JoC Topic 4: Computational Social Science

Joshua Chia
u5024740@anu.edu.au
Computational Social Science

• Relevance of CS to other disciplines
• Introduce concept of Sentiment Analysis
• Teach CS fundamentals:
  – Data structures (map, set)
  – Hash functions
  – Peek at related concepts
Making sense of social media

• #sandy Tweets from 15 Oct – 12 Nov 2012
Making sense of social media

- #sandy Tweets from 15 Oct – 12 Nov 2012
Tutorial Overview

• Analyse the sentiment of #sandy tweets by looking up words in a sentiment lexicon:
  – Implement hash set insert, contains operators
  – Implementing hashing scheme for strings

• For advanced students:
  – Remove operator, hash map, rehashing
Demo

• Student support code
• Visualisation of sentiment polarity
Challenges

• Handling 1.5M tweets
  – Pre-processing, Cleaning
  – Reverse geocoding
  – Extending Snap! With Regular Expressions
  – Web service

• Working with limited memory
  – Sentiment lexicon
  – Visualisation
Reducing Lexicon Size

- Limited to approximately ~24 words/sec
- From 6789 to 288 words
Challenges

- Clear and well motivated explanations
  - Analogies
  - Discussion Questions
- Simple and elegant code
  - Higher-order functions
  - Object-Orientated Design
  - Multithreading
Closing

• Reflection of my own learning experiences and skills gained at ANU
• Opportunity to give back
• Questions?