Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

FAKULTÄT FÜR
GEISTESWISSENSCHAFTEN

HOW TO MAKE MANUAL ANNOTATION MORE EFFICIENT

ENSEMBLE DEPENDENCY PARSING AS A PREPROCESSING STEP TO OBTAIN HIGH-QUALITY ANNOTATION IN AN EFFICIENT WAY

TASK

GOALS

- Creating a syntactically annotated textbook corpus for linguistic research
- This poster: Combining automatic and manual dependency annotation to reduce manual workload

DEPENDENCY PARSING

```
1  Das            das            ART  ART    _  2  SUBJ  _  _
2  ist            sein           V    VVFIN  _  0  S     _  _
3  ein            ein            ART  ART    _  4  DET   _  _
4  Beispielsatz   Beispielsatz    N    NN     _  2  PRED  _  _
5  .              .              $.   $.     _  0  ROOT  _  _
```

Fig.1: Dependency parse CONLL format

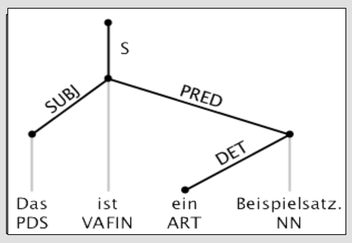

Fig.2: Example dependency parse (HDT AnnoViewer)

THE "MULTISTRATEGY APPROACH"

- A method of tagging texts by combining the output of a cluster of taggers
- Tagger models result from training different learning algorithms on the same data
- Different taggers create their analyses in different ways such that their errors are uncorrelated
- Hypothesis: A reasonable weighted combination of the tagger choices can obtain better results than the individual taggers do (van Halteren et al. 2001, p. 201)
- Related work: Søgaard (2010), Rehbein et al. (2014)

RESEARCH QUESTIONS

i. What accuracy do the parsers have on the training domain?
ii. What accuracy do the parsers have on the target domain?
iii. What accuracy does our ensemble of parsers achieve on the target domain?
iv. What kind of sentences does the ensemble fail to parse correctly?
v. To what extend can the ensemble parser support manual annotation?

THE TRAINING CORPUS

- Hamburg Dependency Treebank (HDT): Part A
- Contains 101,999 sentences produced by manual annotation and subsequent cross-checking for consistency with DECCA (Dickinson & Meurers 2003; Foth et al. 2014).
- Available free of charge for academic use
- For more information visit: https://nats-www.informatik.uni-hamburg.de/HDT

THE TEST CORPUS

- Texts from the Jena Textbook Corpus (unpublished)
- Three geography textbooks: three double pages
- Average of 54 sentences per text
- Overall 144 sentences parsed

OUR ENSEMBLE OF PARSERS

- MALT parser (Nivre et al. 2006): Transition-based parsing
- MATE parser (Björkelund et al. 2010): Second order maximum spanning trees-based parsing
- JWCDG parser (The CDG Team 1997-2015): Weighted hand-written rules; developed on the basis of HDT

QUANTITATIVE RESULTS

TRAINING ON THE HDT: CROSS-VALIDATION

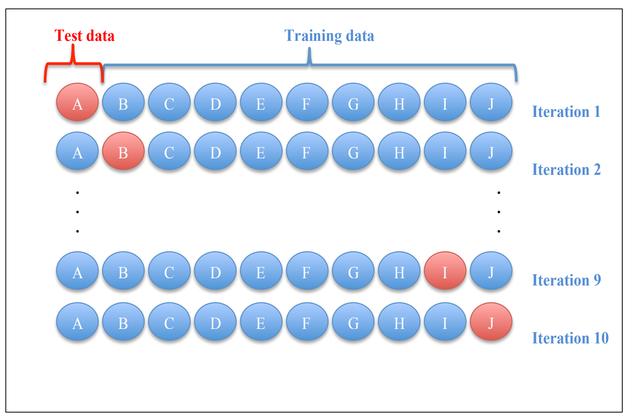

Fig.3: 10-Fold cross-validation

DEVELOPING A GOLD STANDARD ON THE TEST CORPUS

- Two annotators annotated independently
- IAA: UAS= 0.95 (±0.01), LAS= 0.93 (±0.01)
- α=0.93 (±0.02) according to Skjærholt (2014)

ENSEMBLES

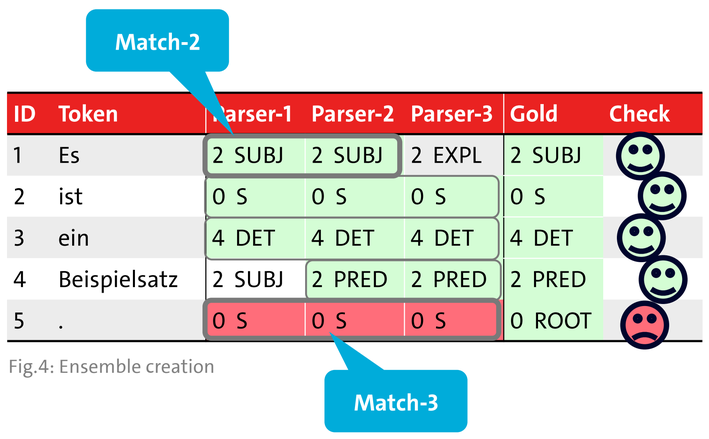

Fig.4: Ensemble creation

- Ens-1: Majority vote (Match-3 /-2), default MATE
- Ens-2: Majority vote (Match-3), default MATE except MATE = S or OBJA, then JWCDG

EVALUATION

- Scores
  - Unlabeled Attachment Score (UAS)
  - Labeled Attachment Score (LAS)
- Settings
  - HDT: Cross validation of 2 statistical parsers
  - Gold Standard: Accuracy of all methods
  - Match-3 instances: Accuracy of parser ensemble

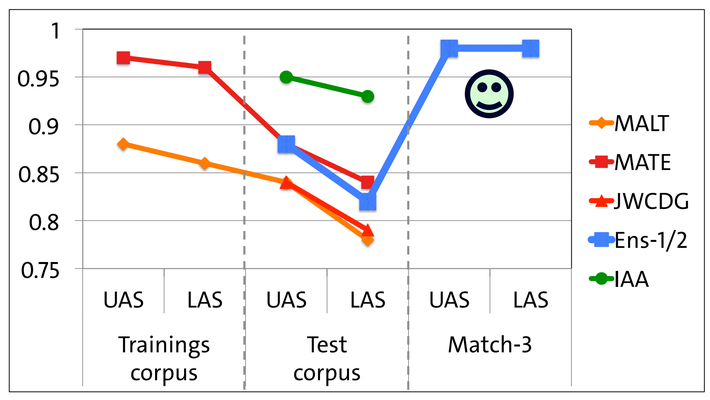

Fig. 5: Accuracy evaluation

- Match-3
  - Ensembles agreed on
    - on 71 % (±0.10) of all tokens
    - on about 15 % of the sentences / text
  - 21 out of 22 agreed-on sentences were correct

QUALITATIVE RESULTS

GENERAL CHALLENGES IN DEPENDENCY PARSING

- Valency decision: OBJP vs. PP (depends on training data)
- Attachment ambiguities (semantic decisions)

PARSER-SPECIFIC CHALLENGES

Parser	Overgenerates		Comment
JWCDG	APP: S:	26% (68/262) 15% (39/262)	default attachment fragments
MALT	ROOT:	64% (214/335)	incomplete analyses
MATE	S: OBJA:	13% (30/225) 12% (26/225)	fragments confusion SUBJ/OBJA

Table 1

DOMAIN-SPECIFIC CHALLENGES

- Incomplete sentences (lists, exercises)

(1) *M4 Auswirkungen des Klimawandels am Beispiel "Starkregen"*

(2) *check-it:*
- *Merkmale einer thematischen Karte - hier Bodennutzung - kennen*
- *[...]*

- Complex coordinations

(3) *Als praktisch sicher gilt, dass es über den meisten Landflächen wärmere und weniger kalte Tage und Nächte sowie wärmere und häufiger heiße Tage und Nächte geben wird.*

CONCLUSIONS

- Pro: High reliability of ensemble majority vote
- Con: Ensembles do not outperform best individual parser
- Application: Ensemble support of manual parser correction
  - Skipping of perfect labels (AUX, AVZ, DET, GMOD, OBJA, SUBJ) and complete sentence matches
  - Highlighting of OBJP and PP
- In addition: Highlighting of lists and exercises
- Future work: Domain adaptation by including the gold standard in the training data

REFERENCES

- Björkelund, Anders; Bohnet, Bernd; Hafdell, Love & Nugues, Pierre (2010): A high-performance syntactic and semantic dependency parser. In Coling 2010 Demonstration Volume.
- Brill, Eric & Wu, Jun (1998): Classifier combination for improved lexical disambiguation. In Proceedings of ACL.
- Dickinson, Markus & Meurers, Detmar 2003: Detecting Errors in Part-of-Speech Annotation. In Proceedings of EACL.
- Foth, Kilian; Köhn, Arne; Beuck, Niels & Menzel, Wolfgang (2014): Because Size Does Matter: The Hamburg Dependency Treebank. In Proceedings of LREC.
- Nivre, Joakim; Hall, Johan & Nilsson, Jens (2006): MaltParser: A Data-Driven Parser-Generator for Dependency Parsing. In Proceedings of LREC.
- Rehbein, Ines; Schalowski, Sören & Wiese, Heike (2014): The KiezDeutsch Korpus (KiDKo) Release 1.0. In Proceedings of LREC.
- Skjærholt, Arne (2014): A chance-corrected measure of inter-annotator agreement for syntax. In Proceedings of the ACL.
- Søgaard, Anders (2010): Simple semi-supervised training of part-of-speech taggers. In Proceedings of the ACL.
- The CDG Team (1997-2015): https://gitlab.com/nats/jwcdg.
- Van Halteren, Hans; Daelemans, Walter & Zavrel, Jakub (2001): Improving Accuracy in Word Class Tagging through the Combination of Machine Learning Systems. In Computational Linguistics 27 (2).

JESSICA SOHL

E-MAIL: JESSICA.KATHARINA.SOHL@STUDIUM.UNI-HAMBURG.DE

HEIKE ZINSMEISTER

E-MAIL: HEIKE.ZINSMEISTER@UNI-HAMBURG.DE