

Computational Intelligence 696i

Language

Lecture 5

Sandiway Fong

Administriva

- Reminder:
 - Homework 1 due next Tuesday (*midnight*)
- Sessions can be saved and reloaded
 - see History menu
- Capturing the display:
 - screen snapshot
 - History => Print
 - generated postscript file is saved under /tmp/,
 - e.g. pappi_history-**1062**.ps
 - the number will vary: **1062** here is the PAPPI process id

Last Time

- we looked at PAPPI: a computational instantiation of the principles-and-parameters model

sentence

syntactic representations

The screenshot shows the PAPPI interface. At the top, the title bar reads "Principles-and-Parameters Parser". Below it, a text input field contains the Korean sentence: "[154] minswu-nun younghee-ka mwues-ul mekessta-ko sayngkakha-ni". To the right of the input field are buttons for "Info ..." and "Demo ...". Below the input field is a menu bar with "Run", "Language", "Theory", "Parsers", "History", and "Options". The main window displays a syntactic tree for the sentence. The tree is rooted at C2 and branches into I2 and C. I2 branches into NP[1] (minswu) and VP. VP branches into V[5] (sayngkakha) and I[AGR][1] (ni). NP[1] branches into I2, which branches into NP[2] (mwues) and I2. NP[2] branches into NP[3] (younghee) and VP. VP branches into VP and I[AGR][3] (mekessta). The tree is labeled "LF (2):" and "2 parses found". On the right side of the interface is a "Filters" panel with a list of linguistic principles and parameters, including "Theta Criterion", "D-structure Theta Condition", "Subjacency", "Wh-movement in Syntax", "S-bar Deletion", "Case Filter", "Case Condition on ECs", "Coindex Subject", "Condition A", "Condition B", "Condition C", "ECP", "Control", "License Clitics", "License Object pro", "ECP at LF", "Fi: License operator/variables", "Fi: Quantifier Scoping", "Fi: Reanalyze Bound Proforms", "License Clausal Arguments", "License Syntactic Adjuncts", "Wh Comp Requirement", "Generators", "Parse PF", "Parse S-Structure", "Assign Theta-Roles", "Inherent Case Assignment", "Assign Structural Case", "Trace Theory", "Functional Determination", "Free Indexation", "Expletive Linking", and "LF Movement". A blue box labeled "user's viewpoint" points to the "Filters" panel. A blue box labeled "(Korean)" points to the Korean sentence. A blue box labeled "parser operations corresponding to linguistic principles (= theory)" points to the "Generators" section of the "Filters" panel.

- user's viewpoint
- (Korean)
- parser operations corresponding to linguistic principles (= theory)

Last Time

- a system (*for parsing only*) with
 - **32 parser operations** \leq linguistic principles
 - X'-theory, Move- α , Subjacency, Binding, Case, Theta, ECP, LF operator-variable licensing...
 - Note: *there is no explicit gap-filler rule*
 - **12 parameters**
 - word order, strong/weak agreement, negation, preposition stranding, case adjacency, subject drop, *wh-movement*, bounding nodes
 - theory implemented is
 - **logically consistent**
 - **flexible enough** to implement language fragments from
 - Arabic, Turkish, Hungarian, Chinese, Japanese, Korean, Dutch, German, French, Spanish, Bangla, English

Last Time

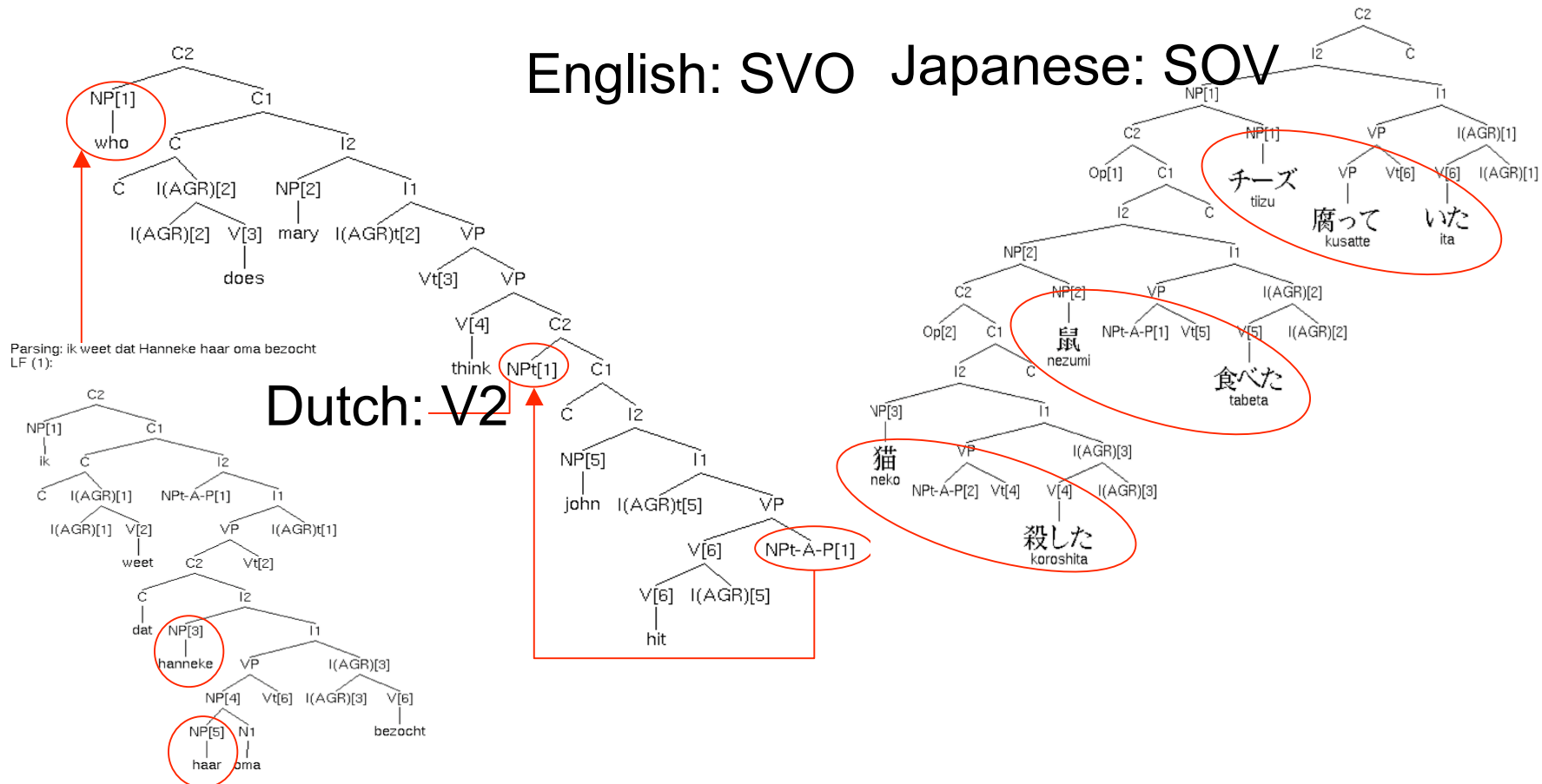
- uses a generate-and-test computational strategy:
 - **generators:** X'-theory, Move- α , free coindexation
 - **filters:** *construction-independent* linguistic constraints



Last Time

- We demoed PAPPI on a variety of examples

English: SVO Japanese: SOV



Case Study: *Gap Filling*

- we've been focusing on **gap filling** as an example of a (*non-trivial*) problem that the language processor must solve
- **Examples used include:**
 - Which report did you file [the report] without [you] reading [the report]?
 - *Which book did you file the report without reading [the book]?
 - *you filed the report without [you] reading [the report]
 - you filed the report without reading it
 - Who does Mary think [who] John hit [who]?
 - *Who does Mary wonder why John hit [who]?
 - John is too stubborn [someone] to talk to [John]
 - John is too stubborn [John] to talk to Bill

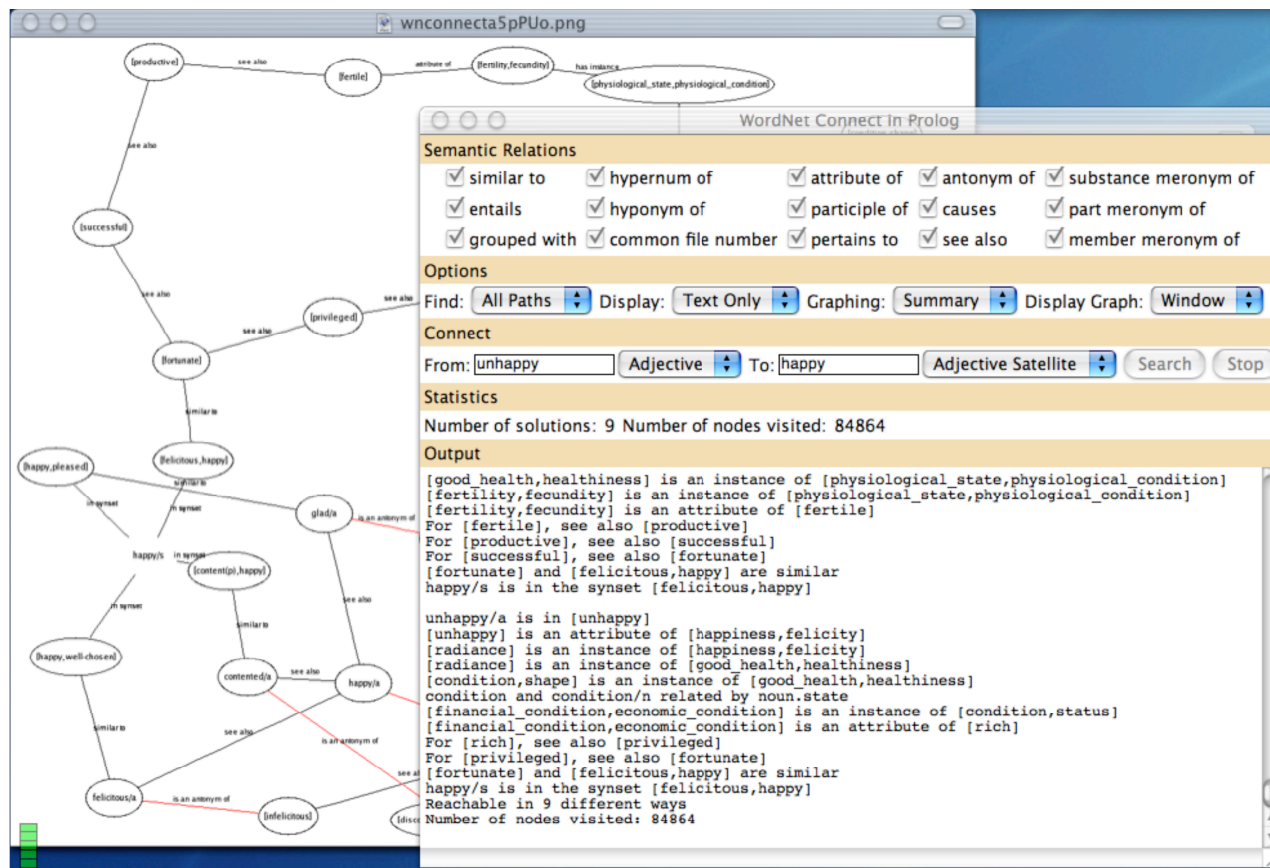
Case Study: *Gap Filling*

- hope I've convinced you all that **gap filling** is
 - something you do correctly without even thinking about it
 - all speakers agree on (some of) the rules
- plus
 - you were never explicitly taught the rules
 - there is a lot of surface complexity, the data is complicated
- **Poll:**
 - who believes there is a UG?
 - who believes there is no UG, language is induced from available data?

Administrivia

- you'll need to download software for Homework 2
- *Homework 2 will be discussed on Tuesday*
- **wnconnect:** *WordNet connect*
 - available for MacOS X
 - available for Linux
 - available for Windows
- download from
 - (soon)
 - <http://dingo.sbs.arizona.edu/~sandiway/wnconnect/>

Administrivia



New Topic

- **Semantic Inference and Language**
- computation using
 - WordNet (Miller @ Princeton University)
 - handbuilt network of synonym sets (**synsets**) with semantic relations connecting them
- *compare with statistically determined co-occurrence vectors from corpora*

Two Problems

- *linguistically relevant puzzles*
- *outside syntax*

1. Semantic Opposition
2. Logical Metonymy

Semantic Opposition

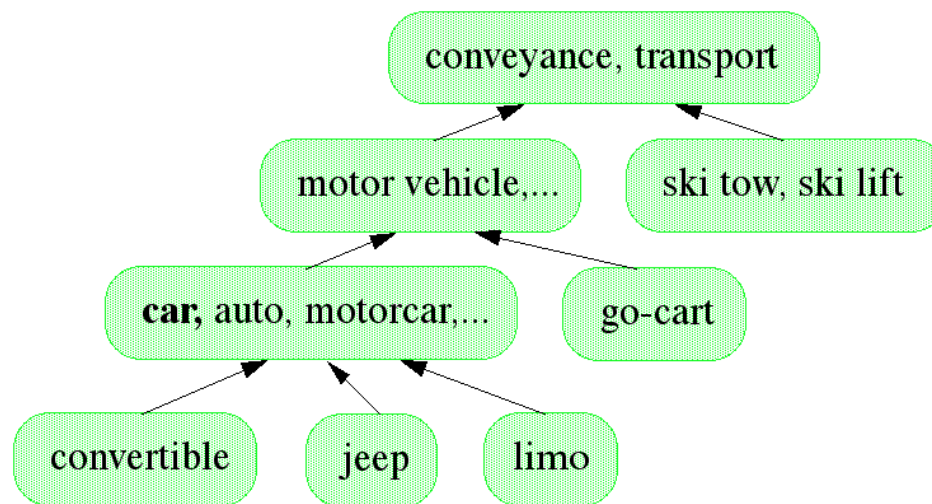
- *Event-based Models of Change and Persistence in Language* (Pustejovsky, 2000):
 - John mended the **torn** dress
 - John mended the **red** dress
 - what kind of knowledge is invoked here?
- from Artificial Intelligence (AI):
 - an instance of the **frame problem**
 - *aka* the **update problem**
 - computation:
 - *what changes in the world and what doesn't?*

Quick Introduction to WordNet

- also see
 - *5 Papers on WordNet*
 - from the Princeton team
 - [5papers.pdf](#)
 - on the language section of the course homepage:
 - <http://dingo.sbs.arizona.edu/~sandiway/ling696/>

WordNet

- **What is it?**
 - Synonym set (**synset**) network for nouns, verbs, adjectives and adverbs
 - Synsets connected by semantic relations (**isa**, **antonymy**,...)
 - 139,000 entries (**word senses**), 10,000 verbs (polysemy 2), 20,000 adjectives (1.5)
 - Originally designed as a model of human semantic memory (Miller, 1985)

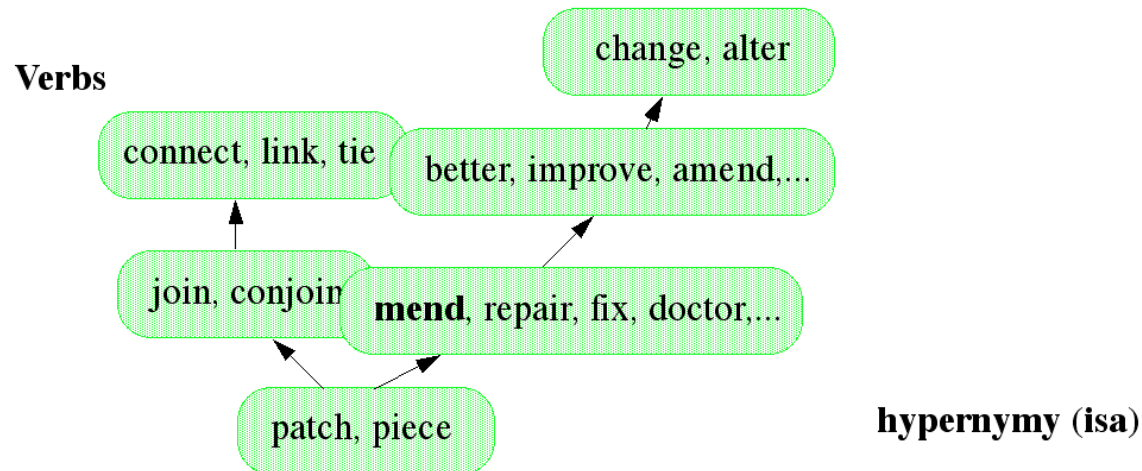


hypernymy (isa)

WordNet

- **What is it?**

- Synonym set (**synset**) network for nouns, verbs, adjectives and adverbs
- Synsets connected by semantic relations (**isa**, **antonymy**,...)
- 139,000 entries (**word senses**), 10,000 verbs (polysemy 2), 20,000 adjectives (1.5)
- Originally designed as a model of human semantic memory (Miller, 1985)



WordNet

- **What is it?**
 - Synonym set (**synset**) network for nouns, verbs, adjectives and adverbs
 - Synsets connected by semantic relations (**isa**, **antonymy**,...)
 - 139,000 entries (**word senses**), 10,000 verbs (polysemy 2), 20,000 adjectives (1.5)
 - Originally designed as a model of human semantic memory (Miller, 1985)

Adjectives

“Dumbell Model”

dinky
elfin
gnomish
micro
minuscule
puny
undersized

satellites

small

large

antonymy

ample
bulky
colossal
gigantic
humongous
spacious
titantic

WordNet

- **What is it?**
 - Synonym set (**synset**) network for nouns, verbs, adjectives and adverbs
 - Synsets connected by semantic relations (**isa**, **antonymy**,...)
 - 139,000 entries (**word senses**), 10,000 verbs (polysemy 2), 20,000 adjectives (1.5)
 - Originally designed as a model of human semantic memory (Miller, 1985)
- **Extremely Popular**
 - Free (\$3M)
 - EuroWordNet (EWN), ItalWordNet, Tamil WordNet, Estonian WordNet,...
 - Conferences
 - ACL Workshop (1998)
 - NAACL Workshop (2001)
 - 1st & 2nd Global WN Conference (2002, 2004)
 - 2 LREC Workshops (May 2002)

WordNet Relations

Relation	Description	Example
x HYP y	x is a hypernym of y	x: <i>repair</i>, y: <i>improve</i>
x ENT y	x entails y	x: <i>breathe</i> , y: <i>inhale</i>
x SIM y	y is similar to x (A)	x: <i>achromatic</i> , y: <i>white</i>
x CS y	y is a cause of x	x: <i>anesthetize</i> , y: <i>sleep</i>
x VGP y	y is similar to x (V)	x: <i>behave</i> , y: <i>pretend</i>
x ANT y	x and y are antonyms	x: <i>present</i>, y: <i>absent</i>
x SA y	x, see also y	x: <i>breathe</i> , y: <i>breathe out</i>
x PPL y	x participle of y	x: <i>applied</i> , y: <i>apply</i>
x PER y	x pertains to y	x: <i>abaxial</i> , y: <i>axial</i>


Back to Semantic Opposition...

Persistence and Change of State Verbs

- Event-based Models of Change and Persistence in Language (Pustejovsky, 2000):
 - John mended the **torn** dress
 - John mended the **red** dress
- **Verb Classes: Aspectual Classes** (Vendler 1967)
 - Mary cleaned the **dirty** table Change of State
 - The waiter filled every **empty** glass
 - Mary fixed the **flat** tire
 - Bill swept the **dirty** floor Activity
 - Bill swept the **dirty** floor clean Accomplishment
 - Nero built the **gleaming** temple Creation
 - Nero ruined the **splendid** temple Destruction

Event Representation

- **Change of State Verbs:**

- John mended the **torn/red** dress
 - **mend**: x CAUS y BECOME <STATE (**mended**)>
 - John CAUS the **torn/red** dress BECOME <STATE (**mended**)>
- 

- **antonym** relation between adjective and the end state

Using wnconnect

- Find shortest link between **mend** and **tear** in WordNet:

The screenshot displays the 'WordNet Connect in Prolog' application window. The 'Semantic Relations' section lists various relations with checkboxes, including 'is an antonym of'. The 'Options' section shows 'Find: Shortest Paths Only', 'Display: Text Only', and 'Graphing: Per Solution'. The 'Connect' section has 'From: mend' and 'To: tear'. The 'Statistics' section shows 'Number of solutions: 1' and 'Number of nodes visited: 82'. The 'Output' section contains the following text:

```
Sandiway Fong, University of Arizona. Working directory (for graphs): /Users/sa
WordNet Connect is built on WordNet 1.7.1, Sicstus Prolog 3.10.1, Tck/Tk 8.4.4

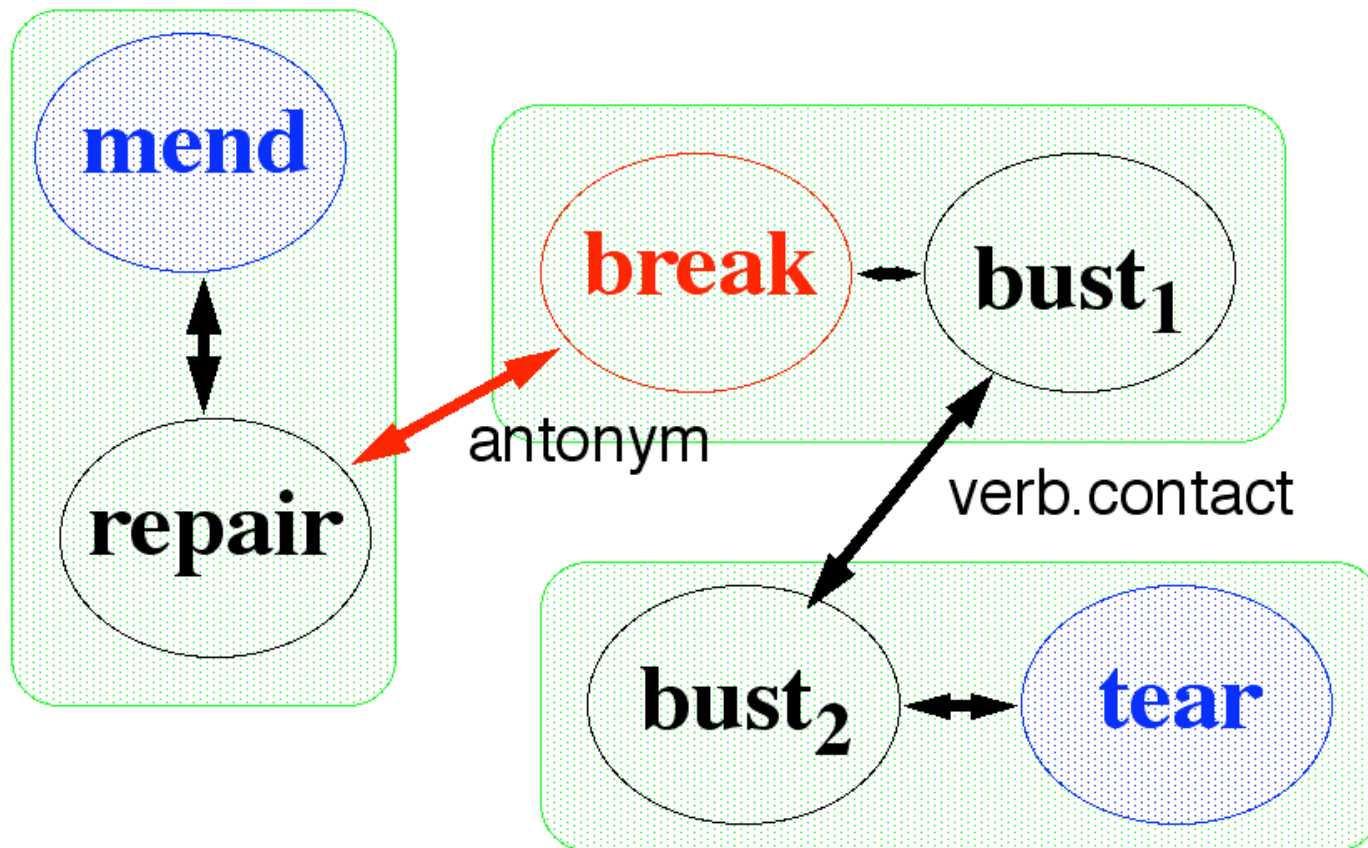
mend/v is in [repair,mend,fix,bushel,doctor,furbish_up,restore,touch_on]
repair and break/v are antonyms
bust in [break,bust] and bust/v related by verb.contact
tear/v is in the synset [tear,rupture,snap,bust]
Unique path
Number of nodes visited: 82
```

Overlaid on the application is a graph window titled 'wnconnectuHxug5.png'. It shows a vertical path of four nodes: 'mend/v' at the top, followed by 'break/v', then 'bust/v', and 'tear/v' at the bottom. The edges are labeled: 'is an antonym of' between 'mend/v' and 'break/v', 'verb.contact' between 'break/v' and 'bust/v', and 'in synset' between 'bust/v' and 'tear/v'.

Using wnconnect

- Find shortest link between **mend** and **tear** in WordNet:
 - **mend/v** is in [repair,mend,fix,bushel,doctor,furbish_up,restore,touch_on]
 - **repair** and **break/v** are antonyms
 - **bust** in [break,bust] and **bust/v** related by verb.contact
 - **tear/v** is in the synset [tear,rupture,snap,bust]

Using wnconnect



two senses of **bust**: (1) to ruin completely,
(2) to separate or cause to separate abruptly

Using wnconnect

- John CAUS the **red** dress BECOME <STATE (**mended**)>



mend/n is in [**mend**,patch]

[mend,patch] is an instance of [sewing,stitchery]

[sewing,stitchery] is an instance of [needlework,needlecraft]

[needlework,needlecraft] is an instance of [creation]

[creation] is an instance of [artifact,artefact]

[artifact,artefact] is an instance of [object,physical_object]

[object,physical_object] is an instance of [entity,physical_thing]

[causal_agent,cause,causal_agency] is an instance of [entity,physical_thing]

[person,individual,someone,somebody,mortal,human,soul] is an instance of [causal_agency]

[disputant,controversialist] is an instance of [person,individual,someone,somebody,mortal,human,soul]

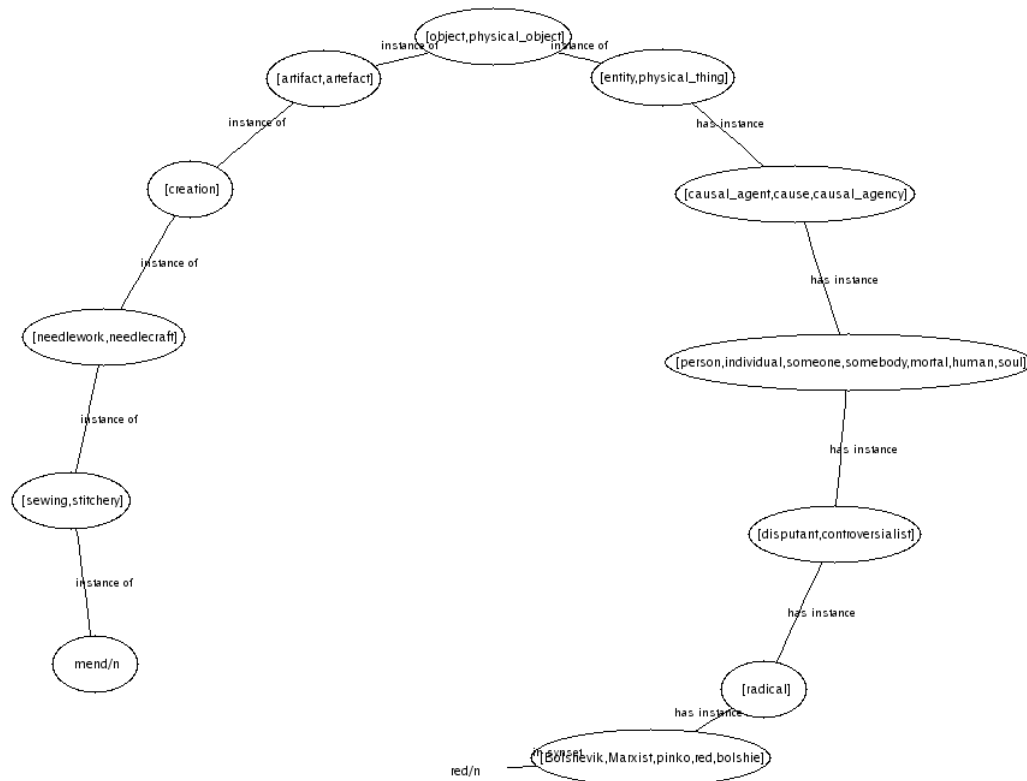
[radical] is an instance of [disputant,controversialist]

[Bolshevik,Marxist,pinko,**red**,bolshie] is an instance of [radical]

red/n is in the synset [Bolshevik,Marxist,pinko,**red**,bolshie]

Using wnconnect

- John CAUS the **red** dress BECOME <STATE (mended)> ←

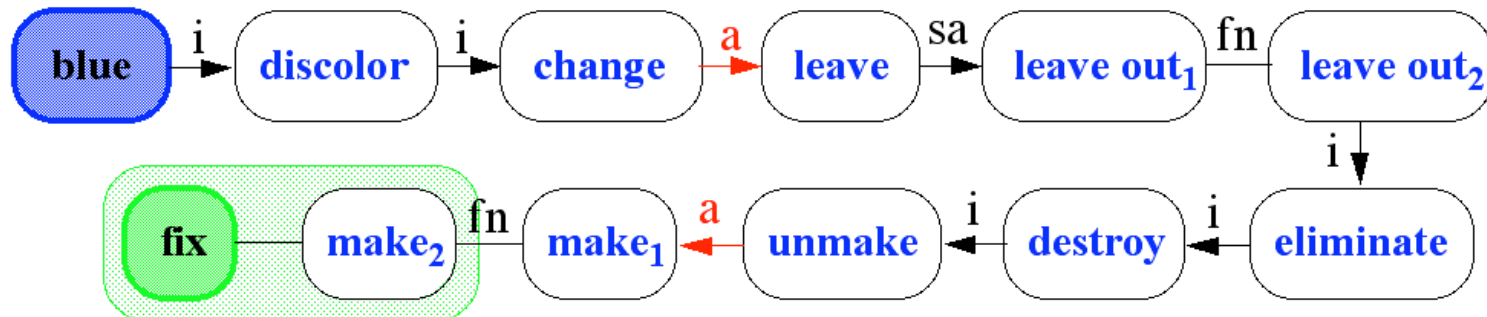


Results

Candidate Pair	Shortest Chain	Semantic Opposition	Search Space
mend-torn	5	Yes	1261
mend-red	-	No	11974
fix-leaky	5	Yes	12167
fix-blue	11	No	14553
fix-flat	-	No*	12286
mix-powdered	6	Yes	11931
comfort-crying	9	Yes	11359
blue-white	-	No*	24431
rescue-drowning	13	Yes	9142
clean-dirty	1	Yes	61
fill-empty	1	Yes	48

Thresholding

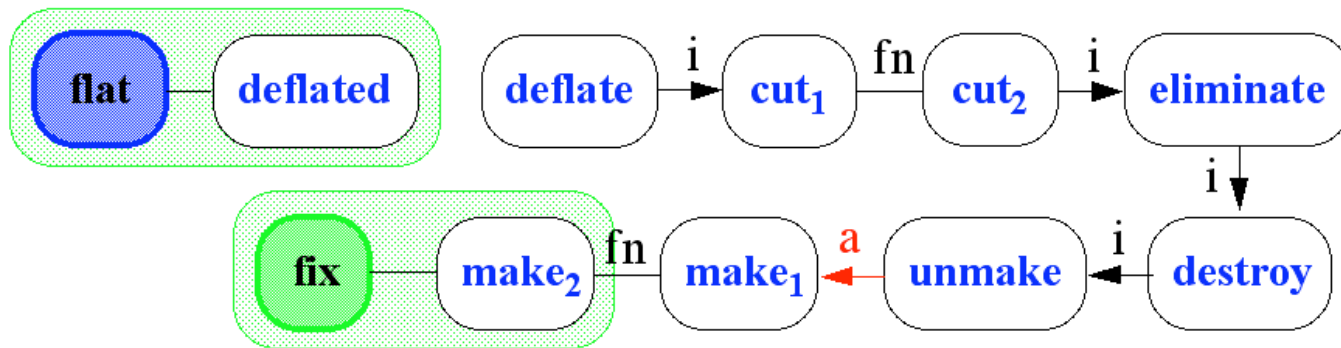
- *No upper limit on the length of the shortest chain*
- **Example:**
 - fix–blue: 11 links (*no semantic opposition*)



cf. *rescue–drowning*: 13 links (*semantic opposition*)

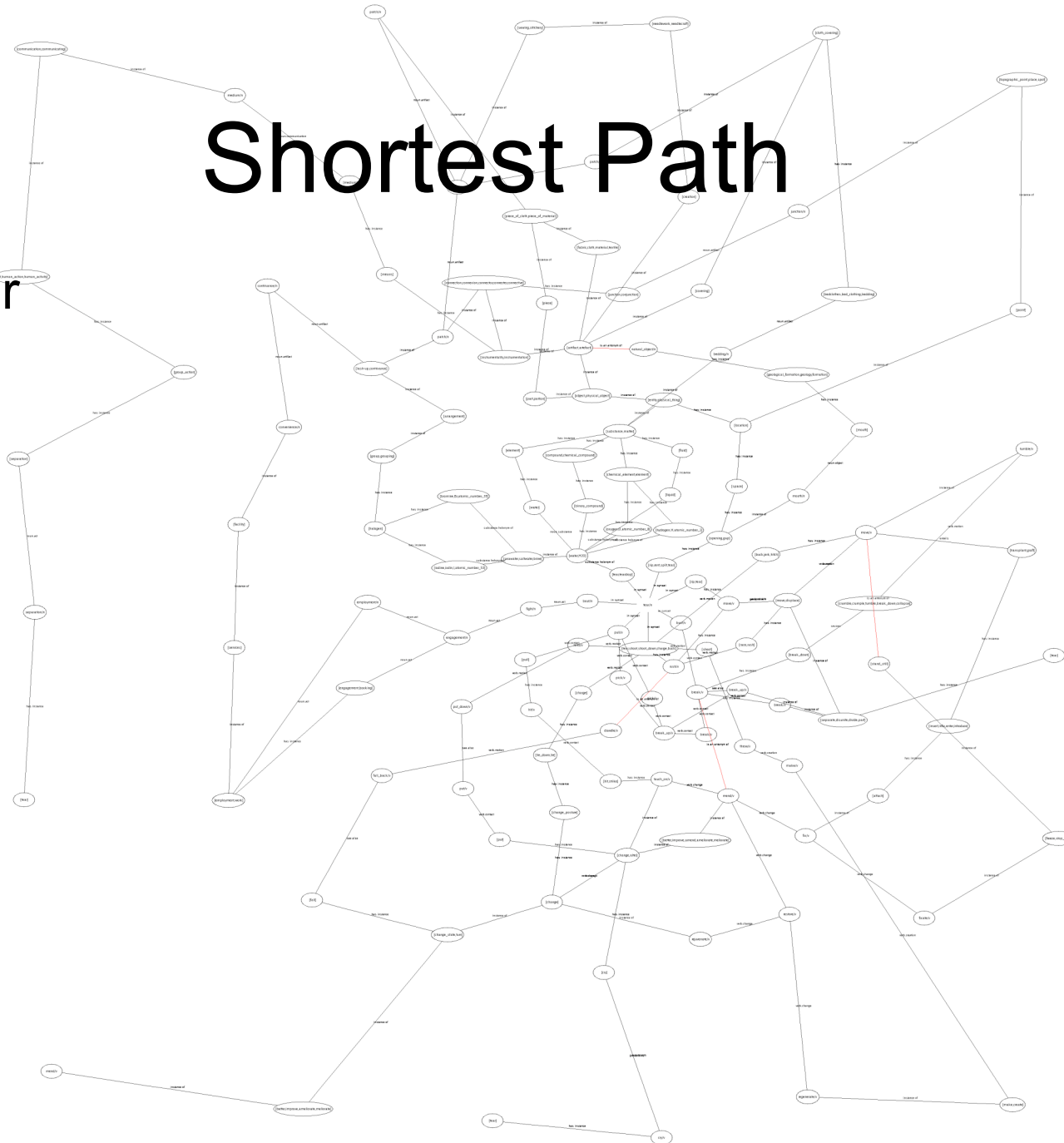
Shortest Path Criterion

- *Take the shortest chain*
- **Example:**
 - fix–flat: no chain found



Shortest Path

mend-tear
all paths



Color

- *WordNet organizes color by chromaticity*
- **Example:**
 - blue–white: *no semantic opposition found*

Color

argent blue-black charcoal gray hueless neutral white

achromatic

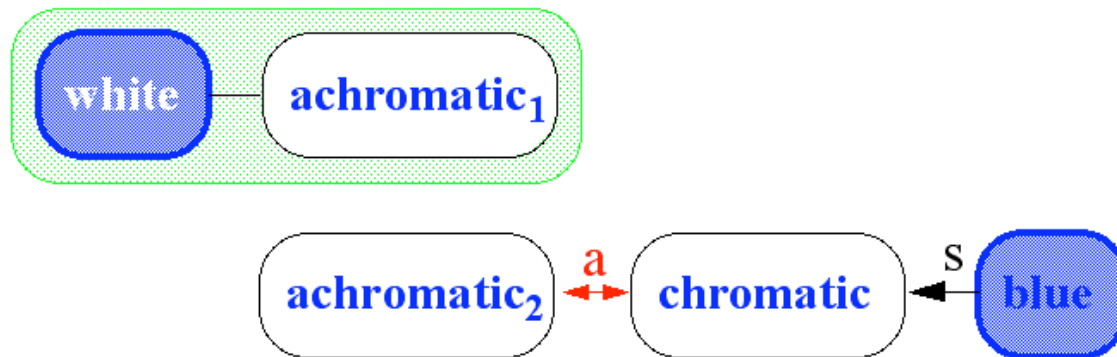
↕ antonym

chromatic

amber azure blue brown dun green red ... yellow

Color

- *WordNet organizes color by chromaticity*
- **Example:**
 - blue–white: *no semantic opposition found*



Color

- *WordNet organizes color by chromaticity*
- **Example:**
 - blue–white: *no semantic opposition found*
- Both chromatic:
 - John painted the red door blue
- Both achromatic:
 - Mary painted the white tiles grey

Two Problems

- *linguistically relevant puzzles*
- *outside syntax*

1. Semantic Opposition

2. Logical Metonymy

... Next Time

Discussion



Discussion

- **Issues**

- **knowledge representation for inferencing**
- *do we have a network like WordNet?*
- *how is it built?*
- *is it wholly outside the “language faculty”?*
- **interaction with language**
 - involves word meanings
 - aspectual verb classes are relevant
 - change-of-state, activity, accomplishments, achievements, etc.
 - states, causation, entailments