بسم الله الرحمن الرحيم

ICS 482 Natural Language Processing

Lecture 24: Project Ideas + Students Presentations

Husni Al-Muhtaseb

Students Presentations

- Evaluation at WebCT
- Presentation file and Document word file
  - Submit through webct assignment
    - ‘Presentation Material’
  - as soon as you present your lecture

Student Presentation Files

- Summary of your lecture
- References
- Obstacles faced
- Things learned and skills practiced
- recommendation
- Three true/ false questions
  - addressing the main issues with answers for possible inclusion in the exam

Students Presentations

- Evaluation at WebCT
- Presentation file and Summary word file
  - Submit through webct assignment
    - 'Presentation Material'
  - as soon as you present your lecture
- Sunday, May 13
  - Al-Elaiwi Moh'd - Diacritization: A Challenge to Arabic Treebank Annotation and Parsing
  - Naif Al-Abdulhay - The Challenge Of Arabic For NLP/MT
  - Abdul Rahman Al Khaldi - Statistical Transliteration for English-Arabic Cross

Online Instructor Evaluation

- Reminder:
  - Online Course and Instructor Evaluation

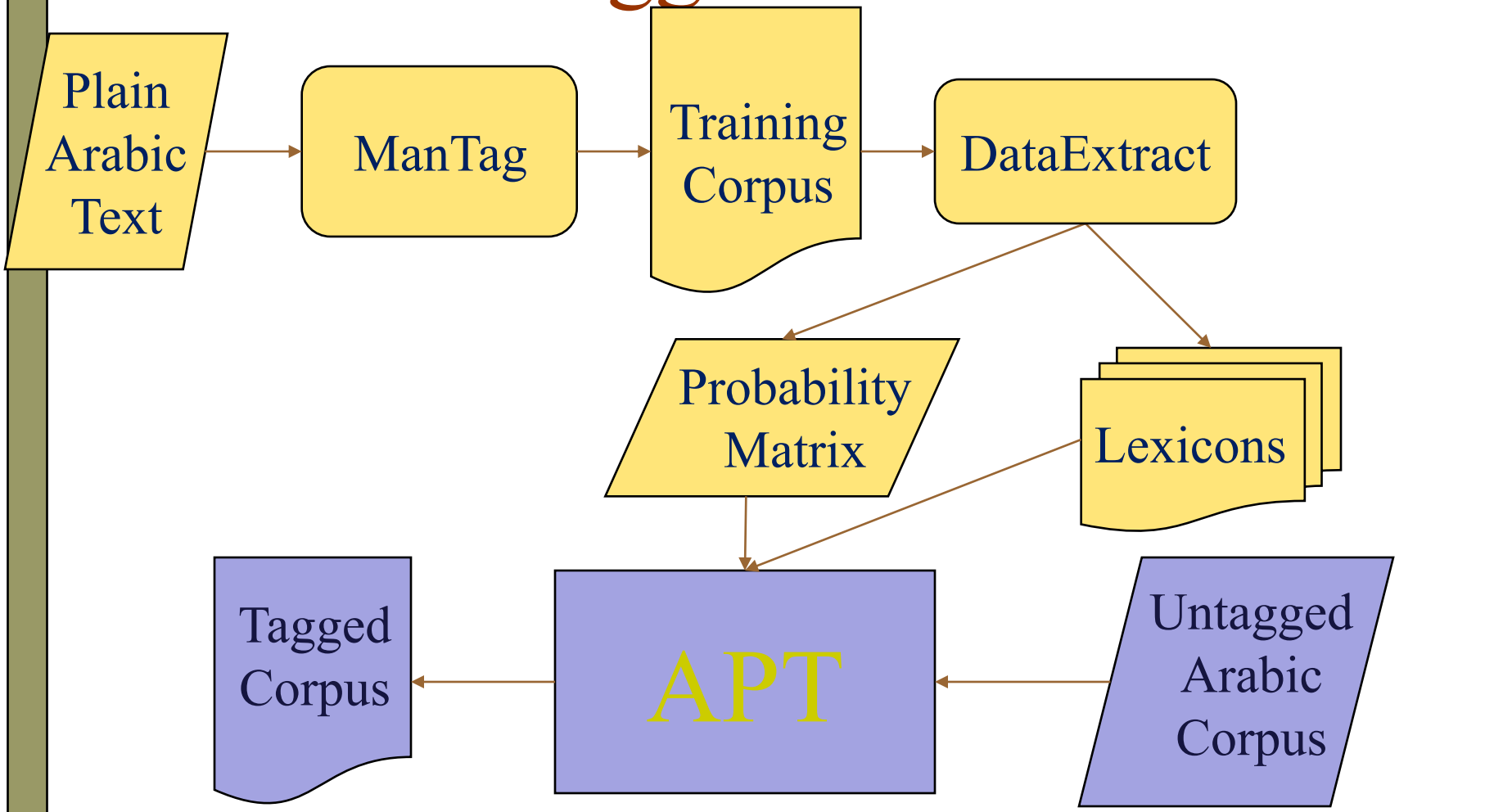
Some Projects

- Arabic POS Tagger
- Specific Information Picker
- An Arabic morphological analyzer
- An Arabic Spell checker w/ morphology analysis
- An Arabic Syntax analyzer
- Random syntactically-correct Arabic sentence generator

Some more Projects

- An English to Arabic machine translation using word re-ordering
- Moh'd Ones on WebCt

Arabic POS Tagger



Specific Information Picker

- Given a list of links to English newspapers/ sites, find all pages that are talking about Saudi Arabia and send them to an automatic translator

An Arabic morphological analyzer

- Invited Lecture by Yousuf Al-Uraini
- Implementation of Buckwalter's AraMorph on C#
- Code in Java is available at supplementary material of the website

An Arabic Spell checker

- Spell checker for Arabic text including morphology analysis

An Arabic Syntax analyzer

- We Have comprehensive rules such as

الجملة البحتة الاسمية البسيطة (ب ا ب):

تعـريـف: هي جملة صحيحة التركيب تبدأ باسم وتخلو من أية أداة وتشكل جملة واحدة فقط

التــركــيـب:

<ب ا ب> ← <مبتدأ (حالة ، جنس)> <خبر حالة جنس)>

<أسلوب إغراء>

<أسلوب تحذير>

Syntax analyzer – Example of rules

<حالة = ح> ←

<إفراد>

<تثنية>

<جمع>

<جنس = س> ←

<تذكير>

<تأنيث>

Syntax analyzer – Example of rules

<مبتدأ ح،س> ←

<اسم علم ح،س> <مبتدأ ح،س غير علم>

<اسم إشارة ح،س> <مبتدأ ح،س غير إشارة>

<اسم ح،س معرف يأل> <مبتدأ ح،س>

<مضاف> <مبتدأ ح،س>

<ضمير مفرد ح،س> <مبتدأ ح،س غير ضمير>

<اسم من الأسماء الستة ح،س> <مبتدأ ح،س ليس من الأسماء الستة>

<اسم عدد من ثلاثة إلى عشرة> <تمييز جمع>

Random syntactically-correct Arabic sentence generator

- Randomly generate Arabic sentences that have correct syntax according to the given grammar

English to Arabic machine translator

- Given an English sentence translate to Arabic using lexicon and word ordering technique
- Example next page

English to Arabic Machine Translation

- Salma came
- Lexicon
  - Salma:.. سلمى، اسم علم، مؤنث، مفرد،
  - Came: ... جاء، فعل، ماض، متعادل
- Word to word: سلمى جاء
- Needed Translation: جاءت سلمى
- Modification Rules
  - Exchange the positions of subject and verb
  - If the gender is feminine the verb should be the same

Moh'd Ones on WebCt

- Generates complete well-formed sentences from a user input of uninflected content words .e.g. like you give the system verb, noun, ...etc and the system will construct the sentence
- Makes use of a bilingual corpus to perform word sense disambiguation
- Detects proper nouns, nouns, verbs, etc... in text(s)
- Intelligent Word Prediction uses knowledge of syntax and word frequencies to predict the next word in a sentence as the sentence is being entered, and updates this prediction as the word is typed

Thank you

- السلام عليكم ورحمة الله