

**Mark Lee**  
*mgl@cs.bham.ac.uk*

<http://www.cs.bham.ac.uk/~mgl/student-projects.html>

I'm available to supervise both UG/MSc. projects

### **Preferred Project Areas**

- Natural Language Processing
- E-learning Support & Computer Aided Learning
- Automated Theorem Proving

NLP =/ Prolog!

NLP > Parsing!

# Applied NLP

- Many applications process Text
  - Websearch
  - Text classification
  - Text prediction/Spelling Correction
  - Automatic Indexing/Intelligent Data Fusion
- Evaluation Competitions
  - International Challenges
  - Clear evaluation of techniques
  - Plenty of useful data and tasks
- Toolkits
  - Why re-invent the wheel when instead we can do something useful?
  - NLTK – open source Python toolkit

# Opinion Mining

Most products are now reviewed on the web

Can we automatically classify opinions in terms of emotion?

Can we Summarize what people are saying about a product?

A naïve approach would be just to count emotional words

(e.g. good/bad etc.)

“Nigerian Hostages feared dead rescued”

“We were looking forward to this hotel as we have been to Egypt several times and looked forward to see the temples etc. recreated but it turned out to be more of a tomb.”

Opinion Mining is currently an *Industry hot topic*

# Question Answering

Websearch = query + search engine → documents

Often we just want an answer – not more documents!

- Who won the last world cup?
- How many Oscars did the English Patient win?
- Which is the largest landlocked country?

Question Answering = Question + Text Collection (internet) → Answer

- Heavily evaluated via research competitions
- Much current commercial internet
- Interesting Applications (e.g. mobile phones)

## Web 2.0 (plus) & NLP

Many Web 2.0 applications are text based

facebook

twitter

digg

RSS feeds etc.

Web 3.0 promises much more (if you're a believer)

I'm interested in projects which use NLP techniques to do interesting things with this data

Social networks analysis

Gossip generators etc.

Data Fusion & Twitter etc.

# Anything Else?

If you have any other ideas then it's best to email me

Please describe the project - if you can't write a paragraph on it then it's probably not worth doing.

*mgl@cs.bham.ac.uk*