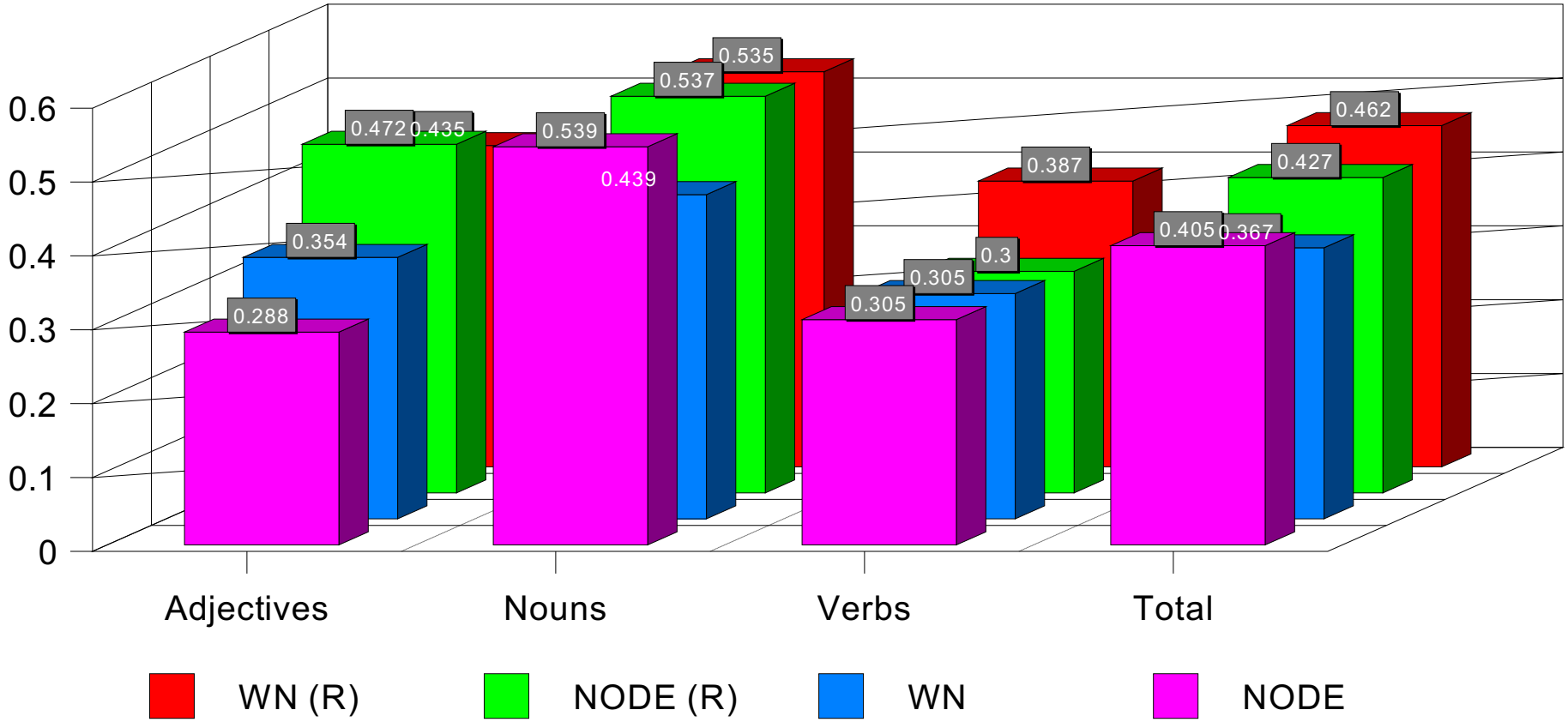
Sense Information for Disambiguation: Confluence of Supervised and Unsupervised Methods

Ken Litkowski
CL Research
ken@clres.com
http://www.clres.com
http://www.clres.com/siglex02-wsd/index.html

CL Research WSD System

- Unsupervised, dictionary-based (WordNet or any MRD)
  - Conversion to DIMAP format, with fields for part of speech, definition, features (all elements of WN or an MRD - New Oxford Dictionary of English), definition parsing to create WordNet style relations
  - Special dictionaries for multiword units (used first in WSD)
- Full sentence parsing (partial parses for irregular input)
- WSD performed on parse output
  - Evaluates each sense of target word using available information (subject labels, subcategorization patterns, selectional preferences, form restrictions, grammatical roles, collocational patterns, contextual clues)
- NODE WSD required mapping to WordNet to select senses
  - Automatic mapping achieved 89% coverage, with 70% accuracy (inaccuracies did not affect results)
  - WSD against NODE mapped into WordNet is minimum; WSD in NODE **must** be higher (e.g., 100 idioms not in WordNet)

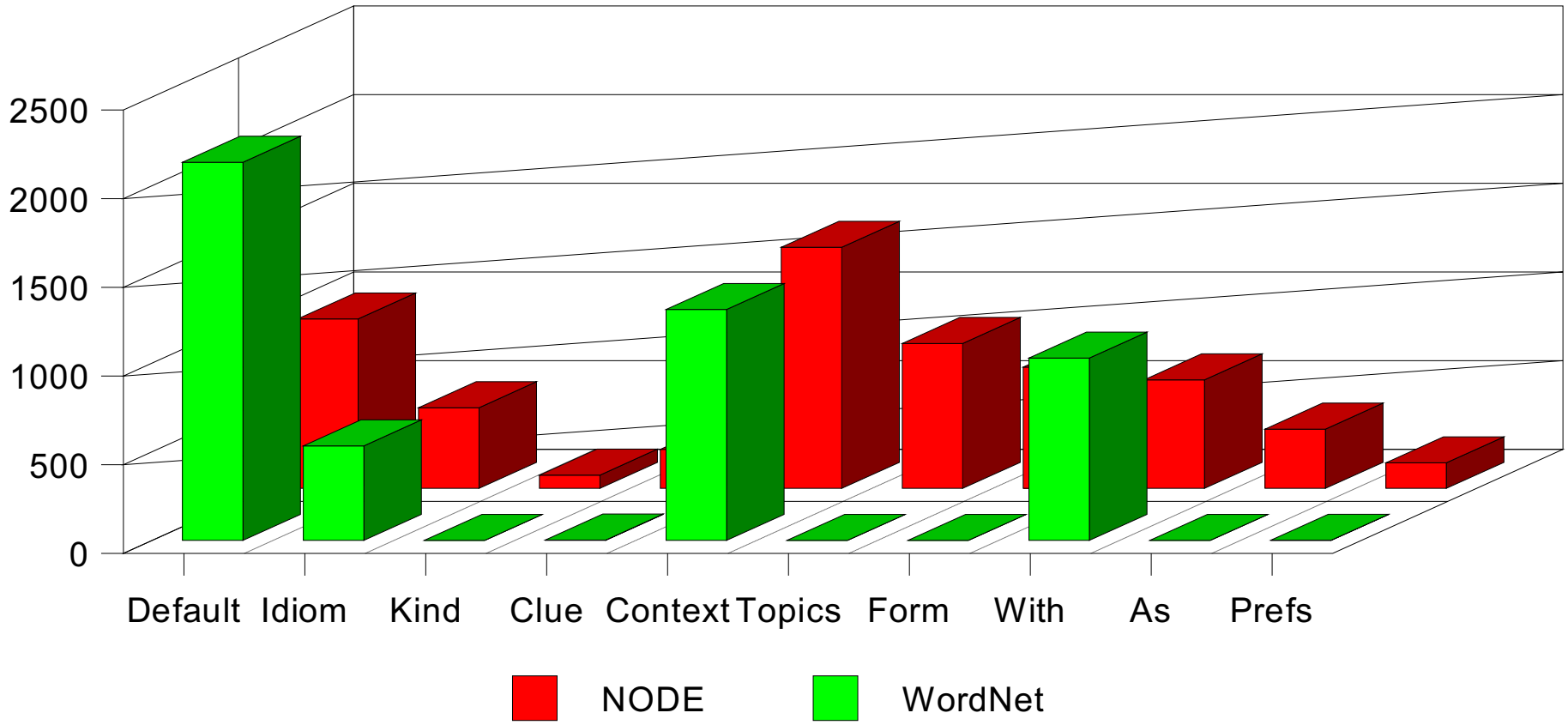
Lexical Sample Coarse-Grained Precision



Features Used in Disambiguation

- Default sense selection (no positive information)
- Idiomatic (phrasal) usages: multiword main entries, phrasal runons, collocational patterns (bolding in NODE examples)
- Verb subcategorization patterns; nouns “with” noun modifier
- Form (plurals, present or past participles)
- Usage (e.g., noun “as” noun modifier or verbs “as” adjectives)
- Lexical preferences (subject or object of verbs, modificand of adjectives)
- Context (overlap with definitions and examples)
- Subject labels

Comparative Analysis of Features Used in WordNet and NODE Disambiguation



Overall Observations About Features

- Feasibility of exploiting a considerable amount of sense information from an MRD
- Very large proportion of default sense selections
- Highest importance given to:
  - Context
  - Idioms
  - Subject labels
  - Form
  - Subcategorization patterns
- No change in distribution for correct vs. incorrect WSD
  - Many cases where NODE gave correct WSD against its inventory, but judged incorrect

Default Sense Selection

- Indicates absence of positive information
  - Very significant for WordNet (almost 50%)
  - Quite significant for NODE (almost 25%)
- Results may be even larger, considering ubiquitous "with" **object** of transitive verbs
- Also may indicate shortcomings of CL Research's WSD system (not making use of available information)
- Indicates inadequate articulation of sense distinctions in the inventories
  - Have lexicographers captured distinctions in a way that can be used in WSD

Multiword Units and Collocations

- NODE shows at least 17.5% use of “set phrases”
- Phrasal headwords (idioms, multiword units) are important
  - Significant mismatch between WordNet and NODE (2000 entries involving lexical sample words)
  - Many hyponyms in WordNet (e.g., “apricot bar” - ?coarse grain of “bar”)
  - Only verbs with particles in WordNet (e.g., no “call it a day”)
- Collocation “clues” (patterns from Hector in Senseval-1)
  - Captures extensively “variable” idioms and phrases
  - Allows syntactic and semantic specifications

Local and Global Contexts

- Local context provided by definitions and examples useful in 30% of cases
  - Importance of good definitions and examples, with perhaps more possibility in MRDs
  - Upper limit of how well examples can cover sense distinctions
- Subject fields important in 20% of cases with NODE
  - Shows value of research into domain-specificity (help is on the way)

Subcategorization Patterns

- Important primarily for verbs (55% in WordNet, 32% in NODE)
  - More of a screening mechanism than a disambiguation criterion
- Can also be useful for nouns (e.g., “metal fatigue”)
- Unexplored territory - “variables” in definitions (?internal arguments)
  - **nature**: “the basic or essential qualities of **something**” (subcat for an “of” PP)
  - **carry**: “take (an idea) to a **specified point**” (requires an adverbial)

Word Forms and Usage

- Word **form** (capitalization, plural, passive) useful in 16% of the case with NODE
- Syntactic usage (noun **as** modifier, verbs **as** adjectives, adjectives **as** nouns) (7.7% for nouns and 10.3% for verbs)
- Useful sense distinction where present, but not very prominent

Lexical Preferences

- Verb subjects and objects and adjective modificands (of small value thus far, but often present)
- Requires semantic testing
  - Used literals and WordNet synonyms and hypernyms
  - Limited implementation, showing only viability
- Unexplored territory - “variables” in definitions
  - **carry**: “support the weight of” (object is something that has weight - “the bridge carries heavy loads”, where “load” has hypernym “weight”)

General Observations About Feature Analysis

- Considerable variation in importance of various features by part of speech and by lexical item
- Technique helps identify differences in sense inventories and where sense distinctions are not well drawn
  - When combined with automatic and hand mapping of sense inventories, particularly highlights problematic areas
- Feature analysis frequently identified instances where answer key was clearly incorrect

Comparison with Features Identified in Other Disambiguation Systems

- Mihalcea & Moldovan: Active features identify part of speech, word form, collocations, nouns before and after, prepositions before and after correspond to **idioms, clues, form, "with"**
- WASPBench (Tugwell & Kilgarriff): Grammatical relations (bare-noun, plural, passive, ing-complement, noun-modifier, PP-comp) correspond to **form, clue, "with", "as"**
- Pedersen: Bigrams correspond to **context**

Future of Feature Analysis

- Working with data sets by WordNet sense in answer keys
  - Developing “feature signatures”
  - Identifying features that can be put into MRDs
  - Improving design of MRDs for WSD
- Enables improvement of WSD system
  - Focusing in on sense distinctions
  - Identifying features that can be generalized across lexical items

Conclusions

- Mapping and feature analysis have identified many difficulties with WordNet sense information and distinctions
  - Affects the quality of the WSD exercise
  - Points up features that should be included in sense inventories
- Is WSD fully-tested without a carefully drawn sense inventory?
- Have we fully tested MRDs? Can we add more to MRDs to make them more useful?
- Can the community find a way to pool resources to come up with a sense inventory? Using the ANC and BNC?