AI Related Ruby Extensions

June 10, 2008

This page will maintain list of AI related libraries for the [Ruby] programming language. Please contact me if you know something I missed.

While maintaining this list, I have 1 induced a theory: If a project’s first public appearance is documentation without code, code will not appear before the heat death of the universe.

1 Natural Language / Linguistic Tools

These libraries are hard to put into categories; if you can’t find what you’re first looking for, check related categories. NLP tools are for English unless otherwise noted.

1.1 Morphology, Tagging and Parsing

- [English] - a nice collection of morphological tools, including Porter stemming, computing phonetic similarity, obfuscating and handling inflection.

- [English tokenizer]

- [English Lemmatizer]

- [Linguistics 1.0.5] —the module contains an interface to the parser and “it includes an English-language module with pluralization, conjunctions, indefinite articles, present participles, ordinal numbers, numbers to words, general quantification, integration with WordNet and

---

1 Induction by analogy. It’s a corollary of the “Webpage Under-construction Law” from Eric Raymond’s [HTML-Hell]
CMU’s LinkGrammar, as well as a framework for providing modules for other languages.” Documentation [here](#).

- **Ruby Utilities** a version of the Brill part of speech tagger with a topic-jister /categorization tool. Version 0.1.1, GPL.

### 1.2 Machine Translation

- **NLP Software** —Several programs and tools for MT (IBM Model1, min error rate, corpora alignment, max entropy modeling)

### 1.3 Semantic Parsing

- **Shalmaneser** —a shallow semantic parser that provides a toolchain (tagging => parsing => pred-arg extraction => semantic role labeling) for labeling sentences semantic roles to FrameNet’s frame-semantic corpus. English and German.

### 1.4 Lexical-Semantic Corpora

- **Ruby-WordNet 0.04** —interface for WordNet database. Requires Berkeley DB (2,3,4) and database interface, Ruby-BDB (special rpm for Red Hat).

- **RubyCon** —A port of ConceptNet semantic network that permits saving concepts.

### 2 Machine Learning

Most machine learning techniques are resource intensive, so you’ll probably want to use Ruby to talk to a binary. If you can’t find what you’re looking for, consider creating a Ruby wrapper for an exciting library using SWIG, putting it online and letting me know.

- **Decision Tree Learning** [An good explanation](#) and sample code for learning decision trees, which can be thought of learning categories where it is important to keep track of the order of the features. The AI4R package contains the ID3 decision tree learning algorithms, and will soon include Ross Quinlan’s C4.5.

- **Clusterer** —“Implements various clustering algorithms for text mining, such as K-Means, Hierarchical clustering, etc.”

• **Ruby SVM**—Ruby bindings to *libsvm* support vector machine library, for classification.

• **Bayesian Classification**—There’s a stemmer/LSA classifier for classifying natural language (with a usage example), and a more general classifier, Bishop related

• **Singular Value Decomposition (SVD)**, a linear algebraic technique for reducing the number of dimensions in your data.

• **Neural Networks**—[Ruby-FANN](http://en.wikipedia.org/wiki/Ruby-FANN) or [AI4R](http://en.wikipedia.org/wiki/AI4R)

3 **Optimization and Search**

Genetic algorithms, searching trees, etc

• **Gecoder/R** is a Ruby interface to [Gecode](http://www.gecode.org), an open source constraint programming library.

• **Directed Ruby Programming (DRP)** is a generative programming technique that uses grammatical evolution (genetic algorithms with a grammar) to solve problems.

• **AI4R** supports genetic algorithms.

• **AI Apps** - Ruby implementations of various AI algorithms, including GA, n-Queens, hill climbing, neural nets with backpropagation, an expert system and bigram parsing.

4 **Knowledge Representation and the Semantic Web**

• **ActiveRDF** is a library for accessing RDF data and can be used as a data layer in Ruby on Rails. Supports SPARQL endpoints, Redland and RDFLib; (native) adapters for Jena, Sesame2 and YARS are in development.

5 **Other Useful Libraries**

• **Ruby GSL** - Ruby GNU Scientific Library that includes loads of mathematical and statistical tools.

• **Data Structures and Algorithms in Ruby**
6 Resources

Helpful websites:

- **RubyForge** - Ruby SourceForge
- **Ruby-doc.org** - a portal for things about Ruby.
- **Ruby-lang.org** - the official Ruby headquarters, lots of resources.
- **RubyCocoa Resources** - a website tutorial of programming is the OS X Cocoa object layer.
- **ConceptNet** - my research group's semantic knowledge base, representing common sense assertions as binary relations. Example Ruby use:

```ruby
require 'rubygems'
require 'active_record'
ActiveRecord::Base.establish_connection(
  :adapter => "postgresql",
  :host => "your.server.edu",
  :username => "",
  :password => "",
  :database => "commons"
)

class Predicate < ActiveRecord::Base; ; end

class Stem < ActiveRecord::Base; ; end
```