

Information retrieval with context-recall in human-computer conversations

Heutzutage sind Instant-Messaging-Anwendungen insbesondere durch die Einführung von Smartphones immer populärer. Chatbots, die Antworten automatisch erstellen, sind oft auf solchen Applikationen aufgebaut.

In dieser Arbeit liegt der Fokus darauf, einen Chatbots aufzubauen, der die Fähigkeit hat, Fragen über Filme zu beantworten. Die Informationen hierfür werden aus abgerufen. Wir untersuchen die wesentlichen Elemente, die für die Konstruktion eines Chatbots benötigt werden und analysieren verschiedene Ansätze, wie die Erinnerung des Gesprächskontexts gewährleistet werden kann und die auf regelbasierte Lösungen aufgebaut werden können.

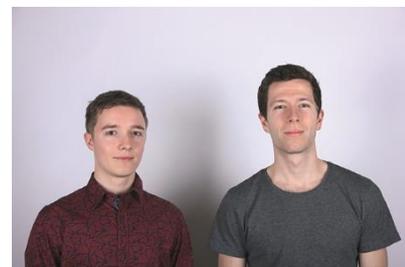
Das Abrufen relevanter Informationen aus der relationalen Movie-Datenbank ist eine anspruchsvolle Aufgabe, da die natürliche Sprache auf eine SQL-Abfrage abgebildet werden muss. Dafür wird die relationale Datenbank in ein RDF-Format umgewandelt. Dies ist eine graphenorientierte Datenbank, die semantische Sachverhalte enthält. Durch diese Umwandlung werden die Daten in einer Form repräsentiert, die der natürlichen Sprache ähnlicher ist. Dies macht das Abfragen einfacher.

Des Weiteren wird eine Lösung zur Erinnerung des Gesprächskontexts einschliesslich des vorherig migrierten RDF-Formats vorgeschlagen. Die Metadaten der semantischen Datenbank werden benutzt, um Wörter eines neuen Satzes dem aktuellen Kontext zuzuordnen.

Die Kombination dieser zwei Lösungen erlaubt das Generieren von qualitativ besseren Abfragen.

Aufgrund zeitlicher Einschränkungen und eines Mangels an geeigneten komparativen Datenbeständen war es nicht möglich, die vorgeschlagene Lösung ausführlich zu evaluieren. Trotzdem war es möglich, festzustellen, dass durch die Nutzung von Daten einer spezifischen Domäne in Form von Tripeln sowohl das Verständnis der Wörter als auch die Einsetzung der Ansätze zur Erinnerung des Gesprächskontexts vereinfacht wird. Ausserdem zeigte die Arbeit, dass die Auswertung von komplexen Verhältnissen in einem Graph rechenintensiv sein kann.

Um eine umsetzbare Lösung zu finden, ist eine Weiterentwicklung der Software notwendig.



Diplomierende
Sandro Panighetti
Matteo Togni

Dozent
Mark Cieliebak

I want to see the
trailer of Inception

Here is the trailer of
Inception (2010)



Ausschnitt einer Konversation mit
unserem Chatbot