

# Homework Project #1: I.I.D. vs. 1st-order Markov Models

---

Due Date: 10/9/2014

# Markov Model for Text Compression

---

- The data source is the text file, jokes.txt, available from the E3 course website
    - The alphabet of the source is composed of 26 letters, A-Z, and the space character.
  - Follow Example 2.3.1 and write a program to estimate the entropies of the source based on:
    - I.I.D. model
    - First-order Markov model
  - Compare the estimated entropies with the average bits/symbol achieved by a "zip" or "rar" compression program
-

# Hand-in for the Homework

---

- Please write a report telling me what you have learned from this experiment. The report should be in pdf format, up to 4 pages.
    - You shall email me the report directly to [cjtsai@cs.nctu.edu.tw](mailto:cjtsai@cs.nctu.edu.tw).
    - Please don't email me your program.
  - Grading is based on
    - Your understanding of the impact of models on estimated entropy.
    - Any bright ideas you can think of to extend this simple experiment.
    - The organization of your report.
-