Master’s Thesis Presentation

A User Interface for Semantic Full Text Search

by Florian Bäurle
Outline

- motivation
- search content and queries
- the user interface
- realization
- summary
Motivation

- search engines like Google, Yahoo or Bing are very popular
- queries are a number of keywords and results contain these keywords
- consider more intricate query: “Movies directed by Steven Spielberg that are about one of the world wars”
there also exist search engines that can process semantic information in queries:

- ontology only: powerful and efficient but limited to fact retrieval (e.g. RDF-3X, Sesame)
- combination ontology and full-text search: enable search in document collections with semantic queries (e.g. ESTER, CONTENTUS)
Motivation

- semantic query creation is a challenge: queries need a format that enables a search engine to understand the semantics
  - natural language queries: most intuitive and convenient for users but complicated and difficult to implement
  - special query languages (e.g. SPARQL): powerful but unintuitive and need special knowledge
  - special user interfaces: hide complex details without limiting the capabilities of the search engine

- we implemented a new special user interface
Outline

- motivation
- search content and queries
  - search content
  - query language
  - query results
- the user interface
- realization
- summary
search content: the type of data a search engine is designed to search in

dictates the kind of information that needs to be displayed to the user

our search engine: semantic full-text search in a text collection that is linked with an ontology

current prototype: English Wikipedia as text collection and YAGO as linked ontology
Query Language

- our queries can be described as trees with the following types of nodes:
  - class, entity, cooccurrence, relation, value, word
- some rules define how these types of nodes can be connected
- example query tree:
Query Results

- search engine can deliver two types of results for a query:
  - normal hits that consist of ontology facts and text documents
  - proposals for the different tree nodes for the query creation
Outline

- motivation
- search content and queries
- the user interface
  - overview
  - main features
  - live demo with examples
- realization
- summary
UI - Overview

Your Query:

- movie
  - directed (reversed)
  - Steven Spielberg
  - occurs-with
  - world war

Hits:

1 - 20 of 87 in 7 groups

- Saving Private Ryan
  - Saving Private Ryan is a movie and movie
  - Steven Spielberg directed Saving Private Ryan
  - Saving Private Ryan is a 1998 American war film set during the invasion of Normandy in World War II

- Saving Private Ryan [113]
  - For years now I've been looking for the right World War II story to shoot, and when Robert Redford wrote Saving Private Ryan, I found it

- Empire of the Sun (film)
  - Empire Of The Sun Film is a movie and movie
  - Steven Spielberg directed Empire Of The Sun Film

- Empire of the Sun (film) [114]
  - Other topics that Spielberg previously dealt with, and are presented in Empire of the Sun, include a child being separated from his parents (The Sugarland Express, Close Encounters of the Third Kind, E.T. the Extra-Terrestrial, and Poltergeist) and World War II topics

- Schindler's List
  - Schindler's List is a movie and movie
  - Steven Spielberg directed Schindler's List

UI – Main Features

- interactive and proactive:
  - rich internet application that uses JavaScript and Ajax
  - no search button, input is processed automatically
  - asynchronous reloading of information

- the proposal boxes:
  - the key why users do not need knowledge about the underlying ontology
  - context sensitive to the current query
  - can be filtered with the help of the input field
  - color-coded by type
UI – Main Features

- the query panel:
  - our advanced breadcrumbs display
  - displays the current query tree
  - can be used to refine the query (add new nodes, remove nodes, replace nodes …)

- the hits area:
  - displays the hits for the current query
  - groups hits by entities if possible
  - shows a Wikipedia article image for each group if possible
UI – Live Demo

- example query:
  “Movies directed by Steven Spielberg that are about one of the world wars”
Outline

- motivation
- search content and queries
- the user interface
- realization
- summary
Realization

- implemented with Google Web Toolkit (GWT)
- using GWT applications are programmed in Java and compiled into JavaScript
- advantages:
  - fully object oriented programming
  - comfortable programming and debugging with any Java IDE
  - code optimizations at compile time
  - code reuse for java server applications and easy client-server communication
Realization

- three-tier client-server architecture:

![Diagram of three-tier client-server architecture]

- Browser Client
  - JavaScript Application (Front End)
    - GWT Ajax Engine
  - GWT RPC Request
  - HTTP Transport
  - GWT RPC Response

- Java Servlet (Middle End)
  - Search Query
  - XML Result

- Search Server (Back End)

**Server-Side Systems**
Outline

- motivation
- search content and queries
- the user interface
- realization

- summary
Summary

- motivation why special semantic search user interfaces are needed
- the search content and queries for which the user interface was built
- overview of the created user interface including its functionality and a live demo
- basic realization of the user interface using GWT
The End

- thank you for your attention!

- any questions?