EBERHARD KARLS UNIVERSITÄT TÜBINGEN

FACULTY OF HUMANITIES

SFB 833
Construction of meaning

Association metrics in neural transition-based dependency parsing

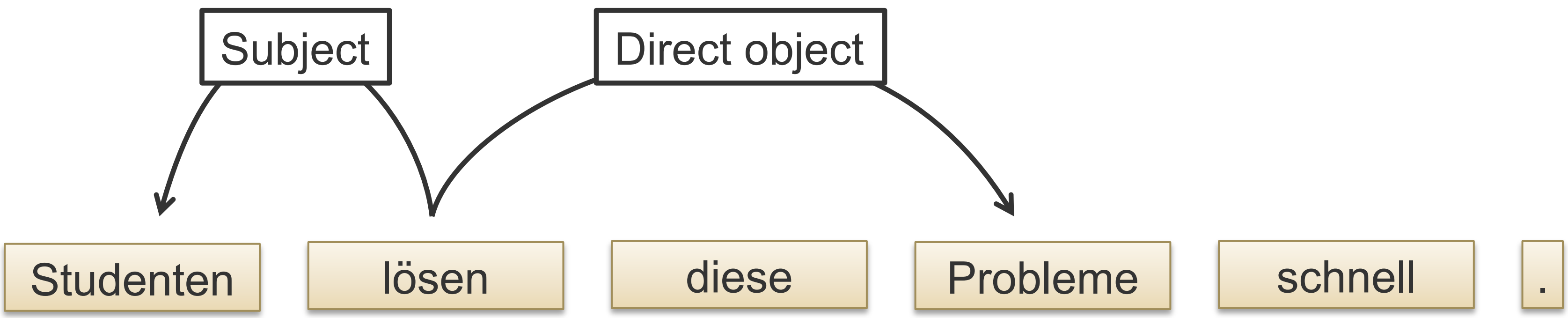
Patricia Fischer, Sebastian Pütz and Daniël de Kok
Tübingen University, SFB 833/**A3**

Content

1. Ambiguity resolution with association metrics
2. Association metrics in transition-based dependency parsing
3. Parser results
4. Outlook

Ambiguity resolution with association metrics

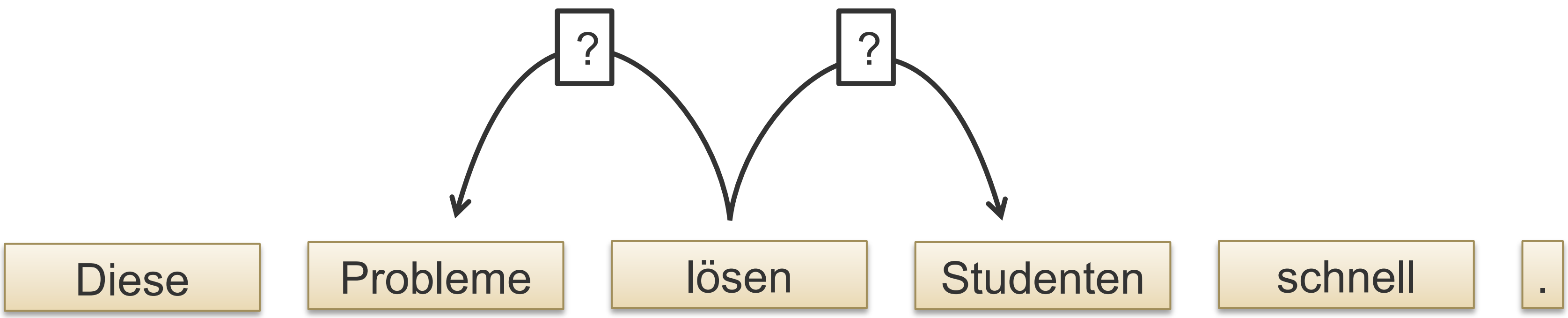
Ambiguity resolution in German



students *solve* *these* *problems* *quickly* *.*

‘Students solve these problems quickly.‘

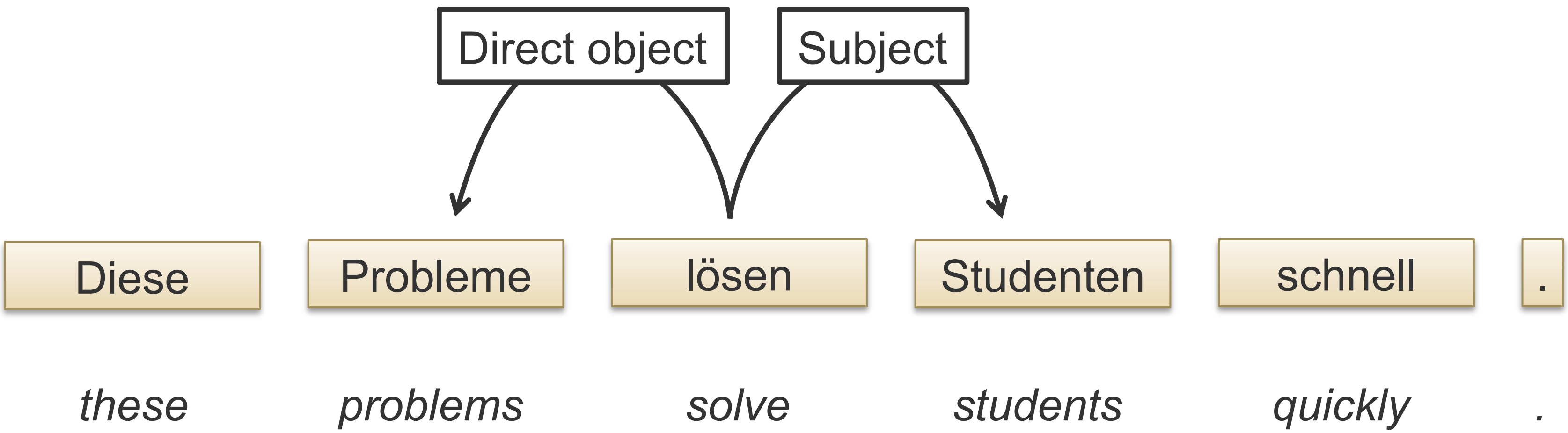
Ambiguity resolution in German



these problems solve students quickly .

‘Students solve these problems quickly.‘

Ambiguity resolution in German



‘Students solve these problems quickly.‘

Summary

Lexical preferences improve:

✓ neural transition-based dependency parsing by 0.26 LAS

✓ accuracy on two ambiguity solving tasks by 2.33 points

The baseline parser: De Kok and Hinrichs (2016)[1]

- Parser configuration
  - Words
  - Part-of-speech tags
  - Dependency relations
  - Prefix and suffix characters
  - Topological field information

- Trained on TüBa-D/Z, release 10 (1.8M tokens, 96K sentences), manually labeled, non-gold PoS, gold topological fields

- Accuracy: 92.01 LAS

[1] De Kok and Hinrichs (2016). Transition-Based Dependency Parsing with Topological Fields. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics: Volume 2, Short Papers*, p.1–7.

Frequent errors by the De Kok and Hinrichs (2016) parser

Relation	Error count	Percent of all errors
Prepositional phrase/object	7,037	24.68%
Accusative object	1,796	6.30%
Subject	1,670	5.85%

Pointwise mutual information[2,3]

- Procedure
  - Retrieve all head-dependent pairs with dependency relation
  - Calculate the PMI for all triples with minimum frequency *n*
- Data
  - Berliner Tageszeitung *taz and* German Wikipedia (1.2B tokens, 62.7M sentences)

[2] Hindle and Rooth. 1993. Structural Ambiguity and Lexical Relations. *Computational Linguistics*, 19(1):103–120.

[3] Van Noord. 2007. Using Self-Trained Bilexical Preferences to Improve Disambiguation Accuracy. *Proceedings of the 10th Conference on Parsing Technologies*:1–10.

Pointwise mutual information

- PMI variants
  - Normalized PMI: NPMI$(h \xrightarrow{r} d)$ [-1, 1]
  - Positive NPMI: PNPMI$(h \xrightarrow{r} d)$ [0, 1]

- Examples (NPMI)
  - lösen $\xrightarrow{OBJ}$ Probleme: 0.484 ✓
    'solve problems'
  - lösen $\xrightarrow{OBJ}$ Studenten: - ✗
    'solve students'

Dependency embedding-based association scores

✗ Sparsity of PMIs

Dependency embeddings[4] predict probabilities for dependency triples $h \xrightarrow{r} d$

✓ Separate embeddings for h and $\xrightarrow{r} d$

Data:

- *taz*, German Wikipedia and europarl (1.25B tokens and 64.9M sentences), annotated by baseline parser
- non-projective / pseudo-projectivized

[4] Levy and Goldberg (2014). Dependency-Based Word Embeddings. *Proceedings of the 52nd Annual Meeting for the Association for Computational Linguistics, Volume 2: Short Papers*:302–308.

Dependency embedding-based association scores

- For each head candidate ***h***, dependent candidate ***d*** and any possible dependency relation ***r***, compute:

$$assoc_{dep}(h \xrightarrow{r} d) = \sigma\,(\,W_h \cdot C_{d,r}\,) \qquad (0,1)$$

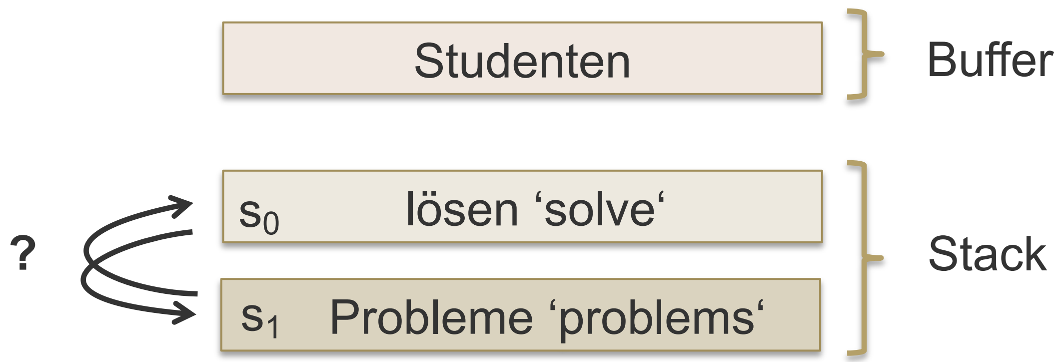
with
- ***W*** : word embedding matrix
- ***C*** : context embedding matrix

- Examples
  - lösen $\xrightarrow{OBJ}$ Probleme: 0.999 ✓
  - lösen $\xrightarrow{OBJ}$ Studenten: 0.400 ✓

Association metrics in parsing

EBERHARD KARLS UNIVERSITÄT TÜBINGEN

An example: Parsing with embedding-based scores



Diese Probleme lösen Studenten schnell.

‘Students solve these problems quickly.‘

Transitions: LEFTARC ($s_0 \rightarrow s_1$) and RIGHTARC ($s_1 \rightarrow s_0$), SHIFT

Dependencies: SUBJ, OBJ, PP

An example: Parsing with association scores

LEFTARC

s_0	lösen ‘solve‘
s_1	Probleme ‘problems‘

$$v_{assoc} = [assoc(\boldsymbol{s_0},\boldsymbol{s_1},r) \mid \forall r \in R]$$

with R = {SUBJ, OBJ, PP}

assoc (lösen, Probleme, SUBJ)	assoc (lösen, Probleme, OBJ)	assoc (lösen, Probleme, PP)

An example: Parsing with association scores

s_0	lösen ‘solve‘	RightArc
s_1	Probleme ‘problems‘	

$$v_{assoc} = [assoc(\mathbf{s_1},\mathbf{s_0},r) \mid \forall r \in R]$$

with R = {SUBJ, OBJ, PP}

assoc (Probleme, lösen, SUBJ)	assoc (Probleme, lösen, OBJ)	assoc (Probleme, lösen, PP)

An example: Parsing with association scores

s_0 lösen ‘solve‘

LEFTARC

RIGHTARC

s_1 Probleme ‘problems‘

$$v_{assoc} = [assoc(s_0,s_1,r), assoc(s_1,s_0,r) \mid \forall r \in R]$$

with R = {SUBJ, OBJ, PP}

assoc (I., P., SUBJ)	assoc (I., P., OBJ)	assoc (I., P., PP)	assoc (P., I., SUBJ)	assoc (P., I., OBJ)	assoc (P., I., PP)

Binary indicators as default markers

- **PMIs**
  - *binary*

$\text{assoc}_{\text{PMI}}$ / 0.0	$\boldsymbol{PMI} \in$ table?

- **Embedding-based scores**
  - *binary*

$\text{assoc}_{\text{dep}}$ / 0.5	$\boldsymbol{h} \in \boldsymbol{W}$? $\xrightarrow{r} \boldsymbol{d} \in \boldsymbol{C}$?

  - *triple-binary*

$\text{assoc}_{\text{dep}}$ / 0.5	$\boldsymbol{h} \in \boldsymbol{W}$?	$\xrightarrow{r} \boldsymbol{d} \in \boldsymbol{C}$?	$\boldsymbol{d} \in \boldsymbol{W}$?

Parser results

Parsing with PMIs ($n = 5$)

Model	LAS	UAS	Inversion accuracy	Preposition accuracy
De Kok and Hinrichs (2016)	92.01	93.88	81.03%	77.80%
+ NPMI	**92.27**	**94.01**	81.93%	78.60%
+ NPMI, binary	92.18	93.94	**82.57%**	**78.78%**
+ PNPMI	92.21	93.99	82.09%	78.44%

Parsing with embedding-based scores

Model	LAS	UAS	Inversion accuracy	Preposition accuracy
De Kok and Hinrichs (2016)	92.01	93.88	81.03%	77.80%
+ projectivized	92.23	**93.97**	82.57%	78.55%
+ projectivized, binary	**92.24**	**93.97**	**83.36%**	78.47%
+ projectivized, triple-binary	92.16	93.88	**83.36%**	**78.62%**

Parsing with association scores

Model	LAS	UAS	Inversion accuracy	Preposition accuracy
De Kok and Hinrichs (2016)	92.01	93.88	81.03%	77.80%
+ NPMI	**92.27**	**94.01**	81.93%	**78.60%**
+ projectivized, binary	92.24	93.97	**83.36%**	78.47%

Outlook

Future work

- Trilexical preferences
- Make alternative attachments available, e.g. via constrained beam search
- Compatibility modeling
- Extend experiments to a language similar to German, e.g. Dutch

THANK YOU

BONUS MATERIAL

PMI variants

- PMI variants
  - PMI: $PMI = \frac{p(h \xrightarrow{r} d)}{p(h \xrightarrow{r}\ )\ p(\ \xrightarrow{r} d)}$ [-∞, ∞]
  - Normalized PMI (NPMI): $PMI_{norm} = \frac{PMI(h \xrightarrow{r} d)}{-\log p(h \xrightarrow{r} d)}$ [-1, 1]
  - Positive NPMI (PNPMI): $PMI_{pos\_norm} = max(0, NPMI)$ [0, 1]

Parsing with PMIs

Model	LAS	UAS	Inversion accuracy	Preposition accuracy
De Kok and Hinrichs (2016)	92.01	93.88	81.03	77.80
+ NPMI, minfreq 5	**92.27**	**94.01**	81.93	78.60
+ NPMI, minfreq 50	92.14	93.92	82.25	78.29
+ NPMI, minfreq 100	92.16	93.92	80.72	78.56
+ NPMI, minfreq 5, bin	92.18	93.94	**82.57**	**78.78**
+ NPMI, minfreq 50, bin	92.16	93.93	81.93	78.35
+ NPMI, minfreq 100, bin	92.18	93.96	81.67	78.29
+ PNPMI, minfreq 5	92.21	93.99	82.09	78.44
+ PNPMI, minfreq 50	92.19	93.95	81.46	78.66
+ PNPMI, minfreq 100	92.17	93.94	82.25	78.57

Parsing with embedding-based scores

Model	LAS	UAS	Inversion accuracy	Preposition accuracy
De Kok and Hinrichs (2016)	92.01	93.88	81.03	77.80
+ proj, typed	92.23	**93.97**	82.57	78.55
+ proj, typed, bin	**92.24**	**93.97**	**83.36**	78.47
+ proj, typed, triple-bin	92.16	93.88	**83.36**	**78.62**
+ proj, semi-typed	92.11	93.94	80.66	77.99
+ proj, semi-typed, bin	91.98	93.89	80.61	77.71
+ proj, semi-typed, double-bin	92.07	93.93	81.93	77.98
+ non-proj, typed	92.17	93.93	81.46	78.17
+ non-proj, typed, bin	92.22	**93.97**	82.20	78.45
+ non-proj, typed, triple-bin	92.08	93.86	82.99	78.26

Binary indicators as default markers

- Association score = default score?
- **PMIs**: *binary*

$\text{assoc}_{\text{PMI}}$	$\boldsymbol{PMI} \in$ table?

- **Embedding-based scores**
  - *binary*

$\text{assoc}_{\text{dep}}$	$\boldsymbol{h} \in \boldsymbol{W}$? $\boldsymbol{d} \in \boldsymbol{C}$?

  - *double-binary*

$\text{assoc}_{\text{dep}}$	$\boldsymbol{h} \in \boldsymbol{W}$?	$\boldsymbol{d} \in \boldsymbol{C}$?

  - *triple-binary*

$\text{assoc}_{\text{dep}}$	$\boldsymbol{h} \in \boldsymbol{W}$?	$\boldsymbol{\{d,r\}} \in \boldsymbol{C}$?	$\boldsymbol{d} \in \boldsymbol{W}$?

Parsing with PMIs on gold PoS tags

Model	LAS	UAS	Inversion accuracy	Preposition accuracy
De Kok and Hinrichs (2016)	92.49	94.19	81.46	78.35
+ NPMI, minfreq 5	**92.77**	**94.37**	82.68	79.05
+ NPMI, minfreq 50	92.65	94.31	81.56	78.90
+ NPMI, minfreq 100	92.57	94.21	**83.26**	78.66
+ NPMI, minfreq 5, bin	92.64	94.24	82.83	78.85
+ NPMI, minfreq 50, bin	92.64	94.25	82.68	**79.21**
+ NPMI, minfreq 100, bin	92.64	94.24	83.21	78.84
+ PNPMI, minfreq 5	92.59	94.23	81.62	78.84
+ PNPMI, minfreq 50	92.69	94.31	81.99	79.10
+ PNPMI, minfreq 100	92.65	94.29	81.56	78.92

Parsing with embedding-based scores on gold PoS

Model	LAS	UAS	Inversion accuracy	Preposition accuracy
De Kok and Hinrichs (2016)	92.49	94.19	81.46	78.35
+ proj, typed	**92.67**	94.29	82.52	78.96
+ proj, typed, bin	92.65	94.25	82.78	78.86
+ proj, typed, triple-bin	92.54	94.16	82.68	78.68
+ proj, semi-typed	92.64	94.28	82.15	78.75
+ proj, semi-typed, bin	92.64	**94.31**	80.29	78.72
+ proj, semi-typed, double-bin	92.63	94.29	82.89	**78.99**
+ non-proj, typed	92.64	94.23	82.36	78.85
+ non-proj, typed, bin	92.64	94.24	82.99	78.87
+ non-proj, typed, triple-bin	92.58	94.21	**83.31**	78.81

