Knowing a Word(sense) by the company it keeps

Martha Palmer

Workshop on Unsupervised and Minimally Supervised Learning of Lexical Semantics

University of Colorado

June 5, 2009

Computational Language and EducAtion Research

C L E A R

WSD and Natural Language Processing (NLP)

- Increasing evidence that word sense disambiguation (WSD), determining the meaning a word bears in its given context, can improve NLP applications such as
  - Machine translation, (Carpuat and Wu, 2007; Chan, Ng and Chiang, 2007)
  - Information retrieval, (Gonzalo *et al.*, 1998; Sanderson, 2000 ; Stokoe 2003;).

Word sense in Machine Translation

- Different syntactic frames
  - *John left the room*

    Juan saiu do quarto. (Portuguese)
  - *John left the book on the table.*

    Juan deizou o livro na mesa.
- Same syntactic frame? Same sense?
  - *John left a fortune.*

    Juan deixou uma fortuna.

Word sense in Machine Translation – not just syntax

- Different syntactic frames
  - *John left the room*

    Juan saiu do quarto. (Portuguese)
  - *John left the book on the table.*

    Juan deizou o livro na mesa.
- Same syntactic frame? Same sense?
  - *John left a fortune to the SPCA.*

    Juan deixou uma fortuna.

Automatic Word Sense Disambiguation – Supervised Approach

- Manually annotated training data based on a pre-existing sense inventory
- Train Machine Learning classifiers
- Run on new data
- Evaluate against Gold Standard Test data

- *Which Sense Inventory?*

Outline

- Sense Distinctions
- Annotation
  - Sense Inventories created by human judgments
    - WordNet
    - PropBank and VerbNet
    - Mappings to VerbNet and FrameNet
    - Groupings of WordNet senses
    - Hierarchical model of sense distinctions
- OntoNotes – based on groupings
- A note about human judgements
- Automatic Word Sense Disambiguation
- What is a word(sense)’s company?

WordNet – Princeton

(Miller 1985, Fellbaum 1998)

On-line lexical reference (dictionary)

- Nouns, verbs, adjectives, and adverbs grouped into synonym sets
- Other relations include hypernyms (ISA), antonyms, meronyms
- Typical top nodes - 5 out of 25
  - *(act, action, activity)*
  - *(animal, fauna)*
  - *(artifact)*
  - *(attribute, property)*
  - *(body, corpus)*

WordNet – Princeton – *leave, n.4, v.14*

(Miller 1985, Fellbaum 1998)

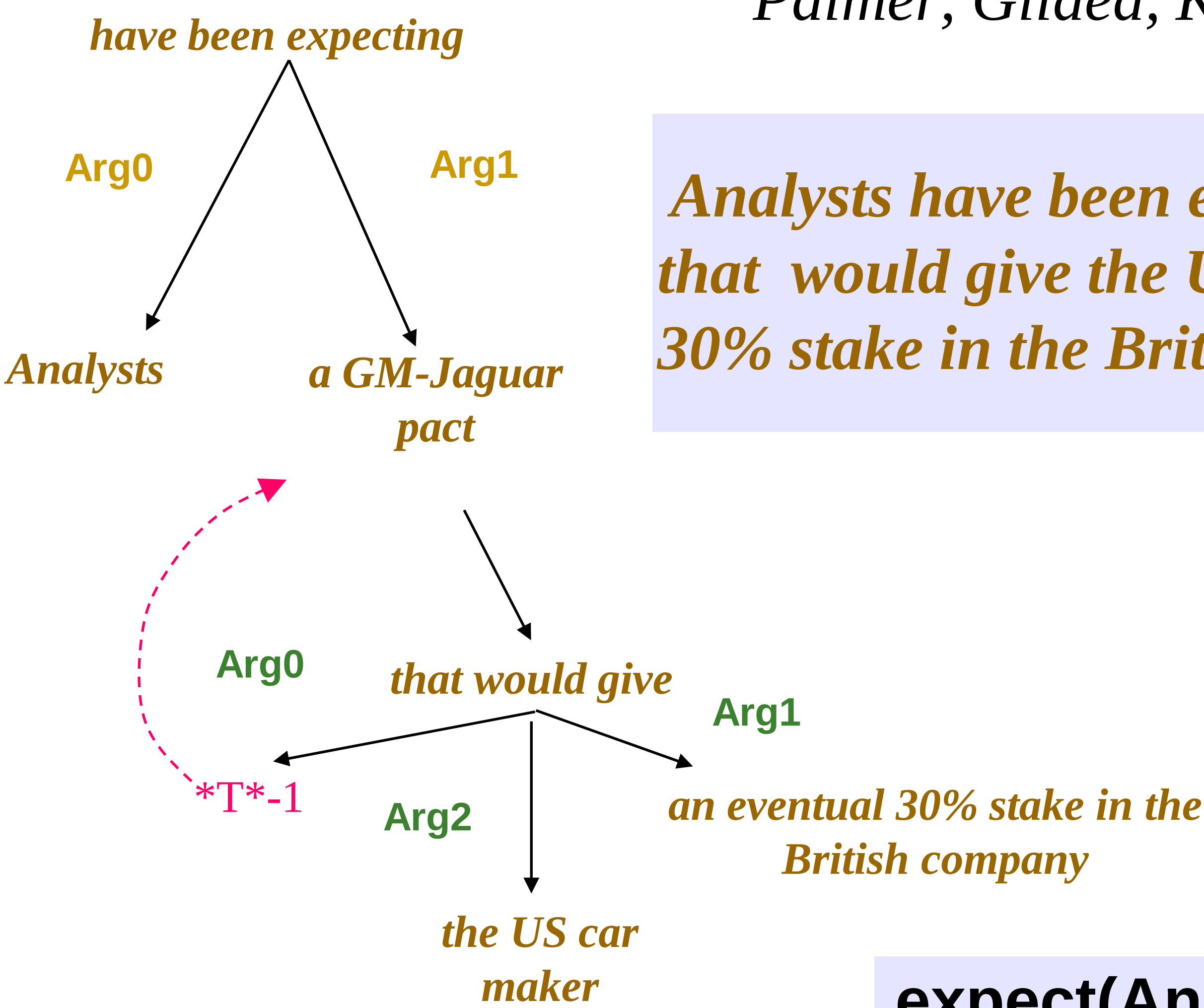
- Limitations as a computational lexicon
  - Contains little syntactic information
  - No explicit lists of participants
  - Sense distinctions very fine-grained,
  - Definitions often vague
- Causes problems with creating training data for supervised Machine Learning – SENSEVAL2
    - Verbs > 16 senses (including *call*)
    - Inter-annotator Agreement ITA 71%,
    - Automatic Word Sense Disambiguation, WSD 64%

Dang & Palmer, SIGLEX02

PropBank – WSJ Penn Treebank

Palmer, Gildea, Kingsbury., CLJ 2005

Analysts have been expecting a GM-Jaguar pact that would give the U.S. car maker an eventual 30% stake in the British company.



expect(Analysts, GM-J pact)
give(GM-J pact, US car maker, 30% stake)

Lexical Resource - Frames Files: *give*

Roles:

Arg0: giver

Arg1: thing given

Arg2: entity given to

Example: double object

The executives gave the chefs a standing ovation.

Arg0:	*The executives*
REL:	*gave*
Arg2:	*the chefs*
Arg1:	*a standing ovation*

Word Senses in PropBank

- Orders to ignore word sense not feasible for 700+ verbs
  - *Mary left the room*
  - *Mary left her daughter-in-law her pearls in her will*

Frameset **leave.01** "move away from":
Arg0: entity leaving
Arg1: place left

Frameset **leave.02** "give":
Arg0: giver
Arg1: thing given
Arg2: beneficiary

How do these relate to word senses in WordNet, VerbNet and FrameNet?

Limitations to PropBank

- Sense distinctions are very coarse-grained – only 700 verbs
  - High ITA, > 94%, High WSD,> 90%
- Args2-4 seriously overloaded, poor performance
  - VerbNet and FrameNet both provide more fine-grained role labels
- WSJ too domain specific,
  - Additional Brown corpus annotation & GALE data
  - FrameNet has selected instances from BNC

Levin classes as a Sense Inventory? – (Levin, 1993)

- Verb class hierarchy: 3100 verbs, 47 top level classes, 193
- Each class has a syntactic signature based on alternations.
  John broke the jar. / The jar broke. / Jars break easily.
  change-of-state

  *John cut the bread. / *The bread cut. / Bread cuts easily.*
  change-of-state, recognizable action,
  sharp instrument

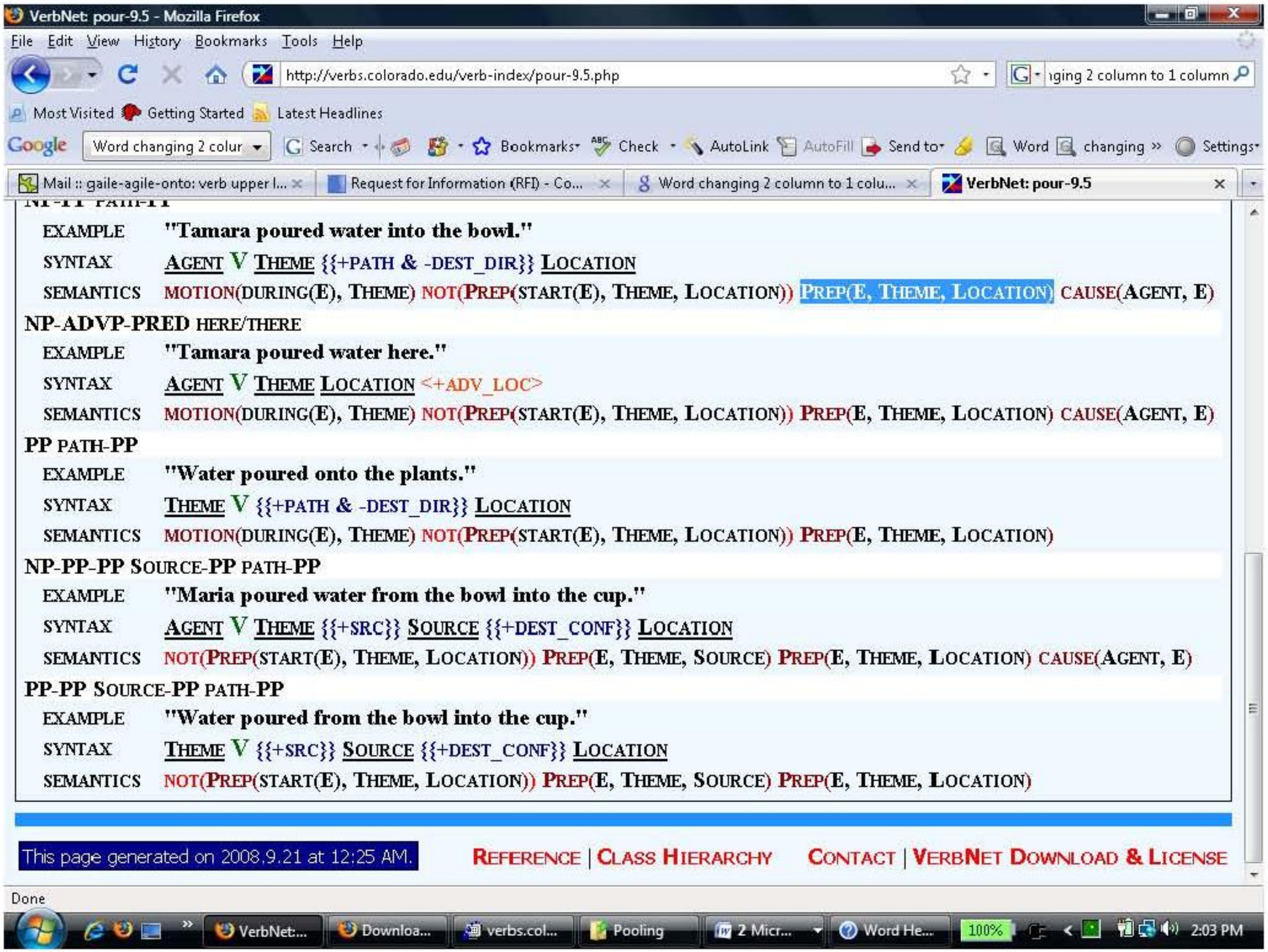
  *John hit the wall. /   *The wall hit. / *Walls hit easily.*
  contact, exertion of force

VerbNet – *Karin Kipper*

- Class entries:
  - Capture generalizations about verb behavior
  - Organized hierarchically
  - Members have common semantic elements, semantic roles and syntactic frames
- Verb entries:
  - Refer to a set of classes (different senses)
  - each class member linked to WN synset(s) (not all WN senses are covered)

VerbNet example – *Pour-9.5*

VerbNet *Pour-9.5* (cont.)



EXAMPLE "Tamara poured water into the bowl."
SYNTAX AGENT V THEME {{+PATH & -DEST_DIR}} LOCATION
SEMANTICS MOTION(DURING(E), THEME) NOT(PREP(START(E), THEME, LOCATION)) PREP(E, THEME, LOCATION) CAUSE(AGENT, E)

NP-ADVP-PRED HERE/THERE
EXAMPLE "Tamara poured water here."
SYNTAX AGENT V THEME LOCATION <+ADV_LOC>
SEMANTICS MOTION(DURING(E), THEME) NOT(PREP(START(E), THEME, LOCATION)) PREP(E, THEME, LOCATION) CAUSE(AGENT, E)

PP PATH-PP
EXAMPLE "Water poured onto the plants."
SYNTAX THEME V {{+PATH & -DEST_DIR}} LOCATION
SEMANTICS MOTION(DURING(E), THEME) NOT(PREP(START(E), THEME, LOCATION)) PREP(E, THEME, LOCATION)

NP-PP-PP SOURCE-PP PATH-PP
EXAMPLE "Maria poured water from the bowl into the cup."
SYNTAX AGENT V THEME {{+SRC}} SOURCE {{+DEST_CONF}} LOCATION
SEMANTICS NOT(PREP(START(E), THEME, LOCATION)) PREP(E, THEME, SOURCE) PREP(E, THEME, LOCATION) CAUSE(AGENT, E)

PP-PP SOURCE-PP PATH-PP
EXAMPLE "Water poured from the bowl into the cup."
SYNTAX THEME V {{+SRC}} SOURCE {{+DEST_CONF}} LOCATION
SEMANTICS NOT(PREP(START(E), THEME, LOCATION)) PREP(E, THEME, SOURCE) PREP(E, THEME, LOCATION)

This page generated on 2008.9.21 at 12:25 AM.

REFERENCE | CLASS HIERARCHY CONTACT | VERBNET DOWNLOAD & LICENSE

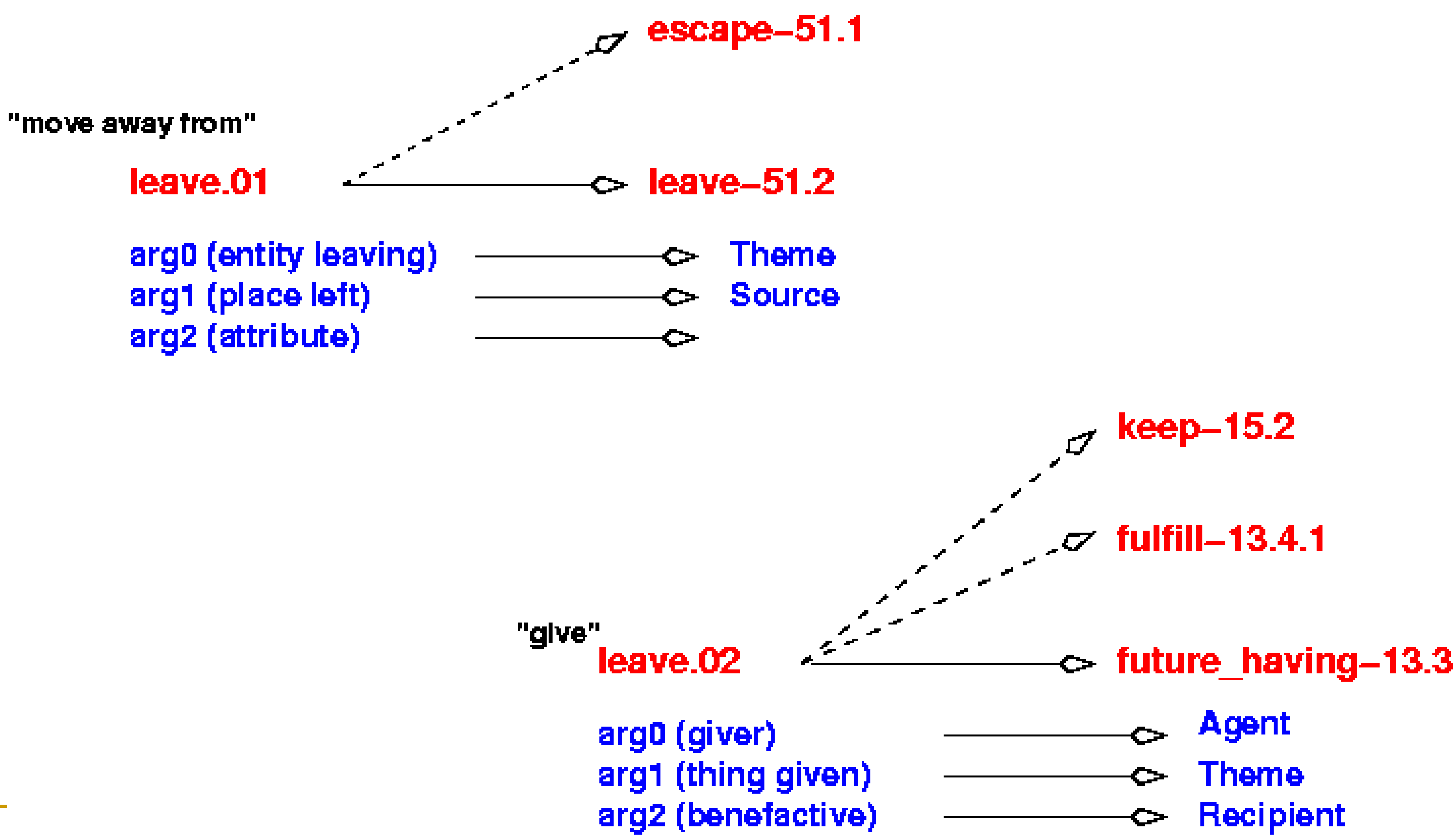
Mapping from PropBank to VerbNet (similar mapping for PB-FrameNet)

Frameset id = ***leave.02***	Sense = *give*	VerbNet class = ***future-having 13.3***
Arg0	Giver	Agent/Donor*
Arg1	Thing given	Theme
Arg2	Benefactive	Recipient

*FrameNet Label

Baker, Fillmore, & Lowe, COLING/ACL-98
Fillmore & Baker, WordNetWKSHP, 2001

Mapping from PB to VerbNet verbs.colorado.edu/~semlink



Mapping PropBank/VerbNet
http://verbs.colorado.edu/~mpalmer/verbnet

- Extended VerbNet (100+ new classes from *(Korhonen and Briscoe, 2004; Korhonen and Ryant, 2005))* now covers 91% of PropBank tokens. *Kipper, et. al., LREC-04, LREC-06, LREJ-08, NAACL09 Tutorial*
- Semi-automatic mapping of PropBank instances to VerbNet classes and thematic roles, hand-corrected.
- VerbNet class tagging as automatic WSD
- Run SRL, map Arg2 to VerbNet roles, Brown performance improves *Yi, Loper, Palmer, NAACL07*

Limitations to VN/FN as sense inventories

- Concrete criteria for sense distinctions
  - Distinct semantic roles
  - Distinct frames
  - Distinct entailments

- But….
- Limited coverage of lemmas
- For each lemma, limited coverage of senses

Sense inventory desiderata

- Coverage of WordNet
- Replicable Sense distinctions captured by concrete differences in underlying representations as in VerbNet and FrameNet
  - Distinct semantic roles
  - Distinct frames
  - Distinct entailments
- Start with WordNet and be more explicit
- Groupings

WordNet: - leave, 14 senses, grouped

WN1, WN5,WN8

Depart, a job, a room, a dock, a country

WN6 WN10 WN2 WN 4 WN9 WN11 WN12

WN14 Wnleave_off2,3 WNleave_behind1,2,3

Leave behind, leave alone

WNleave_alone1 WN13

WN3 WN7

Create a State

WNleave_off1

"leave off" stop, terminate

WNleave_out1, Wnleave_out2

exclude

WordNet: - leave, 14 senses, groups, PB

WN1, WN5,WN8

4

Depart, a job, a room, a dock, a country (for X)

1

2

WN6 WN10 WN2 WN 4 WN9 WN11 WN12

WN14 WNleave_off2,3 WNleave_behind1,2,3

Leave behind, leave alone

WNleave_alone1 WN13

5

WN3 WN7

Create a State /cause an effect:
Left us speechless, leave a stain

WNleave_off1

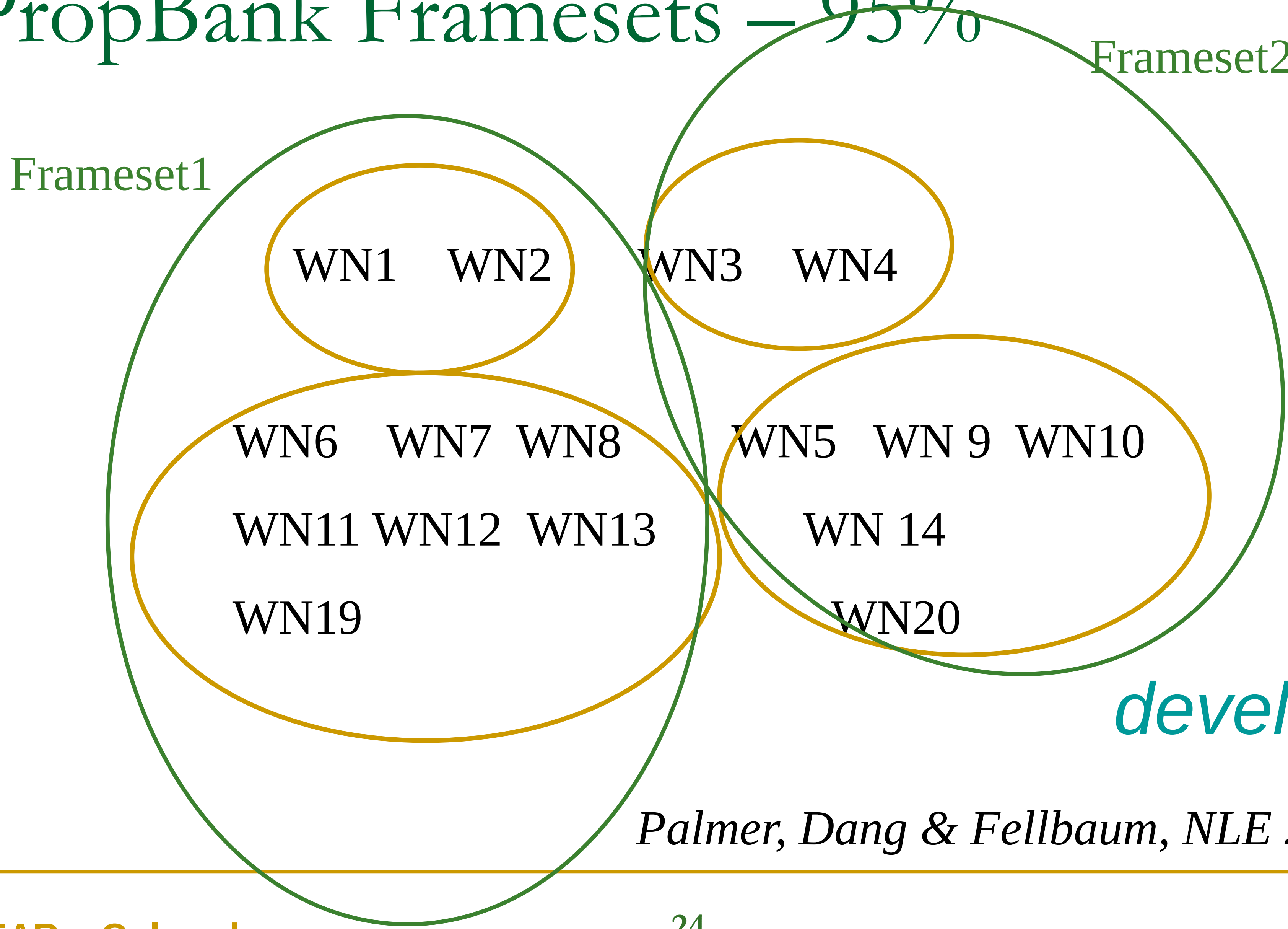
WNleave_out1, WNleave_out2

stop, terminate:
the road leaves off, not leave off your jacket, the result

exclude

3

Overlap between Groups and PropBank Framesets – 95%



develop

Palmer, Dang & Fellbaum, NLE 2007

Sense Hierarchy

(Palmer, et al, SNLU04 - NAACL04, NLE07, Chen, et. al, NAACL06)

- PropBank Framesets – ITA >90%
  coarse grained distinctions
  20 Senseval2 verbs w/ > 1 Frameset
  Maxent WSD system, 73.5% baseline, 90%

  - Sense Groups (Senseval-2) - ITA 82%
    Intermediate level
    (includes Levin classes) – 71.7%

    - WordNet – ITA 73%
      fine grained distinctions, 64%

Tagging w/groups,
ITA 90%, 200@hr,
Taggers - 86.9%
Semeval07

Chen, Dligach & Palmer, ICSC 2007

Groupings Methodology – Human Judges (w/ Dang and Fellbaum)

- Double blind groupings, adjudication
- Syntactic Criteria (VerbNet was useful)
  - Distinct subcategorization frames
    - *call him a bastard*
    - *call him a taxi*
  - Recognizable alternations – regular sense extensions:
    - *play an instrument*
    - *play a song*
    - *play a melody on an instrument*

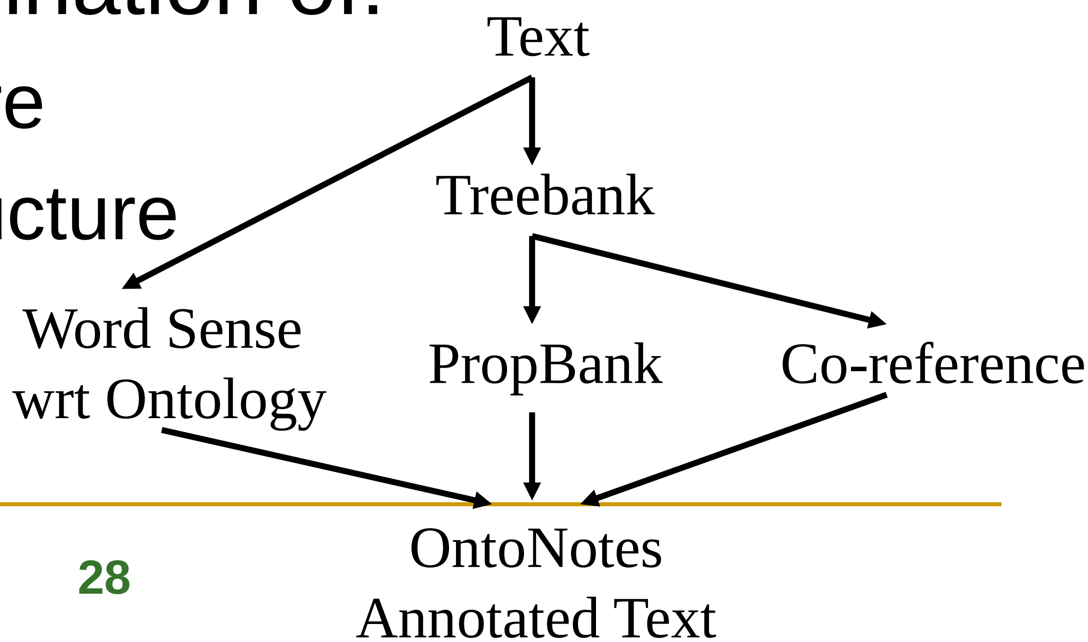
SIGLEX01, SIGLEX02, JNLE07, Duffield, et. al., CogSci 2007

Groupings Methodology (cont.)

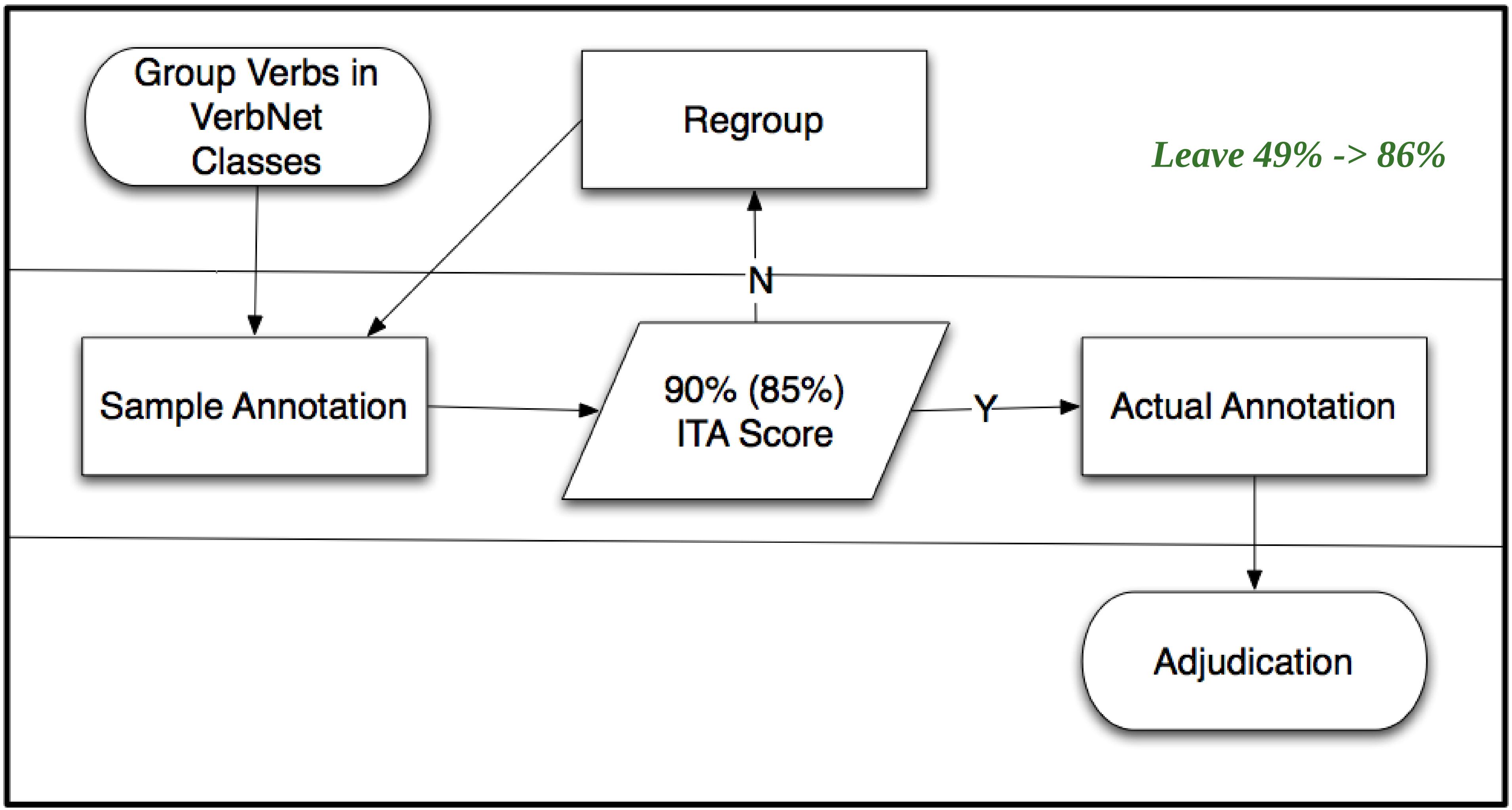
- Semantic Criteria
  - Differences in semantic classes of arguments
    - Abstract/concrete, human/animal, animate/inanimate, different instrument types,…
  - Differences in the number and type of arguments
    - Often reflected in subcategorization frames
    - *John left the room.*
    - *I left my pearls to my daughter-in-law in my will.*
  - Differences in entailments
    - Change of prior entity or creation of a new entity?
  - Differences in types of events
    - Abstract/concrete/mental/emotional/….
  - Specialized subject domains

OntoNotes Goal: Modeling Shallow Semantics DARPA-GALE

- AGILE Team: BBN, Colorado, ISI, Penn
- Skeletal representation of literal meaning
- Synergistic combination of:
  - Syntactic structure
  - Propositional structure
  - Word sense
  - Coreference



Empirical Validation – Human Judges the 90% solution (1700 verbs)



Creation of coarse-grained resources

- Unsupervised clustering using rules (Mihalcea & Moldovan, 2001)
- Clustering by mapping WN senses to OED (Navigli, 2006).
- OntoNotes - Manually grouping WN senses and annotating a corpus (Hovy et al., 2006)
- Supervised clustering WN senses using OntoNotes and another set of manually tagged data (Snow et al., 2007) .

OntoNotes Status

- More than 2,000 verbs grouped
- Average ITA per verbs = 89%
- http://verbs.colorado.edu/html_groupings/
- More than 150,000 instances annotated for 1700 verbs
- WSJ, Brown, ECTB, EBN, EBC
- Training and Testing
- *Are we headed in the right direction?*

Leave behind, leave alone…

- ***John left his keys at the restaurant.***

  We left behind all our cares during our vacation.

  They were told to leave off their coats.

  Leave the young fawn alone.

  Leave the nature park just as you found it.

  I left my shoes on when I entered their house.

  When she put away the food she left out the pie.

  Let's leave enough time to visit the museum.

  He'll leave the decision to his wife.

  When he died he left the farm to his wife.

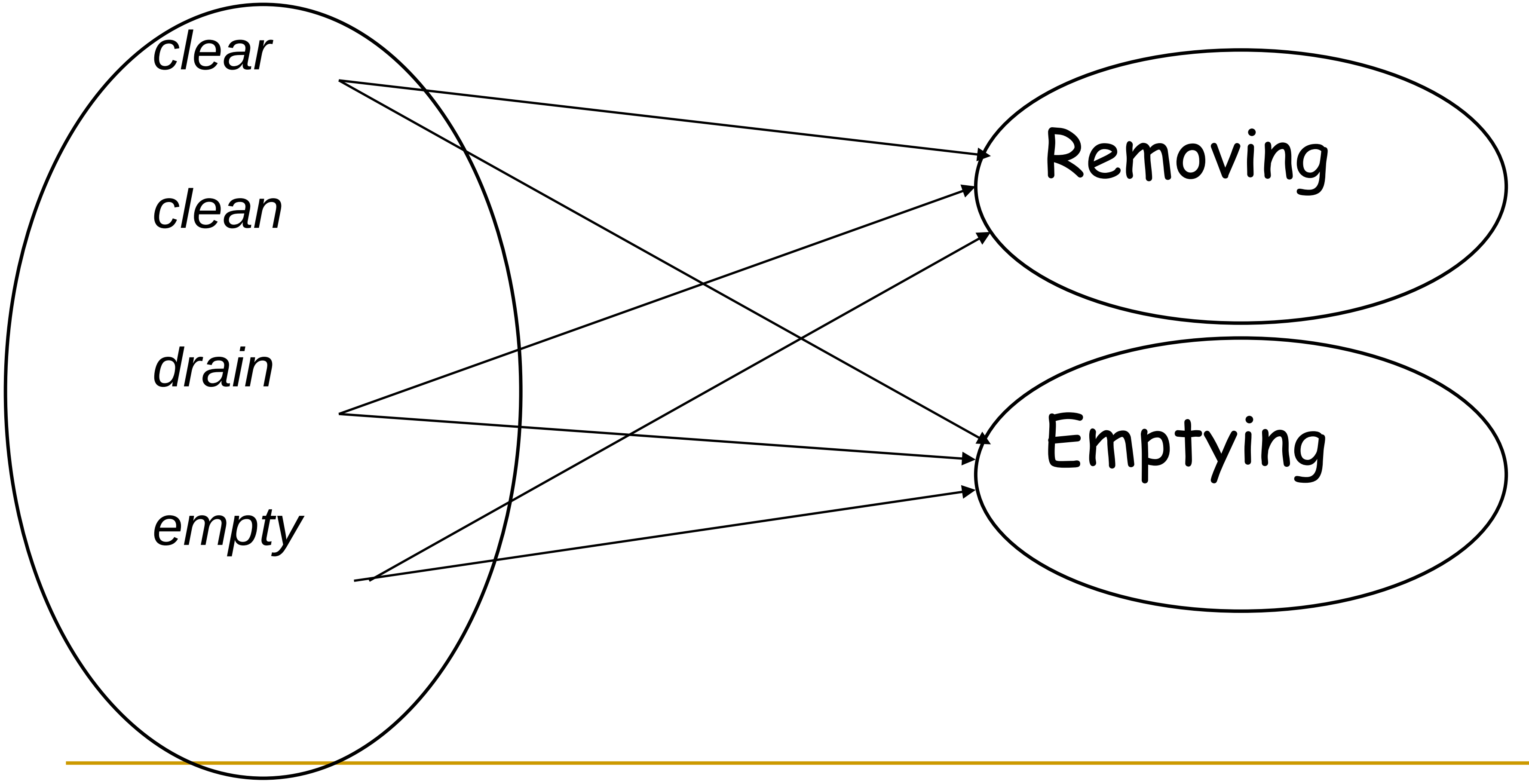
  I'm leaving our telephone and address with you.

FrameNet: Telling.*inform*

Time	In 2002,
Speaker	the U.S. State Department
Target	INFORMED
Addressee	North Korea
Message	that the U.S. was aware of this program , and regards it as a violation of Pyongyang's nonproliferation commitments

Mapping Issues (2)
VerbNet verbs mapped to FrameNet

- *VerbNet clear-10.3*
- FrameNet Classes



Mapping Issues (3)

VerbNet verbs mapped to FrameNet

VN Class: *put 9.1*

Members: *arrange*, immerse, lodge, mount, sling***

Thematic roles:

- agent (+ animate)
- theme (+ concrete)
- destination (+ loc, – region)

Frames:

- …

*different sense

** not in FrameNet

FrameNet frame: *place*

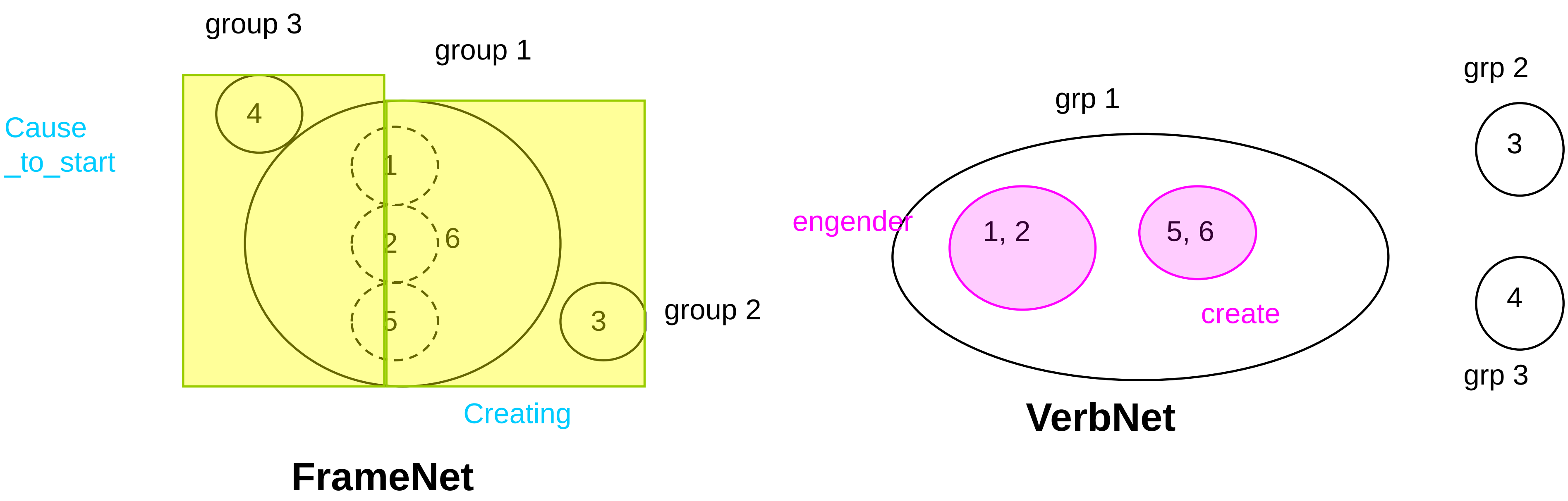
Frame Elements:

- Agent
- Cause
- Theme
- Goal

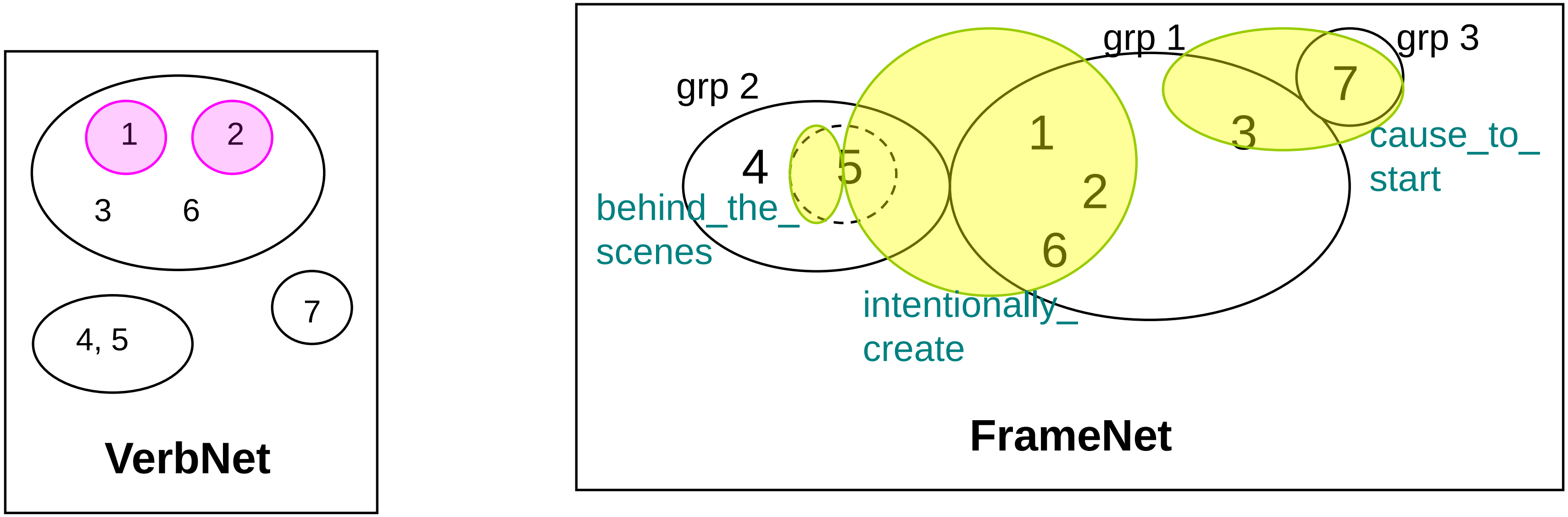
Examples:

- …

Class formation Issues: *create*
Susan Brown



Class formation Issues: *produce*
Susan Brown



Question remains: What is the “right” level of granularity?

- “[Research] has not directly addressed the problem of identifying senses that are distinct enough to warrant, in psychological terms, a separate representation in the mental lexicon.” (Ide and Wilks, 2006)
- Can we determine what type of distinctions are represented in people’s minds?
- Will this help us in deciding on sense distinctions for WSD?

Sense Hierarchy

- PropBank Framesets – ITA >90%
  coarse grained distinctions
  20 Senseval2 verbs w/ > 1 Frameset
  Maxent WSD system, 73.5% baseline, 90%
  - Sense Groups (Senseval-2) - ITA 82%
    Intermediate level
    (includes Levin classes) – 71.7%
    - WordNet – ITA 73%
      fine grained distinctions, 64%

Computational model of the lexicon based on annotation

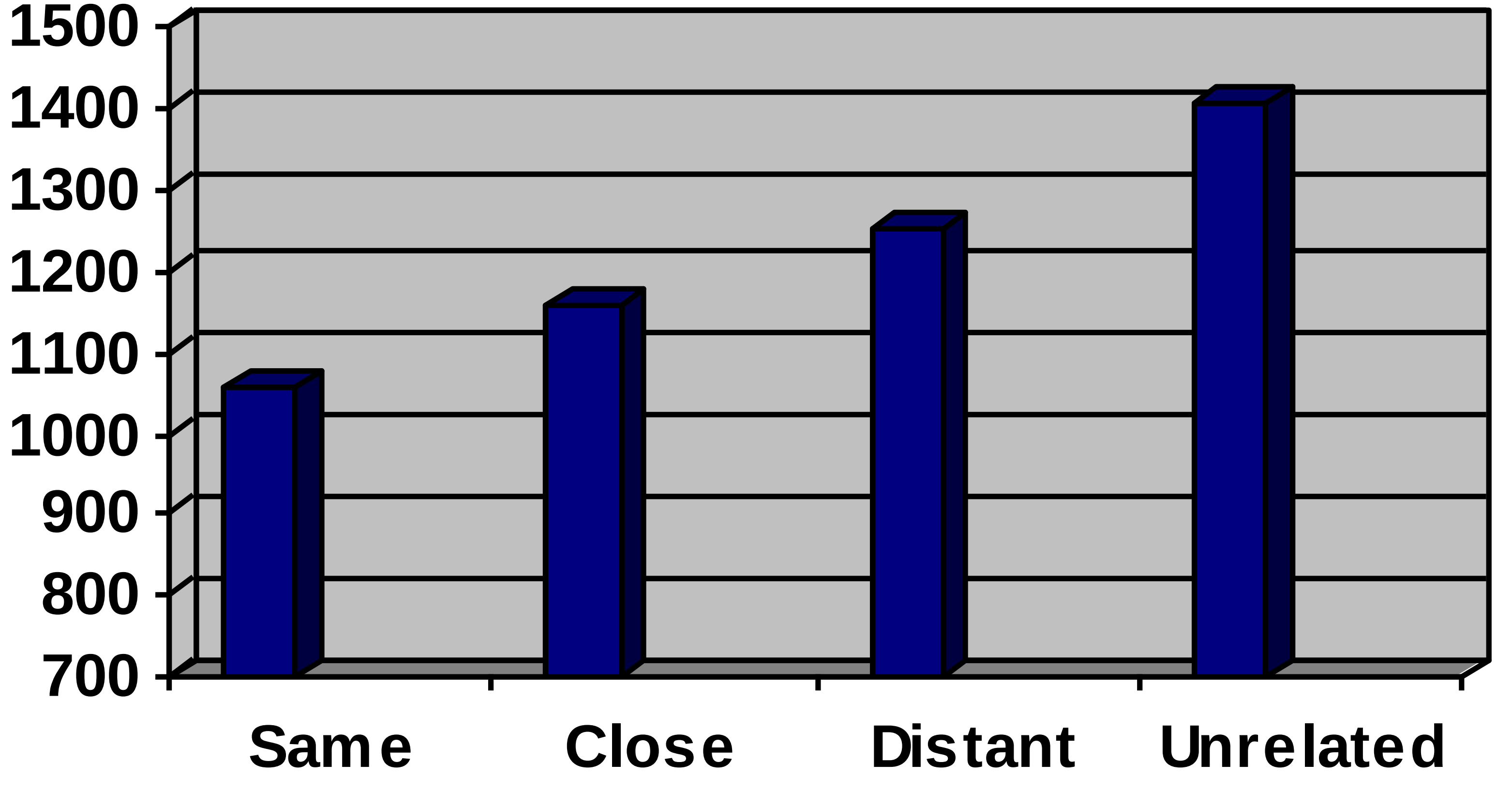
- Hypothesis: Syntactic structure overtly marks very coarse-grained senses
- Subsequently subdivided into more and more fine-grained distinctions.
- A measure of distance between the senses
  - The senses in a particular subdivision share certain elements of meaning.
  - There are many alternative subdivisions

Procedure – Susan Brown

- Semantic decision task
- Judging semantic coherence of short phrases
  - banked the plane "makes sense"
  - hugged the juice   doesn't "make sense"
- Pairs of phrases with the same verb
- Primed with a sense in the first phrase
- Sense in the second phrase was one of 4 degrees of relatedness to the first

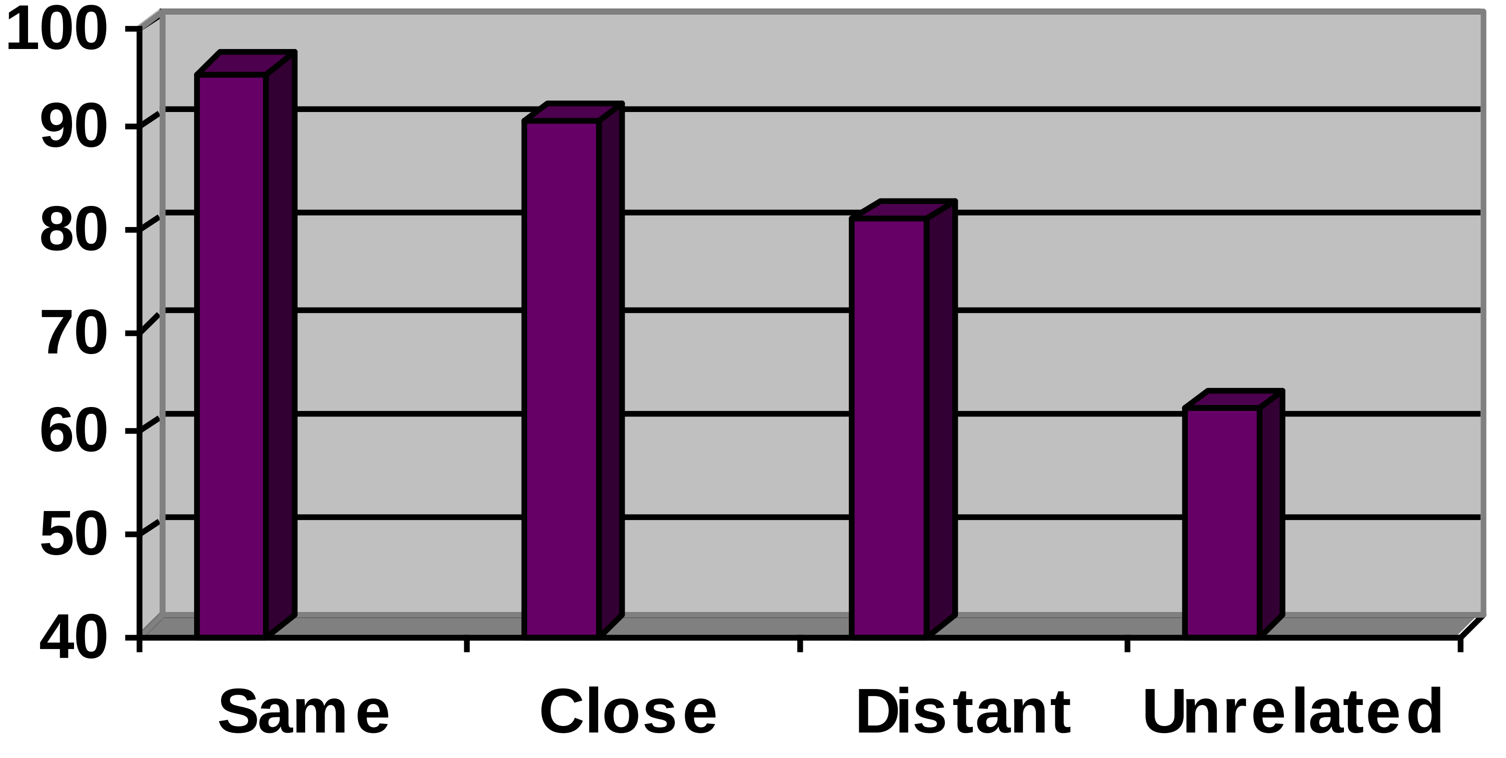
Brown, ACL08

Mean response time (in ms)



Brown, ACL08

Mean accuracy (% correct)



Significant distinction between literal and abstract usages

Brown, ACL08

Implications for WSD

- Enumerating discrete senses may be a convenient (necessary?) construct for computers or lexicographers
- Little information loss when combining closely related senses
- Distantly related senses are more like homonyms, so they are more important to keep separate

Augmenting Features for Automatic Word Sense Disambiguation (Verbs)

- Lexical Features
  - Words and POS tags
- Syntactic Features (*Chen & Palmer, 2009*)
  - Subject/Object (headwords + pos tags)
  - Passive/Active
  - Presence of a Subordinate Clause
  - Presence of a PP adjunct/Preposition/Preposition's argument
  - Path Features
    - Same as in SRL
  - Subcat Frame
    - E.g. VPD-PP-NP for *The lawyers went to the courthouse*
- Semantic Features (*Chen & Palmer, 2009*)
- Classifiers - MaxEnt or SVM

Semantic Features for WSD

Dligach & Palmer, ACL08

- Verb *prepare*

Sense Number	Definition	Example
1	To put together, assemble, concoct	He is going to prepare breakfast for the whole crowd. I haven't prepared my lecture yet.
2	To make ready, fit out	She prepared the children for school every morning

- Important for making sense distinctions
- WordNet
  - Hypernyms and synonyms
- NE Data
  - Person, Organization, Location, Date, Time, Money, etc.

Problems with WordNet and NE

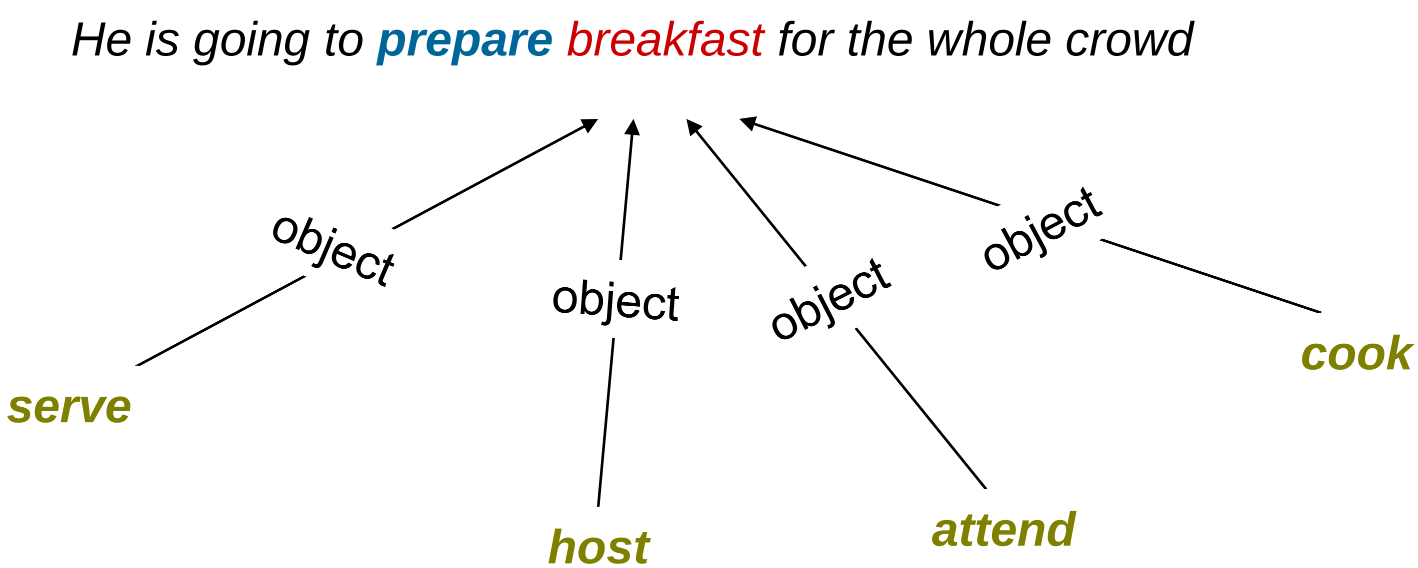
Dligach & Palmer, ACL08

Sense Number	Definition	Example
1	To put together, assemble, concoct	He is going to prepare breakfast for the whole crowd. I haven't prepared my lecture yet.
2	To make ready, fit out	She prepared the children for school every morning

- Both *breakfast* and *lecture* are unrelated according to WordNet
  - Multiple semantic properties grouped into finite classes
- However, both are social events
  - Can be *attended, hosted, delivered, given, held,* and *organized*
  - Belong to the same sense of *prepare*

Dynamic Dependency Neighbors (DDNs)

Dligach & Palmer, ACL08



Dynamic Dependency Neighbor Extraction – Dmitriy Dligach

- Unsupervised Extraction
- English Gigaword
  - Newswire text from 5 sources (New York Times, Associated Press, etc.)
  - 5.7M News Articles
  - 2.1B Words
- Preprocess with MaltParser (Nivre, 2007)
- Index
  - verbs
  - their subjects/objects
  - frequencies

DDNs for some nouns

Dligach & Palmer, ACL08

dinner		breakfast		lecture		child	
verb	freq	verb	freq	verb	freq	verb	freq
have	4100	have	1428	give	1877	have	25967
attend	2236	eat	991	deliver	911	raise	5553
eat	1239	serve	301	attend	483	protect	4457
host	1039	attend	299	get	144	teach	3632
cook	499	make	201	hold	98	help	3606
make	472	skip	117	have	64	adopt	2239
serve	437	offer	115	present	46	educate	1746
get	305	cook	112	organize	28	lose	1585
enjoy	218	provide	76	host	17	want	1420
organize	114	host	75	begin	17	abuse	1402

Discussion

Dligach & Palmer, ACL08

- DDNs beneficial to Verb Sense Disambiguation
  - Decrease in error rate 3 – 15%
- DDNs outperform WordNet + NE
  - 3% decrease in error rate for WordNet + NE
  - 6% decrease in error rate for DDNs
  - Same performance with or without WordNet + NE (1% absolute improvement over just WN + NE)
- Can be important in resource-poor domains
- Potentially useful in SRL, Entailment, etc.

Relevant Work

Dligach & Palmer, ACL08

- Distributional Similarity (Harris, 1968)
  - Similar words occur in similar contexts
- Context ranges from bag-of-words to more structured approaches
  - Schutze (1998), Purandare and Pedersen (2004) experiment with first and second-order bag-of-words
  - Hindle (1990), Lin (1998) grouped nouns into thesaurus-like lists
    - Our approach similar but no static categories
- DDNs viewed as a form of world knowledge
  - Schubert (2003)
  - Lin and Pantel (2001)
    - DIRT system for detecting paraphrases

Current results on WSD - Dligach

	Baseline	ITA	System
Set 1 (217)	.68	.825	.83
Set 2 (200)	.81	.93*	.91

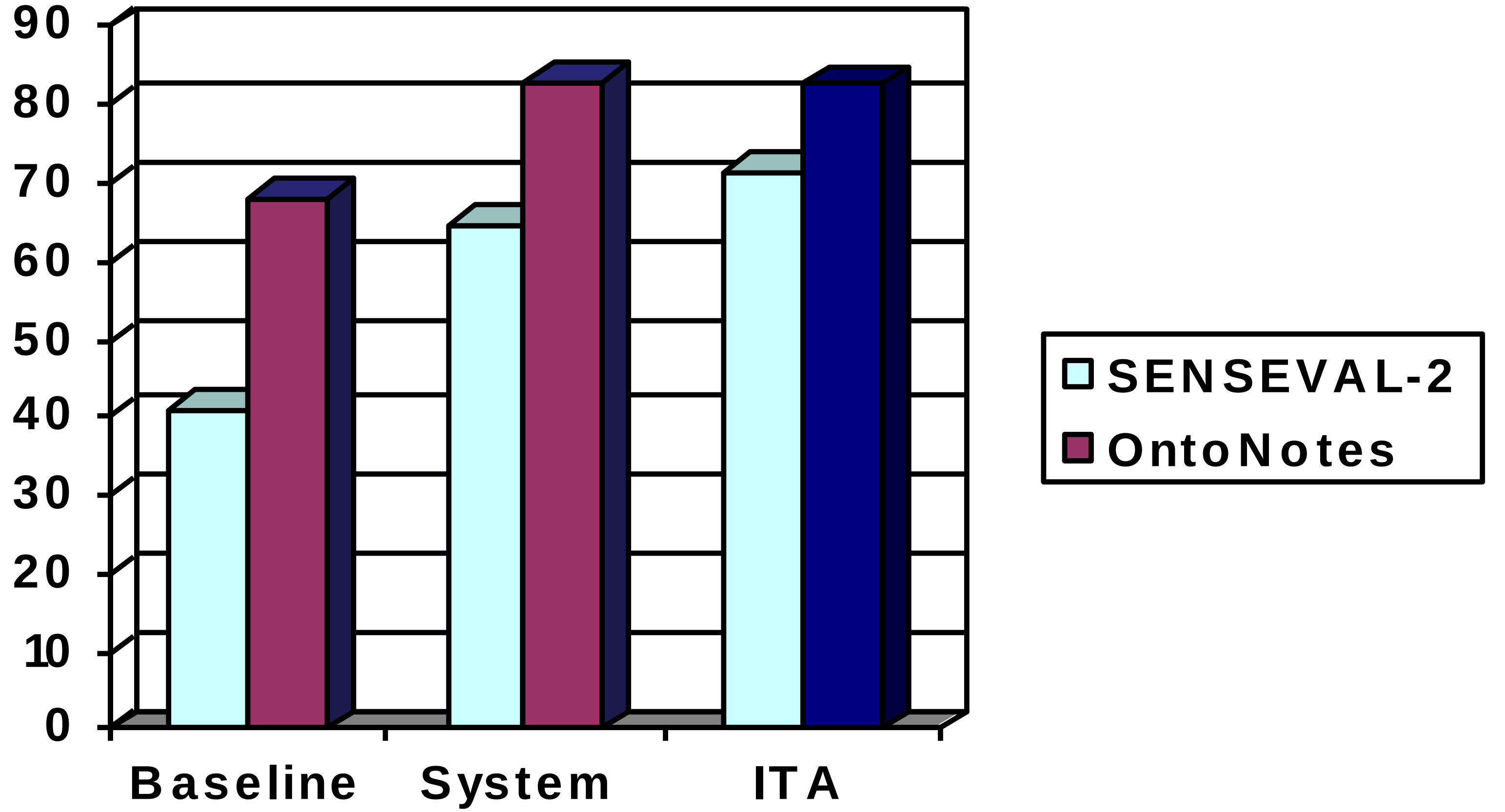
*200 most frequent verbs with ITA > 85%:

Baseline – Most Frequent sense

ITA – InterTaggerArgreement

System - 5-fold cross validation accuracy

Fine-grained vs. coarse-grained senses



Correlations for WSD

	Polysemy	Num. of Instances	Ave. Num. of Instances	Baseline	ITA	Sense Entropy
System Accuracy	-0.4518	0.0228	0.2753	0.7462	0.5758	-0.8093
System_ Acc_Imp	0.2389	0.0124	-0.1316	-0.8555	-0.1675	0.7158

Knowing a Word(sense) by the company it keeps

- Lexical Co-occurrences
- Related senses – literal/abstract?
- Syntactic and semantic dependencies
- Dynamic dependency neighbors
- Entropy of sense distribution

Need more feedback - and you can give it to us

- On VerbNet classifications
- On FrameNet classifications
- On OntoNotes groupings vs WN vs PB
- On usefulness of the distinctions made by all of the above

Acknowledgments

- We gratefully acknowledge the support of the National Science Foundation Grant NSF-0415923, Consistent Criteria for Word Sense Disambiguation and DARPA-GALE via a subcontract from BBN.
- We thank Walter Kintsch and Al Kim for their advice on the psycholinguistic experiments.
- We also thank Rodney Nielsen and Philipp Wetzler for parsing English Gigaword with MaltParser.