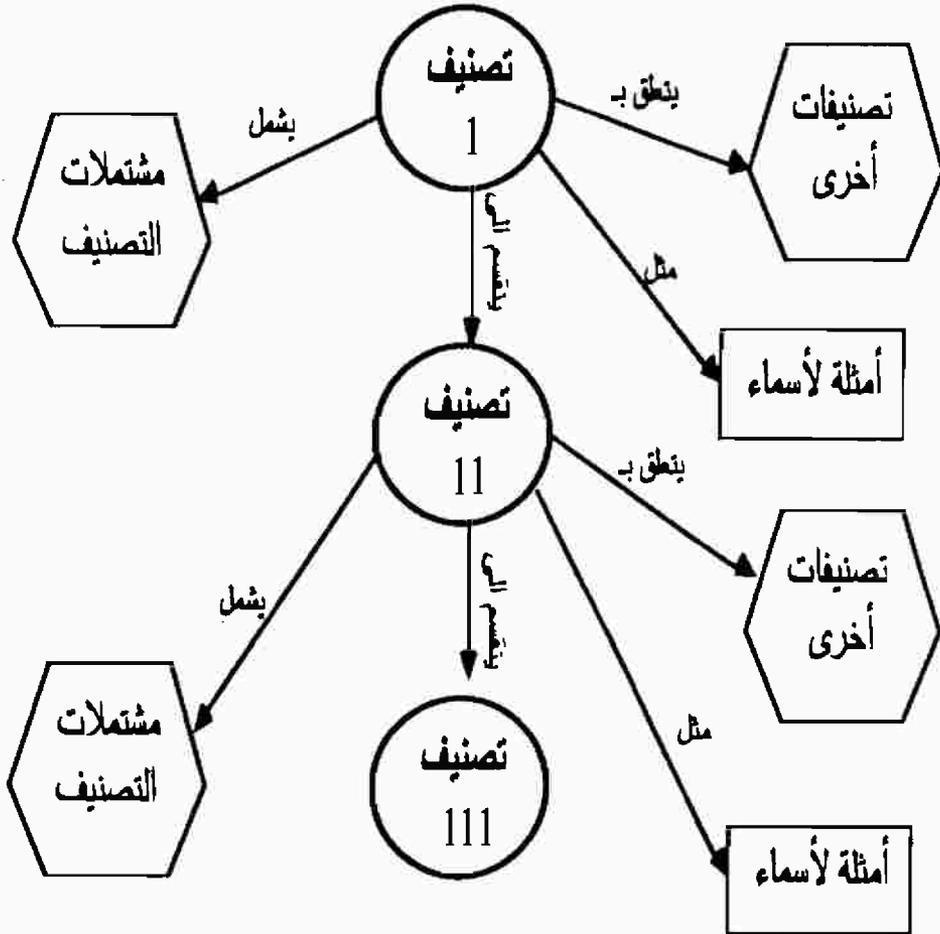


## THE SMANTIC NETS OF THE SYSTEM



### **3. Why RDBMS To Build The Expert Tool?**

In the heart of an expert system lies the inference engine. For this engine work, both general knowledge or rules and specific declarative knowledge or data are needed.

Early ES were mostly prototypes, characterized by a large variety of predicates required, and a small population of specific knowledge. Therefore inefficiency of data handling was not a critical issue. Consequently, with very few exceptions, a little attention has been given to the handling of very large population.

So to build a real Arabic search tool, we must have a large volume of words and features so the database management system is the best mechanism to store and manipulate these facts with fast interactive access through various access routes to be able to react quickly to the different environmental changes.

So some researchers suggested to enhance ES with database management system facilities.

Beside the RDBMS is a good environment for any large volume of data.

Also there are many tools for building very friendly forms to the user to access and manage the data, through SQL, very fast.

So, these relational operations, which support breadth-first search, can handle a high-speed forward chaining to search into the system for an existing information, or for learning new information and.

## **1. The Requirement Resources:**

The proposed tool can run on any platform that support Oracle database. For example on PCs the required resources for the system to work:

- [386, 486 or Pentium].
- 8 MB RAM.
- 100 - MB hard disk.
- Oracle database version 6 or higher.
- DOS 3,1 or higher

## **2. Why An Expert System For The Proposed Search Tool:**

The stored data into a traditional systems is not enough to provide all the search paths needs so, a good search tool also must include many procedures that can use an inference engine to manipulate the stored knowledge to satisfy these requirements.

Also, because of the very large volume of subjects in the DEWEY, in different fields, there are need for an intelligent DB that can learn and accept new subjects and books to improve the capability of the search.

In our proposed lexicon there are procedures uses the SQL language, which include an inference engine, to access, deduce or update the facts, information and rules, which stored in the system in a structural way.

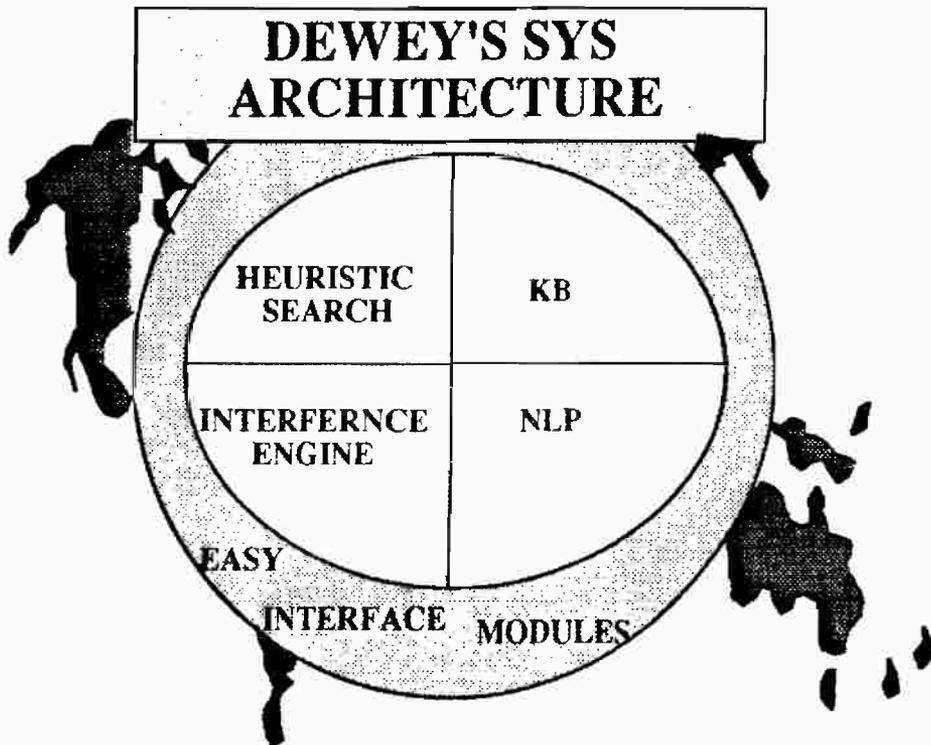
The direction of inference engine, in the proposed lexicon, is forward chaining using the breadth-first search technique which supported by the relational operations in the relational database management system which the proposed tool based on.

## ELEMENTS OF THE PROPOSED TOOL

The whole structure of the expert lexicon is based on the relational database management system (RDBMS) which guarantees an efficient management of the large amount of the knowledge and a high speed of search into DB, using the SQL, to retrieve this knowledge with easy, fast and efficient way.

Also the user has the capability to enlarge the KB through an easy user interface.

The good usage of the capabilities of the SQL in the search can solve the problem of heuristic search without need to any additional routines or any complicated data structures.



A digit stands for a Main Class; that is, 5 stands for the sciences. Strictly and mathematically speaking, the numbers denoting the Main Classes are decimal numbers and should have been written as. Generalities, 4 Language, 5 Natual sciences and mathematics, and so on. The 5 is extended, like a decimal fraction, to indicate the subdivisions of the sciences; that is 59 stands for zoology, 599 stands for mammals, 599.8 stands for primates, 599.88 stands for apes, 599.884 stands for the great apes, and 5998846 stands for the gorilla.

The decimal fractions reveal the intellectual hierarchy of the subject. But for brevity, simplicity of notation, and ease of reading, the decimal point is placed after the third digit in the full DDC number, hence, 599, 884, 599, 884.6, etc.....



from subpath to another according a scientific bases using a knowledge base for subjects calssification.

Dewey decimal classification was choosed from various available systems to be used as knowledge base since it is the most effective classification and the most spread all over the world.

## **WHAT IS DEWEY'S SYSTEM**

### **1. Terminology:**

The Decimal Classification was originally designed in 1873 for arranging the books and the catalog of the library of Amherst College (Amherst, Massachusetts). It was published in 1876 for wider use in other libraries. The word "Dewey" in its current name is that of Melvil Dewey (1851-1931), who devised the Classification. The second word, "Decimal", refers to its base-ten notation that is used to denote and relate subjects. It employs Indo-Arabic numerals treated like decimal fractions.

### **2. Use of DDC:**

DDC is te most widely used library classification system in the world. A survey conducted in the mid-1970s revealed that over 85 percent of the librarians in the United States and Canada were DDC users. DDC has been translated into more than 30 languages (Arabic and Italian being the two most recent official translations or adaptations). It is a general classification scheme that aims to classify books and other material on all subjects in all languages in every kind of library, the lingua franca of labrarians.

### **3. Basic Plan of the Classification:**

DDC is a universal scheme that treats knowledge as a whole, dividing it into ten mutually exclusive Main Classes that are denoted by Arabic numerals used as decimal fractions.

## CHAPTER 16

# AN ARABIC SEARCH ENGINE BASED ON DEWY DECIMAL CLASSIFICATIONS

By

Dr. M. Kouta & Eng. W. Hassan

### ABSTRACT

Most of current library software searches the following domains, title, author and Subject.

Searching the title requires the exact matching between both the author and the researcher terms otherwise the information retrieved would be inexact which ends with a bad search performance such as low hit ratio, high miss ratio, and low relevancy. Using Thesauri (which show the terms similarity, vocabulary etymology, general terms, and narrower ones) enhances the search performance.

Due to the rapid scientific development, the synonym relation, generalization, and specification of many terms have not been established in Arabic, which explains why the scientific Arabic thesauri are rare.

The shortage of a comprehensive scientific Arabic thesauri has been lead to immense shortage in a strong Arabic Search engines.

The available search engines do not offer a search dynamics which allow either search expansion or contraction.

This work tries to get over the shortage of Arabic thesauri. It introduces a pattern for establishing a base for human cognition in Arabic. It offers an Arabic effective search engine which enables the researcher to search dynamically. It also enables him to track the correct search path and to switch