(to) Recipient	transfer_info(during(E),Ag,Rec,Top)∧ cause(Ag,E)

Class	Transfer_mesg-37.1-1		
Parent	Transfer_mesg-37.1		
Members	dictate(2), quote(1), read(3)		
Themroles			
Selrestr			
Frames	Name	Syntax	Semantic Predicates
	Dative (ditrans variant)	Agent V Recipient Topic	transfer_info(during(E),Ag,Rec,Top)∧ cause(Ag,E)

Figure 3.5: Example entries for the *Transfer of a Message* - levels 1 and 2 classes

Size of VerbNet (Kipper, 2005)

- Size
 - 5,200 verbs
 - 3,800 lemmas
 - 237 top-level classes
 - 197 new subclasses
- 23 thematic roles plus 36 selectional restriction types
- 94 semantic predicates
- 357 syntactic frames
- Hierarchy of prepositions (57 entries)

Mappings: PropBank, WN, VN, FrameNet

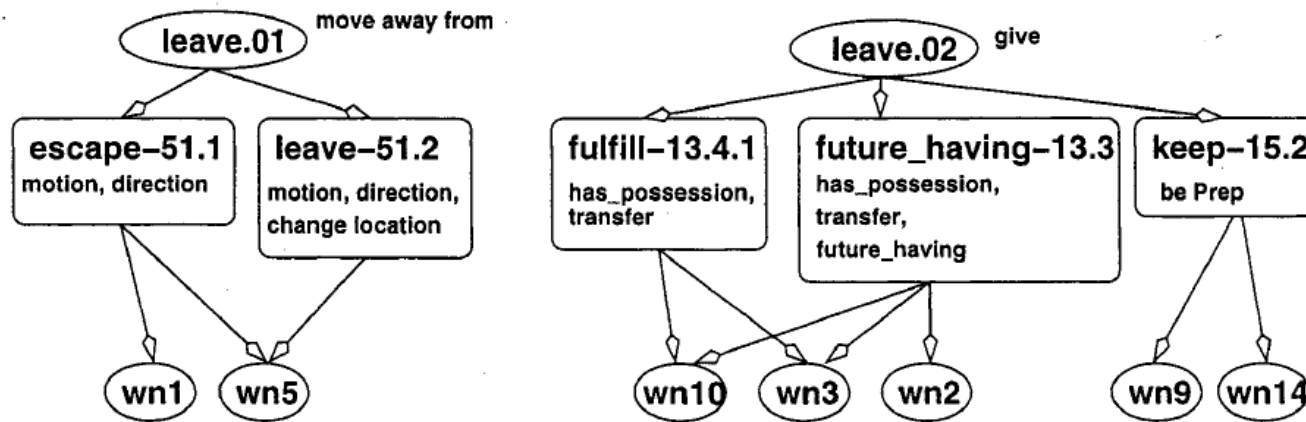
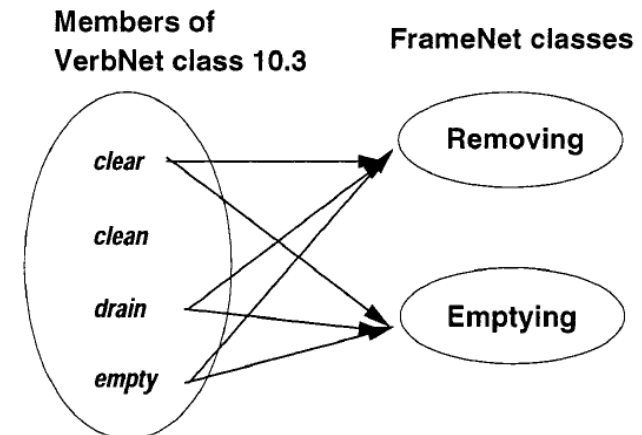


Figure 6.2: Implicit mappings between PropBank and WordNet verb senses



Literatur

- Karin Kipper-Schuler (2005): VerbNet: A Broad-Coverage, Comprehensive Verb Lexicon. PhD Dissertation, University of Pennsylvania.
- Kipper, K., Korhonen, A., Ryant, N., and Palmer, M. (2008). A Large-Scale Classification of English Verbs. *Journal of Language Resources and Evaluation*, 42(1):21-40.
- Verb Index: <http://www-personal.umich.edu/~jlawler/levin.html>
- Kipper-Schuler (2006): VerbNet: extensions and mappings to other lexical resources, SALSA-Workshop: http://www.coli.uni-saarland.de/projects/salsa/workshop/contents/workshop_slides/slides4.pdf
- SemLink: Mapping to PropBank: <http://verbs.colorado.edu/semlink/semlink1.1/vn-pb/README.TXT>

Edward Loper, Szu-ting Yi and Martha Palmer, 2007. *Combining Lexical Resources: Mapping Between PropBank and VerbNet*. *Proc. of the 7th International Workshop on Computational Linguistics, Tilburg, the Netherlands*.

- Colin Bannard, Chris Callison-Burch: *Paraphrasing with Bilingual Parallel Corpora*, ACL 2005.
- Alessandra Zarcone and Alessandro Lenci: Computational Models for Event Type Classification in Context, LREC 2008.
http://www.lrec-conf.org/proceedings/lrec2008/pdf/315_paper.pdf

Projekt 5

Extraktion kausaler Relationen

Extraktion kausaler Relationen

Extraktion kausaler Relationen

Why do babies cry?



Hunger *If your baby is **hungry** her **cry** will be loud and insistent, often accompanied by rooting and finger sucking*

<http://ukfamily.co.uk/ages-stages/baby-toddler/baby-behaviour/crying/cry-baby.html>

Lernen auf Daten mit Relationsindikatoren

- Eduardo Blanco, Nuria Castell, Dan Moldovan (2008): "**Causal Relation Extraction**", *Proceedings of LREC 2008*, Marrakech.
- Schritt 1. Erstellung und Evaluierung von syntaktischen Patterns

relator	Example	
	encoding causation	not encoding causation
after	<i>Marty stood with his mouth hanging open foolishly after it had happened.</i>	<i>The executions took place a few hours after the radio announced their conviction.</i>

Relator	Occurrences encoding causation	Causations signaled
after	15.35 %	6.85 %
as	11.21 %	7.34 %
because	98.43 %	73.39 %
since	49.61 %	12.52 %

Lernen auf Daten mit Relationsindikatoren

- Eduardo Blanco, Nuria Castell, Dan Moldovan (2008): "**Causal Relation Extraction**", *Proceedings of LREC 2008*, Marrakech.
- Schritt 1: Erstellung und Evaluierung von syntaktischen Patterns
- Schritt 2: Erstellen eines annotierten Korpus
- Schritt 3: Merkmalsextraktion und Training eines Klassifikators (Weka)
 - Semantische Verbklassen
 - Modifikatoren
 - ...

Lernen auf Daten mit Relationsindikatoren

- Eduardo Blanco, Nuria Castell, Dan Moldovan (2008): "**Causal Relation Extraction**", *Proceedings of LREC 2008*, Marrakech.
- Schritt 1: Erstellung und Evaluierung von syntaktischen Patterns
- Schritt 2: Erstellen eines annotierten Korpus
- Schritt 3: Merkmalsextraktion und Training eines Klassifikators
- Schritt 4: Evaluierung

Class	Precision	Recall	F-Measure
<i>cause</i>	0.955	0.842	0.895
$\neg$ <i>cause</i>	0.869	0.964	0.914

Table 6: Results obtained during testing.

Projektaufgabe

Reimplementierung/Adaptation für das Deutsche

- Annotierte Korpora
 - TIGER, TüBa-D/Z
- Web Corpora: deWAC
- Syntaktische und Semantische Annotation
- Erstellung eines Goldstandards (!) und Evaluation