Starkie Enterprises

Next Generation Search Research

Brad Starkie

HCSNet Priority Area Workshop on Next-Generation Search Technology

21st-22nd September 2006

Who Is Starkie Enterprises?

- Starkie Enterprises is Brad Starkie
  - 16 Years R&D experience, numerous patents, international & national product awards
- Founded December 24 2005
- Products
  - Speech Recognition Development Tools
  - License Management
  - Image Processing
  - Question Answering Systems (under development)

http://www.starkieenterprises.com

Starkie QA System

- Based upon reduction systems
- Sentences are reduced to normal forms
- Ideally two sentences reduce to the same normal form iff they have the same meaning.

Research Objectives

- Construct reduction systems via machine learning
- Make them efficient (Scalable to 10 billion pages)

Potential For Collaboration

- Software Components
  - Goodwood parser
- Knowledge representation schema

Good Wood Parser

- Developed a method of inferring Deterministic Context-Sensitive Grammar from treebanks in linear time.
- Parses text in linear time
  - Word sense disambiguation
- 700× Faster than Charniak Parser
  - Parse 10 billion pages in 3 months on 16 pcs.

Knowledge-Representation

- Modified Higher Order Logic
- Associative database of answers, questions and named entities
- Database insertion & extraction times same as for bag of words
- Viewed as Trees, Terms or Objects, depending upon node labels

Example Fact

John ate pasta with a spoon.

1. Sentence

```
(EVENT
  (PERSON_MALE
    (FIRSTNAME_MALE John))
  (VERB
    (VBD ate))
  (NP
    (NN pasta))
  (PP
    (NTERM_IN with)
    (NP
      (DT a)
      (NN spoon)))
  (NTERM_DOT .))
```

2. Parse Tree

```
(EVENT
  (subject
    (OBJECT
      (first_name john)
      (is_a noun.person.man … …
          noun.Tops.entity)
      (id X2)))
  (VERB
    (root ate)(tense past))
  (object
    (OBJECT
      (is_a noun.food.pasta …)(id
X2))
    (with
      (OBJECT
        (is_a noun.artefact.spoon …)
        (id X3)))))
```

3.Knowledge Representation

Example Queries

If database contains the fact ..

John ate ravioli with a spoon.

..then questions like..

Did john eat pasta?

Did john eat pasta or pizza?

What did john eat?

..can be answered by a simple efficient database query.

Summary

- Standardised knowledge representation
- Transformation of existing corpora or creation of new corpora
- Re-use of software components
- Sub contracting of development tasks

Key problem – Equivalence of Sentences

- Equivalence of words
- Resolving tense and aspect
- Arbitrary syntactic order
- Matching positive & negative expressions
- Passive ↔ active
- Contracted ↔ uncontracted
- Declarative ↔ questions
- Idioms
- Removing unnecessary information

Formalisms

- Deterministic Context-Sensitive Grammars
- Syntax Directed Translation Schema
  - Unification Grammars
- Term Rewriting Systems