

Auswertung und Wettbewerb

Praktikum Verarbeitung natürlicher Sprache

Richard Mörbitz

09.07.2019

Was wurde implementiert?

Pflicht: Korpustransformationen

- Debinarisierung
- triviales Unking

Wahlpflicht: Korpustransformationen

- Markovisierung (3x)
- Smoothing (3x)

Wahlpflicht: Optimierungen des Parsers

- Pruning (5x)
- k -best-Parsing (1x)
- A*-Suche

Anmerkungen: Korpustransformationen (Pflicht)

Debinarisierung

- rekursive Implementierung (meist) kein Problem
- Ansätze: neuen Baum erstellen, in-Place-Änderung, Hybrid
- Laufzeit: in-Place < neuer Baum < Hybrid

Triviales Unking

- Korpustransformation meist problemlos, Achtung: off-by-one-Fehler
- neuer Baum vs. in-Place wie oben
- Reflektierung im Parser problemlos (wenn man daran denkt)

Anmerkungen: Korpustransformationen (Wahlpflicht)

Smoothing

- leichte Erweiterung auf Basis des Unking
- Achtung: off-by-one-Fehler bei Satzbeginn (SC vs. C)

Binarisierung

- rekursive Implementierung (meist) kein Problem
- Unklarheiten bei Parent-Annotation
- Laufzeit: in-Place < Hybrid
- Keine Einsendung mit ermittelten guten Werten für h und v

Anmerkungen: Optimierungen des Parsers (Wahlpflicht)

Reihenfolge der Iterationen (und Parallelisierung)

Favorisierte Variante (alle Einreichungen):

- 1 Größe des Spans
- 2 Anfang des Spans
- 3 Split-Position
- 4 Regeln

Ansätze für Parallelisierung:

- auf Satzebene
- auf Zellebene (2. Iteration)

Wettbewerb – Baseline

Daten

- Trainingsdaten: erste 100 Sätze eures Trainingskorpus / gesamter Trainingskorpus
- Testdaten: erste 100 Sätze eures Trainingskorpus / erste 100 Sätze WSJ Sec. 22

F1-Score: ca. 81 % / 72 % bei allen

Laufzeit:

Abgabe	kleine Grammatik		große Grammatik	
	real	user+sys	real	user+sys
Rust # 1	3.652s	53.828s	55.307s	6m58.372s
C++ # 1	2.143s	3.196s	22.716s	40.184s
C++ # 2	11.610s	11.604s	2m52.398s	2m52.384s
Rust # 2	3m38.500s	3m38.484s	–	–
Python	91m55.054s	91m54.064s	–	–
Rust # 3	1.947s	1.940s	30.457s	30.452s

Wettbewerb – Threshold

Daten: gesamter Trainingskorpus bzw. erste 100 Sätze WSJ Sec. 22

Abgabe	Threshold 10^{-5}			Threshold 10^{-4}		
	f1	real	user+sys	f1	real	user+sys
Rust # 1	72	21.850s	3m30.524s	67	15.661s	2m49.556s
C++ # 1	72	25.654s	42.948s	68	24.353s	36.692s
C++ # 2	72	1m47.223s	1m47.212s	68	1m43.273s	1m43.26s
<hr/>						
	Rank 500			Rank 100		
	f1	real	user+sys	f1	real	user+sys
Rust # 1	72	22.901s	3m40.488s	64	14.109s	2m36.556s
C++ # 1	72	1m13.757s	7m12.088s	68	26.394s	42.504s
C++ # 2	72	6m30.936s	6m30.912s	69	3m07.073s	3m07.056s

Prüfung

Grundlage der Bewertung ist die abgegebene Implementierung.

Formal (Modulbeschreibung): **mündliche Prüfung**

Inhalt: im wesentlichen Plausibilitätsprüfung

Dauer: 15 min

Termin: Mi, 17.07.2019

- 14:00 Uhr
- 14:20 Uhr
- 14:40 Uhr
- 15:00 Uhr
- 15:20 Uhr