

Master- oder Diplomarbeit



**Programmieren in natürlicher Sprache:
Auflösen von Referenzen in englischen Texten**

Ausgangssituation/Kontext

An unserem Lehrstuhl werden Verfahren zur natürlichsprachlichen Programmierung entwickelt. Benutzer sollen mit unseren Systemen auf Englisch interagieren und ihre Anliegen beschreiben, z.B. das Erzeugen von 3D-Animationen oder das Steuern von Haushaltsrobotern.

Hierbei tritt immer wieder das Problem auf, dass Referenzen in Texten nicht richtig aufgelöst werden können (z.B. Pronomen wie in „The dog eats the schnitzel. Then *it* barks.“). Gängige computerlinguistische Analysen betrachten nur syntaktische Informationen und führen in unserem Kontext oft zu falschen Ergebnissen (im Beispiel wäre dann *it* das *schnitzel*).

Aufgabenstellung/Ziel

In unserem Kontext verfügen wir über Zusatzinformationen (z.B. welche Klasse verfügt über welche Methoden), sogenanntes Domänenwissen. Dieses Wissen wird derzeit jedoch nicht für die Referenzauflösung genutzt.

Das Ziel der Arbeit ist eine wissensgestützte Referenzauflösung für Englisch. In Ihrem ersten Schritt recherchieren Sie die verfügbaren computerlinguistischen Referenzauflösungssysteme. Dann beschäftigen Sie sich damit, wie die Verfahren mithilfe von Domänen- und/oder Weltwissen verbessert werden können, so dass sie in der natürlichsprachlichen Programmierung eingesetzt werden können.

Voraussetzungen

Um für die Anfertigung dieser programmierlastigen Arbeit gerüstet zu sein, verfügen Sie über Programmiererfahrung in Java und/oder Python. Außerdem sollten Sie über ein gutes Sprachgefühl verfügen und sich mit Englisch auskennen. Sie haben keine Angst vor Arbeiten im Team und scheuen sich nicht, neue, Ihnen unbekannte Techniken einzusetzen. Sie zögern nicht, eine E-Mail zur Vereinbarung eines ersten Gesprächs zu schreiben, in welchem wir Ihnen unverbindlich Details und einen persönlichen Eindruck unserer Arbeit geben werden.

Informatikerfreundliche Arbeitsumgebung

- Redundante Kaffeemaschinenanbindung
- Klimatisierter Poolraum
- Gut ausgebaute Süßigkeiteninfrastruktur
- Remote-Zugang zum Arbeiten von überall

Betreuer

Mathias Landhäußer, Raum 343, landhaeusser@kit.edu, doodle.com/mathias.landhaeusser
Sebastian Weigelt, Raum 346, sebastian.weigelt@partner.kit.edu, doodle.com/sebastian.weigelt
Sprechzeiten nach Abstimmung

