Semantic Technologies for Automatic Question Answering

Siddharth Patwardhan
IBM T. J. Watson Research Center

Description:

This course will teach participants about question answering technologies like IBM Watson. We will cover in-depth the architecture of the system, and details of the key components within it. We will describe a general procedure for development of such a system, and implementation details. Students will gain an appreciation for working on a QA research project, and building a Watson-like system from scratch.

Program:

Day 1:

Overview and The QA System Architecture (DeepQA). We will briefly cover some background material, and dive right into the architecture of DeepQA system. We will discuss the various design decisions made along the way, and study the pros and cons of these decisions.

Day 2:

QA System Components. In this session, we will describe the various components running inside QA system. We will discuss, in-depth, the NLP and machine learning algorithms that power these components, and also study the impact of some of these on the QA system.

Day 3:

Implementing the System with UIMA and Scaling Out. This session will cover implementation details of the QA system. We will follow the process of building a QA system from scratch using the open source Apache UIMA framework. We will also discuss a strategy for leverging Apache UIMA-AS for scaling out the system over 1000s of cores.

Day 4:

Methodology for QA Research, and Adapting QA to New Domains. This class will describe a general research methodology that enables the analysis and development of QA system. It will go further and cover details of steps involved in adapting the system to new domains.

Day 5:

Advanced QA topics, Handling Special Questions and Commercial Applications. In the final session of this course, we will cover advanced topics, such as question decomposition, and handling special questions. A discussion of commercial applications of the technology will conclude the course.