**Technology:** Question Answering

1. **Brief Overview**

Question answering (QA) can be defined as the task of retrieving concise information that is relevant to a specific question posed in natural language. English QA systems date as far back as to 1960, but despite the fact that the concept of QA is quite old, work in that area is very much an active one with many challenges that need to be overcome. The main challenge in QA lies in understanding what is meant by a question and mapping that to a piece or to pieces of information that corresponds to an answer. The holy grail of information retrieval is being able to pose a question in natural language and to make use of all available resources to get the correct answer to that question in a timely fashion. Getting that answer may entail collecting information from various resources and integrating them intelligently. This goal however is yet to be realized (Roussinov et al, 2008). In the past QA systems were confined to a closed domain, but recent work aims to avail these systems in open domains such as the worldwide web. In an effort to advance the field, the Text Retrieval Conference (TREC) created a special track for QA in the late 1990s. The track’s focus is on open-domain based QA and last ran in 2007.

1. **State of the Art (For Latin Languages)**
   1. **Technology**

The technology used for QA draws from many other fields including but not limited to:

* Natural language processing (part-of-speech tagging is one of the most common components of a QA system)
* Traditional information retrieval (open systems are relying more and more on existing IR engines for retrieving a set of results that are then later filtered using more sophisticated techniques)
* Information Extraction (named entity recognition is being recognized as an essential part of any QA system)
* Text and Knowledge Mining
* The Semantic Web

* 1. Applications and Reported Performance

There are way too many Latin based QA applications to report in this pre-swot report. Recent work on English QA has focused on open domain, web based QA systems. The START system developed by MIT is one of the first web based question answering systems (START, 2006). First launched in 1993, the system has continued to evolve with its last publication being made in 2007 (a list of all START related publications can be found in its website). The START system uses natural language annotations to answer factoid like questions. When tested using the TREC-8’s 200 questions the NIST score of START as reported in (Zheng, 2002) was 14.5 % and its average response time was 9.84 seconds. AnswerBus (AnswerBus, 2008) is another an open-domain question answering system developed at the University of Michigan which accepts questions in English, German, Spanish, French, Italian and Portuguese and extracts possible answers from the Web. When tested using TREC-8’s questions, the NIST score of AnswerBus was 64.18% and its average response time was 7.2 seconds (Zheng, 2002).

Another interesting QA system that is currently being developed is QANUS which is being developed by the National University of Singapore (QANUS, 2010). The reason this QA system is interesting is that it is perhaps the only known open source question answering system and can thus serve as a possible starting point for the development of more sophisticated QA systems (will explore this further). There are number of other commercial web based QA system available examples of which include:

* Answers.com: <http://www.answers.com/>
* Blurtit: <http://www.blurtit.com/>
* Ask (formerly known as Ask Jeeves): <http://www.ask.com/>
* askED: ttp://asked.jp

So far, the author of this document has been able to find any performance measures or comparisons between these systems.

1. **State of the Art (For Arabic Language)**
   1. Technology and Future Trends

The technology used for building an Arabic QA systems is more or less the same as that used for latin languages. However each technology that is employed has to be adjusted to work with Arabic.

* 1. Current and Envisioned Applications and Market Priorities

Current work on Arabic QA systems can be considered as lagging behind compared to its English counterparts. AQAS is one of the earliest Arabic QA systems (Mohammed et al, 1993). The system is a knowledge based one. It accepts Arabic queries that follow pre-defined rules and matches these against frames in a knowledge base. The system was applied to the radiation field and is thus a closed domain one. The architecture of the system is similar to early English based QA systems. No experimental results have been reported for this system.

QARAB (Hammo el al, 2002) is a more recent QA system that uses text extracted from the Arabic newspaper “Al-Raya” to answer natural language queries posed by users. As such it can be considered a closed system. QARAB retrieves short passages that are likely to contain an answer to the user’s query rather than extracting the exact answer. It also assumes that the answer to a query posed by the user is contained in its text collection and that the answer can only be found in one document in that collection. These assumptions would not be valid in an open domain. The system uses a combination of IR and NLP techniques to achieve its task. Again, no experimental results have been presented for this system.

ArabiQA (Benajiba et al, 2007) is a factoid centered Arabic QA system. The system makes use of the Java Information Retrieval System, a Passage Retrieval system, a Named Entities Recognition module and an Answer Extraction module. The authors of this system report a precision of 83.3% over a manually created test dataset the details of which are not given. The system appears to have been designed for an open domain, but has not been tested in such an environment.

In (Abouenour el al, 2008), the authors present a method for expanding natural language queries using Arabic wordnet in order to improve the results of a QA system. Rather than actually using a QA system in their evaluation, the authors feed in the expanded and un-expanded forms of the query into the Google search engine, and demonstrate that using their query expansion module, the accuracy of the returned results are improved. The statistical significance of the returned results has not been tested. In (Abouenour el al, 2009), the same authors build on this work and on the work reported in (Benajiba et al, 2007) again to improve IR results. To test their system, they’ve used a set of translated TREC and CLEF questions. In their work, they demonstrate improved accuracy over available keyword based search engines. However the response time of their system compared to these engines is not reported and neither is the statistical significance of the improvements.

(Kanaan et al, 2009) present another Arabic QA system which makes use of data redundancy rather than complicated linguistic analyses of either questions or candidate answers, to achieve its task. The author test their system using a collection of 25 documents gathered from the Internet, 12 queries. It is not clear in their work, why and how these documents and queries, which are considerably small in size, have been selected. When analyzing their results, the authors state that these are close to the results of the traditional vector space model, which really negates the value of the presented work. The authors have not compared their results to any previously developed Arabic QA system.

At the end of 2009, Google launched its Beta Arabic question answering system Ejabat (Google’s Ejabat, 2009). The move came after observing that “many of its Arabic users' searches failed to turn up relevant results. The Mountain View, Calif.-based company estimates that less than 1 percent of information online is in Arabic” (Google Unveils Ejabat, 2009). Details of this system are not known, and neither is its performance.

Because of the immaturity of the technology, there are little commercial applications of Arabic QA. The application of QA within companies, or organizations with large document bases can be of great value. The Egyptian government, army, and judicial sector with their huge Arabic document bases are examples of potential clients for such an application. Arabic OCR might be needed to convert these documents to a textual database. Building an Arabic Web based QA system is yet another potential application. Further investigation of potential applications should be carried out in the SWOT analysis.

1. **Language Resources**
   1. Available Resources (English, Arabic)

* Arabic and English NLP tools, specially part of speech taggers and stemmers.
* Arabic Named Entity recognizers. Work on the area of Arabic NER has matured during the past couple of years, and there are a number of approaches that can be carried out for accurate NER including the use of gazetteers. Dr. Samir Abd El-Rahman is also developing a new approach the preliminary results of which look very promising.
* Over the years, TREC has accumulated many English datasets specifically for the task of QA. The datasets contain questions of varying types, topics, complexity and difficulty. However QA datasets specific to the Arabic language are not available.
* Arabic word net.
* A set of 200 Cross Language Evaluation Forum (CLEF) questions that were translated into Arabic and their answers. Questions are classified into different domains (sport, geography, politic, etc.) and different types (questions seeking for time answers, persons, places, etc). These are available from: <http://www.dsic.upv.es/~ybenajiba/downloads.html>
  1. Needed Resources (English, Arabic)

There is a need for a bigger collection of Arabic questions and their answers as well as for information extraction and named entity recognition tools.

1. **Strengths, weaknesses, opportunities and threats** 
   1. Strengths

There is a reasonable pool of capable researchers in Egypt that can participate in the development of Arabic QA applications.

* 1. Weaknesses

QA is not an easy task, and building a good QA system that responds to queries in real time, will be a challenging and expensive task.

* 1. Opportunities

The market is wide open for powerful Arabic QA systems as there isn’t a single Arabic, tested, open domain QA system available as of yet.

* 1. Threats

Google’s Arabic QA system or any emerging Arabic QA system can mature rapidly and dominate the field.

1. **Suggestions for points to be further surveyed**
   1. What types of Question classes are there?
   2. What types of Question processing is needed?
   3. What methods are there to identify the Context of a question?
   4. What are the available Data sources for Arabic QA?
   5. What are the available techniques for mapping a question to a data resource?
   6. What are the available means for Answer extraction?
   7. Has any work been done on answer formulation, where a complete answer is formulated by fusing information from one source? If so, what kind of techniques were used.
   8. What methods have been applied to evaluate QA?
   9. Are there any Egyptian/Arab companies/organizations that may be interested in domain specific QA systems. If so, who are they and how much will they be willing to spend on this technology.
2. **List of people/organizations pioneers in QA related areas**

These are only the ones who have been identified thus far, but the list should be expanded in the future:

* + Boris Katz: MIT Computer Science and Artificial Intelligence Laboratory
  + John D. Burger: MITRE cooperation
  + Lahsen ABOUENOUR, and Karim Bouzouba: Mohammadia School of Engineers, Morroco
  + Paolo Rosso: Natural Language Engineering Laboratory, Universidad Politécnica Valencia, Spain
  + Yassine Benajiba: Columbia University

1. **Key persons in each application area (on technical/LR levels)**

To be identified at a later stage

1. **Suggestions for Language Resources (specific to the application area) if ALTEC would like to start collection immediately.**

There is a dire need for a collection of Arabic questions and their answers that can serve as an evaluation dataset for Arabic QA systems.

1. **Summary**

Research and development in the area of Arabic QA is seriously lagging behind compared to similar work on English systems. The available publications on work carried out in this area does not contain any plausible evaluation experiments and is usually very narrow in its scope. Due to the fact that Arabic still forms a small percentage of Web content, Arabic QA can have even more power is combined with machine translation capabilities that enables entry of Arabic natural language queries and the retrieval of answers from multi-linginal resources on the web, but finally displaying the results in Arabic (cross language QA). There is an increasingly growing number of researchers in Egypt that can aid in the building of such a system. But to be competitive, fast and to allow for improvements, a system like that needs careful design.

**References**

1. Abouenour L., Bouzoubaa K., Rosso P. 2008. Improving Q/A Using Arabic Wordnet. In: Proc. The 2008 International Arab Conference on Information Technology (ACIT'2008),Tunisia.
2. Abouenour L., Bouzoubaa K., Rosso P., 2009. “Three-level approach for Passage Retrieval in Arabic Question /Answering Systems”. In Proc Of the 3rd International Conference on Arabic Language Processing CITALA2009, Rabat, Morocco, May, 2009.
3. AnswerBus, 2008. <http://www.answerbus.com/index.shtml>
4. Benajiba Y., Rosso P., Lyhyaoui A. 2007. Implementation of the ArabiQA Question Answering System's components, In: Proc. Workshop on Arabic Natural Language Processing, 2nd Information Communication Technologies Int. Symposium, ICTIS-2007, Fez, Morroco, April 3-5.
5. Elkateb, S., Black W., Vossen P., Farwell D., Rodríguez H., Pease A., Alkhalifa M. 2006. “Arabic WordNet and the Challenges of Arabic”. *In proceedings of Arabic NLP/MT Conference*, London, U.K,.
6. Google Unveils Ejabat, 2009. <http://news.ebrandz.com/google/2009/2849-google-unveiled-egabat-arabic-version-of-question-answer-service-for-the-middle-east.html>
7. Google’s Ejabat, 2009. <http://ejabat.google.com/ejabat/>
8. Hammo, B., H. Abu-Salem, S. Lytinen and M. Evens, 2002. QARAB: A question answering system to support the Arabic language. Proceedings of the 40th Association for Computational Linguistics on Computational Approaches to Semetic Languages, ACL’02 University of Pennsylvania, PA, USA, 55-65.
9. Kanaan, G., Hammouri, A., Al-Shalabi, R. and Swalha, M. 200. “A New Question Answering System for the Arabic Language”. American Journal of Applied Sciences 6 (4): 797-805
10. Mihai Surdeanu, Rada Mihalcea, Roxana Girju,Vasile Rus, Finley Lacatusu, Paul Morarescu and Razvan Bunescu. 2001. Answering Complex, List and Context Questions with LCC’s Question- Answering Server. *Tenth Text Retrieval Conference (TREC-10).* Gaithersburg, MD.
11. Mohammed F.A., Nasser K., Harb H.M. 1993. “A knowledge-based Arabic Question Answering System (AQAS)”. In: ACM SIGART Bulletin, pp. 21-33.
12. QANUS, 2010. <https://junbin.com/qanus/>
13. Roussinov D., Fan W., Robles-Flores J. 2008. “Beyond Keywords: Automated Question Answering on the Web”; Communications of the ACM; Vol. 51 No. 9; 60-65.
14. START, 2006. <http://start.csail.mit.edu/>
15. TREC, 2010. <http://trec.nist.gov/>
16. Yarmohammadi Mahsa A., Shamsfard M., Yarmohammadi Mahshid. A., Rouhizadeh M. 2008. “Using WordNet in Extracting the Final Answer from Retrieved Documents in a Question Answering System”. *In : Proc. The Fourth Global WordNet Conference*, Szeged, Hungary,
17. Zheng, Z. 2002. “AnswerBus Question Answering System”. In  *Proceedings of HLT Human Language Technology Conference (HLT 2002).* San Diego, CA.