

# Presentations & Assignments

## ICS 482 Natural Language Processing



**LECTURE 16: PRESENTATIONS &  
ASSIGNMENTS**  
**HUSNI AL-MUHTASEB**

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ

# ICS 482 Natural Language Processing



**LECTURE 16: PRESENTATIONS &  
ASSIGNMENTS**  
**HUSNI AL-MUHTASEB**

# Assignment 4 – A Paper on Arabic NLP



- search for a recent research paper on Natural Arabic processing
- published within the previous 3 years
- pass the title, author, abstract and references to instructor
- When approved announce your topic under the discussion list ‘Presentations’.
  - Paper title
  - outline
  - abstract
- Submit the same on the assignment section

# Assignment 4 – A Paper on Arabic NLP



- The paper could be on any topic of Arabic NLP:
  - Arabic Morphology
  - Arabic N-Grams
  - Arabic Parts of Speech
  - Arabic Syntax & grammar Parsing
  - Arabic Semantic analysis
  - English – Arabic Machine Translation
  - Arabic Information Extraction
  - Arabic web search
  - Arabic NLP Applications

# Assignment 4 – A Paper on Arabic NLP



- Get the reference in hand
- A reference you can understand
  - A person cannot present a topic that he does not comprehend
- The easier the topic, the more interesting the presentation
- Related to Arabic NLP
- Recent paper
  - published 2003 or later
- If describing a tool, get the tool and try it
- Prepare an outline for your presentation

# Assignment 5 -Corpora Bigrams Statistics



- Complete corpora for targeted size
- A program to produce bigram frequencies and probabilities for all tokens in your corpora
- Each entry of the table should look like:
  - Token PreviousToken frequency Probability

# Assignment 5 -Corpora Bigrams Statistics



- **Submit**
  - the program
  - the resulted frequency-probability table in a file
  - corpora statistics including
    - ✦ total number of tokens
    - ✦ vocabulary size (number of types)
    - ✦ most used 100 words in your corpora along with frequencies and probabilities
- A document file describing your corpora and results
- A sample program for assignment 3 is included
  - Thank you Ahmad.

# Assignment 6 - Arabic Part of Speech Tagging - Give it a thought



- Planning to use Al-Usaimi's tagset
  - given in the lecture area of webct
  - to tag each word in your corpora by its corresponding part of speech (POS)
  - Manually and/ or Semi Manually and/ or Automatically
- **DON'T Start tagging now 😊**



# Assignment 6 - Arabic Part of Speech Tagging - Give it a thought



- review tagging procedures
- suggest a suitable way that we can use for tagging
  - How could we benefit from individual work for this purpose
- Submit a document file that includes your suggestions, ideas, and plans.