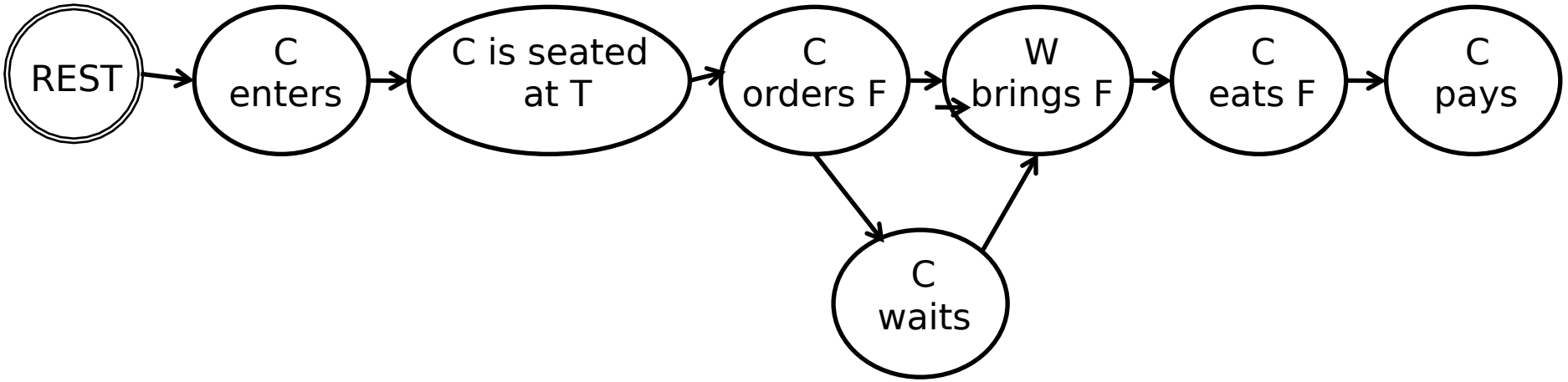
Selected Topics in Semantics and Discourse I

Manfred Pinkal – Saarland University – WS 2015/16

Major Sub-Areas

- Acquisition of semantic knowledge from NL data
  - Sentence-level extraction of paraphrase and inference patterns
  - Discourse-level acquisition of script knowledge
- Compositional/ Denotational distributional semantics
- Enriching logic-based interpretation with statistical information
- Learning semantic interpretation from NL data

An Example Script



Combined Distributional and Logical Semantics (Lewis and Steedman 2013)

Mark Steedman: “A distributional theory of content for NLP”

Language Science Colloquium

Thursday, Nov 12, 4 p.m.

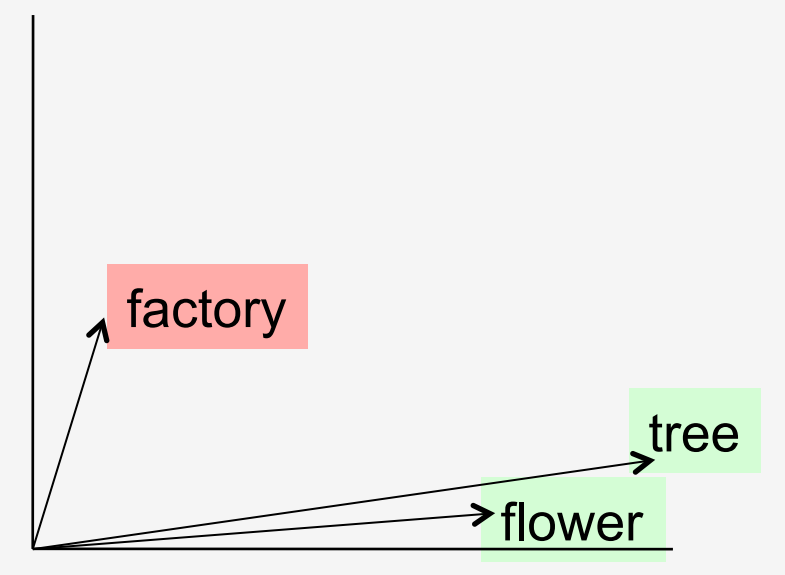
Question Answering: Semantic Problems

Did Google buy YouTube?

(1) Google purchased YouTube
(2) Google's acquisition of YouTube
(3) Google acquired every company
(4) YouTube may be sold to Google
(5) Google will buy YouTube and Microsoft
(6) Google will buy YouTube or Microsoft
(7) Google took over YouTube
(8) Google didn't take over YouTube

Distributional Semantics

$$sim\ (a,b) = \cos(\vec{a},\vec{b})$$



	factory	flower	tree	water	fork
...	...	...	...	...	...
grow	15	147	330	106	3
garden	5	200	198	118	17
worker	279	0	5	18	0
production	102	6	9	28	0
wild	3	216	35	30	0
...	...	...	...	...	...

Question Answering: Semantic Problems

Did Google buy YouTube?

(1) Google purchased YouTube

(2) Google's acquisition of YouTube

(3) Google acquired every company ???

(4) YouTube may be sold to Google

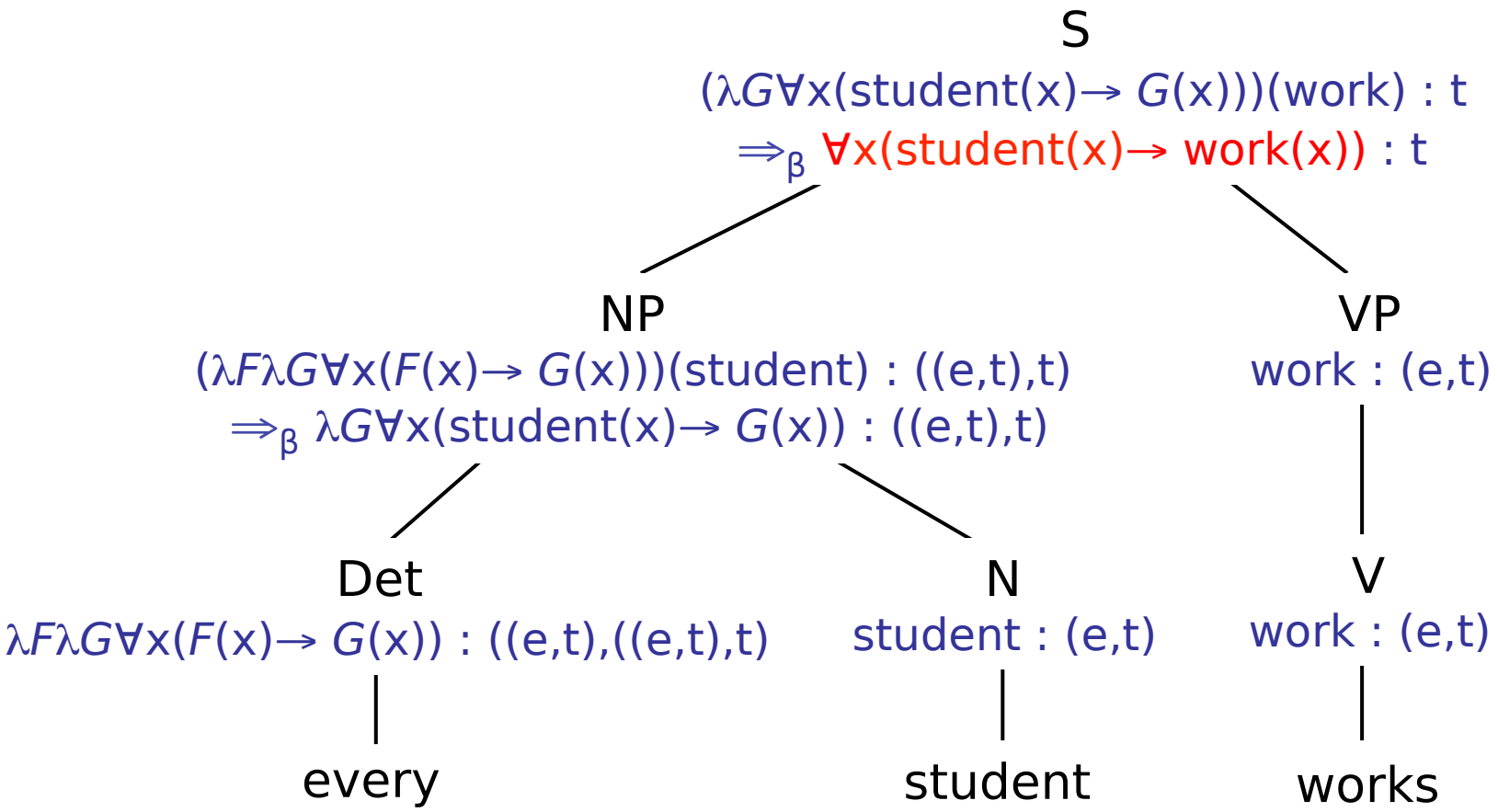
(5) Google will buy YouTube and Microsoft

(6) Google will buy YouTube or Microsoft ???

(7) Google took over YouTube

(8) Google didn't take over YouTube ???

Typed Lambda Calculus for Semantic Composition



Question Answering: Semantic Problems

Did Google buy YouTube?

(1) Google purchased YouTube ?? (WordNet)

(2) Google's acquisition of YouTube ???

(3) Google acquired every company !!!

(4) YouTube may be sold to Google ?? (FrameNet)

(5) Google will buy YouTube and Microsoft

(6) Google will buy YouTube or Microsoft !!!

(7) Google took over YouTube ?? (WordNet)

(8) Google didn't take over YouTube !!!

Lewis/Steedman's Approach, Step 1

"Shakespeare wrote Macbeth"

"Shakespeare is the author of Macbeth'

⇓

Initial semantic analysis

(map CCG-parser output to a deterministic logical form)

$write_{arg0,arg1}(shakespeare, macbeth)$

$author_{arg0,argOf}(shakespeare, macbeth)$

CCG (Combinatory Categorial Grammar)

Logical Form Derivation

Semantic Roles: PropBank

CCG

- CCG: Steedman (2000)
- CCG Parser: C&C, Clark&Curran (2004)
- CCGBank: Hockenmaier&Steedman (2007)
  - Contains a short overview of CCG including semantics
  - See also Jurafsky&Martin, Ch. 12.7 (dependency grammar and categorial grammar)

- Semantic Roles and PropBank
  - See Jurafsky&Martin, Ch. 19.4
  - PropBank: Palmer et al. 2005

	Word	Category	Logical form
Auto.	author	$N/PP[of]$	$\lambda x.\lambda y.author_{arg0,argOf}(y,x)$
	write	$(S\backslash NP)/NP$	$\lambda x.\lambda y.write_{arg0,arg1}(y,x)$
Man.	every	$NP^{\uparrow}/N$	$\lambda p.\lambda q.\forall x[p(x) \rightarrow q(x)]$
	not	$(S\backslash NP)/(S\backslash NP)$	$\lambda p.\lambda q.\neg p(x)$

Challenges

Disambiguation/ Word-sense induction:

- *wear a suit / file a suit*
- *charge a fee/ charge a battery*

Paraphrase identification:

- *Shakespeare wrote Macbeth*
- *Shakespeare is the author of Macbeth*

Disambiguation/ Word-sense induction

Idea: Identification of word meaning via argument type:

One type (signature) – one meaning

e.g.: $\text{suit}_{\text{Arg1}}$:*CLOTHES* *vs.* $\text{suit}_{\text{Arg1}}$:*LEGAL*

$\text{file}_{\text{Arg0}}$:*PERSON,*$_{\text{Arg1}}$ *:LEGAL*

$\text{file}_{\text{Arg0}}$:*PERSON,*$_{\text{Arg1}}$ *:DOCUMENT*

Lewis/Steedman's Approach, Step 2

"Shakespeare wrote Macbeth"

"Shakespeare is the author of Macbeth"

$\Downarrow$

Initial semantic analysis

(map CCG-parser output to a deterministic logical form)

$write_{arg0,arg1}(shakespeare, macbeth)$

$author_{arg0,argOf}(shakespeare, macbeth)$

$\Downarrow$

Entity typing

(apply LDA to assign types)

$write_{arg0:PER,arg1:BOOK}(shakespeare{:}PER, macbeth{:}BOOK)$

$author_{arg0:PER,arg1:BOOK}(shakespeare{:}PER, macbeth{:}BOOK)$

Lewis/Steedman's Approach, Step 2

"Shakespeare wrote Macbeth"

"Shakespeare is the author of Macbeth"

⇓

Initial semantic analysis

(map CCG-parser output to a deterministic logical form)

$write_{arg0,arg1}(shakespeare, macbeth)$

$author_{arg0,argOf}(shakespeare, macbeth)$

⇓

Entity typing

(apply LDA to assign types)

$write_{arg0:PER,arg1:BOOK}(shakespeare{:}PER, macbeth{:}BOOK)$

$author_{arg0:PER,arg1:BOOK}(shakespeare{:}PER, macbeth{:}BOOK)$

Topic models

LDA

Word-sense induction

Lewis/Steedman's Approach, Step 3

Initial semantic analysis

(map CCG-parser output to a deterministic logical form)

$write_{arg0,arg1}(shakespeare, macbeth)$

$author_{arg0,argOf}(shakespeare, macbeth)$

$\Downarrow$

Entity typing

(apply LDA to assign types)

$write_{arg0:PER,arg1:BOOK}(shakespeare{:}PER, macbeth{:}BOOK)$

$author_{arg0:PER,arg1:BOOK}(shakespeare{:}PER, macbeth{:}BOOK)$

$\Downarrow$

Distributional semantic analysis

(cluster typed predicates)

$relation37(shakespeare{:}PER, macbeth{:}BOOK)$

Clustering, Based on Distributional similarity information

Topics

Extraction of Inference Rules (Week2)

- Lin&Pantel 2001
- Bhagat et al. 2007
- Berant et al. 2011

Compositionality in Distributional Semantics (Week 3+4)

- Mitchell&Lapata 2008
- Erk&Pado 2008
- Dinu&Lapata 2010 (Dinu 2010)
- Baroni&Zamparelli 2010
- Grefenstette et al. 2013
- Mikolov et al. 2013

Topics

Extraction of Discourse-level Script Knowledge (Weeks 5+6)

- Chambers&Jurafsky 2008
- Chambers&Jurafsky 2009
- Regneri et al. 2010 (Regneri et al. 2011, Regneri 2013)
- Modi&Titov 2014
- Pichotta&Mooney 2014

Extensions of the logical framework)Week 6)

- Bos&Markert 2005, (Bos&Markert 2006, Dagan et al. 2006)
- MacCartney&Manning 2007 (MacCartney 2009)

Truth-Conditionality and Distributional Semantics (Week 7)

- Young et al. 2014
- Kruszewski et al. 2015

Schedule for Weeks 2 + 3

03.11.	CCG	
	Lin&Pantel (DIRT)	
04.11.	Bhagat et al. (2007)	
	Berant etal. (2011)	
10.11.	Mitchell&Lapata (2008)	
	Erk&Pado (2008)	
11.11.	Dinu&Lapata (2010)	
	General Discussion	