Hapless Path XPath Evaluator

Eliza Brock

May 16, 2008

Premise

Pedagogical tool to illustrate the basic principles of querying XML using XPath.

Premise

Pedagogical tool to illustrate the basic principles of querying XML using XPath.

- Browser based

Premise

Pedagogical tool to illustrate the basic principles of querying XML using XPath.

- Browser based
- Dynamic visualization based on sub-queries

Premise

Pedagogical tool to illustrate the basic principles of querying XML using XPath.

- Browser based
- Dynamic visualization based on sub-queries
- Written in Haskell

Tools

HAppS "Haskell Application Server" Features auto recompilation, optional state storage, and is designed for the "pattern of developing static web pages and using AJAX to populate them with dynamic content"

Tools

HAppS "Haskell Application Server" Features auto recompilation, optional state storage, and is designed for the "pattern of developing static web pages and using AJAX to populate them with dynamic content"

Parsec "industrial strength, monadic parser combinator library for Haskell. " [2] i.e. entirely bad-ass library to facilitate writing parsers in Haskell.

Tools

HAppS "Haskell Application Server" Features auto recompilation, optional state storage, and is designed for the "pattern of developing static web pages and using AJAX to populate them with dynamic content"

Parsec "industrial strength, monadic parser combinator library for Haskell. " [2] i.e. entirely bad-ass library to facilitate writing parsers in Haskell.

AJAX XMLHttpRequest. Magic.

Workflow

Client	Server
User selects XML file to use	
	Graphical representation of the XML is generated
User inputs XPath, which is sent to the server on a per-token basis	
	Identifiers of elements selected by individual subexpression are returned to the client
The browser highlights the elements and corresponding expressions	

Approach

- 3 pass approach
  1. Parse XML
  2. Parse XPath
  3. Evaluate them together

Approach

- 3 pass approach
  1. Parse XML
  2. Parse XPath
  3. Evaluate them together
- Not as bad as it sounds
  - Haskell is lazy
  - Haskell remembers things

Papers/Resources


Scott Boag.
Building a tokenizer for XPath or XQuery.
Technical report, W3C, April 2005.


Daan Leijen.
Parsec, a fast combinator parser.
University of Utrecht, October 2001.


W3C.
XML Path Language (XPath), 1.0 edition, November 1999.


W3C.
Extensible Markup Language (XML) 1.0, 4.0 edition, September 2006.

Implemented Features

- XML
  - elements with children
  - empty elements
  - comments
  - doctype declarations
  - mixed content
  - attributes

Implemented Features

- XML
  - elements with children
  - empty elements
  - comments
  - doctype declarations
  - mixed content
  - attributes
- XPath
  - abbreviated location paths (except //)
  - function calls
  - predicates

Implemented Features

- XML
  - elements with children
  - empty elements
  - comments
  - doctype declarations
  - mixed content
  - attributes
- XPath
  - abbreviated location paths (except //)
  - function calls
  - predicates
- Evaluation
  - location steps
  - predicate tests
  - a subset of the function calls that result in nodes or node-sets

Demonstration