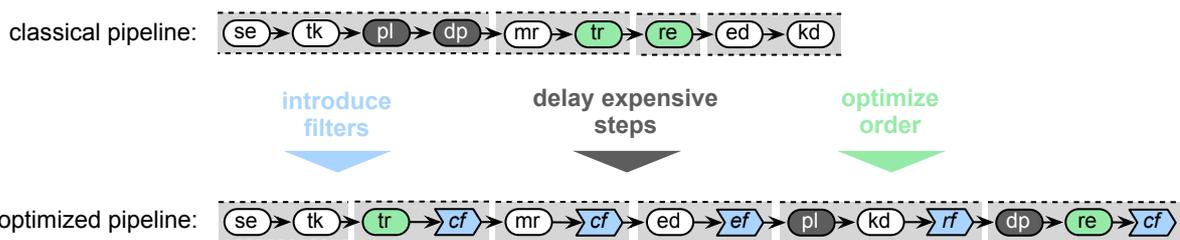


Masterarbeit

Heuristische Suchverfahren zur Laufzeitoptimierung von Information-Extraction-Pipelines

Motivation

Information-Extraction-Pipelines analysieren automatisch natürlichsprachigen Text, um strukturierte Informationen über festgelegte Typen von Entitäten, Beziehungen zwischen diesen und umgebende Ereignisse zu identifizieren und verarbeitbar zu machen. Dadurch lassen sich die Möglichkeiten gegenwärtiger Internet-Suchanwendungen deutlich erweitern. In der Praxis sind Pipelines dafür jedoch häufig zu langsam, vor allem da sie berechnungsintensive Analysen auch auf irrelevanten Textpassagen durchführen. In einem aktuellen Forschungsprojekt wird daher an der Laufzeitoptimierung solcher Pipelines gearbeitet, was ein Scheduling der Pipelines auf Basis verschiedener Kriterien beinhaltet. Ein Ansatz, dieses Optimierungsproblem effizient zu lösen, führt über die Kernalgorithmen der künstlichen Intelligenz: heuristische (informierte) Suchverfahren.



Aufgabenstellung

Ziel der Arbeit ist es, das Pipeline-Scheduling-Problem als heuristisches Suchproblem zu modellieren und dessen Eigenschaften abstrakt zu untersuchen. Davon ausgehend soll ein geeignetes Suchverfahren ausgewählt werden, mittels dessen sich das Problem möglichst effizient und optimal lösen lässt. Dieses Verfahren soll konzeptuell entwickelt und implementiert werden. Zur Evaluierung des Verfahrens wie auch der Laufzeit resultierender Pipelines dient ein Projekt des Bundesministeriums für Bildung und Forschung; in diesem wurde eine Information-Extraction-Pipeline zur Erkennung und Verarbeitung textueller Umsatzprognosen entwickelt, die zahlreiche Analyseschritte beinhaltet und den Nutzen von State-of-the-Art-Algorithmen der künstlichen Intelligenz im industriellen Einsatz zeigt.

Voraussetzungen

Sicherer Umgang mit den Inhalten der Master-Veranstaltung „Heuristische Suchverfahren“

Bearbeiter
Tobias Melzner

Wissenschaftliche Leitung
Prof. Dr. G. Engels

Ansprechpartner

s-lab – Universität Paderborn
Henning Wachsmuth
Raum E4.107
Tel: +49 (0) 5251 / 60-3359
Email: hwachsmuth@s-lab.upb.de