

ADVENTURES IN MEANING BANKING

Johan Bos

ADVENTURES IN MEANING BANKING

- ① A new adventure
- ② Choice of currency
- ③ Preparation
- ④ Off we go!
- ⑤ Help from strangers
- ⑥ Well on our way
- ⑦ Sweet memories
- ⑧ Bumps on the road
- ⑨ Lost in translation



Chapter I

A NEW ADVENTURE

WHERE ARE WE NOW?

Computational semantics is now at a point where computational syntax was twenty (thirty?) years ago:

- ▶ robust semantic parsing only for small domains
- ▶ semantic tasks often carried out in isolation
- ▶ lack of large semantically annotated linguistic resources

OUR GOAL

- Syntactic parsing made a breakthrough by developing treebanks in the 1990s
- We want to develop a large semantically annotated resource and ...
- ... cause a similar breakthrough in computational semantics!

THE MOTLEY CREW

Kilian
Evang



Valerio
Basile



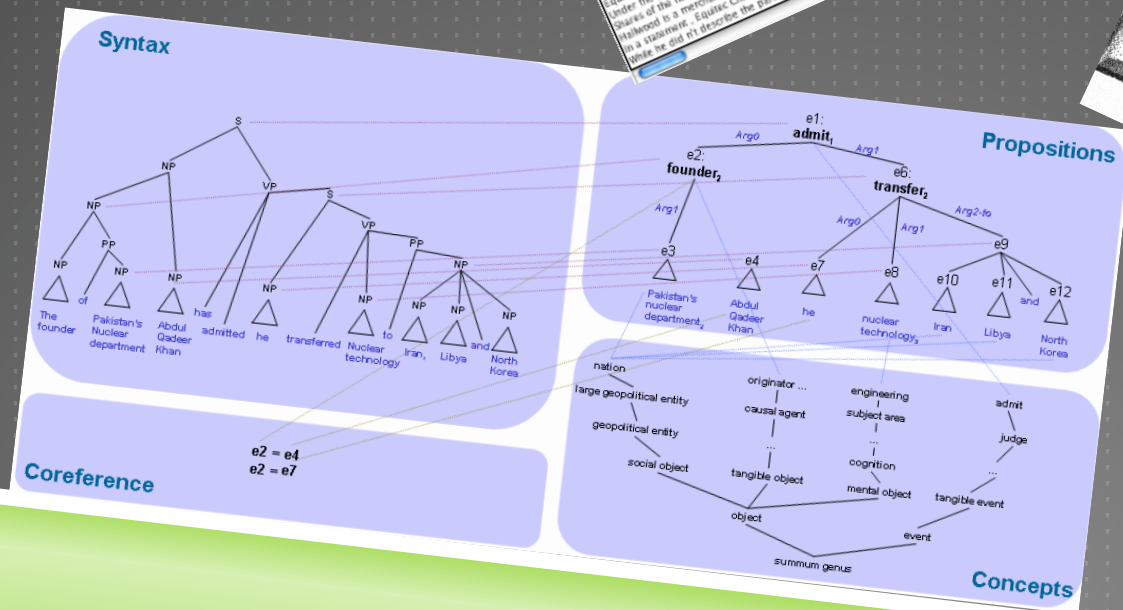
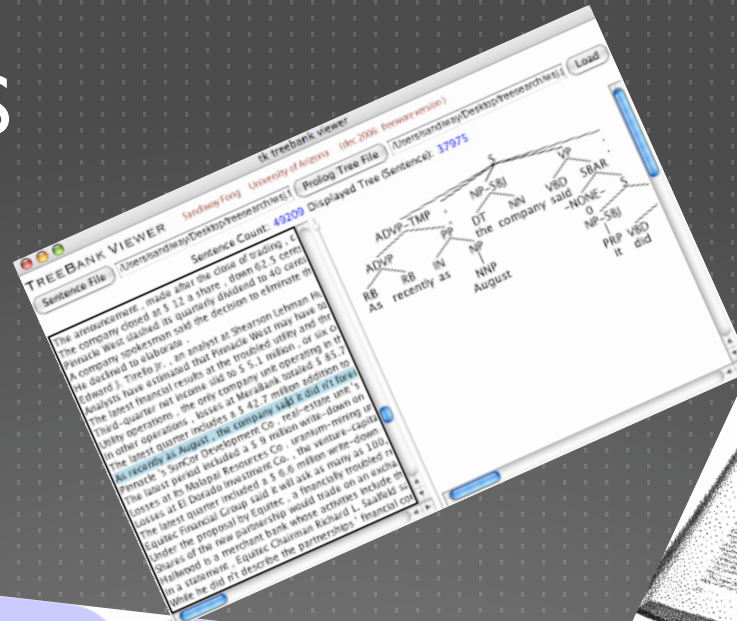
Noortje
Venhuizen



Johannes
Bjerva



PREVIOUS ADVENTURES



THE GRONINGEN MEANING BANK

- ▶ large (English) corpus of public domain texts
- ▶ annotated with deep (formal), not shallow semantics
- ▶ largest annotated units are texts (not sentences)
- ▶ integrates various phenomena in one formalism
- ▶ automatically produced, manually corrected



Groningen

OBJECTIVES OF THE GMB

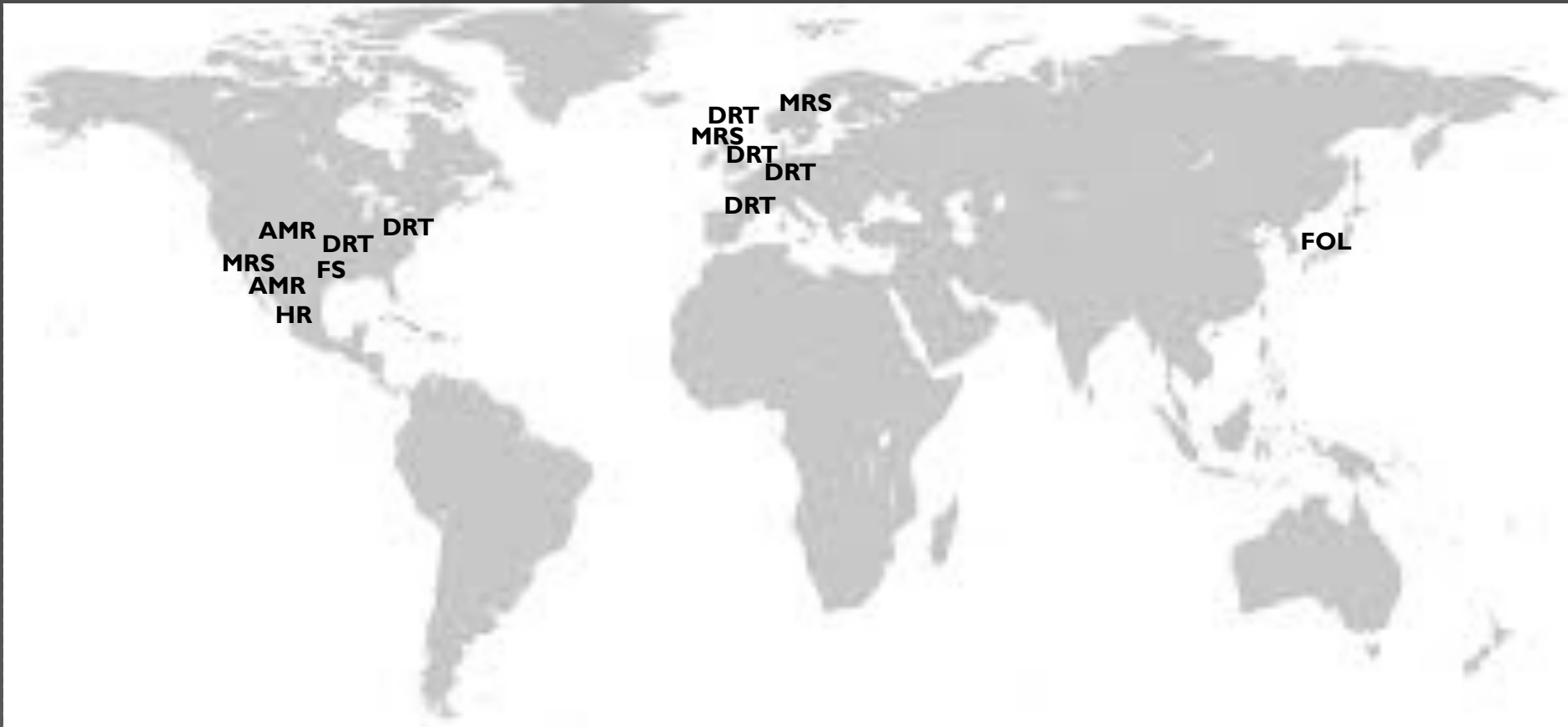
- ▶ develop, refine, and evaluate wide-coverage semantic parsers
- ▶ develop NLG components from formal meaning representations
- ▶ revise semantic theory where needed



Chapter 2

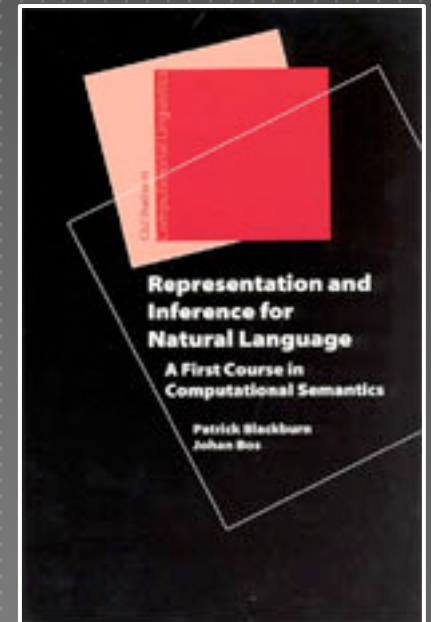
CHOICE OF CURRENCY

DISTRIBUTION OF REPRESENTATION EMPIRES



REPRESENTATION AND INFERENCE

- ▶ syntactic representation serve to build meaning representations
- ▶ meaning representations serve to perform inferences



Blackburn & Bos 2005

EXAMPLE: VENICE

Turn left and or right to reach San Marco.



EXAMPLE: BOLT

Bolt is faster than everyone else. **YES**

Bolt is in last position. **NO**



EXAMPLE: PETS

No pets are allowed in this area.

All pets must be on the leash in this area.



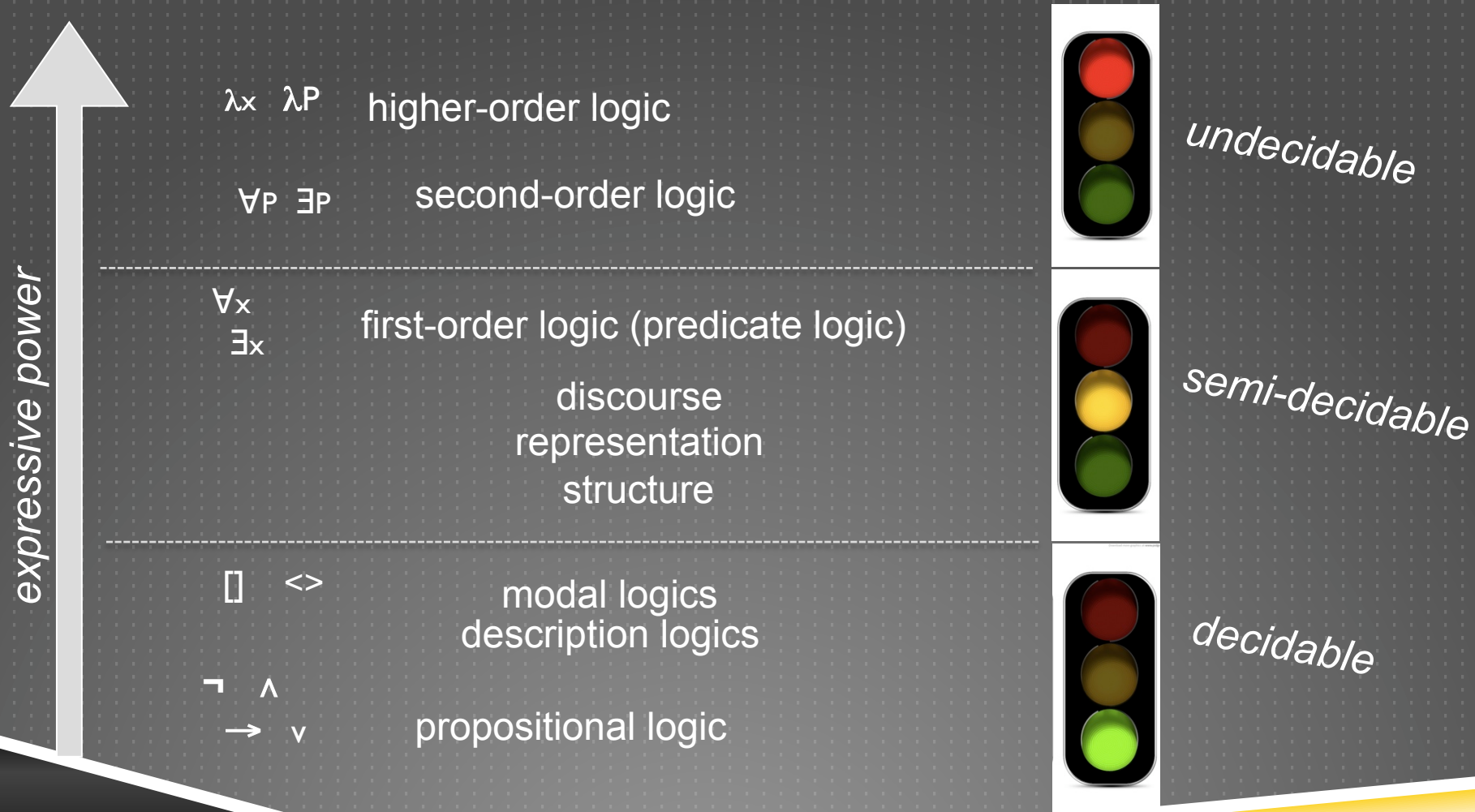
EXAMPLE: BUSH

“... when there's more trade, there's more commerce.”

**not
informative**

George W. Bush, at the Summit of the Americas in Quebec City,
April 21, 2001 (source: Language Log 24/10/2004)

CONTROLLING INFERENCE



Butch stole a chopper.
It was parked in a garage.

no events

x y u v
Butch(x) chopper(y) stole(x,y) u=y garage(v) parked-in(u,v)



Davidsonian

x y e u v e'
Butch(x) chopper(y) stole(e,x,y) u=y garage(v) parked(e',u) in(e',v)



Hobbsian

x y e u v e' a b c d
Butch(x) chopper(y) stole(e,x,a,b) agent(e,x) u=y garage(v) parked(e',u,c,d) in(e',v)



neo-Davidsonian

x y e u v e'
Butch(x) chopper(y) stole(e) agent(e,x) theme(e,y) u=y garage(v) parked(e') theme(e',u) location(e',v)



Chapter 3

PREPARATION

SURVIVAL TOOLS

- ▶ Good tools are needed to reach wide coverage
- ▶ Tokenisation, Tagging, Parsing
- ▶ Compositional semantics, semantic resolution: Boxing

BOXER

- ▶ Input: syntactic tree (categorial grammar derivation)
CCG: see Steedman
- ▶ Output: resolved semantic representation
(box, i.e. a Discourse Representation Structure)

Boxer works with
the C&C parser and easyCCG

Online demo:
<http://www.let.rug.nl/basile/gsb/webdemo/demo.php>



Chapter 4

OFF WE GO!

ITERATIVE DEVELOPMENT

- ▶ Annotations are produced by machines, corrected by humans (crowdsourcing + experts)
- ▶ A lot of it is done by (statistical) NLP tools:
C&C tools + **Boxer**



MEANING BANK DEVELOPMENT

Document **28 of 10110**, ID: /

size: 3 sentences, 21 tokens

last processed: 21 November 2013, 05:08:49

C&C tools/Boxer revision: 2415

[< first](#) [<< previous](#) [next >>](#) [last >](#) [random](#)

Status: [accepted](#) [history](#)

[report issue](#)

Change to: Comment:

[metadata](#) [raw](#) [tokens](#) [sentences](#) [discourse](#) [1 bit of wisdom](#) [0 warnings](#)

Show: ☒ POS ☐ lemmas ☐ namex ☐ numex ☐ timex ☐ animacy ☐ senses ☐ roles ☐ relations ☐ coreference ☐ syntax

- 1

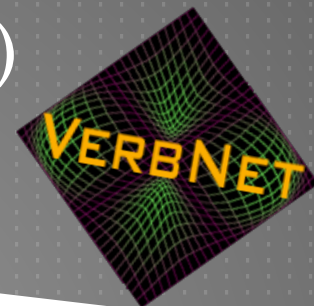
We	're	sorry	.
PRP	VBP	RB	.
- 2

You	have	reached	an	imaginary	number	.
PRP	VBP	VRN	DT	JJ	NN	.
- 3

Please	rotate	your	phone	90	degrees	and	try	again	.
VB	VB	PRP\$	NN	CD	NNS	CC	VB	RB	.

LAYERS IN THE MEANING BANK

- ▶ PTB part of speech tags
- ▶ Animacy classes
- ▶ Named entity types
- ▶ WordNet senses
- ▶ VerbNet roles
- ▶ CCG derivations (syntax)
- ▶ DRT boxes (semantics)
- ▶ SDRT rhetorical relations (discourse)



ANNOTATION LAYERS

Document 28 of 10110, ID: /

size: 3 sentences, 21 tokens
last processed: 21 November 2013, 05:08:49
C&C tools/Boxer revision: 2415

[< first](#) [<< previous](#) [next >>](#) [last >](#) [random](#)

Status: [accepted](#) [history](#)

[report issue](#)

Change to: Comment:

[metadata](#) [raw](#) [tokens](#) [sentences](#) [discourse](#) [1 bit of wisdom](#) [0 warnings](#)

Show: ☒ POS ☐ lemmas ☐ namex ☐ numex ☐ timex ☐ animacy ☐ senses ☒ roles ☐ relations ☐ coreference ☒ syntax

1	<div><div>+</div></div>	<div>We PRP NP</div>	<div>'re VBP (S[dc]\NP)/(S[ad]\NP)</div>	<div>sorry RB [Patient] S[ad]\NP</div>	<div>.</div>						
2	<div><div>+</div></div>	<div>You PRP NP</div>	<div>have VBP (S[dc]\NP)/(S[pt]\NP)</div>	<div>reached VBN [Theme,Agent] (S[pt]\NP)/NP</div>	<div>an DT NP/N</div>	<div>imaginary JJ N/N</div>	<div>number NN N</div>	<div>.</div>			
3	<div><div>+</div></div>	<div>Please VB [Theme,Agent] (S[b]\NP)/(S[b]\NP)</div>	<div>rotate VB [Theme,Agent] (S[b]\NP)/NP</div>	<div>your PRP\$ NP/N</div>	<div>phone NN N/N</div>	<div>90 CD N/N</div>	<div>degrees NNS N</div>	<div>and CC conj</div>	<div>try VB [Experiencer] S[b]\NP</div>	<div>again RB (S[b]\NP)/(S[b]\NP)</div>	<div>.</div>

THE COMPOSITIONAL SEMANTICS LAYER

100

N/N

$\lambda v1. \lambda v2. (\quad ; (v1 @ v2))$

$|v2| = 100$

injured

N/N

$\lambda v1. \lambda v2. (\quad s1 \quad ; (v1 @ v2))$

Topic(s1, v2)
injured(s1)

passengers

N

$\lambda v1. \quad$

passenger(v1)

injured passengers

N

$\lambda v1. \quad$

s1

Topic(s1, v1)
injured(s1)
passenger(v1)

100 injured passengers

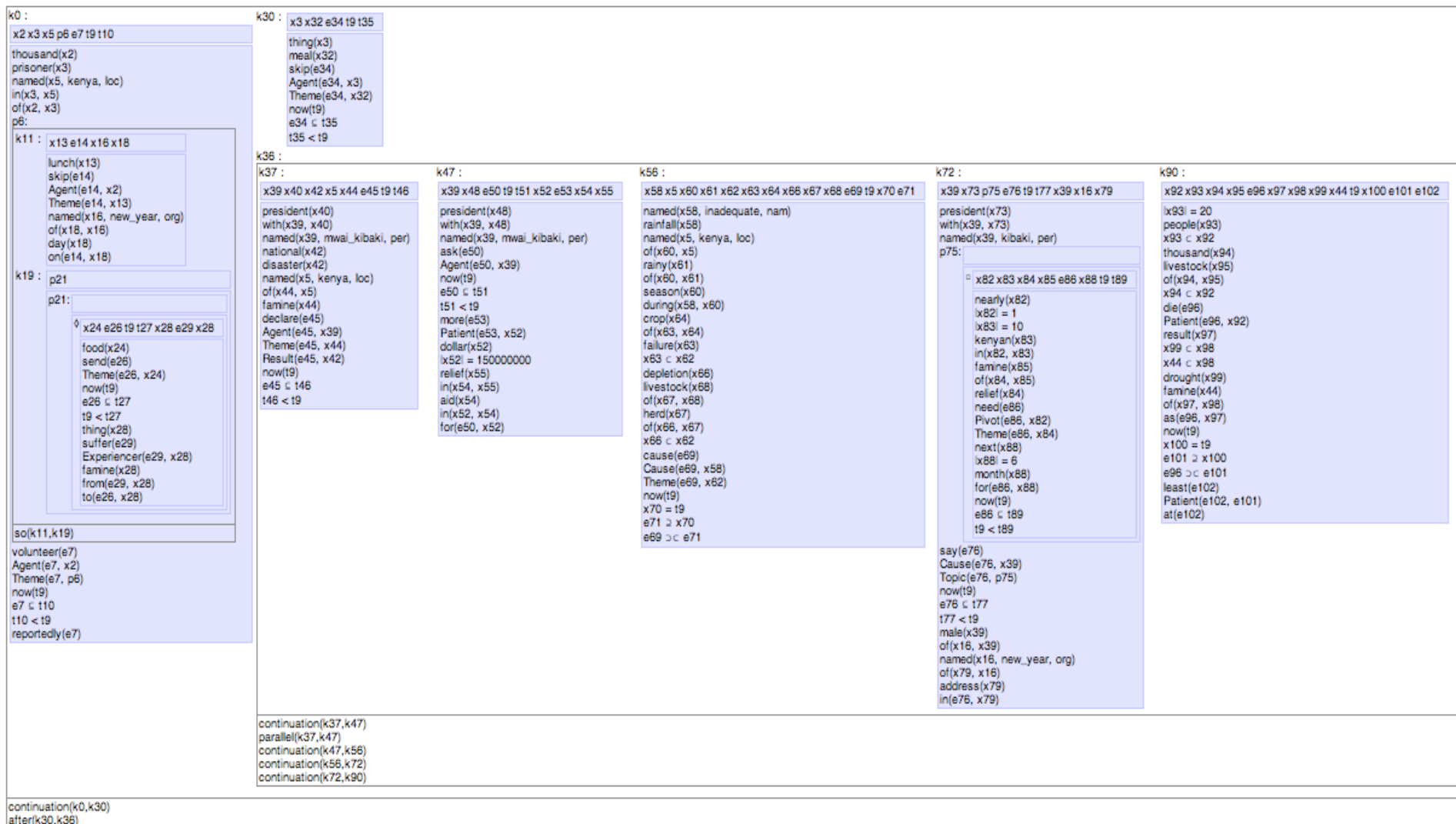
N

$\lambda v1. \quad$

s1

$|v1| = 100$
Topic(s1, v1)
injured(s1)
passenger(v1)

SEGMENTED DISCOURSE REPRESENTATION STRUCTURE (SDRS)

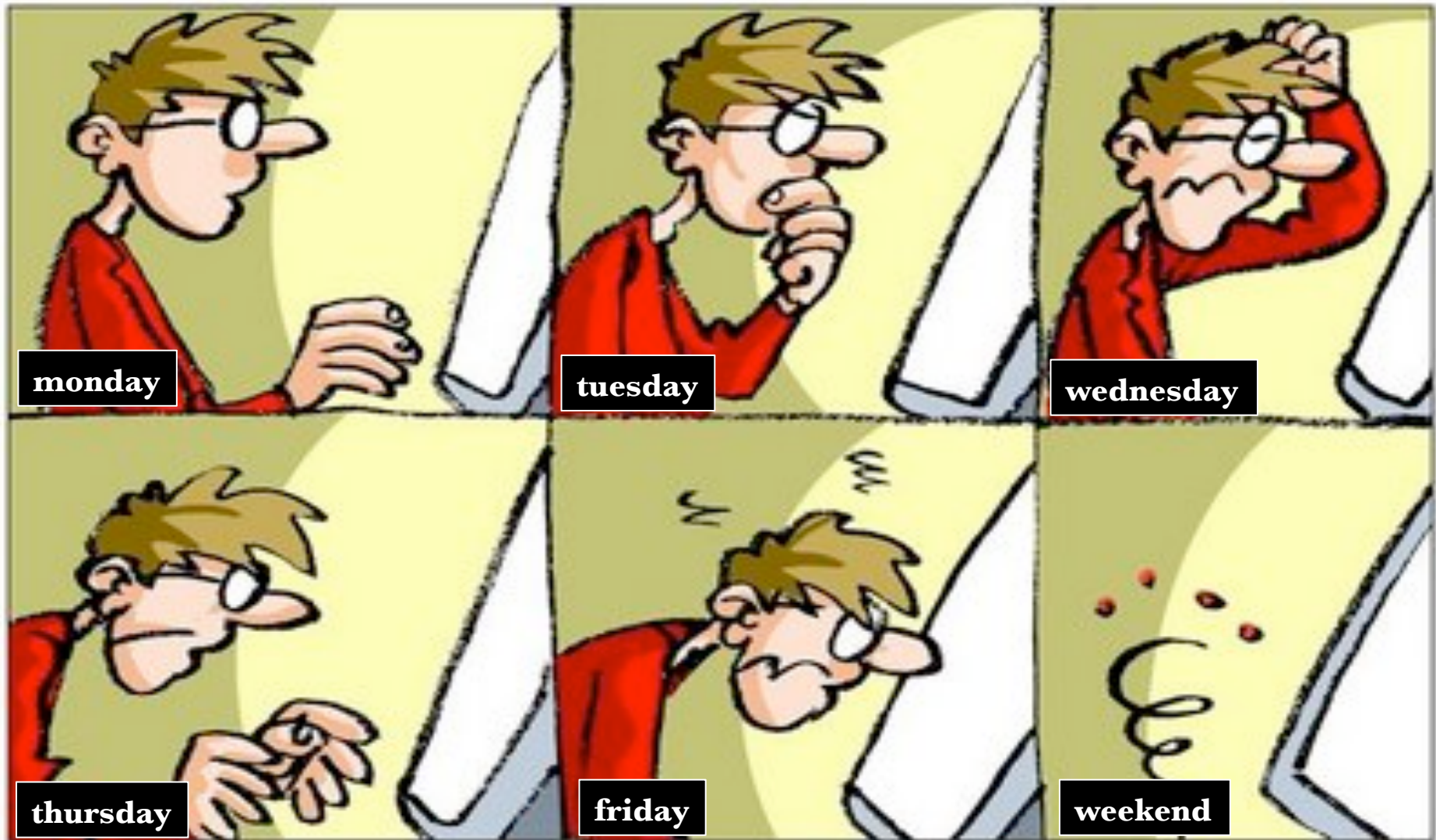




Chapter 5

HELP FROM STRANGERS

TRADITIONAL ANNOTATION



AMAZON'S MECHANICAL TURK

CROWDSOURCING

- ▶ Outsourcing tasks to a distributed group of people
- ▶ The internet provides infrastructure
- ▶ The GMB has two crowds:



Expert Annotators

Edit



GWAP players

?
wordrobe
play what you mean

GWAP (GAME WITH A PURPOSE)



**Play the Games,
Change the Web.**

When you play a game at Gwap,
you aren't just having fun.

[Learn More](#) [Register](#)



Gender Guesser **Prizes** **Best Images**

Matchin
A Question of Taste

Do you prefer babies over
flowers? Score high by reading
your partner's mind and matching
on the same images.

8750 Matchin 3:07

Which image do you prefer?

A  

PLAY

With **Duolingo** you learn a language for free
while helping to translate the web

Enjoy a language game with Wordrobe!

There are many games to choose from!

PLAY TWINS

the other games the other games



Play with words, play Wordrobe!



Become a new Wordrobe wizard!



Fascinated by language? Play!

Synchronize wordrobe with your social network

Facebook sync

Twitter sync

Play Wordrobe *play*

Top scores



PLAY TWINS

Is it a noun or a verb? Easy game for beginners. Double the trouble, twice the fun!



PLAY SENSES

Identify the correct sense of a word. Quite a challenging game. Does it make sense?

1 Aristotle | 2791 points

1 Kilian | 602 points

Choose another game

1		Aristotle	4526 points
2		Kilian	2572 points
3		Nynke	1542 points
4		Valerio	1083 points
5		Leo van Maanen	1024 points
6		Potato	631 points
7		Noortje	581 points
8		MichaelHahn	503 points

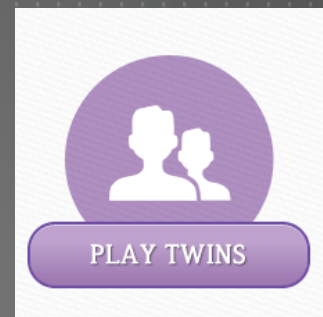
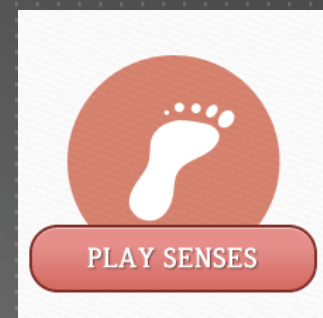
?
wordrobe

play what you mean

www.wordrobe.org

WORDROBE PHILOSOPHY

- ▶ not a single game, but a series of games that share same structure and scoring strategies
- ▶ each linguistic phenomenon that requires annotation corresponds to a different game
- ▶ every game consists of multiple-choice questions
- ▶ each question is presented by a text snippet plus a (small) number of possible answers
- ▶ these questions (and answers) are automatically generated from the corpus



“TWINS” (HOMOGRAPHS)



PLAY TWINS



Twins Questions left until drawer is completed: 2



U.S.-led forces are hunting down remnants of Afghanistan's ousted Taliban regime who frequently carry out hit and run attacks on coalition and Afghan government **forces**, mainly in the country's eastern and southern regions.

- ☒ noun
☐ verb

Place your bet: low



high

answer

skip

“SENSES” (WORDNET)



PLAY SENSES



Senses Questions left until drawer is completed: 1



gameimage

Russian **officials** say at least five more people have died from a wave of extremely cold weather gripping the nation, bringing the death toll to 43 in the past week.

- ☒ a worker who holds or is invested with an office
- ☐ someone who administers the rules of a game or sport

Place your bet: low



high

answer

skip

“POINTERS” (PRONOUNS)



PLAY POINTERS

 Pointers Questions left until drawer is completed: 3



Venezuela's *president* is urging *President Bush* to use **his** second term in office to strengthen ties with Latin American nations.

- ☐ 1) Venezuela
- ☐ 2) president
- ☒ 3) President Bush









Place your bet: low  high

answer

skip



SCORING IN WORDROBE

- ▶ Every answer increases the score of a player
- ▶ The more overlap of a player's answer with other players, the higher the score
- ▶ Total scores for a game are calculated over answers given in the last N days (N=50)

1		Aristotle		4186 points
2		Kilian		2399 points
3		Nynke		1473 points
4		Valerio		1058 points
5		Leo van Maanen		1004 points
6		Potato		618 points
7		Erik		471 points
8		Noortje		466 points

A SCORE WITH A TWIST

- ▶ In addition, Wordrobe players can take risks and bet on the correctness of an answer
- ▶ The higher the bet, the more points you can win (or loose)

 Senses Questions left until drawer is completed: 4 


One Afghan soldier was killed and four others, **including** a U.S. soldier, were injured in the fighting.

☐ have as a part, be made up out of

☒ consider as part of something

☐ add as part of something else – put in as part of a set, group, or category

☐ allow participation in or the right to be part of – permit to exercise the rights, functions, and responsibilities of (synonyms: admit, let in)

Place your bet: low  high

answer

skip

WHY DO PEOPLE PLAY WORDROBE?

- unlocking achievements
- outperforming other players
- learning about language
- help computational linguists

Kilian



[Facebook page](#)

Personal information

Name: Kilian

Completed drawers: 33

Senses: 10

Pointers: 13

Twins: 10





Chapter 6

WELL ON OUR WAY

SEMANTIC PHENOMENA IN THE GMB

Included

- ▶ Scope (negation, quantifiers)
- ▶ Anaphoric pronouns
- ▶ Word senses (WordNet)
- ▶ Thematic Roles (VerbNet)
- ▶ Presuppositions
- ▶ Tense and aspect
- ▶ neo-Davidsonian events
- ▶ Rhetorical relations

Planned

- ▶ Collective vs distributive NPs
- ▶ Reflexive pronouns
- ▶ Verb phrase ellipsis
- ▶ Background knowledge

GMB: CORPUS SIZE

	genre	texts	sentences	words	s/t	w/s
Voice of America	newswire	9,207	57,174	1,238,576	6.2	21.7
CIA world factbook	almanac	514	4,436	112,516	8.6	25.4
Aesop's Fables	narrative	224	949	23,105	4.2	24.3
jokes	humor	122	443	7,531	3.6	17.0
MASC		35	291	6,985	8.3	24.0
RTE		1,338	1,537	29,854	1.1	19.4
		11,440	64,830	1,418,567	5.7	21.9

INTEGRATION OF SEMANTICS

k1 :

x1 e1 t1 t2

named(x1, smith, per)
frown(e1)
Agent(e1, x1)
now(t1)
 $e1 \subseteq t2$
 $t2 < t1$

k2 :

x1

male(x1)

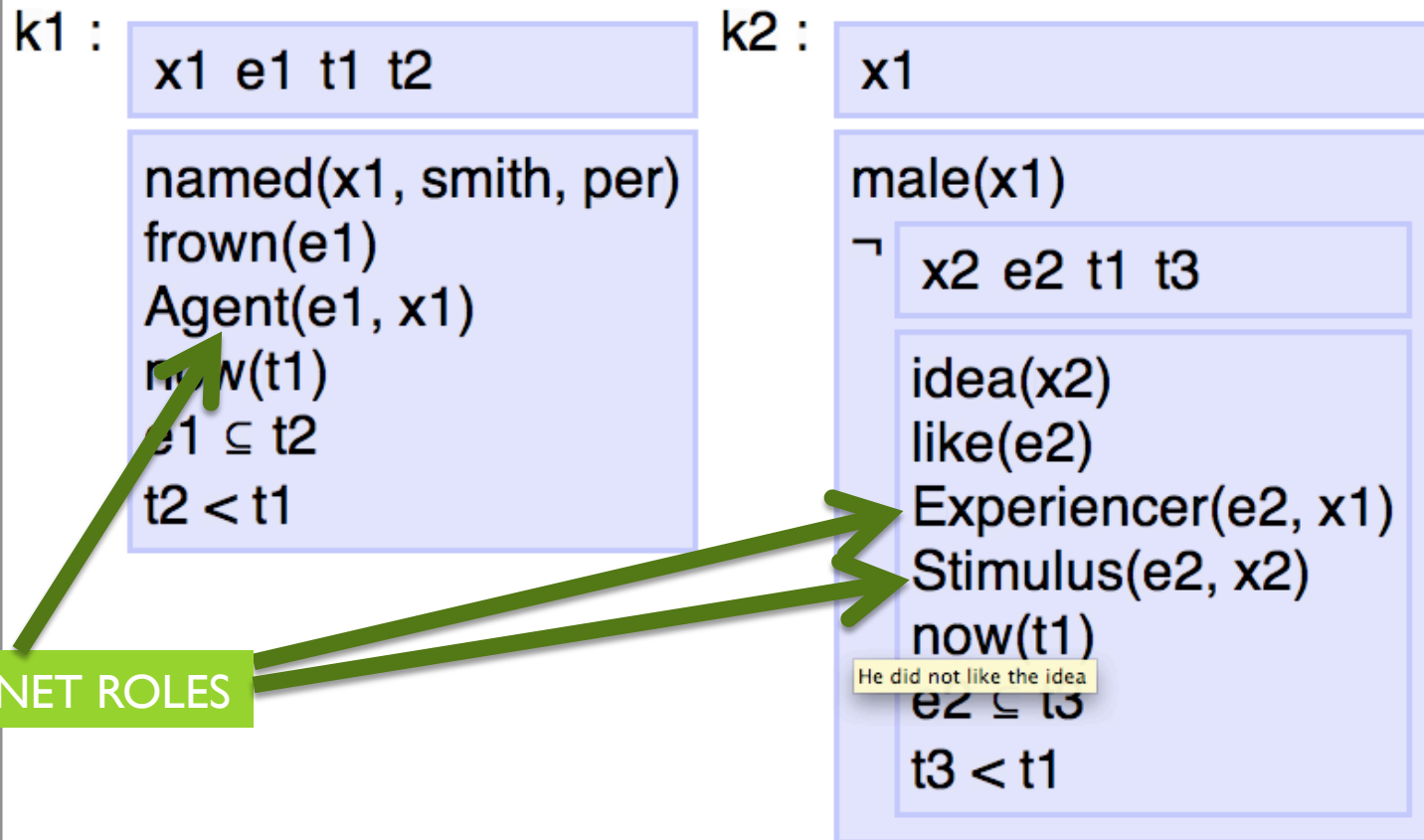
\neg x2 e2 t1 t3

idea(x2)
like(e2)
Experiencer(e2, x1)
Stimulus(e2, x2)
now(t1)
 $e2 \subseteq t3$
 $t3 < t1$

He did not like the idea

continuation(k1,k2)

INTEGRATION OF SEMANTICS



VERBNET ROLES

continuation(k1,k2)

INTEGRATION OF SEMANTICS

k1 :

x1 e1 t1 t2

named(x1, smith, per)
frown(e1)
Agent(e1, x1)
now(t1)
e1 \subseteq t2
t2 < t1

k2 :

x1

male(x1)

\neg x2 e2 t1 t3

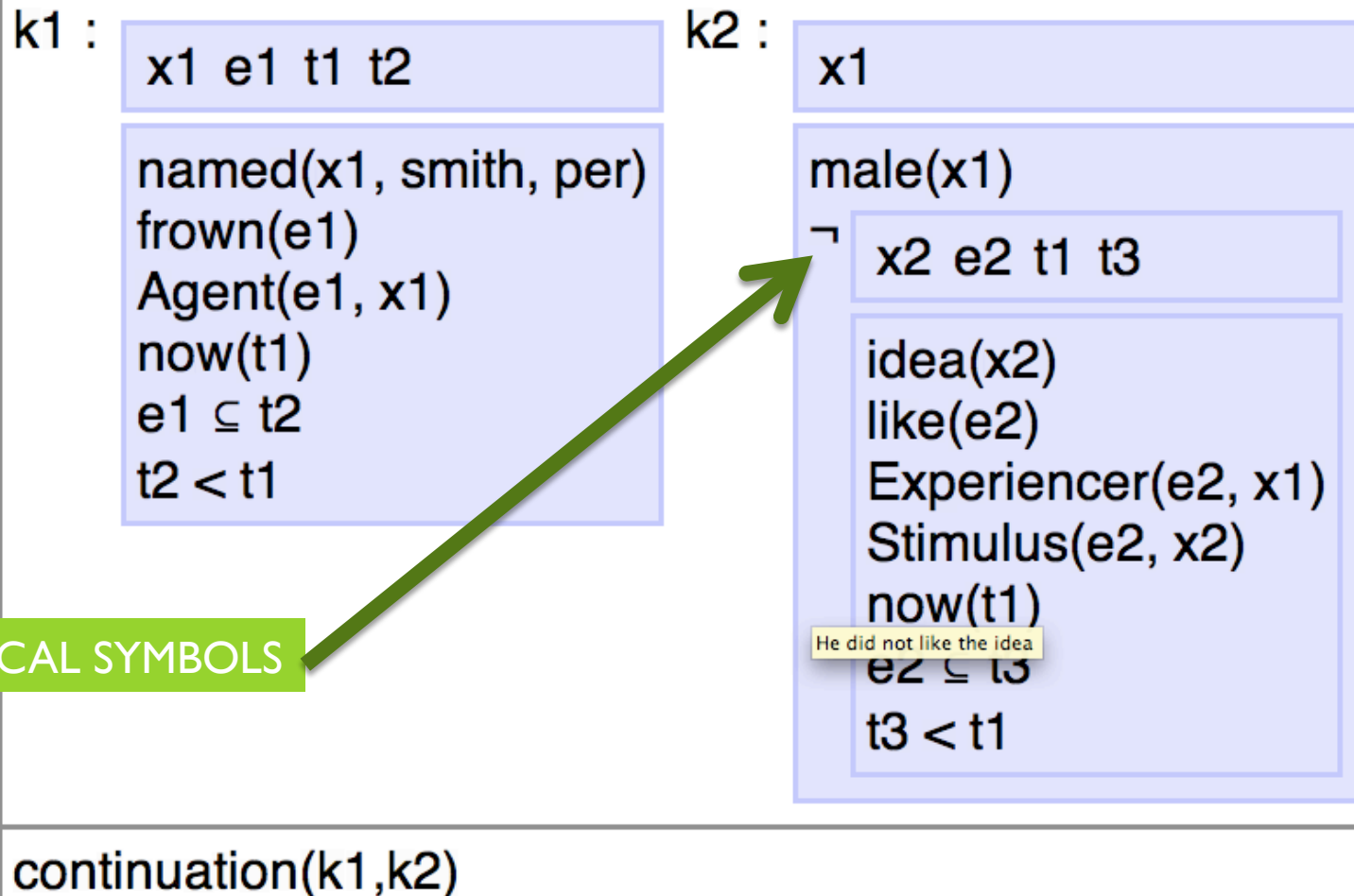
idea(x2)
like(e2)
Experiencer(e2, x1)
Stimulus(e2, x2)
now(t1)
e2 \subseteq t3
t3 < t1

NAMED ENTITIES

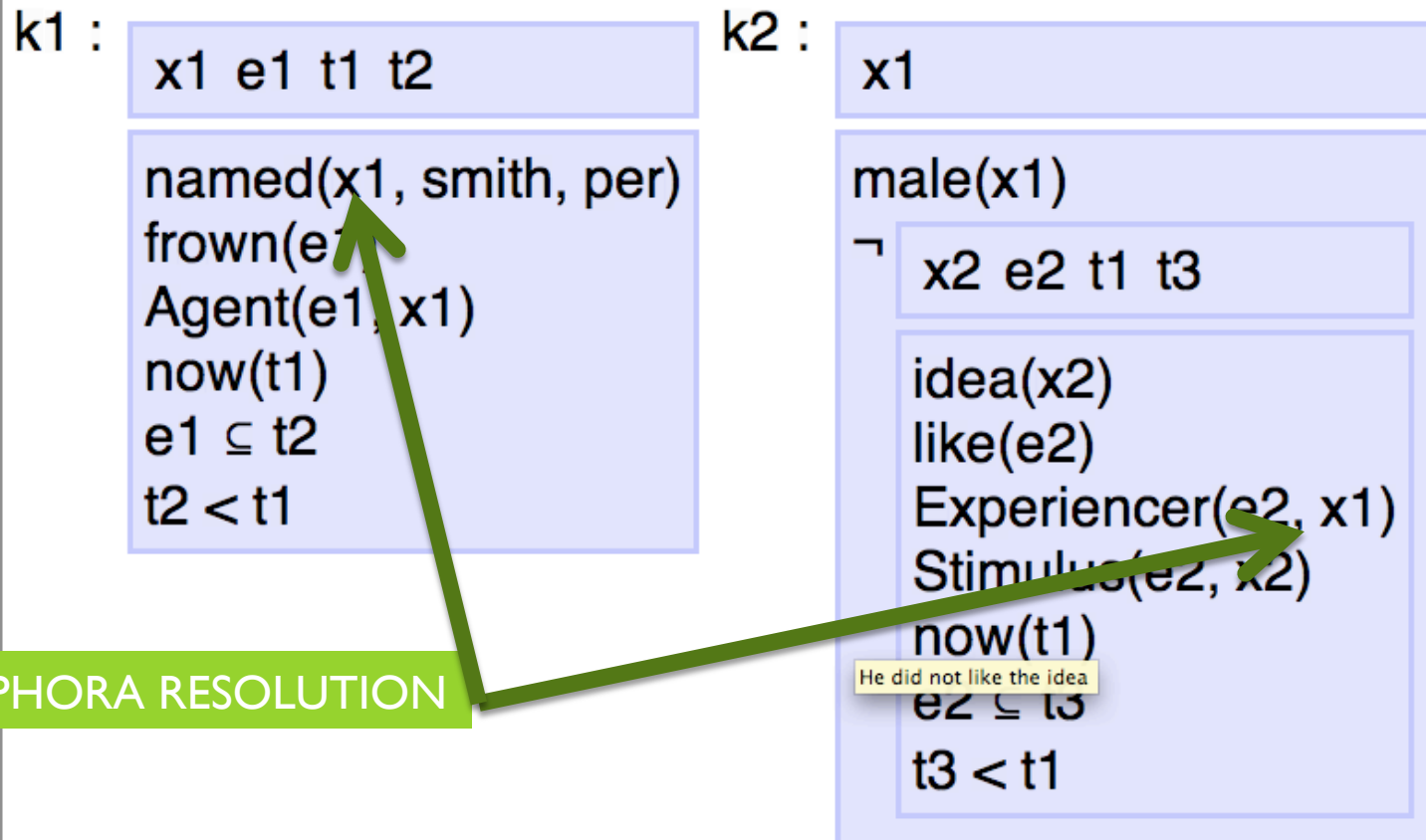
He did not like the idea

continuation(k1,k2)

INTEGRATION OF SEMANTICS



INTEGRATION OF SEMANTICS



ANAPHORA RESOLUTION

He did not like the idea

$\text{continuation}(k1, k2)$

INTEGRATION OF SEMANTICS

k1 :

x1 e1 t1 t2

named(x1, smith, per)
frown(e1)
Agent(e1, x1)
now(t1)
 $e1 \subseteq t2$
 $t2 < t1$

k2 :

x1

male(x1)

\neg x2 e2 t1 t3

idea(x2)
like(e2)
Experiencer(e2, x1)
Stimulus(e2, x2)
now(t1)
 $e2 \subseteq t3$
 $t3 < t1$

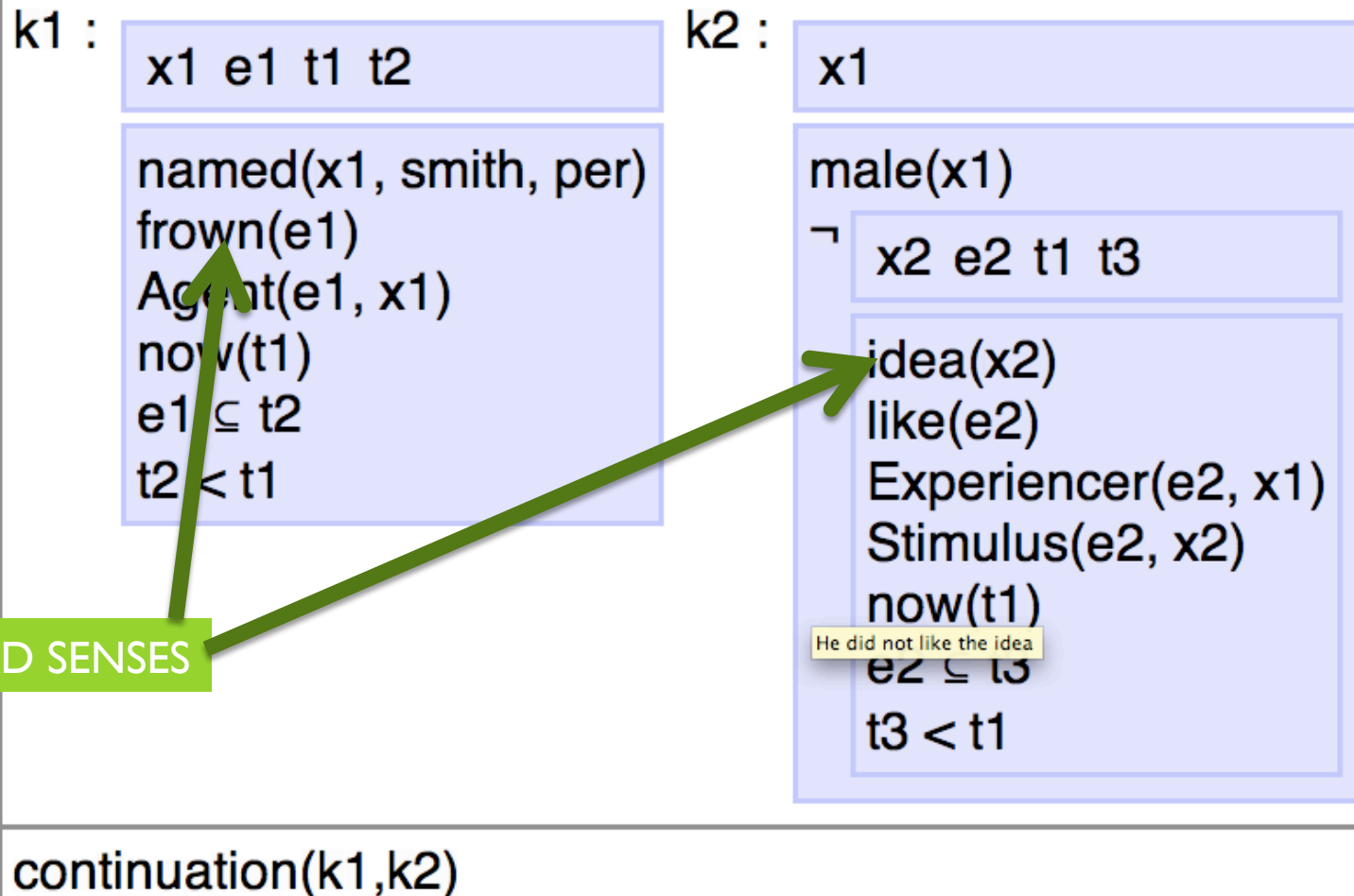
DISCOURSE RELATIONS



continuation(k1,k2)

He did not like the idea

INTEGRATION OF SEMANTICS



INTEGRATION OF SEMANTICS

k1 :

x1 e1 t1 t2

named(x1, smith, per)
frown(e1)
Agent(e1, x1)
now(t1)
 $e1 \subseteq t2$
 $t2 < t1$

k2 :

x1

male(x1)

\neg x2 e2 t1 t3

idea(x2)
like(e2)
Experiencer(e2, x1)
Stimulus(e2, x2)
now(t1)
 $e2 \subseteq t3$
 $t3 < t1$

TENSE & ASPECT

He did not like the idea

continuation(k1,k2)

THE GMB IN NUMBERS

- ▶ 1,285 wordrobe players
- ▶ 62,102 answers (6,982 “bits of wisdom”)
- ▶ 30 expert annotators
- ▶ 50,013 “bits of wisdom”
- ▶ 5 stable releases
- ▶ 233 downloads



Chapter 7

SWEET MEMORIES

FEATURE #1: DECOMPOSITION

34

N/N

$\lambda v0. \lambda v1. (\quad ; (v0 @ v1))$

$|v1| = 34$

Egyptians

N

$\lambda v0. \quad x2$

person(v0)
named(x2, egypt, gpe)
from(v0, x2)

34 Egyptians

N

$\lambda v0. \quad x2$

$|v0| = 34$
person(v0)
named(x2, egypt, gpe)
from(v0, x2)

FEATURE #2: NOUN-NOUN COMPOUNDS

unemployment

N/N

$\lambda v0. \lambda v1. ($ x3 $;$ $(v0 @ v1))$

unemployment(x3)
against(v1, x3)

insurance

N

$\lambda v0.$
insurance(v0)

unemployment insurance

N

$\lambda v0.$ x2
unemployment(x2)
against(v0, x2)
insurance(v0)

FEATURE #3: LONG-DISTANCE DEPENDENCIES

Two	documents	that	a	former	FBI	agent	said	he	took
N/N	N	(NP\NP)/(S[dcI]/NP)	NP/N	N/N	N/N	N	(S[dcI]\NP)/S[dcI]	NP	(S[dcI]\NP)/NP

p4: x9 x10 s11 x12 e13 p14 t16 t17 x18 s19 p20 e21 t23 x24 e25

|x9| = 2
document(x9)
Topic(s11, x10)
former(s11)
named(x12, fbi, org)
of(x10, x12)
agent(x10)
say(e13)
Cause(e13, x10)
Topic(e13, p14)

p14: x10 x27 e28 t30 t31 x33 x34 x35 x37 x38 x40 x42 x43 x44 x45 x46

male(x10)
x9 = x27
take(e28)
Agent(e28, x10)
Theme(e28, x27)



Chapter 8

BUMPS ON THE ROAD

BUMP #1: SEGMENTATION ISSUES

New

NNP

(N/N)/(N/N)

λv0. λv1. λv2. (; ((v0 @ v1) @ v2))

named(v2, new, nam)

York-based

VBN

N/N

λv0. λv1. (

e3

 ; (v0 @ v1))

york-base(e3)

Theme(e3, v1)

organization

NN

N

λv0.

organization(v0)

New York-based

N/N

λv0. λv1. (

e3

 ; (v0 @ v1))

named(v1, new, nam)

york-base(e3)

Theme(e3, v1)

New York-based organization

N

λv0.

e2

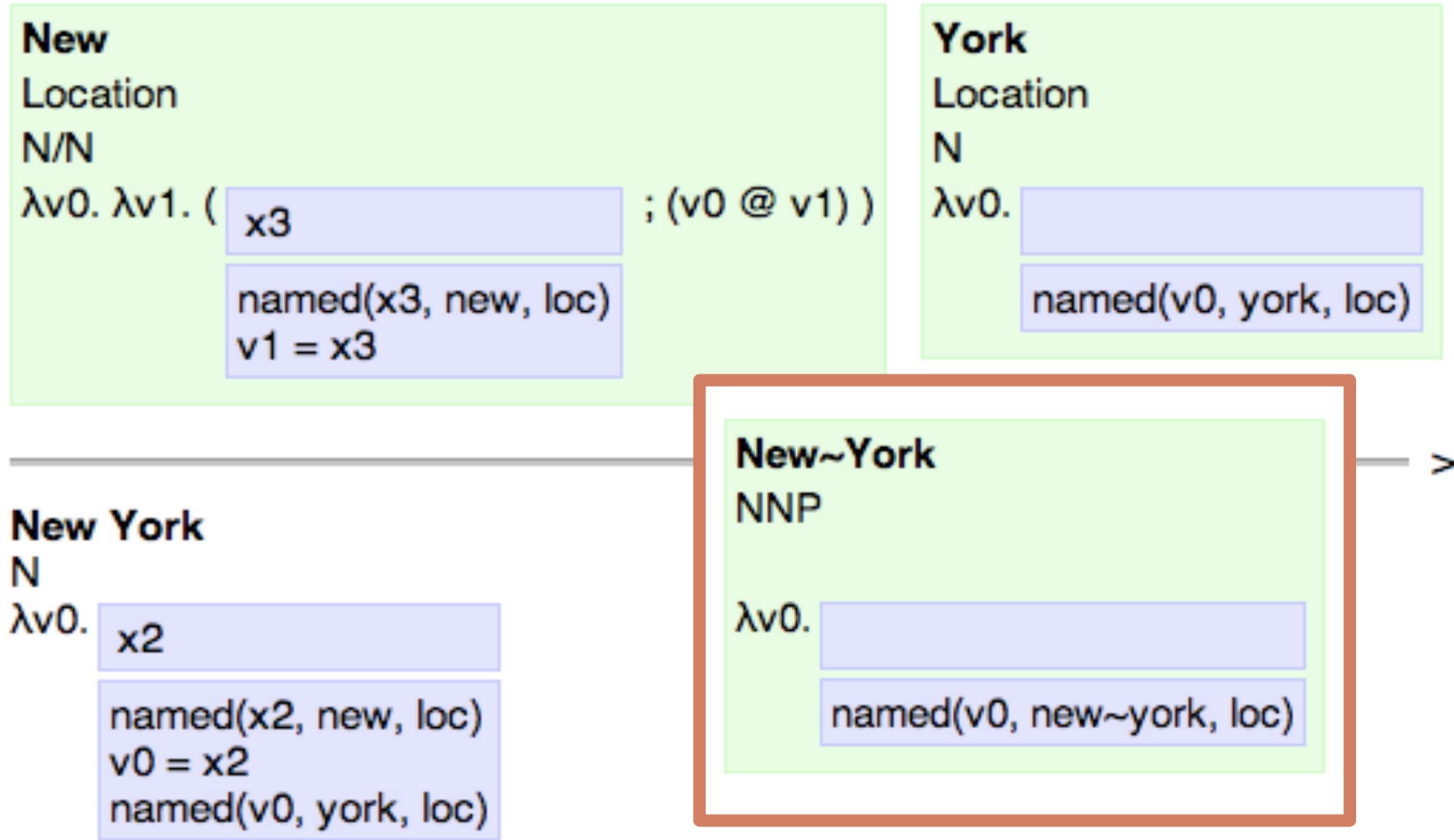
named(v0, new, nam)

york-base(e2)

Theme(e2, v0)

organization(v0)

BUMP #1: SEGMENTATION ISSUES



BUMP #2: COMPOUND CLASS NAMES

great

N/N

$\lambda v0. \lambda v1. ($ s3 $;$ $(v0 @ v1))$

Topic(s3, v1)
great(s3)

whales

N

$\lambda v0.$

whale(v0)

great whales

N

$\lambda v0.$ s2

Topic(s2, v0)
great(s2)
whale(v0)

BUMP #2: COMPOUND CLASS NAMES

common

N/N

$\lambda v0. \lambda v1. ($ s3 $;$ $(v0 @ v1))$

Topic(s3, v1)
common(s3)

tern

N

$\lambda v0.$

tern(v0)

common tern

N

$\lambda v0.$

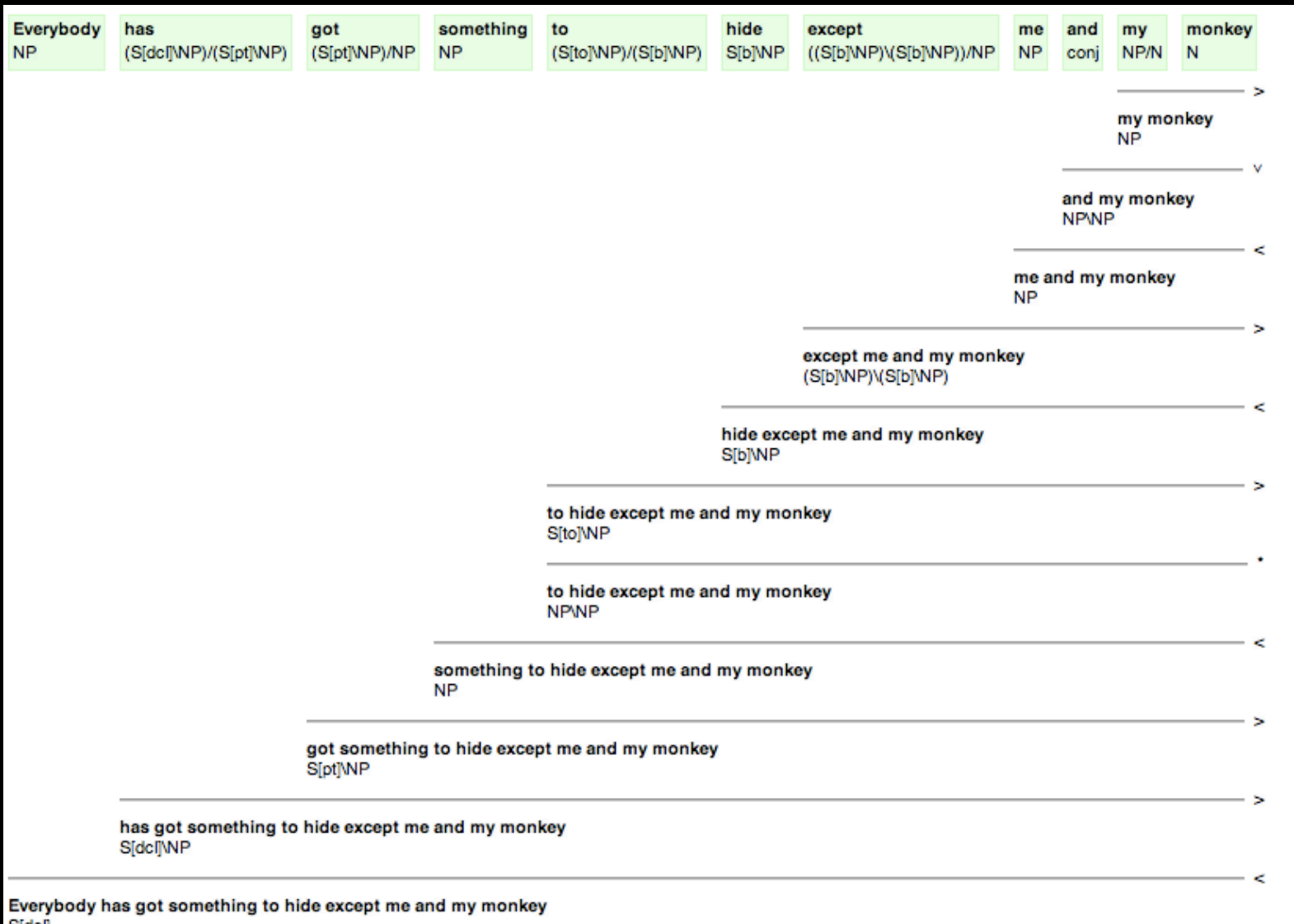
s2

Topic(s2, v0)
common(s2)
tern(v0)

BUMP #3: ANALYSIS OF NAMES OF CREATIVE WORKS

Other N/N	hits N	include (S[dcI]\NP)/NP	Blueberry N/N	Hill N	and conj	Ai N/N	n't N/N	That N/N	a N/N	Shame N
<hr/>			<hr/>			<hr/>				
Other hits			Blueberry Hill						a Shame	
N			N						N	
<hr/>			<hr/>			<hr/>				
Other hits			Blueberry Hill					That a Shame		
NP			NP					N		
							<hr/>			
							n't That a Shame			
							N			
						<hr/>				
						Ai n't That a Shame				
						N				
						<hr/>				
						Ai n't That a Shame				
						NP				
					<hr/>					
					and Ai n't That a Shame					
					NP\NP					
			<hr/>							
			Blueberry Hill and Ai n't That a Shame							
			NP							
									<hr/>	
									include Blueberry Hill and Ai n't That a Shame	
									S[dcI]\NP	
<hr/>										
Other hits include Blueberry Hill and Ai n't That a Shame										
S[dcI]										

BUMP #3: ANALYSIS OF NAMES OF CREATIVE WORKS



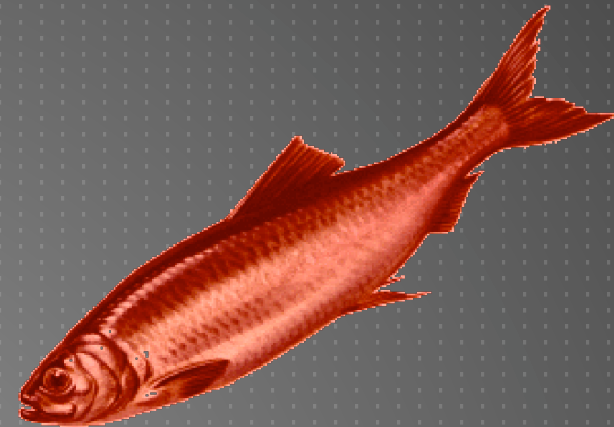
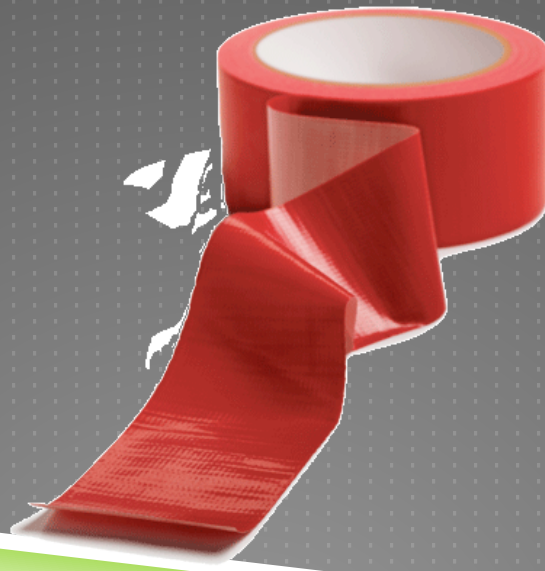
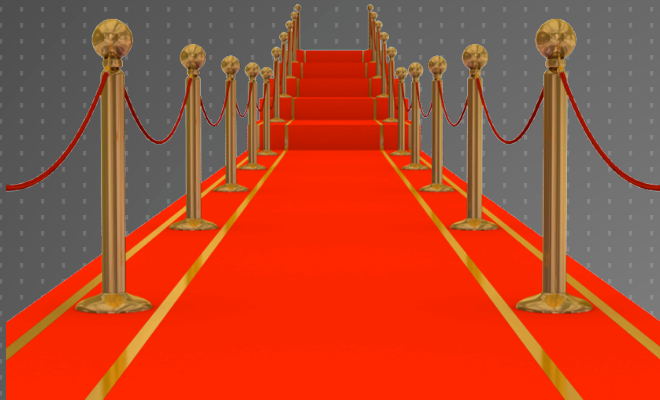
BUMP #4:

INTERNAL STRUCTURE OF NAMED ENTITIES

- ▶ Los Angeles Mayor Antonio Villaraigosa
 - ▶ U.S. Secretary of State Condoleezza Rice
 - ▶ Cobourg Glasgow Rangers Supporters Club
 - ▶ Air Force Brigadier General Thomas Hemingway
 - ▶ Burmese Information Minister Brigadier General Kyaw Hsann
- 

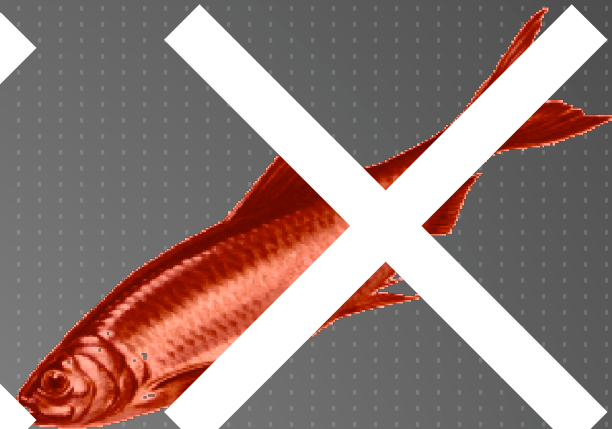
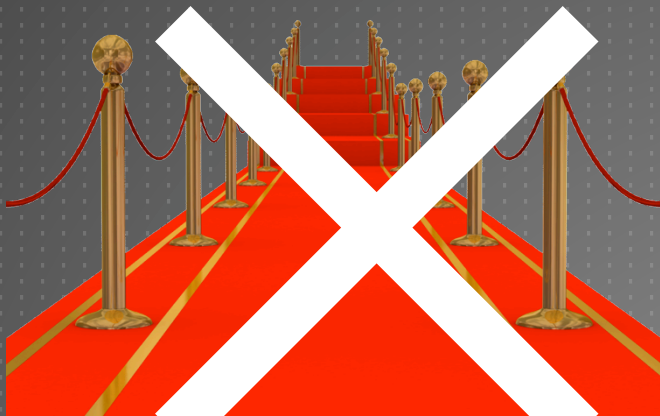
BUMP #5: NON-LITERAL INTERPRETATION

- ▶ Today Argentina gets out the red carpet for the UK Duke of York.
- ▶ The investment climate suffers from red tape and corruption.
- ▶ The plot of the mystery was full of red herrings.



BUMP #5: NON-LITERAL INTERPRETATION

- ▶ Today Argentina gets out the red carpet for the UK Duke of York.
- ▶ The investment climate suffers from red tape and corruption.
- ▶ The plot of the mystery was full of red herrings.






Chapter 9


LOST IN TRANSLATION

PARALLEL MEANING BANKING



- ▶ Input: parallel corpus (text aligned)
- ▶ Idea: translation ought to be meaning-preserving
- ▶ Capture the meaning of a translation by a cross-lingual semantic representation that aligns non-logical symbols of both/all languages
- ▶ Introduce new machinery for dealing with faithful ($=$), loose (\approx), and informative translations (C)

EXAMPLE 1


x p	
$x \mapsto$ “the chance to p”	
CHANCE(x) TO(x,p)	


x p	
$x \mapsto$ “die Gelegenheit zu p”	
GELEGENHEIT(x) ZU(x,p)	





x p	 
CHANCE \equiv GELEGENHEIT (x)	
TO \equiv ZU (x,p)	

EXAMPLE 2


$x \ e \ y$	
$x \mapsto$ "The Magpies" $e \mapsto$ "x have produced y"	
MAGPIES(x) AGENT(e,x) PRODUCE(e) THEME(e,y)	


$x \ e \ y$	
$x \mapsto$ "Die Elstern" $e \mapsto$ "x brachten y hervor"	
ELSTERN(x) AGENT(e,x) HERVORBRINGEN(e) THEME(e,y)	



$x \ e \ y$	 
MAGPIES \equiv ELSTERN (x) AGENT(e,x) PRODUCE \equiv HERVORBRINGEN (e) THEME(e,y)	

EXAMPLE 3

x	
$x \mapsto \text{“player”}$	
PLAYER(x)	

x	
$x \mapsto \text{“Fußballspieler”}$	
FUSSBALLSPIELER(x)	





x	 
FUSSBALLSPIELER \sqsubseteq PLAYER(x)	

The “Magpies” have produced some of Britain’s finest **players**.



Die “Elstern” brachten einige der besten **Fußballspieler** Grossbritanniens hervor.

EXAMPLE 4

e x y 
e \mapsto "taste \dot{x} "
x \mapsto "a pint of \dot{y} "
y \mapsto "beer"
TASTE(e)
THEME(e,x)
PINT(x)
OF(x,y)
BEER(y)

e x y 
e \mapsto " \dot{x} trinken"
x \mapsto "ein Glas \dot{y} "
y \mapsto "Bier"
TRINKEN(e)
THEME(e,x)
GLAS(x)
RELATION(x,y)
BIER(y)



e x y  
TASTE \approx TRINKEN(e)
THEME \equiv THEME(e,x)
PINT \approx GLAS(x)
OF \sqsubset RELATION(x,y)
BEER \equiv BIER(y)



Chapter 10

THE END

CONCLUSION

- The GMB is a public domain corpus with formal semantic annotations for texts (not sentences)
- Anyone can contribute! (experts and street folk)
- Integration of semantic information into one formalism (DRS)
- Everyone can download it (free for research purposes)

<http://gmb.let.rug.nl>

REFERENCES

- ▶ Valerio Basile, Johan Bos, Kilian Evang, Noortje Venhuizen (2012): Developing a large semantically annotated corpus. Proceedings of the Eight International Conference on Language Resources and Evaluation (LREC 2012), pp 3196–3200, Istanbul, Turkey.
- ▶ Noortje Venhuizen, Valerio Basile, Kilian Evang, Johan Bos (2013): Gamification for Word Sense Labeling. Proceedings of the 10th International Conference on Computational Semantics (IWCS 2013), pp 397–403, Potsdam, Germany
- ▶ J. Bos (2014): Semantic Annotation Issues in Parallel Meaning Banking. Proceedings of the Tenth Joint ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-10), pp 17–20, Reykjavik, Iceland.

<http://www.let.rug.nl/bos/publications.html>