



Interaktive Vorleseanwendung auf Basis der gesprochenen Wikipedia

Abschlussarbeit (Bachelor)

Motivation: Gesprochene Inhalte (PodCasts, Audiobücher, ...) eignen sich hervorragend zur Information und Unterhaltung in Situationen in denen Hände und Augen anderweitig beschäftigt sind (zum Beispiel Autofahren). Die gesprochene Wikipedia bietet zusätzlich eine barrierefreie Alternative für Menschen mit Sehbehinderung beim Zugang zu Information. Die Navigation in üblichen, als reine Audiodateien vorliegenden Inhalten ist jedoch ineffizient. Gesprochensprachliche Interaktion wäre intuitiver und effizienter.



Kurzfassung: Die deutschsprachige gesprochene Wikipedia¹ umfasst knapp 900 vorgelesene Artikel (insgesamt 270 Stunden Audiomaterial); gegenüber maschinell vorgelesenen Artikeln² sind sie deutlich verständlicher und angenehmer zu konsumieren. Ziel der Arbeit ist die bessere Erschließung der vorliegenden Daten um dem Nutzer außer dem bloßen Abhören die aktive Navigation und Exploration (Nutzung des Artikel-Inhaltsverzeichnisses, Überspringen von Erläuterungen, Folgen von Links, ...) zu ermöglichen. Die Interaktion mit der Anwendung soll dabei mittels Spracheingabe möglich sein.

Stichworte: Mensch-Maschine-Interaktion, Sprechschnittstelle, Sprachdialogverarbeitung

Ziel der Abschlussarbeit ist es, auf Basis bereits vorliegender Alignments zwischen Wikipedia-Markup und zugehörigem Audio einen *Voice Browser* zu entwickeln, der die Navigation sowohl innerhalb des Artikels als auch über Links zu anderen Artikeln ermöglicht. Die Anwendung soll für nicht gesprochen vorliegende Artikel diese mittels Software synthetisieren.

Der Arbeitsprozess könnte sich in folgende Schritte gliedern:

- Aneignung von Wissen über gesprochene Sprache, Spracherkennung und -synthese (z.B. [1])
- Analyse der vorliegenden Text-Audio-Alignments und Qualitätsabschätzung
- Prototypische Perfektionierung des Alignments (sowie Erstellung weiterer notwendiger Annotationen) für einige wenige Testartikel
- Einarbeitung in Spracherkennung und -synthese für die gewählte Plattform (Android oder Java) sowie in die Wikipedia-API

¹http://de.wikipedia.org/wiki/Portal:Gesprochene_Wikipedia

²<http://www.pediaphon.org>

- Entwurf der Interaktionsmöglichkeiten zur Navigation innerhalb und über gesprochene Artikel hinweg
- Entwurf einer Anwendungsarchitektur die diese Interaktionen ermöglicht und zukünftig flexibel erweitert werden kann
- Evaluation der erweiterten Navigationsmöglichkeiten innerhalb der gesprochenen Artikel in einer Nutzerstudie
- Evaluation des Nutzwerts gesprochener vs. maschinengesprochener Artikel

Die erfolgreiche Bearbeitung des Themas erfordert die Fähigkeit sich in teils umfangreiche vorhandene Softwareprojekte, Datensammlungen und Algorithmen einzuarbeiten, gute Programmierkenntnisse in JAVA sowie Kenntnisse auf der Kommandozeile. Vorwissen aus den Lehrveranstaltungen Sprachsignalverarbeitung, Computerlinguistik und/oder User Interface Design ist nicht notwendig, aber von Vorteil.

Die Abschlussarbeit kann auf Deutsch oder Englisch erarbeitet und verfasst werden; relevante Literatur ist überwiegend auf Englisch verfügbar. gBei jeder Abschlussarbeit streben wir die Veröffentlichung der Forschungsergebnisse auf einer internationalen oder nationalen Konferenz an.

Literatur

- [1] Daniel Jurafsky und James H. Martin. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. 2. Aufl. Pearson International, 2009.

Kontakt

Timo Baumann (*baumann@inf...*), Prof. Wolfgang Menzel

URL dieses Dokuments:

