

Similarity-based Classification in Compositional Semantics

Carla Umbach
Institute of Cognitive Science Osnabrück/Germany

classification processes: category-based vs. similarity-based

natural language semantics:

category-based: lexical rules, composition rules

similarity-based: metaphors

>>> Are there similarity-based mechanism in natural language semantics beyond the lexical level ??

Lund University, September 2005

Cognitive architectures

Cognitive Science PhD Programme Osnabrück:
“Cognitive Architectures. The Integration of Rules and Patterns“

category-based vs. similarity-based classification:
competing models of human cognition ?

Hahn & Chater (1998):

- no decisive evidence in favour of one or the other mechanism
- evidence: both mechanisms co-exist
- interact making use of their complementary strength

psychology, AI, ... natural language semantics ?

2

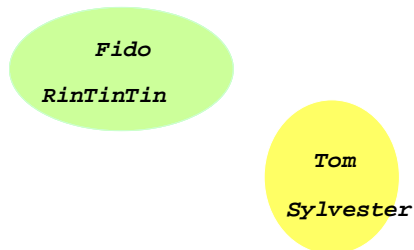
Category- vs. similarity-based classification

▶ category-based “learning by instruction“

if furry(x) and barks(x) then dog(x)



▶ similarity-based “learning by experience“



3

Basic differences

▶ category-based (basic) categories apply to object → true/false

static

▶ similarity-based similarity relation + relevant features align object features with features of previously given objects → add to similar items

dynamic

4

Coordination constraints

Lang (1977):

conjuncts must be parallel (syntax, semantics and prosody)

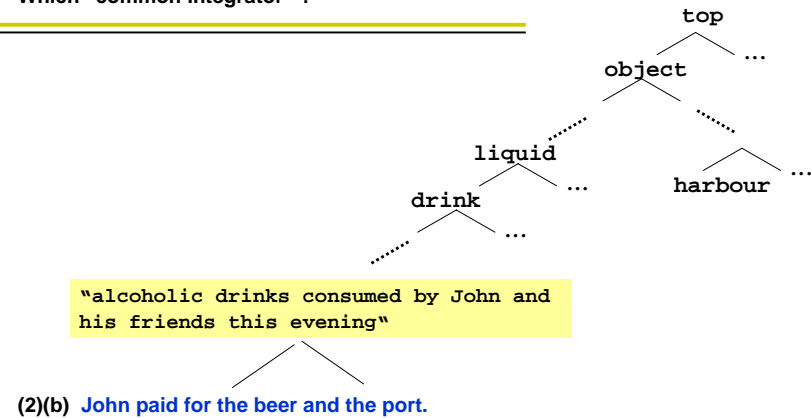
semantics:

conjuncts must be (i) semantically distinct
(ii) subsumed by a "common integrator"

- (1) (a) # John likes animals and birds.
(b) # John had a drink and a martini.
- (2) (a) John has a girlfriend and a computer.
(b) John paid for the beer and the port.
- (3) Göttingen ist bekannt für Professoren, Philister und Vieh. (H.Heine)
'Göttingen is well-known for professors, bourgeois and cattle.'

5

Which "common integrator" ?



→ "common integrator": requirement for similarity

6

Focus semantics

- Focus evokes a set of alternatives: (Rooth 1992, Krifka 1993, ...)
- (4) (a) Bob wants COFFEE. what does Bob want?
ALT (COFFEE) = {coffee, tea, juice, ...}
- (b) BOB wants coffee. who wants coffee?
ALT (BOB) = {Bob, Sue, Mary ...}
- alternativ sets:
quantificational domain for, e.g., *only*
- (5) Bob *only* wants COFFEE.
ALT (COFFEE) = {coffee, tea, ...} (5) → Bob doesn't want tea ...
- ALT (COFFEE) = {coffee, tea, ... sandwiches} (5) → Bob doesn't want sandwiches

7

The problem of restriction

- (6) (... at the party, ...)
John *only* saw MARY / the FRENCH girl. ⚡ John didn't see any other people.
- Rooth (1992):
- | | |
|--|----------------------------------|
| alternative sets | ALT(the FRENCH girl) |
| • have to be of the same semantic type | ⊆ domain of individuals |
| • are restricted by the background | ⊆ girls of various nationalities |
| • are restricted by the context | ⊆ contextually given set C (?) |
- (7) (The research team arrived late at night.)
Ben *only* talked to the DUTCH geologist.
→ ... not to any other team member
→ ∃ non-dutch team members
→ the researchers are geologists
- Umbach (2003):
For focus in definite NPs:
alternative sets are restricted by a bridging antecedent

8

Alternatives comply with Lang's coordination constraints

(8) (a) John only paid for the DRINK (... # not for the martini).

→ martini \notin ALT (DRINK)

(b) John only paid for the BEER (... not for the port).

port \in ALT (BEER) → port \subset drink

→ alternatives must be (i) semantically distinct
(ii) subsumed by a "common integrator"

(no surprise because conjuncts have to be alternatives with respect to each other)

9

Alternatives require similarity

again: which "common integrator" ?

(9) (... at the party, ...)

John only saw MARY in the crowd.

ALT (MARY) = {Mary, ... ??}

... persons/women mentioned before ?
... persons John wanted to meet at the party. " ?
... persons John did not want to meet ?
... etc.

→ persons similar to Mary in this context

(10) John ordered a BEER.

ALT (BEER) = {beer, ... ??}

... drinks/food usually/usually not ordered by
John/someone at this occasion ?

→ items similar to beer in this context

→ alternatives have to be similar (and dissimilar)
with respect to features given by the context

10

Alternatives constructed by similarity

▶ alternative sets evoked by focus

- context-dependent
- ad hoc categories, constructed on the fly
- temporary

→ constructed via similarity to focussed item
(with respect to contextually salient features)

▶ features salient in the context

- provided by
- properties of the focussed item
 - selectional restrictions
 - anaphoricity
 - bridging relations
 - etc.

11

What does similarity buy us?

▶ focus semantics:

- closer examination of the nature of
alternative sets

▶ similarity- vs category-based classification:

- similarity-based classification plays a role
even in compositional semantics
- evidence that natural language semantics involves both
category-based and similarity-based mechanisms.

12

Structure plays an important role

- (11) **John took Mary, Sue and Alice to the party.
But he only danced with SUE.**

ALT (SUE) = {Sue, Mary, Alice}

- (12) **John looked for his girlfriend all over the place.
But he only found a SHOE.**

ALT (SHOE) = {a shoe, John's girlfriend}

Brown Corpus ("only"):

- (13) **He not only knew soldiering, but mathematics, history and literature as well.**
- (14) **Bari was chosen as a depot , not only for its seeming safety , but because of its proximity to airfields.**
- (15) **It never seemed real; never seemed as if it could happen to you; only to the other fellow .**

References

- Gärdenfors, P. (2000) *Conceptual Spaces*. MIT Press.
- Goodman, N. (1972). Seven strictures on similarity. In Goodman, N., editor, *Problems and Projects*, pages 437–447. The BobbsMerrill Company, Indianapolis and New York.
- Lang, E. (1984). *The semantics of coordination*. Studies in Language Companion Series 9. Benjamins, Amsterdam.
- Lang, E. (1991) *Koordinierende Konjunktionen*. In von Stechow, A. & Wunderlich, D. (eds) *Semantik. Ein internationales Handbuch der zeitgenössischen Forschung*. de Gruyter, Berlin, 1991.
- Hahn, U. and Chater, N. (1998). Similarity and Rules: Distinct? Exhaustive? Empirically Distinguishable? *Cognition*, 65:197–230.
- Hartmann, K. (2000) *Right Node Raising and Gapping*. Interface conditions on prosodic deletion. Amsterdam Philadelphia, John Benjamins.
- Lakoff, G. & M. Johnson (1980) *Metaphors we live by*. Chicago, IL.
- Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1(1):75–116.
- Umbach, C. (2003) *Anaphoric Restriction of Alternative Sets: On the Role of Bridging Antecedents*. Proceedings of "Sinn und Bedeutung" 7, Konstanz linguistics working papers. No 114, 310-323.
- Umbach, C. (2004) *On the Notion of Contrast in Information Structure and Discourse Structure*. *Journal of Semantics*, Vol. 21 issue 2.