

Research Projects

Contents

- 1 Exploiting Treebank Resources for Extending a Rule-Based Arabic Parser Coverage
- 2 Using Machine Learning to Improve Arabic Named Entity Recognition
- 3 Arabic English Conceptual Interlingua (AECI)
- 4 Context-Aware Multimedia Information System for e-Health Decision Support
- 5 Arabic Lexical Error Diagnosis of Language Learners in Intelligent Language Tutoring Systems
- 6 Named entity recognition for Arabic (NERA)
- 7 Machine Translation of Spoken Arabic into English
- 8 Expert System Integrated with Multimedia information system for improved Grapes management (GRAPEX)

| | |
|--------------------------------|---|
| <p>Dec. 2011 - August 2014</p> | <p>Principal Investigator</p> <p>Exploiting Treebank Resources for Extending a Rule-Based Arabic Parser Coverage</p> <ul style="list-style-type: none"> ▪ Research Fellow (Sept. 2012-March 2013): Prof. Ali Farghaly ▪ Research Fellow (Feb. 2012-August 2012): Dr. Mohammed Attia (http://www.attiaspace.com/). ▪ Funding Source: National Research Foundation (NRF), UAE. ▪ Amount: 200,000 AED (UAE Dirham) ▪ Description: The aim of this research is to formalize reusable linguistic knowledge of Arabic that can be utilized in a wide range of applications. We will develop a syntactic parser for Arabic within the framework of the constraint-based formalisms of LFG (Lexical Functional Grammar) using XLE (Xerox Linguistics Environment). Our approach is to use the Penn Arabic Treebank (ATB) resources to enrich and guide the manual development of Arabic deep syntactic parser by extracting statistical information on the distribution of syntactic structures as well as increasing the coverage of the morphological analyser by harvesting lexical entries from the ATB. ▪ Outcome: In progress. |
| <p>April 2010 - Sept. 2011</p> | <p>Principal Investigator</p> <p>Using Machine Learning to Improve Arabic Named Entity Recognition</p> <ul style="list-style-type: none"> ▪ Co-Principal Investigator: Dr Sherief Abdallah (http://dis.cs.umass.edu/~shario/), The British University in Dubai, UAE. ▪ Funding Source: The British University in Dubai, UAE. ▪ Amount: 68,000 AED (UAE Dirham) ▪ Description: The objective is to improve the performance of our rule-based Arabic Named Entity Recognition tool by employing a machine |

| | |
|-------------------------|--|
| | <p>learning approach. Another objective is to automate the acquisition of the recognition rules.</p> <ul style="list-style-type: none"> ▪ Outcome: one book chapter (LNCS). |
| April 2010 - Sept. 2011 | <p>Principal Investigator</p> <p>Arabic English Conceptual Interlingua (AECI)</p> <ul style="list-style-type: none"> ▪ Co-Principal Investigator: <u>Dr. Farhad Oroumchian</u>, (http://works.bepress.com/farhadoroumchian/), University of Wollongong in Dubai (UOWD), UAE. ▪ Funding Source: The British University in Dubai, UAE. ▪ Amount: 50,000 AED (UAE Dirham) ▪ Description: A major part of a Cross Language Information Retrieval system is creating a mapping between words of different languages. This mapping enables a search engine to convert a query from a language to another language. In this research, we will create an Arabic-English Conceptual Interlingua (AECI). AECI will contain Arabic and English concepts and relations (i.e., Part-of, kind-of, ISA). Moreover it will contain the similarity relationships among the Arabic and English concepts. We will show the usefulness of AECI in cross language information retrieval. ▪ Outcome: In progress. |
| May 2009 - May 2011 | <p>Co-Principal Investigator</p> <p><u>Context-Aware Multimedia Information System for e-Health Decision Support</u></p> <ul style="list-style-type: none"> ▪ Principal Investigator: <u>Dr. Hissam Tawfik</u>, (http://www.hope.ac.uk/staff/tawfikh.html), Liverpool Hope University, UK. ▪ Funding Source: The UK Prime Minister's Initiative - PM12 Connect: Research Co-operation in the Gulf States. ▪ Amount: £33225 (UK Pound) ▪ Description: The objective is to develop a context-aware multimedia information system to support collaborative medical decision making in an e-Health environment ▪ Outcome: Three international conference publications and one Msc thesis. |
| Feb. 2008 - April 2009 | <p>Principal Investigator</p> <p>Arabic Lexical Error Diagnosis of Language Learners in Intelligent Language Tutoring Systems</p> <ul style="list-style-type: none"> ▪ Funding Source: The British University in Dubai, UAE. ▪ Amount: 41,500 AED (UAE Dirham) |

| | |
|-------------------------------------|---|
| | <ul style="list-style-type: none"> ▪ Description: The objective is to develop an automated lexical error diagnosis system, which is an integral part in the development of intelligent computer-assisted language learning system of Arabic. ▪ Outcome: final project report, one Msc thesis, four international conference publications, and one book chapter. |
| <p>Sept. 2006- August. 2007</p> | <p>Principal Investigator & consultant</p> <p>Named entity recognition for Arabic (NERA)</p> <ul style="list-style-type: none"> ▪ Funding Source: Faculty of Informatics, BUiD, UAE and FAST Search & Transfer Inc., Oslo Atrium, Christian Frederiks plass 6. N-0154 Oslo, Norway. (<u>FAST is a world leading search provider company recently acquired by Microsoft©</u>). ▪ Amount: 45,000 USD ▪ Description: The objective is to recognize names such as person, company and locations names from large scale documents. ▪ Outcome: eleven named entity products, eight technical reports, one master thesis, two international conference publications, and one international journal publication. |
| <p>March 2001 - Nov. 2004</p> | <p>Co-Principal Investigator (Cairo University, Egypt)</p> <p><u>Machine Translation of Spoken Arabic into English</u></p> <ul style="list-style-type: none"> ▪ Co-Principal Investigator (CMU, USA): <u>Dr.Lori Levin</u>, (http://www.cs.cmu.edu/~lsl/), Carnegie Mellon University (CMU). ▪ Funding Source: Joint Fund, Ministry of Scientific Research (Egypt)/Cairo University and NSF foundation/Carnegie Mellon University (CMU) (NSF Project Number: INT-0001613 USA & US-Egypt Joint Science and Technology Board Project Number: INF4001023). ▪ Amount: 50,000 USD ▪ Description: This is a joint project of Carnegie Mellon University (CMU) in the United States and Cairo University in Egypt. The objective of this project is automated computer translation of spoken Arabic into English, French, Italian, German, and Korean, and also translation from these languages into Arabic. Translation is limited to sentences about making travel plans (for example, hotel and plane reservations). ▪ Outcome: four annual progress reports, two Ph. D dissertations, and 5 international conference papers, 1 international Journal, and 1 book chapter. |
| <p>April 1996 - March</p> | <p>Principal Investigator</p> |

1997

Expert System Integrated with Multimedia information system for improved Grapes management (GRAPEX)

- **Funding Source:** Agricultural Technology Utilization and Transfer (ATUT), US-AID, Ministry of Agriculture and Land Reclamation, Egypt.
- **Amount:** 50,000 LE (Egyptian Pounds)
- **Description:** The objective is to develop an expert system integrated with multimedia information system that provides growers with advice and recommendations about grapes and agricultural operations, and enhances the capacity of personal working at these units and assists them in their decisions fastly and accurately.
- **Outcomet:** An expert systems package for improved Grapes management (GRAPEX), a final technical report, a conference paper.