



Johan Bos

University of Groningen

www.rug.nl/staff/johan.bos

COMPUTATIONAL SEMANTICS: DAY 5

Meaning Banking



gmb.let.rug.nl: Groningen Meaning Bank

pmb.let.rug.nl: Parallel Meaning Bank



Part 1

SEMANTIC PARSING WITH BOXER

bin/boxer

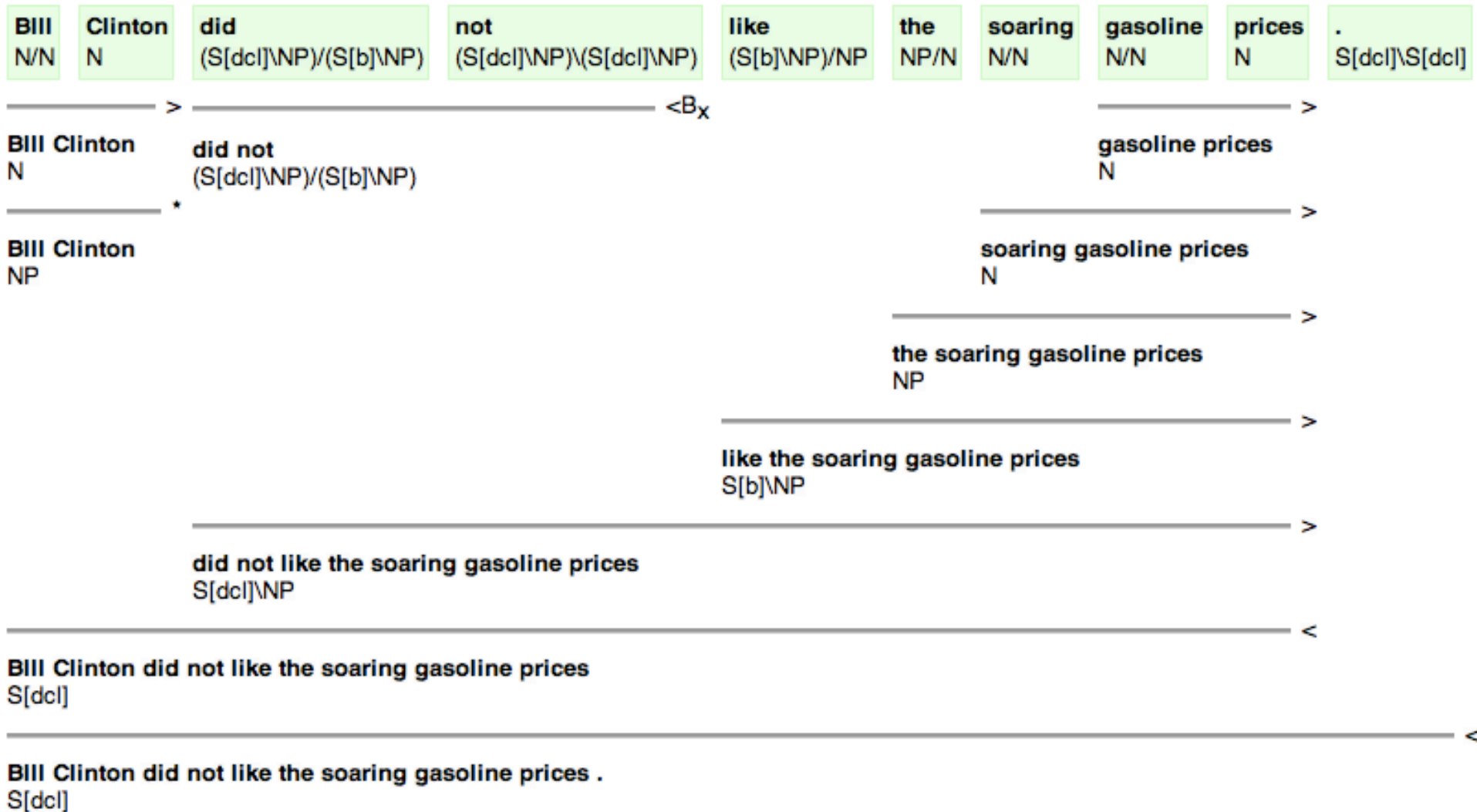
Mr. Johnson was travelling to San Franacie Bay. He went to New York and he smoked.

bin/boxer

Mr. Johnson was travelling to San Franacie Bay. He went to New York and he smoked.

<pre> x1 e1 x2 k1: named(x1, mr.~johnson, per) travel(e1) agent(e1, x1) named(x2, san~franacie~bay, geo) to(e1, x2) _____</pre>	<pre> x1 e2 x3 k2: male(x1) go(e2) agent(e2, x1) named(x3, new~york, geo) to(e2, x3) _____</pre>	<pre> x1 e3 k3: male(x1) smoke(e3) agent(e3, x1) _____</pre>
<pre> continuation(k1, k2) continuation(k2, k3) parallel(k2, k3)</pre>		

CCG Parsing



Compositional Semantics (DRT)

the
NP/N
 $\lambda v1. \lambda v2. ((\boxed{x1} ; (v1 @ x1)) * (v2 @ x1))$

soaring
N/N
 $\lambda v1. \lambda v2. (\boxed{e1} ; (v1 @ v2))$

soar(e1)
Patient(e1, v2)

gasoline
N/N
 $\lambda v1. \lambda v2. (\boxed{x1} ; (v1 @ v2))$

gasoline(x1)
of(v2, x1)

prices
N
 $\lambda v1. \boxed{}$

price(v1)

gasoline prices
N
 $\lambda v1. \boxed{x1}$

gasoline(x1)
of(v1, x1)
price(v1)

soaring gasoline prices
N
 $\lambda v1. \boxed{e1 \ x1}$

soar(e1)
Patient(e1, v1)
gasoline(x1)
of(v1, x1)
price(v1)

the soaring gasoline prices
NP
 $\lambda v1. (\boxed{x1 \ e1 \ x2} * (v1 @ x1))$

soar(e1)
Patient(e1, x1)
gasoline(x2)
of(x1, x2)
price(x1)

Discourse Representation Theory

- **Theory:**
 - Kamp & Reyle, Heim, Asher
 - Van der Sandt, Geurts, Beaver
- **Ingredients:**
 - concepts, events
 - relations, thematic roles
- **Scope:**
 - Boolean and modal operators
 - conjunction implicit
- **Boxes:**
 - can also be named (“hybrid” logic)
 - easier to read than FOL-formulas
 - translation of boxes to (modal) FOL



Hans Kamp

Semantic Parsing with Boxer

Johan Bos

University of Groningen



Interface Formats

--input
syntactic derivation based on
combinatory categorial grammar

--output
discourse representation structure
(DRS), first-order logic (FOL), or
abstract meaning representations
(AMR)

Meaning Frameworks

--theory drt
adopts a version of Kamp's
Discourse Representation Theory

--theory sdrt
outputs meaning representations
based on Asher's Segmented
Discourse Representation Theory

--integrate true
a single meaning representation
spanning all input sentences

Semantic Features

--mwe
dealing with multiword expressions

--nn
processing noun-noun compounds

--resolve
resolving pronouns and names

--roles
proto thematic roles or verbnets

--modal
analysing modal expressions

--tense
additional relations for tense

Works with
C&C as well
as the
EasyCCG
parser!

Ask
for a
demo!

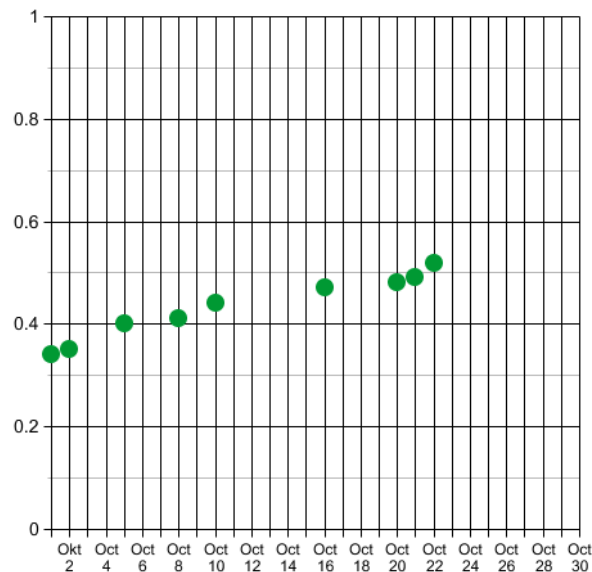
How good is Boxer?

Past results

- Relatedness: 0.83 (3rd/17 SemEval 2014)
- Textual Entailment: 0.82 (5th/18 SemEval 2014)

Work in progress

- Pronoun resolution: 59%
- AMR parsing: $F \approx 0.50$



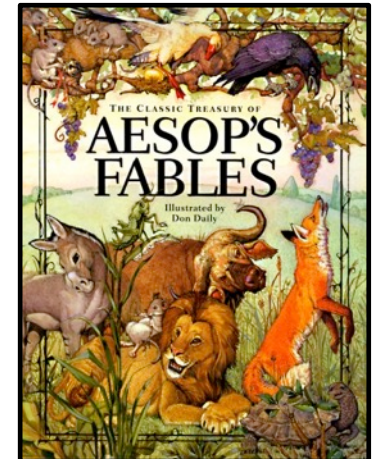
The Groningen Meaning Bank

- Large (English) corpus of public domain texts
- Annotated with meaning representations
 - generated by Boxer
 - corrected by humans (experts and “the crowd”)

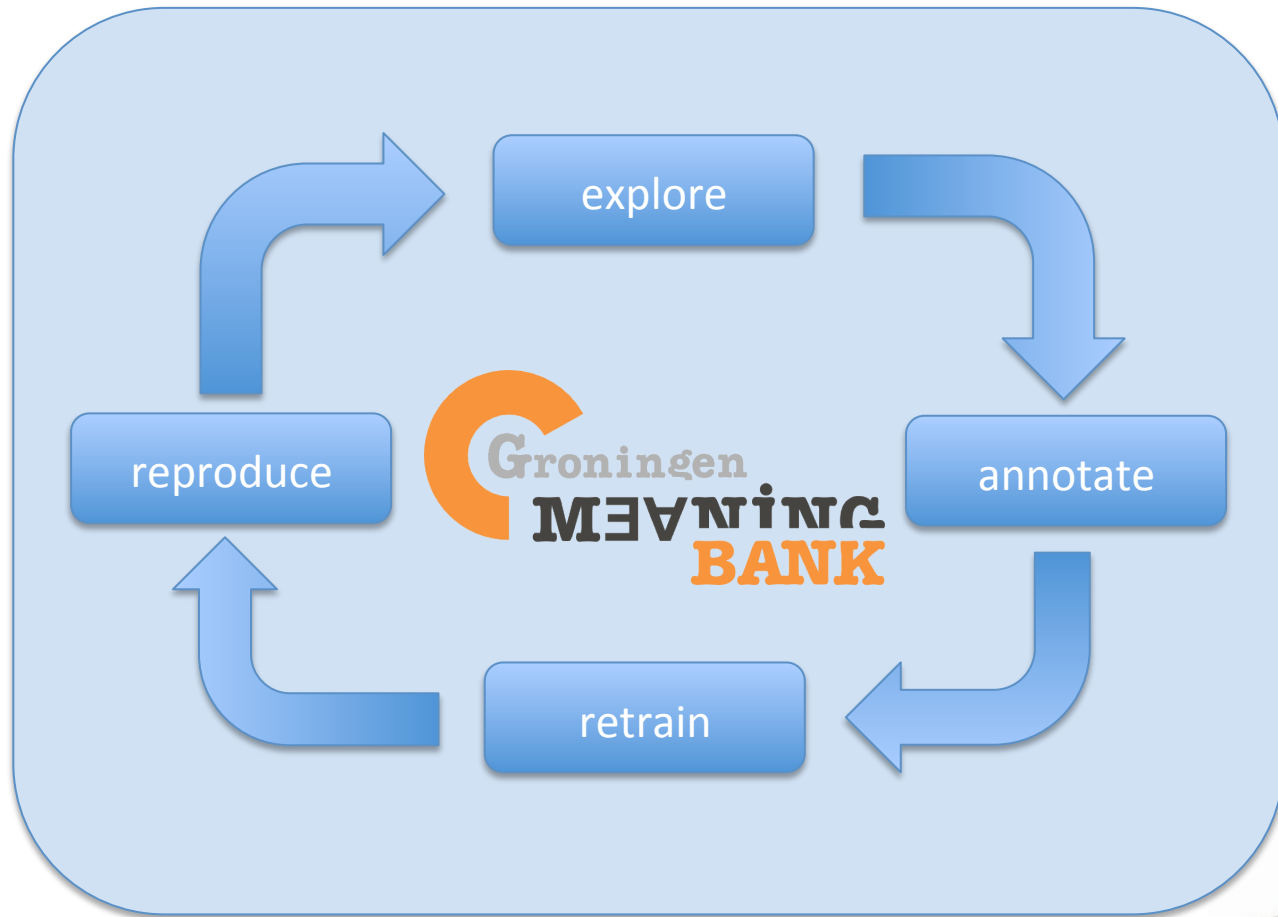


Groningen Meaning Bank: corpus size

	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



Semantic Annotation by Active Learning



The GMB explorer

Document 1 of 10103, ID: /

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

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

Change to: Comment:

size: 1 sentences, 9 tokens

last processed: 04 November 2015, 04:32:59

C&C tools/Boxer revision: 2591

[report issue](#)

[metadata](#) [raw](#) [tokens](#) [sentences](#) [discourse](#) [7 bits of wisdom](#) [0 warnings](#)

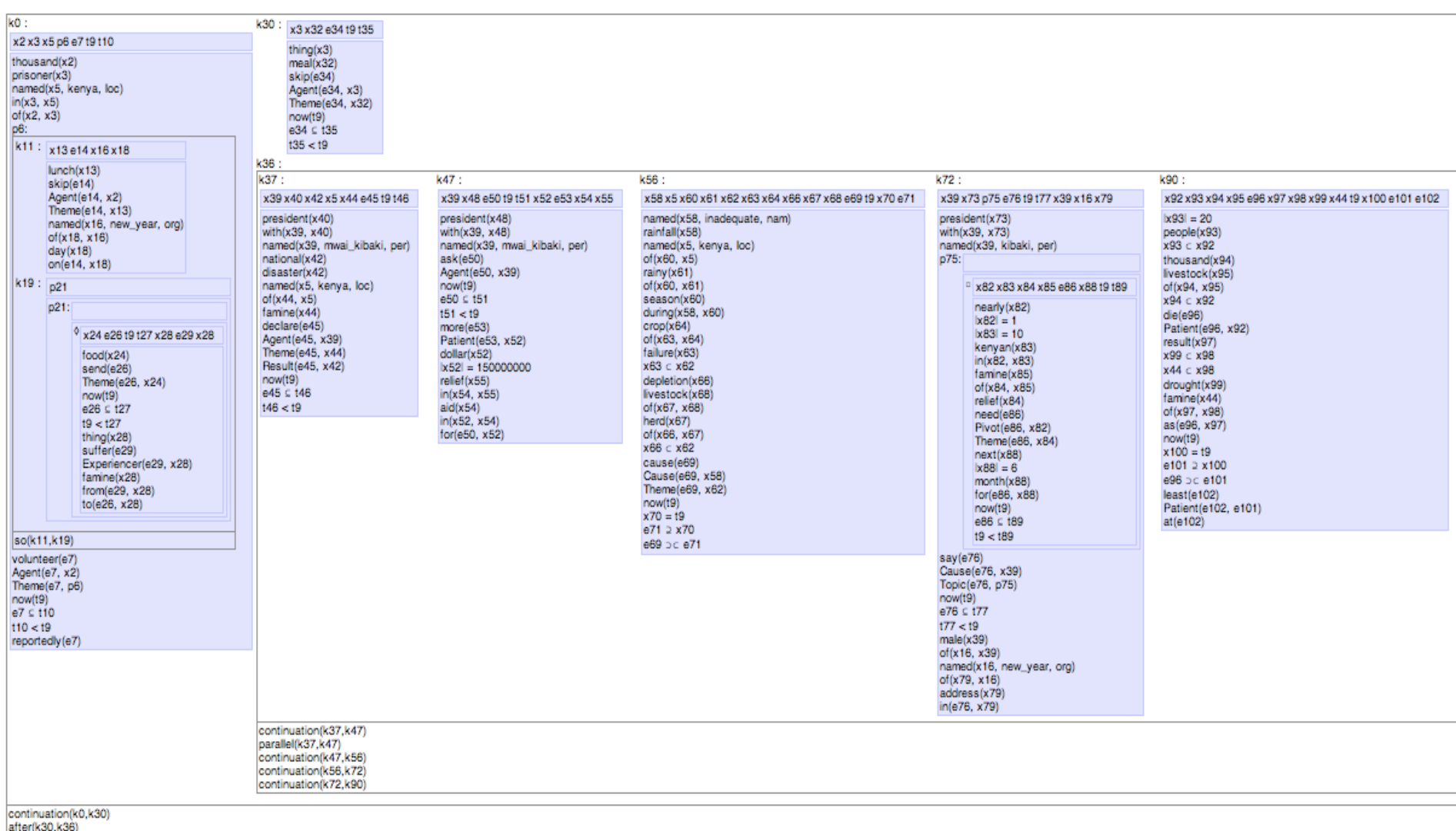
Show: POS lemmas namex animacy senses roles relations scope reference syntax semantics

[Cancel](#)

1 (+)

Officials <input type="text" value="NNS"/> <input type="text" value="official"/> <input type="text" value="O"/> <input type="text" value="1: official functionary"/>	have <input type="text" value="VBP"/> <input type="text" value="have"/> <input type="text" value="O"/> <input checked="" type="text" value="O"/> <ul style="list-style-type: none">1: have have_got hold2: have feature3: experience receive have ...4: own have possess5: get let have	warned <input type="text" value="VBN"/> <input type="text" value="warn"/> <input type="text" value="O"/> <input type="text" value="2: warn discourage admonish ..."/> <input type="text" value="[Recipient,Topic,Agent]"/>	opposition <input type="text" value="NN"/> <input type="text" value="opposition"/> <input type="text" value="O"/> <input type="text" value="1: resistance opposition"/> <input type="text" value="of"/>	activists <input type="text" value="NNS"/> <input type="text" value="activist"/> <input type="text" value="O"/> <input type="text" value="1: militant activist"/>	not <input type="text" value="RB"/> <input type="text" value="not"/> <input type="text" value="O"/> <input type="text" value="1: not non"/>
---	---	--	---	--	--

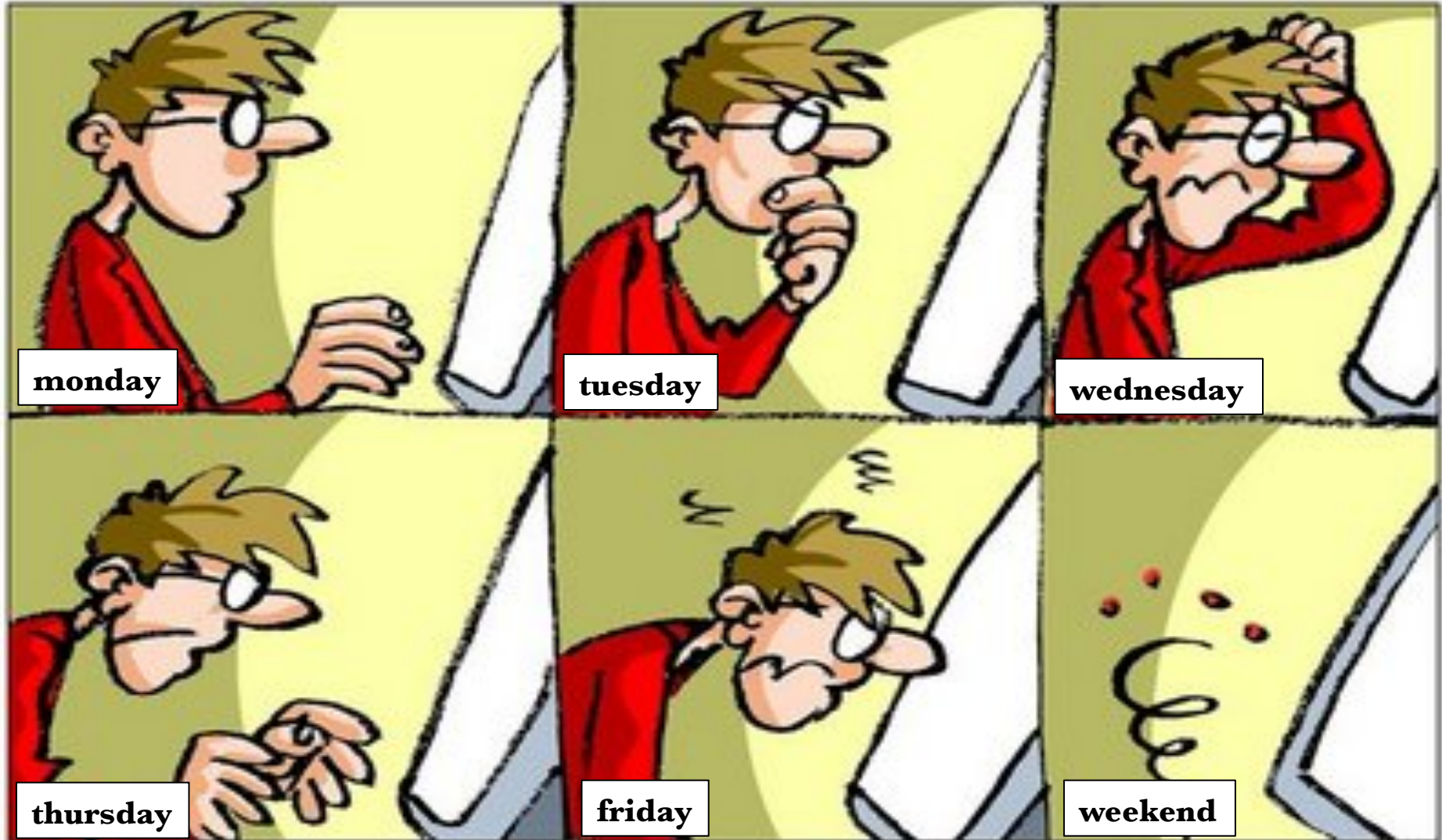
10,000 Discourse Representation Structures



Lots of labelled data needed

- syntactic analysis (pos PTB)
- word senses (wordnet)
- thematic roles (verbnet/lirics)
- antecedents of pronouns
- relations in e.g. compound nouns

Traditional Annotation



Crowdsourcing

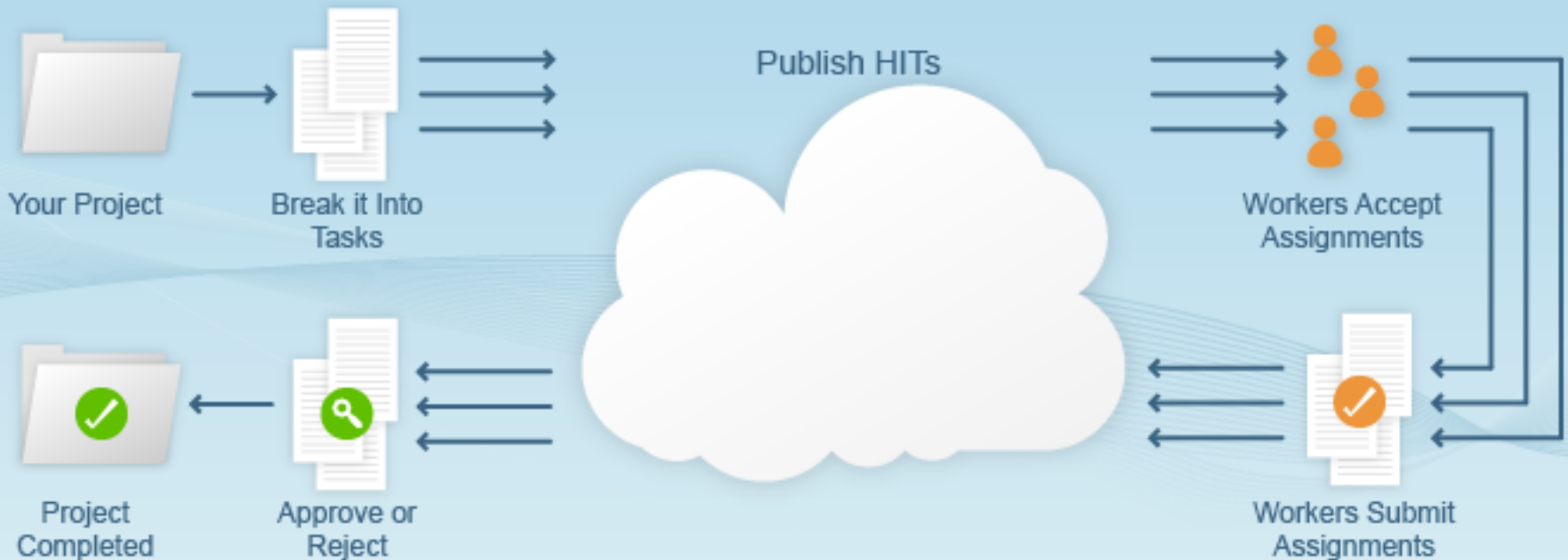
- Outsourcing tasks to a distributed group of people
- The Internet provides the infrastructure
- Two main approaches: MARKETPLACES
GAMIFICATION



Internet Marketplaces



How It Works



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.

Which image do you prefer?

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!

1  Aristotle | 2791 points











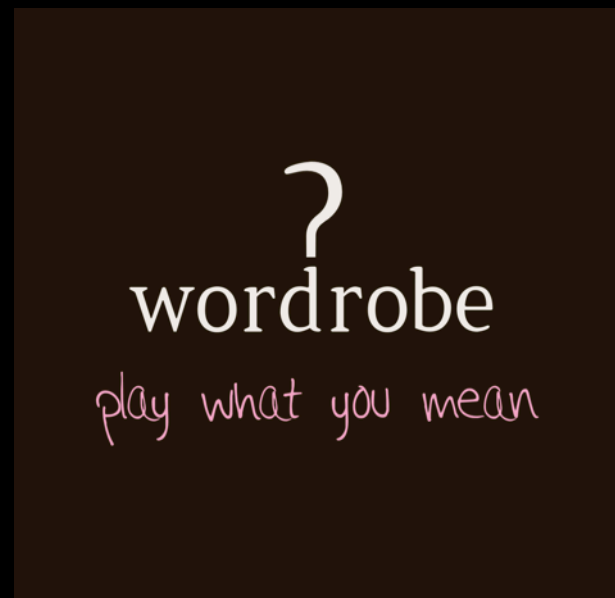
PLAY SENSES

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

1  Kilian | 602 points

Choose another game

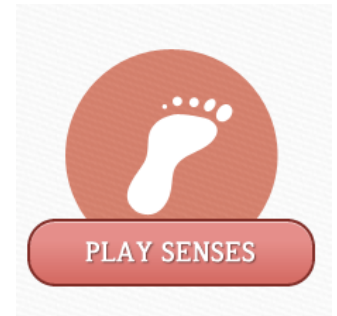
1		Aristotle		4526 points
2		Kilian		2572 points
3		Nynke		1542 points
4		Valerio		1083 points
5		Leo van Maanen		1024 points
6		Potato		631 points
7		Noortje		581 points
8		MichaelHahn		503 points



www.wordrobe.org

Wordrobe Philosophy

- Not a single game, but a series of games that share structure and scoring strategies
- Each semantic 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 Taleban 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” (word ambiguities)



 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” (pronoun interpretation)



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



Parallel Meaning Banking

- Meaning ought to be independent of language: exploit translations!
- Perhaps faster to generate semantic resources for other languages
- Motivation
 - Learn about how human translators work
 - Improve semantic analysis of single languages
 - Verify translations

Meaning Banks

Existing:

- Groningen Meaning Bank (GMB)
- The AMR Bank
- Treebank Semantics (Alistair Butler)

In development:

- Parallel Meaning Bank (PMB)

Preserving Meaning in Translation



Preserving Meaning in Translation



Preserving Meaning in Translation



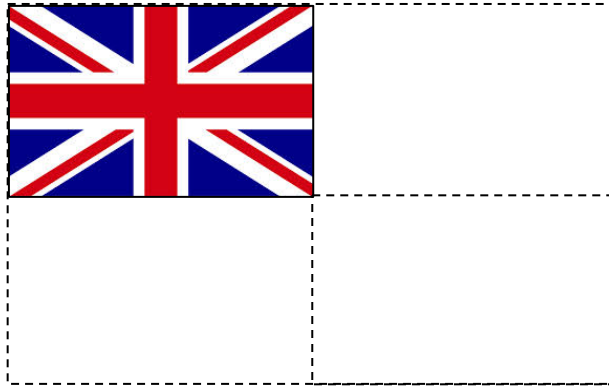
Electrolux

RNB

"Nothing sucks like Electrolux"

An advertisement for Electrolux featuring a yellow and black canister vacuum cleaner. The vacuum is shown from a three-quarter view, highlighting its flexible hose and motor unit. The background is white with a blue vertical bar on the left and an orange vertical bar on the right. The Electrolux logo is at the top left, and the slogan "Nothing sucks like Electrolux" is at the bottom. The code "RNB" is visible near the vacuum's base.

Parallel Meaning Banking



Idea: use parallel corpora (translations) to synchronize meanings

The Parallel Meaning Bank

11,5M word tokens



wiseGEEK



INTERSECT

qtleap

CRPUS



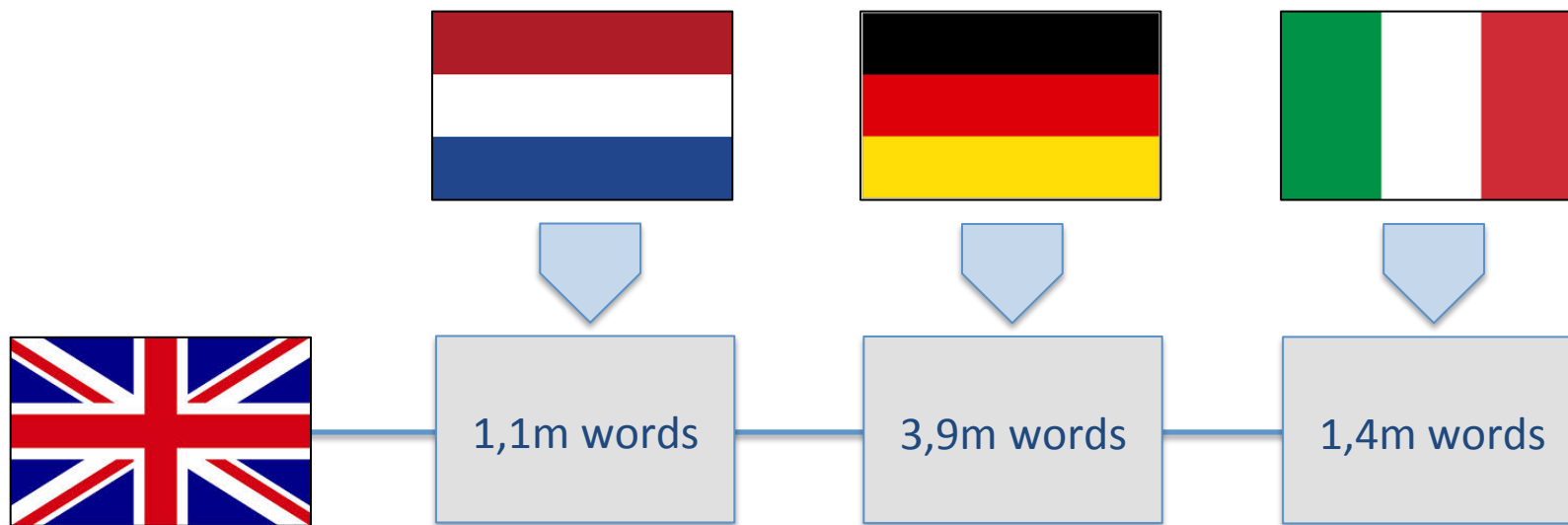
LONWEB
VOLUNTEERS
program

WWW.LONWEB.ORG

 **PASCAL2**
Pattern Analysis, Statistical Modelling and
Computational Learning

QA@CLEF-2004

TED



The Parallel Meaning Bank

English as pivot language (5 million words)

(ca. 10,000 documents for all four languages)

What will be different in the PMB?

GMB

- one language
- POS-tagging
- C&C parser
- Wordnet senses
- neo-Davidsonian events
- lexical rules
- gamification




PMB


- four languages
- Semantic tagging
- EasyCCG parser
- Corpus-driven senses
- hyper-Davidsonian events
- empty elements
- crowd-sourcing





SEMANTIC ALIGNMENT

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)

Human Translators

Explication

English: The “Magpies”, Newcastle United Football Club, have ...

German: Die “Elstern”, **wie der** Newcastle United Football Club **auch genannt wird**, brachten ...

Hyperonym – Hyponym

English: ... have produced some of Britain’s finest **players**.

German: ... brachten einige der besten **Fußballspieler** Großbritanniens hervor.

Co-Hyponym

English: ... the chance to **taste a pint** of beer and have a chat with the locals

German: ... die Gelegenheit **ein Glas** zu **trinken** und mit den Einheimischen zu plaudern.

Numerical Expression [38]

English: That man is **not above** forty. (*e.g.* ≤ 40)

Dutch: Die man is **nog geen** veertig. (*e.g.* < 40)

Simile

English ... passing through the ranks of the Ostyak (...) **like a scythe through standing grain**.

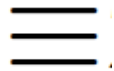
German ... herüberwanderten und Otjaken (...) **buchstäblich niedermähten**.

Anaphoric Expression

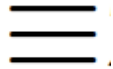
English: Construction of the first floor (...) began on August 9, 1173. **This first floor** is ...

German: Der Bau der ersten Etage (...) begann am 9. August 1173. **Diese Etage** ist ...


THE GOOD, THE BAD, AND THE UGLY




THE GOOD, THE BAD, AND THE UGLY





EXAMPLE 2

$x e y$	
$x \mapsto$ “The Magpies”	
$e \mapsto$ “ \dot{x} have produced \dot{y} ”	
MAGPIES(x)	
AGENT(e, x)	
PRODUCE(e)	
THEME(e, y)	

$x e y$	
$x \mapsto$ “Die Elstern”	
$e \mapsto$ “ \dot{x} brachten \dot{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)	

THE GOOD, **THE BAD,** AND THE UGLY



EXAMPLE 3

X	
X \mapsto “player”	
PLAYER(x)	

X	
X \mapsto “Fußballspieler”	
FUSSBALLSPIELER(x)	



X	 
FUSSBALLSPIELER \sqsubset PLAYER(x)	


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


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

THE GOOD, THE BAD, **AND THE UGLY**



≈

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)

Lexical Relations in Parallel Meanings

- Synonym: ≡
- Hyponym: □
- Co-Hyponym: ≈

- Meronym: <<

possible, but not seen in EN/DE/IT/NL data but seen in English-Japanese data (Francis Bond)

Category Transfer Theory

- **Framework:**
Combinatory Categorical Grammar
- **Given:**
An expression $S (s_1 \dots s_i)$ and its translation $T (t_1 \dots t_j)$ with CCG-derivations $d(S)$ and $d(T)$.
- **Then:**
The translation from S to T is meaning-preserving iff $d(S)$ can be transferred to $d(T)$ closed under CCG's combinatorial rules

Bootstrapping Semantic Parsers

Step 1: Parse source sentence



the	chance	to	taste	a	pint	of	beer		
np/n	n/vp _{to}	vp _{to} /vp _b	vp _b /np	np/n	n/pp	pp/np	np		
	np/vp _{to}	vp _{to} /vp _b	vp _b /np	np/n		pp			
	np/vp _{to}	vp _{to} /vp _b	vp _b /np	np/n	n				
	np/vp _{to}	vp _{to} /vp _b	vp _b /np	np					
	np/vp _{to}	vp _{to} /vp _b	vp _b						
	np/vp _{to}	vp _{to}							
np									

Bootstrapping Semantic Parsers

Step 2: Use word alignments



the	chance	to	taste	a	pint	of	beer		
np/n	n/vp _{to}	vp _{to} /vp _b	vp _b /np	np/n	n/pp	pp/np	np		
die	Gelegenheit			ein	Glas		Bier	zu	trinken



Bootstrapping Semantic Parsers

Step 3: Carry over syntactic categories where possible



the	chance	to	taste	a	pint	of	beer		
np/n	n/vp _{to}	vp _{to} /vp _b	vp _b /np	np/n	n/pp	pp/np	np		
np/n	n/vp _{to}			np/n			np	vp _{to} /vp _b	vp _b /np
die	Gelegenheit			ein	Glas		Bier	zu	trinken



Bootstrapping Semantic Parsers

Step 4: Reverse slashes where needed to match word order



the	chance	to	taste	a	pint	of	beer		
np/n	n/vp _{to}	vp _{to} /vp _b	vp _b /np	np/n	n/pp	pp/np	np		
np/n	n/vp _{to}			np/n			np	vp _{to} /vp _b	vp _b \np
die	Gelegenheit			ein	Glas		Bier	zu	trinken



Bootstrapping Semantic Parsers

Step 5: Infer new categories using CCG's combinators

Here: $n/pp + pp/np \rightarrow n/np$ [$>B$]



the	chance	to	taste	a	pint	of	beer		
np/n	n/vp _{to}	vp _{to} /vp _b	vp _b /np	np/n	n/pp	pp/np	np		
np/n	n/vp _{to}			np/n	n/np		np	vp _{to} /vp _b	vp _b \np
die	Gelegenheit			ein	Glas		Bier	zu	trinken



Bootstrapping Semantic Parsers

np									
np/vp _{to}				vp _{to}					
np/vp _{to}				np			vp _{to} \np		
np/vp _{to}				np/n	n			vp _{to} \np	
np/n	n/vp _{to}			np/n	n/np		np	vp _{to} /vp _b	vp _b \np
die	Gelegenheit			ein	Glas		Bier	zu	trinken



Aligning Meanings



pint :: n/pp	of :: pp/np
$\lambda p.\lambda x.[PINT(x) \ \& \ p(x)]$	$\lambda n.\lambda z.n(\lambda y.CONTAINS(z,y))$
pint of :: n/np	
$\lambda n.\lambda x.[PINT(x) \ \& \ n(\lambda y.CONTAINS(x,y))]$	
Glas :: n/np	
$\lambda n.\lambda x.[GLAS \approx PINT(x) \ \& \ n(\lambda y.CONTAINS(x,y))]$	

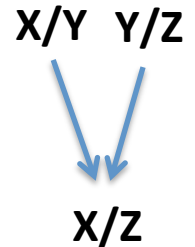


Copy, Merge & Split

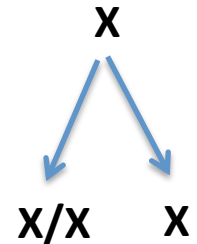
- **Copy:**
transfer of category from source to target



- **Merge:**
two source categories merge into one target category (composition)



- **Split:**
one source category into two target categories (de-composition)



Merge

N/N: **dog** N: **house**

----->

N: **hondelok**

(S\NP)/(S_{to}\NP): **likes** (S_{to}\NP)/(S\NP):**to**

----->B

(S\NP)/(S\NP): **graag**

Split

$S_{\text{adj}} \setminus \text{NP}$: **impossible**

B>

 $(S \setminus \text{NP}) / (S \setminus \text{NP})$: **niet**

$S_{\text{adj}} \setminus \text{NP}$: **mogelijk**

Boxer Learning Dutch

[ze]
 pronpers
 NP
 λv0. (x2 * (v0 @ x2))
thing(x2)

[leest]
 verbpressg
 S/NP
 λv0. λv1. (v0 @ λv2. (x4 e6 ; (v1 @ e6)))
neuter(x4)
read(e6)
agent(e6, v2)
patient(e6, x4)

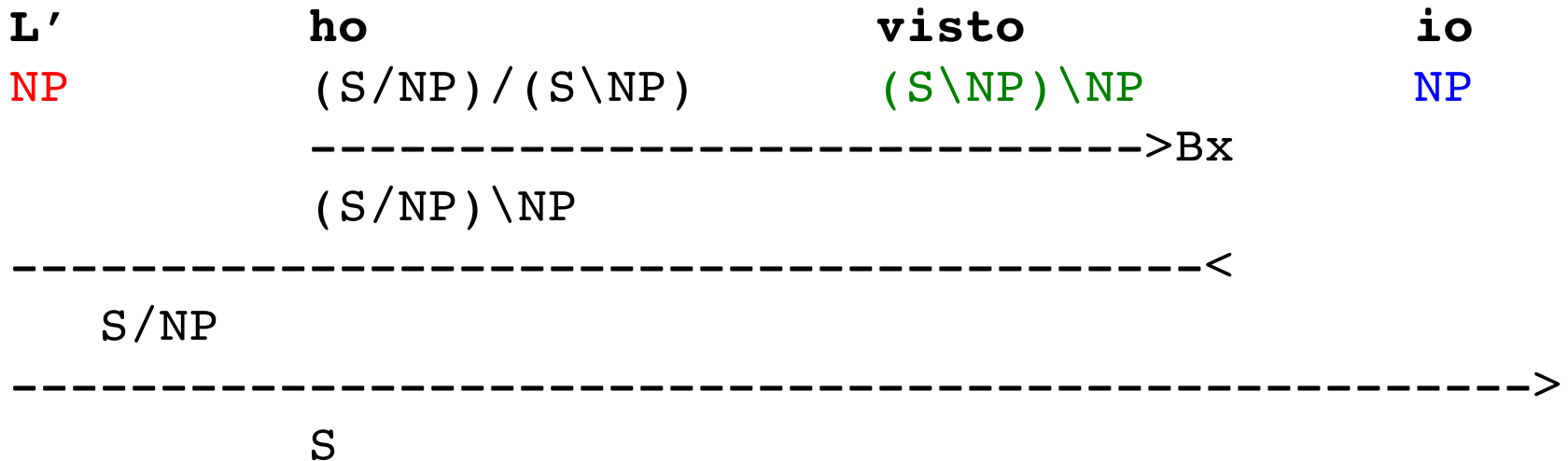
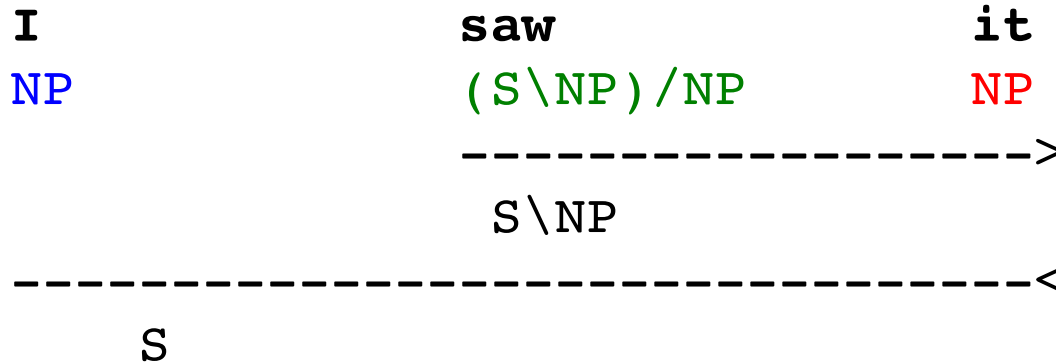
[graag]
 adj
 (S/NP)/(S/NP)
 λv0. λv1. λv2. (v1 @ λv3. (p5 e6 ; (v2 @ e6)))
p5: ((v0 @ λv7. (v7 @ v3)) @ λv8.)
wish(e6)
agent(e6, v3)
theme(e6, p5)

S/NP
 λv0. λv1. (v0 @ λv2. (p4 e5 ; (v1 @ e5)))
p4: x7 e9
neuter(x7)
read(e9)
agent(e9, v2)
patient(e9, x7)
wish(e5)
agent(e5, v2)
theme(e5, p4)

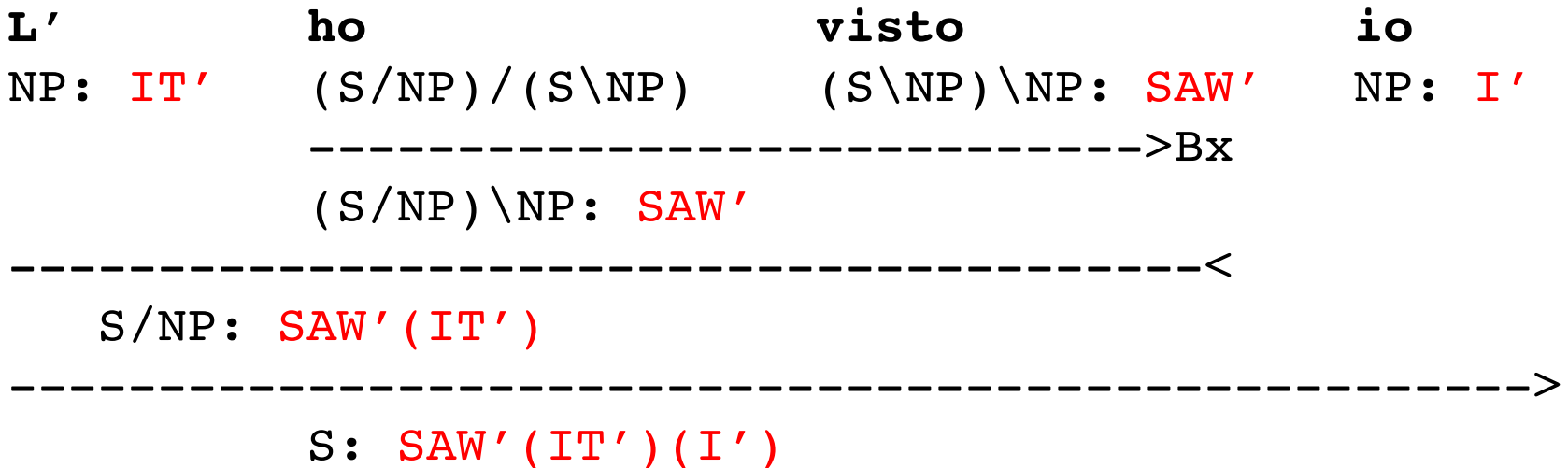
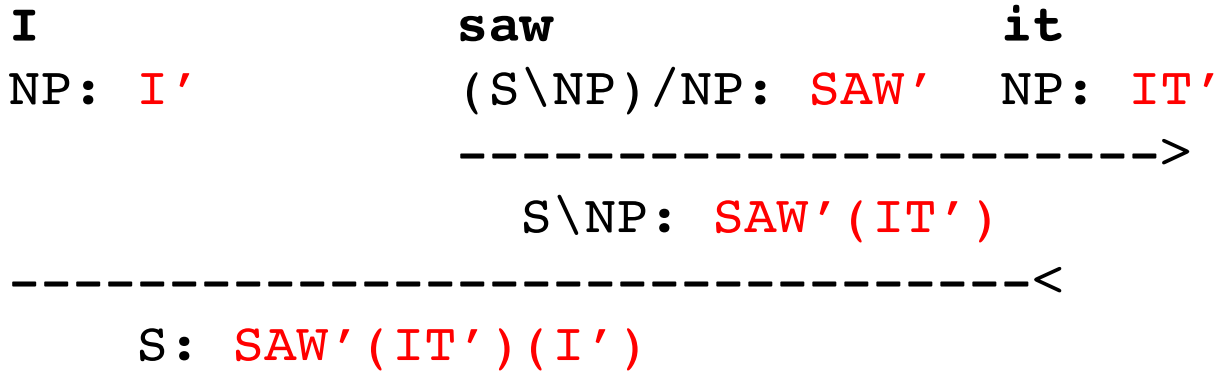
S
 λv0. (x2 p4 e5 ; (v0 @ e5))
thing(x2)
p4: x7 e9
neuter(x7)
read(e9)
agent(e9, x2)
patient(e9, x7)
wish(e5)
agent(e5, x2)
theme(e5, p4)



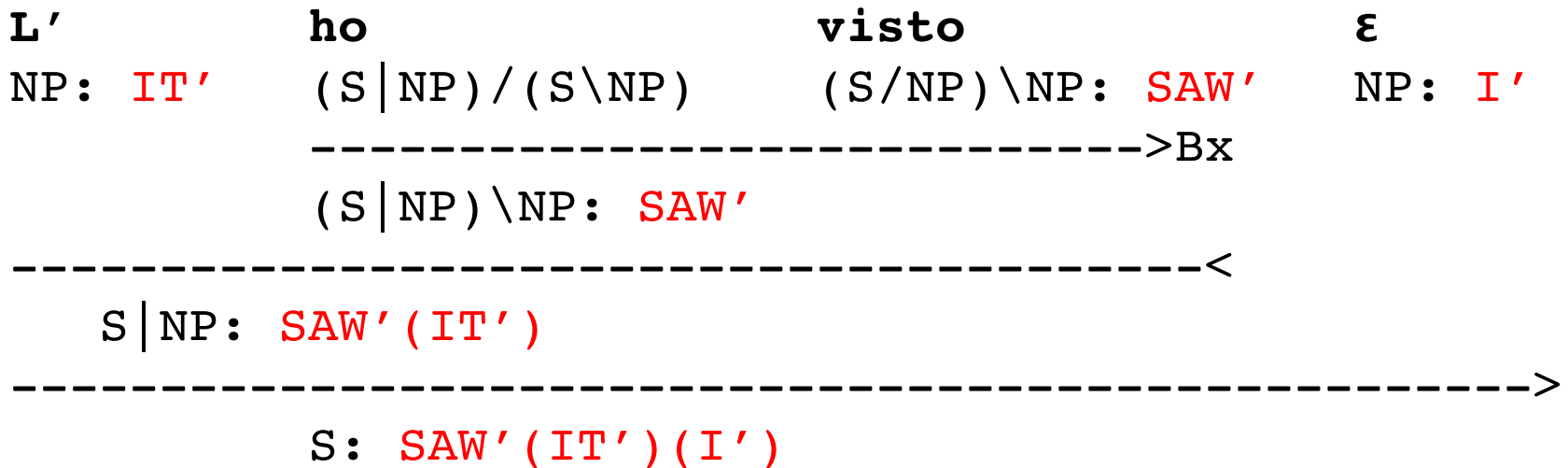
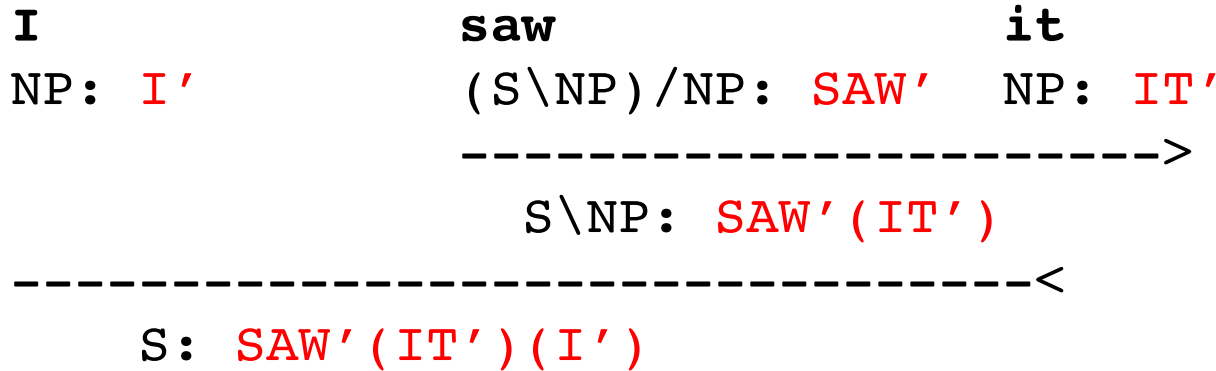
Alignment Example



Transferring Meanings



Transferring Meanings (empty elements)



Progress in parallel meaning banking but a lot left to do...

- Lexical meaning alignment (“good, bad, ugly”)
- Splitting meanings
- Ellipsis (e.g. pro-drop in Italian)
- Cases of mixed split and merge
(is unable to || kan niet)
- Non-literal interpretations



Meaning Banking



gmb.let.rug.nl: Groningen Meaning Bank

pmb.let.rug.nl: Parallel Meaning Bank

Computational Semantics

- Day 1: Exploring Models
- Day 2: Meaning Representations
- Day 3: Computing Meanings with DCG
- Day 4: Computing Meanings with CCG
- Day 5: Drawing Inferences and Meaning Banking