Natural Language Processing
CMSC 723 (spring, 2001)

May 2, 2001

- Chapter 16 (continued)
- Why thematic roles?
- Selectional Restrictions
- Schank
- Jackendoff
- Applications

Why Posit a Thematic Level Distinct from that of Syntactic Subcategorization?

1. Capture similarity between different (but related) uses of same lexical item
2. Obviate need for subcategorization frames: Mapping from syntax to lexical semantics

Selectional Restrictions (Section 16.3)

What are selectional restrictions?

Recall the "Godzilla" example.

How are Selectional Restrictions Implemented?

[Figure 16.10]

Primitive Decomposition

- Jim killed his philodendron

- Jim did something to cause his philodendron to become not alive

Schank's Primitives

[Figure 16.11]

Two Axes of Comparison for Lexical-Semantic Approaches

	Local	Non-local
Predicate-dependent	G, J	–
Predicate-independent	S	S,F

Predicate-Independent vs. Predicate Dependent

- Predicate-Independent: Single set of roles is chosen independent of the type of predicates involved (no reference to type of predicates)

- Predicate-Dependent: Roles identified by particular positions arguments occupy wrt primitive predicates

Local vs. Non-Local Orientation

- Local (Localist Hypothesis): Notions of motion & location are central.
- Non-Local Orientation: Concerned w/ causal dimension (affected obj)

Problems: Local/NonLocal Approach

- Problem w/ local approach
  - How do we assimilate abstract verbs to verbs of motion & location?
  - Sometimes need a dual representation to account for causal dimension.
- Problem w/ non-local approach: Notions of motion/location are not given any special consideration.

Gruber (1965)

Agent	NP expressing will toward action
Theme	Object undergoing motion/or located object (instrument too)
Location	PP location
Source	initial position of theme
Goal	final posn of theme (benefactive too)

Jackendoff (1972, 1976, 1983)

Three semantic functions in *Lexical Conceptual Structure* (LCS):

BE $\rightarrow$ Theme X Location
CAUSE $\rightarrow$ Agent X Theme/Event
CHANGE $\rightarrow$ Theme X Source X Goal
Later: Added STAY

Why Gruber?

Rejects Katz-Postal hypothesis: All semantic information about sentence must be represented at D-structure.

Thematic Hierarchy Constraint

Agent > Instrument > Theme

Predicate-Dependent vs. Predicate-Independent

Example	**Fillmore** Open: O+(I)+(A)	**G/J** Open: BE, CAUSE, CHANGE
The door was open	?	BE(NP1, OPEN)
The door opened	O	CHANGE(NP1, ¬OPEN, OPEN)
John opened the door	O + A	CAUSE(NP1, CHANGE(NP2, ¬OPEN, OPEN))
The wind opened the door	O+I	
John opened the door with a chisel	O+I+A	CAUSE(NP1, CAUSE(NP2, CHANGE(NP3,¬OPEN,OPEN)))

Another Axis of Comparison: Decomposition vs. Non-Decomposition

- Decomposition / Compositional Approach (Schank, Jackendoff) vs. Non-decomposition / Non-compositional Approach (Fillmore)
- Within Compositional approaches: exhaustive (Schank) vs. nonexhaustive (Jackendoff).

Schank: Motivation

- Underlying Motivation: "Strong AI"
- Focus: understanding. Argues that the representation is reversible.
- Rejects syntax during analysis. Allows it during generation.
- Attempts to come up with well-defined system of rules and conceptualizations.

Schank: Motivation (continued)

- Inferences, expectation, syntax, conversational norms, real world.
- Conceptual Structure (CD): Language-independent conceptual level.

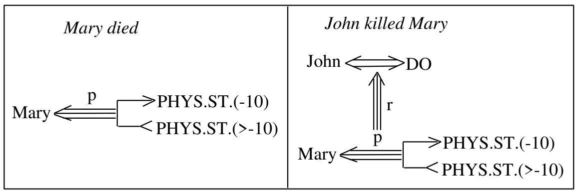
Schank: Primitives

- Four Units
  - PP (N)
  - PA (Adj)
  - AA (Adv)
  - ACT (V)
- Rules for relating these:
  - PP $\Longleftrightarrow$ ACT
  - ACT $\overset{o}{\longleftarrow}$ PP
- Conceptual Cases: CD representation
  - ACT $\overset{o}{\longleftarrow}$ PP
  - ACT $\overset{R}{\longleftarrow} \begin{smallmatrix}\text{PP}\\ \text{PP}\end{smallmatrix}$
  - ACT $\overset{I}{\longleftarrow} \Updownarrow$
  - ACT $\overset{D}{\longleftarrow} \begin{smallmatrix}\text{PP}\\ \text{PP}\end{smallmatrix}$
- Primitive acts:
  - Mental
  - Abstract
  - Physical
  - Others

Schank: Causal Dependencies

- x is dependent on y; or y caused x:
- Notion of dependency is blurry
- Other dependencies: time ($\downarrow$) and place ($\Downarrow$)

Schank: Kill vs. Die



Schank: Problem 1

- "John caused Mary to die" vs. "John killed Mary"
- Identically substitutable?
- Flaw of all compositional approaches of this nature.

Schank: Problem 2

Problem with instrumentality: infinite recursion!

Schank: Problem 3

The decompositions are very complex.

- Too specific.
- Why are these conceptualizations so radically distinct from the syntactic realization?
- Talks CD from NL understanding point of view — what about generation?

Jackendoff's Later Semantic Theory – 1983

- Predicate-centered (Gruber)
- Survey of spatial relations
- Three basic primitives:
  GO
  STAY
  BE
- Test: What happened was __ (event/*state)
- PLACE/PATHs: TO, FROM, etc.

Jackendoff's Later Semantic Theory (continued)

- GO_Ext:
  The road goes to Boston (state, not event)
- ORIENT:
  The sign points to Philadelphia.
- Agentive predicates: CAUSE, LET
  Laura took the bird from the cage.
  Laura released the bird from the cage.
- CAUSE vs. LET

  Apparent counterexample:
  Laura took ... with a hanger.
  Laura released ... with a hanger.

Jackendoff's Fields

- Central notion in J's later work
- Allows many generalizations to be stated
- GO, STAY, BE extend to fields other than the spatial field
- Localist Hypothesis: focuses on motion/location (spatial dimension)

GO	
Possessional	
Identificational	
Circumstantial	
Temporal	

Cross-field Generalizations

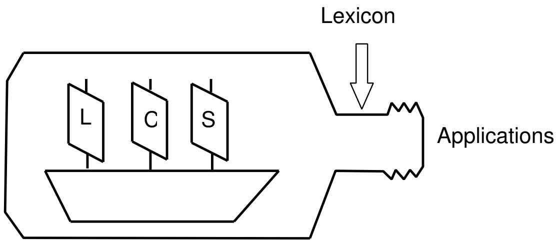
- GO
  Loc: Max came into the room
  Circ: Max came to be called a hero
- LET GO
  Loc: Laura released the bird from the cage
  Circ: Laura released Fred from washing dishes
- GO_Ext:
  Loc: The highway goes from D to B
  Temp: Ron's speech went from 2 to 4
- STAY
  Loc: The iguana stayed in Africa
  Poss: The iguana stayed in Ron's possession
  Ident: The iguana stayed ugly
- CAUSE STAY
  Loc: Bill kept the book on the shelf
  Circ: Bill kept David working
  Temp: Bill kept the meeting at 2pm
- BE
  Loc: Bill is in Africa
  Ident: Bill is happy
  Circ: Bill is working

Decompositionality

- Capturing semantic info: "The hawk flew from the nest"
  [GO (Bird, [FROM (NEST)])]
- − John climbed the mountain (trans)
  − John climbed up the mountain (intrans)
  − John climbed down the mountain (intrans)

  [GO (X, [TO (Y)])]

The NLP Bottleneck:
Acquisition of Computational Lexicons



Toy System → Real World: Why hard?

Automatic Lexical Acquisition
Deriving Meaning (Content)
from Syntactic Patterns (Form)

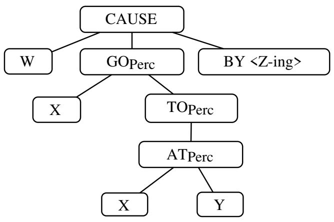
- **Online Dictionaries:** provide purely syntactic information (form)

  Example: Dictionary entry for *coerce*

```
T1-INTO          [np,v,np,pp(into)]
V3               [np,v,np,infinitive]
```

- **Computational Lexicons:** require cohesive semantic classes (content)

  Example: *Coerce Class*: allure, blackmail, bluff, bribe, cajole, coerce, ...



Levin's Introduction (1993)

Syntax	Middle	Conative	Body-Part	Causative
Semantics	COS (0-rel res) vs. no COS (0-rel act)	motion	contact	pure change
Break	Y	N	N	Y
Cut	Y	Y	Y	N
Hit	N	Y	Y	N
Touch	N	N	Y	N

Levin: Basic Tenets

- Verb Behavior determined from meaning
- Alternations can be found in different languages.
- Alternations not always the same in all languages but meaning components are the same.
- Not necessary to store all aspects of verb's behavior in lexical entry (i.e., all alternations).
- Lexical entries factor out predictable information and store idiosyncratic stuff separately.
- After classes are defined, need to determine meaning components.
- Can't rely on intuitions. Need to find "right" meaning components!

Using Linguistic Classes

Claim: The syntactic behavior of a word is fully semantically determined. This property holds across all languages.

Semantic Class: Manner of Motion (Levin, 1993)

Path: The horse ran The horse ran through/into/out of the stream
There Insertion: A horse ran out of the barn There ran out of the barn a horse
Resultative: She ran She ran herself ragged

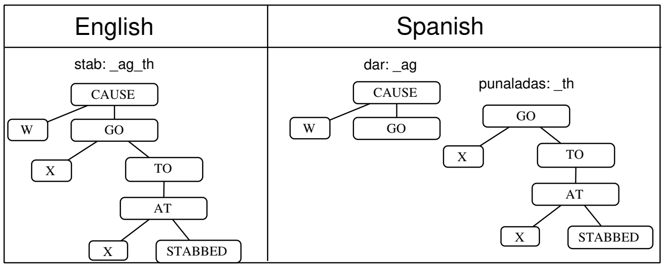
Semantic Class: Disappearance

Path: The horse died ∗The horse died through/into/out of the stream
There Insertion: A horse died ∗There died a horse
Resultative: She died ∗She died herself cold

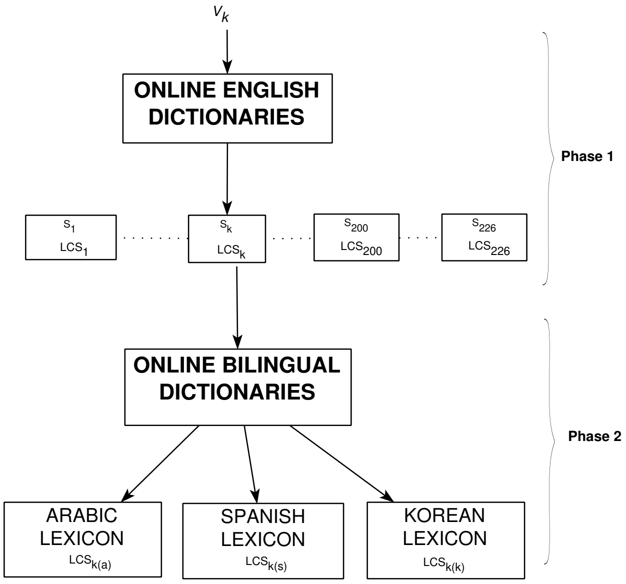
Approximately 200 Semantic Classes.

Semantic Classification: Deriving Basic Meaning Components

- Semantic classes associated with alternations are comprised of “basic meaning components.”
- Cross-linguistic Applicability is brought about by decomposition of “basic meaning components” into primitive units of meaning—these carry over to other languages.

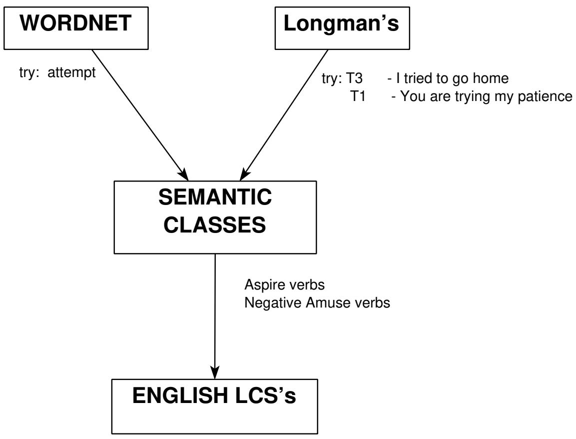


Automatic Lexicon Construction



Phase I: Semantic Classification

- Test syntax ↔ semantics relationship.
- Produce database of English LCS’s.



Syntactic Codes in Longman's

Longman's Code	Example
D1	She **allowed** him some money for expenses
D1-FOR	She **built** him the house
D1-TO	She **brought** him the newspaper
T1	I **pioneered** the new land
T1-FOR	They **admired** him for his stamina
T3	He **tried** to do it
T4	She **tried** eating the new food

Control Verbs from LDOCE not found in Levin

afford, aim, answer, attack, attempt, attend, attract, bar, bear, become, beg, belong, blame, branch, breed, broadcast, bury, calculate, command, consist, curse, damage, dare, deal, deceive, decide, defeat, defend, demand, deserve, do, doubt, dream, educate, equal, expect, experience, fail, farm, force, forget, fulfil, govern, help, include, influence, inform, inquire, insure, interrupt, lack, let, limit, manage, match, mend, mistake, park, participate, pause, permit, persuade, plan, possess, practise, preserve, pretend, prevent, protect, reduce, refuse, repair, reply, request, retreat, safeguard, seem, shade, share, spell, spend, spoil, spring, step, succeed, suit, swear, tend, test, tidy, tour, translate, understand, undo, urge

Correct Assignments to Levin's Classes			
Word	**Assigned Class**	**LDOCE Codes**	**WordNet Synonyms**
afford	13.1	D1 T1 T1-TO T3	render give
blame	33	D1-FOR D1-ON T1 T1-FOR N	impeach fault criticize
branch	23.4	I N [I-FROM]	diverge
breed	22.2	I T1 N [T1-WITH]	mate pair
bury	9.1	T1 X9	situate immerse
bury	16	T1 X9	conceal hide
command	37.2	I T1 T5B T5C V3 N	tell
curse	33	I T1 N [T1-FOR]	abuse
damage	31.1	T1 WV4 N	devastate afflict
deceive	33	I T1 T1-IN T1-INTO WV4	victimize
doubt	29.5	T1 T5A T6A N	suspect
doubt	31.2	T1 T5A T6A N	distrust
fulfil	13.4.1	T1	supply provide
help	13.4.1	I T1 T1-TO T4 V2 V3 V4 N [T1-WITH]	supply provide serve
inquire	31.3	I I-ABOUT T1 T6A	wonder
interrupt	55.1	I T1 [L9]	terminate end
lack	32.1	T1 N	need want
mend	45.4	I T1 N	improve heal ameliorate
park	9.1	I T1 X9 N	position place set put
practise	14	I I-AS T1 T1-ON T1-UPON T4 WV4	read study learn
preserve	13.5.1	T1 T1-FROM WV5 N	keep save
reduce	26.6	I T1 T1-FROM T1-TO	turn change
reply	37.7	I I-TO T5 N	say state
spell	37.1	D1 I T1 T1-WITH N	recite write
spoil	31.1	I T1 T1-OF WV5 N	devastate afflict discourage frustrate baffle
spring	51.3.1	I L7 L9 T1 T1-ON X7 N	bounce move
succeed	47.8	I I-IN I-TO T1	follow
swear	37.7	I I-AT T1 T1-ON T1-TO T3 T5	declare
tour	51.3.2	L9 N	journey travel
translate	26.6	I T1	alter change transform
undo	44	T1	wreck ruin destroy

Correct Assignments to New Classes			
Word	**Assigned Class**	**LDOCE Codes**	**WordNet Synonyms**
attempt	005	T1 T3 T4 WV5 N	try seek
attend	013	I I-ON I-TO I-UPON T1	serve
bear	014	D1 D1-FOR D1-TO I L9 T1 T4 X9 N	take make pay
beg	015	I I-FOR I3 T1 T1-OF T3 T5C V3	pray ask
consist	022	T1 T4 WV6 [L9]	lie
decide	017	I T1 T3 T5A T6A T6B V3	[make get]
defend	021	T1 T1-AGAINST T1-FROM T4	guard keep
demand	017	T1 T3 T5C N	claim
expect	015	T1 T1-FROM T3 T5A T5B V3 WV6 X9	wish hope desire
fail	005	I I3 T1 T3 N	omit
force	002	T1 T1-INTO V3 WV5 X7 N	push
forget	005	I T1 T3 T4 T5A T6A T6B	omit
insure	021	T1 T1-AGAINST T5 [T1-FROM]	guard
manage	005	I T1 T3	watch
persuade	002	D5 T1 T1-INTO T1-OF T1-OUT_OF V3	[make]
plan	005	I I-FOR I-ON T1 T3 N	think intend mean propose
prevent	023	T1 V4 [T1-FROM]	keep
protect	021	T1 T1-AGAINST T1-FROM	guard
refuse	005	D1 I T1 T3 N	decline
request	015	T1 T5 V3 N [T3]	ask
request	017	T1 T5 V3 N [T3]	ask
seem	020	I-IT L1 L1-TO_BE L7 L7-TO_BE L9 L9-TO_BE WV6	appear
shade	021	L9 T1 N [T1-AGAINST] [T1-FROM]	[draw]
share	026	I T1 T1-AMONG T1-BETWEEN T1-WITH N	distribute
tend	005	L9 T1 T3	think watch
understand	018	I T1 T1-BY T5 T6 V3 WV6	feel perceive understand
urge	002	T1 T1-ON V3 X9 N [T1-INTO]	press

New Semantic Classes Acquired			
Class		LDOCE Codes	New Verbs
001	Be-For Verbs	I-FOR	be suited for, be proper for, ...
002	Coerce Verbs	V3 T1-INTO	allure, blackmail, bribe, cajole, coerce, ...
003	Conspire Verbs	T1-AGAINST	legislate, protest, rebel, ...
004	Drop Verbs	L9	let drop, let fall, let flow, ...
005	Aspire Verbs	T3	afford, aspire, attempt, ... try, ...
006	Do Verbs	L7 L9 D1 D1-FOR	do, re-do, ...
007	Exceed Verbs	I-IN	exceed, excel, outbid, outnumber, transcend, ...
008	Impose Verbs	T1-ON T1-UPON	hazard, enforce, exert, gamble, impose, ...
009	Play Verbs	I-WITH I-AGAINST	hock, pawn, play with, ...
010	Become Verbs	L1 L7	become bored with, become impatient with, ...
011	Penetrate Verbs	I-THROUGH WV5-WITH	penetrate, permeate, pervade, ...
012	Sustain Verbs	WV4	keep to, maintain, stick to, sustain, ...
013	Attend Verbs	I-ON	attend, serve, ...
014	Bear Verbs	D1 D1-FOR D1-TO X9	bear, take
015	Beg Verbs	T3 V3	ask, beg, dare, expect, need, pray, request, wish
016	Attract Verbs	L9 WV5 X9	attract, draw, pull
017	Claim Verbs	T3 T5 T5A T5B T5C	ask, claim, decide,demand, pretend, request
018	Perceive Verbs	T5 T5A T5B T5C T6 T6A T6B V3	feel, perceive, understand
019	Let Verbs	V2	let, make
020	Seem Verbs	L1-TO_BE	appear, seem
021	Defend Verbs	T1-AGAINST T1-FROM	defend, guard, insure, protect, shade
022	Consist Verbs	L9 WV6	belong, consist, lie
023	Prevent Verbs	V4 T1-FROM	keep, prevent
024	Admit Verbs	T4 X9	admit, permit, include, allow
025	Spend Verbs	T1-ON T1-FOR	spend, pay
026	Share Verbs	T1-AMONG	distribute, share

Overall Results

82	61% Correct Semantic Class Assignments
29	22% Incorrect Based on Syntactic Omissions
23	17% Incorrect Based on Semantic mismatch (WordNet/Levin)
134	100% Total semantic class assignments

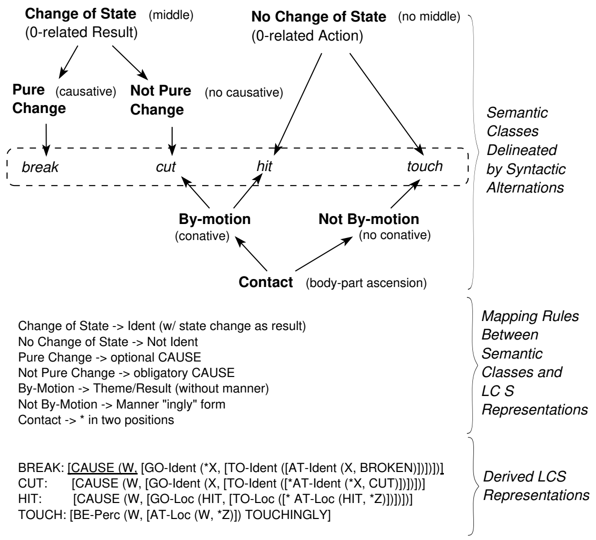
- Identification of new semantic classes for verbs that exhibit "non-canonical" syntactic behaviors with respect to existing classes.

- The assignment of verbs already included in Levin's classes to additional (existing or new) classes.

Second Task in Phase I: LCS Acquisition

Once verbs are classified, how do we derive an LCS representation?

LEXICALL: Language-Independent Mapping Between Verb Classes and Meaning



LCSs Based on Levin's Verb Classification

Category	Verb	Class	LCS
Motion	leave	51.2	[GO_{Loc} (Y, [⟨DIRECTION⟩$_{Loc}$ (Y, [AT_{Loc} (Y, Z)])])]
	run	51.3.1	[GO_{Loc} (Y, [BY ⟨MANNER⟩])]
Placement	fill	9.8	[CAUSE (X, [GO_{Ident} (Y, [$TOWARD_{Ident}$ (Y, [AT_{Ident} (Y, [⟨STATE⟩$_{Ident}$ ([⟨WITH⟩$_{Poss}$ (*HEAD*, Z)])])])])])]
	pour	9.5	[CAUSE (X, [GO_{Loc} (Y)], [BY ⟨MANNER⟩])]
Sound	say	37.7	[CAUSE (X, [GO_{Ident} (Y, [$TOWARD_{Ident}$ (Y, [AT_{Ident} (Y, [⟨STATE⟩$_{Ident}$])])])])]
	shout	37.3	[CAUSE (X, [GO_{Perc} (Y)], [BY ⟨MANNER⟩])]