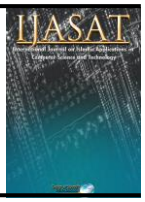

Intelligent Search Engine For The Holy Quran

Arwa Mohamed Basabrain[1,a], Dr. Amr Abdel Aziz Sharaf[2,b], Dr.Ibraheem Abdulmohsen Al-Bidewi[2,c]
[1]Computer Science Department, King Abdulaziz University, Jeddah, Saudi Arabia
[2]Computer Science Department, King Abdulaziz University, Jeddah, Saudi Arabia
[a]arwamb@yahoo.com, [b]amrsharaf@gmail.com, [c]ialbidewi@kau.edu.sa

ABSTRACT
The Holy Quran is a miraculous book with regard to the rhetorical, legislative, scientific, historical and numerical aspects. Muslims are always interested in using information technology to serve the Holy Quran, such as search engines. In spite of the evolution of the current search engines for the Holy Quran, there are some limitations such as ranking and classifying of search results. Also they do not provide accurate and detailed information for the search results, which makes them unsuitable for researchers in the numerical miracles. The main idea of this paper is to build an intelligent search engine for the Holy Quran with full-text features to be a tool that helps the researcher in the Holy Quran fields and the researcher in numerical miracle of Quran. The search engine gives detailed numerical information and offers facilities that help to analyze this information. Furthermore, it solves many limitations of the other search engines. In the first stage, the source data is collected and entered manually in text files, and then it is extracted automatically to fill up the database tables. The relational database contains main tables that holds detailed information about every segment of the Holy Quran. In addition, it contains tables, which provide valuable information about the relations between chapters and verses. These relations are automatically extracted from the main tables. A Graphical User Interface (GUI) application is implemented, which is divided into three sections: browsing and display sections, search sections and numerical miracle analysis sections. The intelligent search engine provides accurate and detailed information for the search results and gives answers for several questions, which range from simple to complex ones.

***Keywords*: Holy Quran, Information Retrieval, numerical miracle and Full-text Search Engine.**

1. Introduction

The scientific facts in the Quran exist in different subjects, including creation, astronomy, the animals and vegetables kingdom, and human bodies. There a large number of Muslims researchers studing the miracles and numerical miracles in the Quran. These researchers need to use full-text search engine for the Holy Quran, that has specials features to help them in their researches. The basic concept of full-text search engine is storing the full-text of all verses in the Holy Quran, so that every word in the verses is searchable and can function as a key for retrieval. In these types of engines, the contents of the verses are analyzed in such a way to allow users to search for any string of the Holy Quran. A high-quality full-text search engine includes some forms of relevance weighting method, so the items that have a higher level of relevance are displayed first. The weighting method can be based on the factors, such as: the number of times the search word occurs in the text, the position of the search word in the documents (Malviya, 1999; Moens, 2002). In addition, it allows search based on multiple criteria. This may consist of either looking for the relationship between two or more search

terms, or looking for terms in the context of verses. The main objective of this work is to use the computer technology for serving Islam by implementing an intelligent search engine for the Holy Quran with full-text features. It will be used as a tool for researchers in numerical miracles of the Holy Quran. In this work, the intelligent search engine has three common elements: relational database, Search Processor and GUI application. The relational database for the whole scripts of Holy Quran is automatically constructed from unstructured data files. It contains indexes that hold detailed information for each parts (*Hezb*), chapters (*Sura*), verses (*Aya*) and words. Also it contains other indexes, which provide valuable information to build the relations between chapters and verses, such as topics and root indexes. Furthermore, the automatically extracted relations between verses and chapters are indexed as well. The full-text search processor with special features to handle problems, which the other search engines have. Besides helping the user navigate through chapters, verses or words, the user-friendly application, which goes beyond simple keyword matching, provides a numerical miracle analysis section, which helps users to compare between chapters and verses.

This paper is organized as follow. Section two shows the related works, Section three discusses the characteristics of Arabic language and the Holy Quran scripts, which has an impact on search engine's processes, Section four explains the methodology of implementation of the Holy Quran search engine. Section five presents some results from experiments and section six concludes with a discussion.

2. Related Work

The project of (Al-Khalifa, Al-Yahya, Bahanshal, & Al-Odah, 2009) uses Semantic Web techniques to automatically identify semantic opposition terms in the Holy Quran. The project presents a Semantic Quran (SemQ) framework for recognizing and identifying semantic opposition terms using Natural Language processing armed with domain ontologies. The main purpose of (Umm-e-Laila, Saeed, Aziz, & Aziz, 2008) research is to develop software that can count the occurrences of specific words in the Holy Quran, which will be helpful in verifying and studying the miracles of the Holy Quran. A simple information retrieval application is implemented, called AWE (Arabic Word Extractor), which extracts specific words from Arabic text and count their occurrences. The work of (Khoja & Garside, 1999) introduced a new algorithm that extracts roots from Arabic words. The algorithm uses stop-words and considers weak letters when returning roots. The algorithm uses lists of valid Arabic roots and patterns. The algorithm of (Al-Shammari & Lin, 2008) introduced a novel lemmatization algorithm for Arabic Language. The new lemmatizer is a part of a comprehensive Arabic tokenization system, with a stop words list exceeding 2200 Arabic words. The stop words are categorized into useful and useless stop words. The work in (Hammo, Sleit, & El-Haj, 2009) research, they propose a passage retrieval approach to search for diacritic and diacritic-less text through query expansion to match the user's query. The Holy Quran scripts are chosen to test the approach for retrieving diacritisized Arabic text. The search process is improved through automatic query expansion using a stemmer and thesaurus. Khoja rule-based stemmer (Khoja & Garside, 1999) is applied and compiled a huge thesaurus for this purpose.

3. The characteristics of Arabic language

Search engine, as a language-dependent operation, is greatly affected by the language of documents and how the search engine handles the characteristics of this language. Linguistic characteristics typically have an impact on the accuracy and relevancy of the search results (Moukdad, 2004). Unlike regular Latin scripts, Arabic scripts are much more complicated. It

is written right-to-left, and the characters are written continuously in a word. The scripts of Holy Quran has two main parts: *Alphabets* (or characters) that represent the consonant sounds, and diacritics (or vowel signs) that represent the short vowels and cause variations in pronunciation (Hammo, et al., 2009). To handle the Arabic script, a number of encoding systems has been developed. The most common of these systems are Arabic (Windows), Arabic (ISMO 708), Arabic (DOS), and Arabic (ISO). In addition, Arabic is covered by the Unicode encoding system (Moukdad, 2004).

3.1 Arabic Alphabet

Alphabet is the most important part of the written language. It conveys the basic meanings of the words. Arabic alphabet consists of 29 characters including the glottal 'hamza' (it sounds like the 'a' in 'answer'), which appears as a separate character in written language, but it rarely used alone. Arabic characters change shape based on their position within words. This extends the Arabic alphabet to ninety different character representations (Tayli & Al-Salamah, 1990). An Arabic letter might have four different shapes: Isolated, initial, medial, and final. In computer encoding systems, the different representations of a character are often mapped to a single base code. For example, the letters "م","مـ","ـمـ" and "ـم" are four different shapes of the same letter pronounced "meem"

3.2 Arabic Diacritics

The diacritics are used in the Holy Quran to clarify the pronunciation of characters within an Arabic word; some can appear with any characters, while others appear only with a limited subset. For example the diacritic hamza "ء"/?/ is used by itself and is also used with the letters "ا", "و" and "ي ". Table 3-1 shows the Arabic Diacritics along with their codes in the Unicode system. Three diacritics are used to represent short vowels that can be used with every consonant character. They mark the consonant to clarify its pronunciation. For example, the consonant "س" /s/ with the diacritic Fath "سَ", is pronounced /sa/, with the diacritic Damm "سُ", is pronounced /su/, and with the diacritic Kasr "سِ"is pronounced /se/. The diacritic shadda is used to marks the double letter (a silent letter followed by a vocalised one) as in " سّ "/ss/ is used to mark the geminate (doubling) of a consonant. For example, in the word "رَدّ" (/radda/<returned>), the diacritic shadda indicates that the letter "د" is found twice in this word and should be stressed. The diacritic sukoon is a small circle that is placed above letters, as in "سْ", indicating a vowel-less consonant. It is used to close an Arabic syllable by marking the closing consonant. Two identical diacritics when placed above or below the last letter of Arabic nouns indicate the sound /n/; this is called tanween. For example the word "قِصَة"<story> is pronounced /qisata/ without tanween, " قِصَةً " with tanween fataha is pronounced /qisatan/, " قِصَةٌ " with tanween Damm is pronounced /qisatun/,and with tanween Kasr " قِصَةٍ " is pronounced /qisaten/. The *tatweel* (also known as *kashida*), "ـ", is a special character that is commonly used in typing Arabic text. This character is not an actual letter, as it is used only for cosmetic purposes (Goweder, Poesio, & Roeck, 2004). It can be inserted between any two concatenating letters. For example, the word "قال"(/qaala/<said>) can be written as "قـال ", and even "قـــــال ".

Table 3.1 Arabic Diacritics.

Diacritics are used with every consonant character		
Diacritic Name	Mark	Unicode(hex)
Fath	ـَ	"064E"
Damm	ـُ	"064F"
Kasr	ـِ	"0650"
Sukun	ـْ	"0652"
Shadd	ـّ	"0651"
Fath above Shadda	ـَّ	"FCF2"
Damm Above Shadda	ـُّ	"FCF3"
Kasr under Shadda	ـِّ	"FCF4"
Madd ـٓ	آ	" 0671"
Diacritics placed above or below the last character of the word		
Tanween Fath	ـً	"064B"
Tanween Damm	ـٌ	"064C"
Tanween Kasr	ـٍ	"064D"
Diacritic placed between the consonant characters		
Tatweel or (kashida)	ـ	" 0640"

3.3 The Holy Quran Scripts

Many Muslims including Arabs and non-Arabs read Holy Quran. Even in Arabic speaking countries, Arabic language is written in different styles, whereas non-Arabic Muslims also use different writing styles for reading the Holy Quran. Due to this fact, the Noble Quran has been published with different scripts or styles. The Othmani script is the most common and famous script of the Holy Quran. In Table 3-2 the first symbol is used to organize the Quran into parts and sections. The second symbol is used as a splitter between verses. In the second part, the symbol is located at the end of verses where prostration is recommended once a Muslim recites any of those verses. In the third part of Table 3-2, Pause Marks are special marks used in the Quran in order to show the person who is reciting the glorious Quran when he can pause, or when he can recite consecutively. All of the previous symbols should be removed during the tokenization process. In the fourth part of Table 3-2, Pronunciation Marks are used to show the reader of the Quran the correct way to pronounce the words. Sometimes, using these marks lead to a different spelling of the same word. The last two parts show different shapes for some symbols.

Table 3.2 Special symbols in the Holy Quran (Othman format).

Diacritic Name	Mark	Example	Unicode(hex)
Diacritics are used to organize the Quran into parts and sections			
Rub-el-hizb **<Start of section >**	۞	أَفَلَا ۞	06DE
Aya **<End of verse >**	۝	ٱلنَّاسِ ۝	06DD
Diacritics for Prostration of recitation			
Sajdah **<Prostration of recitation>**	۩	يَسْجُدُونَ ۩	06E9
Diacritics are used for reciting the Quran (Pause Marks)			
Al-Waslu-aula **< continuation is preferable >**	صلے	شَيْـًٔا صلے	06D6
Al-Waqfu-aula **< pause is preferable >**	قلے	بِهِۦ قلے	06D7
Lazim **< Compulsory Pause >**	م	مَثَلًا م	06D8
La-yuqaf-alaihi < Impermissible Pause>	لا	ٱلْعِلْمِ لا	06D9
Ja'iz **< Permissible Pause >**	ج	ٱللَّهُ ج	06DA
Mu'aniqa < Interchangeable Pause >	∴	لَا رَيْبَ ∴ فِيهِ	06DB
Saktah **<Short Pause >**	س	بَلْ س	06DC
Diacritics are used to indicate different pronunciation			
Long Fath	ا ـَـ	إِلَىٰٓ	FE7E
Long Dammh	ـُـو	إِنَّهُۥ	06E5
Long Kasrh	ـِـے	ظَهْرِهِۦ	06E7
Indicate silent letter	ه	أُو۟لَٰٓئِكَ	0652
Indicate silent letter when continuation	٥	أَنَا۠	06E0
Pronounce "م"/m/ instead of "ن"/n/	م	مَمْنُونٍ	06E2

Different Tanween shapes

Diacritic Name	Mark	Example	Unicode(hex)	Mark	Example	Unicode(hex)
Tanween Fath	ـً	مَسْرُورًا	064B	ـً	يَسِيرًا	0657
Tanween Damm	ـٌ	أَجْرٌ	064C	ـٌ	مَّرْقُومٌ	065E
Tanween Kasr	ـٍ	أَلِيمٍ	064D	ـٍ	لِّنَفْسٍ	0656
Different Alaf shapes						
Alaf "ا"	إ	إِذَا	0625	أ	فَأَيْنَ	0623
	ٱ	وَٱلَّيْلِ	0671	آ	شَآءَ	0622

4. Methodology

Implementing a search engine has gone through several phases: The data-gathering phase is the first activity in the implementation process to collect the data resources and to write it in a proper text file format. Then, in the design phase, a relational database model is drawn. At this point all the required tools are provided to start the implementation phase.

4.1 Data Gathering Phase

In this paper, database is based on three resources of raw data to create the underlying tables. These resources are collected and entered manually in text files:

1) The Holy Quran Script: The famous texts in which the Holy Quran had been written with is the Othmani text. The Holy Quran script had been taken from ("Zade Almoalemh ", 2008) which exists as a text file.
2) The Root Lexicon:Words and their roots are collected from two books of Indexed lexicon for the words of the Holy Quran (Baki, 1938; Sewar & Attar, 1991). They are entered manually in text files by using a special format to be used later in the database.
3) The Topic Lexicon: The topics and numbers of the verses are collected from two books of Indexed lexicon for the Holy Quran (Asor, 1990; Marofe, 2000). They are entered manually in text files by using special format to be used later in the database.

4.2 Design Phase

With the ability to modify and add any other relations, the structure of database model is specified and designed. Detailed information about every segment of the Holy Quran is kept and translated into a table, field or relation for the purpose of future use. According to the tables' resources, the database model is divided into several tables: the Holy Quran script tables, roots table, topics table and automatic extraction tables.

4.2.1 The Holy Quran Script Tables

The main components tables of the model are the Holy Quran script tables: sections table (*HezbInfo*), chapters table (*SuraInfo*), verses table (*VerseInfo*), words table (*WordsInfo*), and word address table (*WordAddress*). The data of these tables and their relations are extracted and derived from the Holy Quran script. These tables are the key components of the database as they contain important and detailed information about all sections of the Holy Quran. Also they are used to conclude the fundamental information to configure the other tables. The Holy Quran script tables are depicted in Figure 4.1

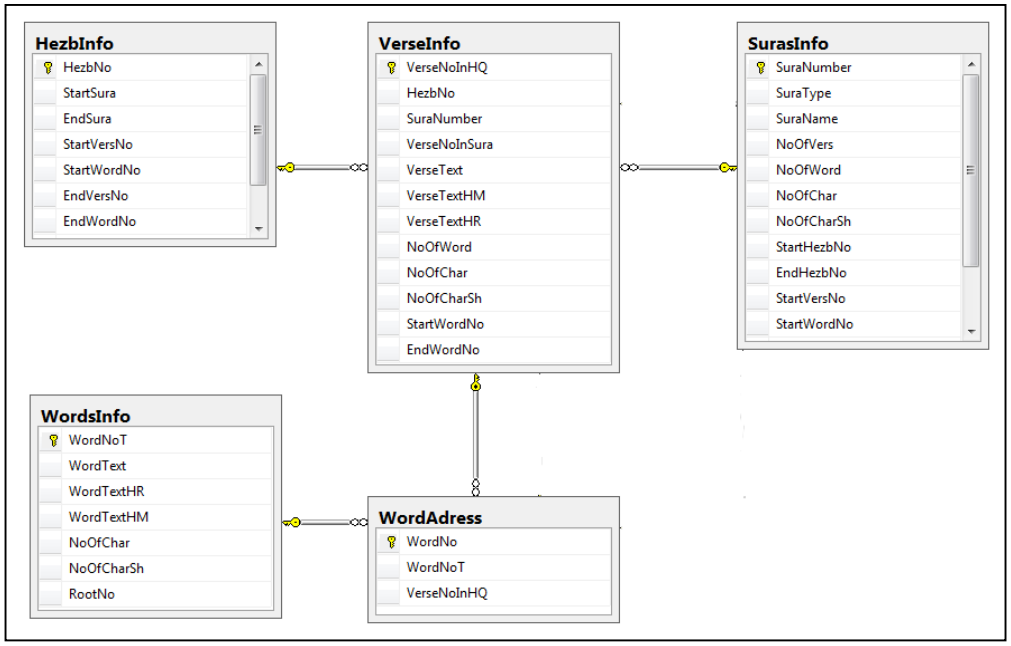

Figure 4.1 The Holy Quran script tables.

4.2.2 The Roots Table

The data of this table and its relations are extracted and derived from the root dictionary, which was collected in the text file. This table contains two field: the (*RootNo)* for the root ID and (*RootText)* for the root name. The words and the root tables are depicted in Figure 4-2

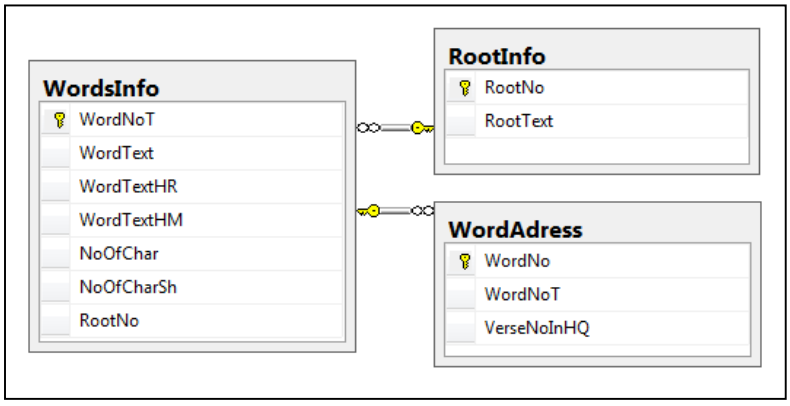

Figure 4.2 The words and the root tables.

4.2.3 The Topics Tables

The data of these tables and their relations are extracted and derived from topics lexicon of the Holy Quran. To represent the topics and their relations to verses, two tables are designed: (*SubjectInfo*) and (*SubjectAndVerse).* The Topic tables are depicted in Figure 4.3.

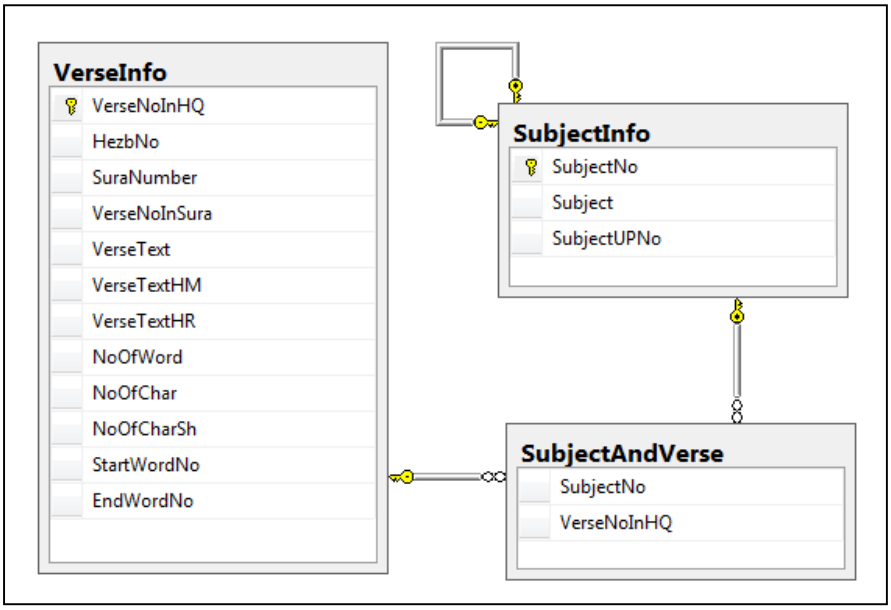

Figure 4.3 The topics tables.

4.2.4 Chapters Relationships Table

The data of these tables and its relations are extracted and derived from the Holy Quran script tables: chapters (*SuraInfo*) table and verses (*VerseInfo*) table.

The idea of this table is based on the numerical miracle in the Holy Quran. This miracle is in the positions of the chapters and the number of their verses. There is a special relation between chapters if the difference between their verses equal to the difference between their positions. For example, there is a relationship between chapter *Al-Falaq* "الفلق" and chapter *An-Nâs*"الناس" as shown in Table 4.1.

Table 4.1 The relationship between Al-Falaq "الفلق" and An-Nâs"الناس".

Chapter	Order	Verses
Al-Falaq "الفلق"	113	5
An-Nâs"الناس"	114	6
Different	1	1

4.3 Implementation phase

The implementation phase of the developed search engine mainly consists of three sections: Implementing database, implementing search processor and implementing graphical user interface (GUI) application.

4.3.1 Implementing Database

After designing and constructing the database using MS SQL Server, the C# codes are required to fill up the database tables automatically through extracted data from the source files and to implement the GUI application. Also Dundas Chart for. NET is used to create charts in the application.

4.3.2 Implementing the Search Processor

The basic responsibilities of the search processor are finding matches and ranking search results, but the search processor in the intelligent search engine must have more responsibilities and capabilities to commensurate with the needs of the target users. The target users of this search engine spread from regular users looking up words in the Quran to the more specialized users of the science of the Quran, or researcher in the numerical miracle of the Quran. Therefore, the search engine was developed to commensurate with the requirements and aspirations of the user by implementing search processor that has a number of features. In the following, there are some of these features:

- Search results are presented in several ways to the same query to help the user finding the proper information.
- Give detailed figures on the number of iterations and the number of verses that appear in the search results.
- Search results are sorted by multiple levels according to the type of search.
- The search results are presented in clustering order to help the user locating relevant information quickly.
- By using a simple query and choosing a proper type of search, the search processor gives answers for several questions, which range from simple to complex ones as in:
  - How many times the name of (Allah) or "الرحمن الرحيم " had been mentioned in the Holy Quran, and what are the Quran verse in which they appear?
  - What are the words which contained a certain letter in the Quran such as "ق" and how many are they?
  - What are the Quran verses that begin with the word "قل" or "سبح"?
  - What are the Quran verses in which the character "و" appeared and how many times had come on them?
  - What are the Quran verses that had a certain topic such as Zekat "الزكاة"?
  - What are the words that it stems are "قول"?
  - How many times that the character "ن" had been mentioned in the Holy Quran?
  - What are the Quran verses in which the intonation ruling appears (the Insertion) "الادغام" between the silent "ن" and "ي"?

4.3.2.1 Specifications of Search Process

Before a query is submitted to the search processor, four specifications must be determined by the user or as a default selection.

- Range: Besides searching in the whole Quran, the user is allowed to specify the range of verses to search in by selecting the start and the end of chapters. This includes parts or sections.
- Text type: The query might be represented by using three types of text: diacritic, diacritics-less or normalized *Alef* "ا".
- Search in: As mentioned before, there are three basics resources in the developed search engine: The Holy Quran, roots or topics. The user must determine the source in which he or she wants to search.
- Search for: The users are allowed to search for words, part of word or part of verses. The regular expression can be used to write the query if a user chooses to search for part of word or part of verses.

4.3.2.2 Query Processor

The query processor extracts the terms of the query submitted by the user and performs some lexical analysis steps. After a query is submitted, the query is tokenized and normalized by query processor to extract its terms depending on the selected specifications. The tokenization is done only if the user chooses to search for words but not for a part of word or a part of a verse. The normalization is done if the selected type of text is diacritics-less. Although the filtrating is done if the selected types of text are diacritics-less or normalized Alef "ا". It should be noted that the filtrating process just removes the diacritics, but it does not remove any of the stop-words considering that every word in the Holy Quran has a meaning and important positions.

4.3.2.3 Searching and Ranking Algorithms

After the query processor isolates and analyses all terms in the query, an appropriate search algorithm is chosen to start searching in database tables, based on the selected specifications. There are five types of algorithms: *SearchInRoot*, *SearchInTopic*, *SearchForPartOfWord*, *SearchForPartOfVerse* and *SearchForWords*.

4.3.3 Implementing the GUI Application

The last phase in the implementation process for search engine is a GUI application, which is considered as the expressing face of the capability and quality of the search engine. With user-friendly interface, the application helps users to navigate through chapters, verses and words, and displays their information. Moreover, the application provides a numerical miracle analysis section. The application is divided into three sections: Browsing and display sections, search sections and numerical miracle analysis sections

4.3.3.1 Browsing and Display Sections

As shown in Figure 4-4, the browsing section helps users to navigate through chapters, verses and words, and the user can use it to display the retrieved verses from searching process. The browsing section contains three display sections: Chapters, verses and words. These sections display chapters, verses or words information in separate windows to help in comparing between them.

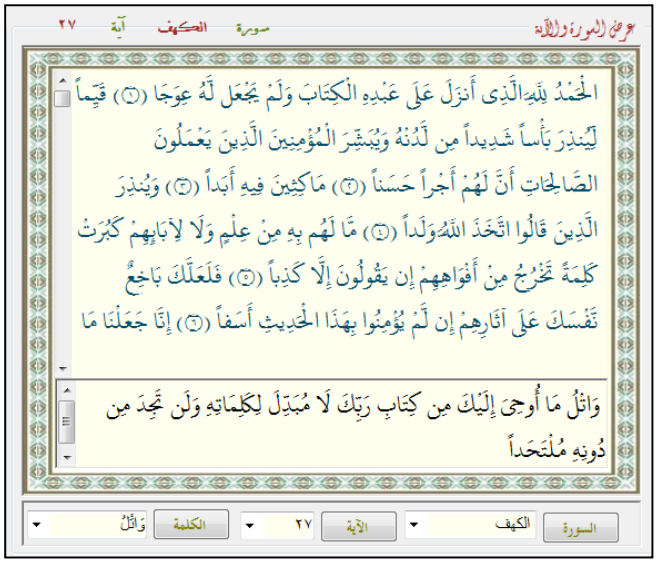

Figure 4.4 Browsing Section.

4.3.3.2 Search Sections

The search section is an important component of GUI application. It consists of two sections: Results section and specifications section. Figure 4-5 shows these sections. In results section, a tree hierarchy view is used to display the search results to help users finding desired information quickly. In specifications section, the user determines the range and text type of search and chooses to search for words, part of word or part of verse as mentioned before. Also the user submits the query in specifications section and determines the source in where he or she would like to search.

Figure 4.5 Search Sections: Results section and Specifications section.

4.3.3.3 Numerical Miracle Analysis Section

One of the objectives of the work is to implement tool for researchers in numerical miracle in the Holy Quran. This section of the application uses all numerical information in database and presents these numbers in a manner to serve the needs of researchers. This section will be used to compare between verses or chapters and to analyze specific numbers in Quran. This section contains two parts: The first part is used to select two chapters or two verses, the second part is used to analyze and compare the numerical information. As shown in Figure 4-6. The second part is consists of four sub-parts:

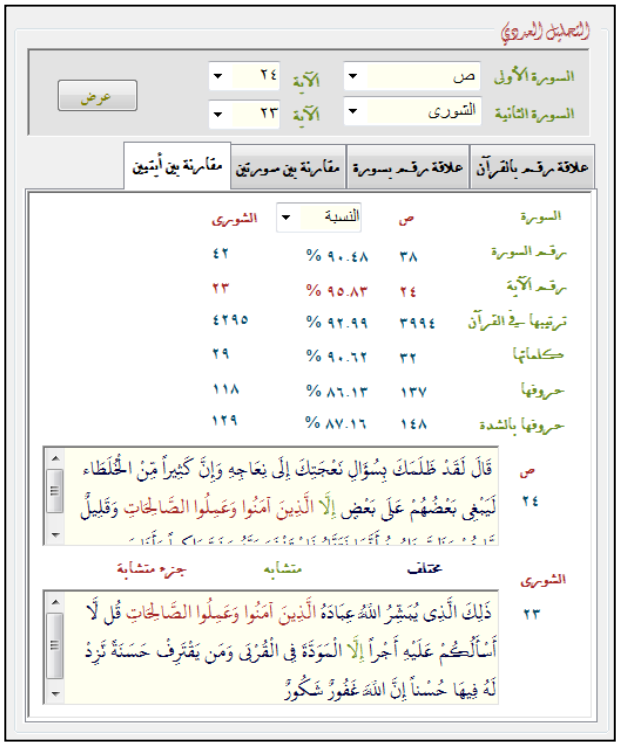

Figure 4.6 Numerical Miracle analysis section

- Number and The Holy Quran**:** This part is designed to help a researcher to the conclusion and analysis of numbers by comparing the numerical information for Quran and the number chosen by the researcher. For example, in this part, by using the basic arithmetic operations, the researcher can study the relation between number (19) and number of verses in the Holy Quran.
- Number and Chapter**:** In the field of numerical Miracles, often mentions relations between the specific numbers and chapter, such as number (7) and chapter Al-Fatiha. This part is designed to help finding such relations by using the basic arithmetic operations between specific number and the numerical information for specific chapter.
- Two Chapters Relation**:** This part is designed to help comparing between chapters. Besides offering the numerical information for each chapter and the arithmetic operations, and it helps to find the common topics. Also it helps the researcher to find the relation between number of their verses and their position in Quran**.**
- Two Verses Relation**:** There are many similar verses in the Quran, which may not differ in some cases except by a single character. Specialists may be interested in studying these verses. This part is designed to help comparing verses by examining similarities and differences between them. Also this part, offers the arithmetic operations for studying the numerical information of verses.

5. Search Engine Testing and Results

The overall testing process of the search engine was divided into two phases. During phase one, a set of 150 words was selected randomly from the thirty parts of the Holy Quran; these words were selected as a set of five words from each part (*Juze*). In this phase, the selected words were submit to (*Iszlám*) search engine (Abdulghany, 2009), then they were submitted to the developed search engine. For each search engine, the results, which include the frequency and the roots of each word, were recorded and documented. In the second phase, the produced results of both search engines were compared and analyzed. This phase aims to provide indications of the performance of the developed search engine.

Table 4-2 shows a part of these selected words along with their produced results. The included results represent the identical results produced by both search engines, the developed search engine for this research and (*Iszlám*) search engine.

Table 4.2 Tested words results

Serial No	Chapter	Veers	Word	Root	Frequency	No of Verses
1	البقرة	2	الكتاب	كتب	176	162
2	البقرة	19	أصابعهم	صبع	2	2
3	البقرة	52	تشكرون	شكر	19	19
4	البقرة	79	فويل	ويل	9	8
5	البقرة	102	يعلمون	علم	91	89
6	البقرة	255	يشفع	شفع	5	4
7	البقرة	265	ينفقون	نفق	22	21

6. Conclusions

The primary contribution of this work is to implement an intelligent search engine for the Holy Quran with full-text features to be used as a tool for researchers who are interested in the Holy Quran and numerical miracles. The paper started by presenting the characteristics of Arabic language, which has an impact on search engine's processes. The existing search engines for the Holy Quran have several shortcomings, which make them not very useful for researchers who are interested in the numerical miracles. These deficiencies include lacking of accurate numerical information about the results, providing the results without any ranking and classifying. This work attempted to overcome these problems by implementing a search processor with different functionalities and features. It includes several searching algorithms to present the search results in several ways and give detailed information about results. Before implementing the search processor, a relational database has been designed and constructed. The Database of the Holy Quran contains main tables that hold detailed information about every segment of the Holy Quran. Moreover, it is based on three resources: the Holy Quran script, topic lexicon, and root lexicon. These resources are transformed manually into text files, from which the data is extracted automatically to fill up the main tables. In addition, the database produced new tables automatically extracted from existing tables, which provide valuable information about the relations between chapters and verses. To achieve the objectives of the work, a GUI application is implemented, which is divided into three sections: Browsing and display section, search section, and numerical miracle analysis section. A user-friendly interface helps users to navigate through chapters, verses and words, and displays their information. In search section, a tree hierarchy view is used to display the search results and its detailed information to help users finding desired information quickly. The numerical miracle analysis section is used to compare verses or

chapters and to analyze specific numbers in Quran. This section uses all numerical information in the database and presents these numbers on a manner to serve the needs of the researchers.

References

Abdulghany, A. (2009). Koran Iszlam. Retrieved 2010, from http://koran.iszlam.net/program.php

Al-Khalifa, H. S., Al-Yahya, M. M., Bahanshal, A., & Al-Odah, I. (2009, 15-16 Dec. 2009). *SemQ: A proposed framework for representing semantic opposition in the Holy Quran using Semantic Web technologies.* Paper presented at the Current Trends in Information Technology (CTIT), 2009 International Conference on the**.**

Al-Shammari, E., & Lin, J. (2008). *A novel Arabic lemmatization algorithm*. Paper presented at the Proceedings of the second workshop on Analytics for noisy unstructured text data **.**

Asor, S. A. R. (1990). *Thematic Dictionary of Quranic verses* (First ed.): Dar Al Fadeelah Publishing, distribution and export**.**

Baki, M. F. A. (1938). *Indexed dictionary of words of the Holy Quran* (First ed.). Istanbul, Turkey: Islamic Library**.**

Goweder, A., Poesio, M., & Roeck, A. D. (2004). *Broken plural detection for arabic information retrieval*. Paper presented at the Proceedings of the 27th annual international ACM SIGIR conference on Research and development in information retrieval **.**

Hammo, B., Sleit, A., & El-Haj, M. (2009). Enhancing Retrieval Effectiveness of Diacritisized Arabic Passages Using Stemmer and Thesaurus *Information Retrieval 12*(3), 300-323**.**

Khoja, S **&** ,**.**Garside, R. (1999). *Stemming Arabic Text.* Lancaster University, Lancaster, UK**.**

Malviya, M. K. (1999). *A Web-based Search engine for Indian Language.* Indian Institute of Technology, Kanpur**.**

Marofe, M. N. (2000). *Indexed lexicon of the Quran topics* (First ed.). Beirut, Lebanon: Al Nafaes Printing, Publishing and Distribution**.**

Moens, M.-F. (2002). *Automatic Indexing and Abstracting of Document Texts*: Kluwer Academic Pub**.**

Moukdad, H. (2004). *Lost In Cyberspace: How Do Search Engines Handle Arabic Queries?* Paper presented at the Access to Information: Technologies, Skills, and Socio-Political Context, University of Manitoba, Winnipeg, Manitoba**.**

Sewar, M. N., & Attar, S. A. E. (1991). *Indexed lexicon the words of the Quran expressing Fatah Al Rahman to student verses of the Quran* (First ed.): Dara Alfager Alislami**.**

Tayli, M., & Al-Salamah, A. I. (1990). Building bilingual microcomputer systems. *Commun. ACM, 33*(5), 495-504**.**

Umm-e-Laila, Saeed, F., Aziz, A., & Aziz, F. (2008). Verification of Occurrence of Arabic Word in Quran. *Journal of Information & Communication Technology 2*(2), 109-115**.**

.Zade Almoalemh (Version 1). (2008): Majid institution of production and distribution**.**