DEVELOPING A NATURAL LANGUAGE PROCESSING (NLP) INTERFACE FOR A BANKING SYSTEM

BY

RHOLAND CHRIS O. CORNELIO

ROANNE L. TORTAL

CHERRYL R. VALDUEZA

SCHOOL OF ARTS AND SCIENCES
ATENEO DE DAVAO UNIVERSITY

MARCH 2003

DEVELOPING A NATURAL LANGUAGE PROCESSING (NLP) INTERFACE FOR A BANKING SYSTEM

A Mini-Thesis

Presented to

The Faculty of the Computer Science Division

Ateneo de Davao University

In Partial Fulfillment

of the Requirements for the Degree

Bachelor of Science Major in Computer Science

By

Rholand Chris O. Cornelio

Roanne L. Tortal

Cherryl R. Valdueza

March 2003

TABLE OF CONTENTS

ABSTRACT

A number of existing Relational Database Management System (RDBMS) Applications that uses the Data Manipulation Language of the Structured Query Language (SQL) is available today. Remembering the syntax is necessary in order to get the intended query results thus memorizing syntax is more focused not on the task itself.

This study aims to address the problem by developing a Natural Language Processing Interface for a Banking System that recognizes conversational English commands. Several natural language processing theories and concepts must be considered. This includes the basics of Relational Databases, SQL, linguistics, grammar rules and parts of speech, lexicon, syntax, semantics, context, Key Word Analyzer, Syntactical and Semantic Analysis, Parsers and Information Extraction.

CHAPTER I

INTRODUCTION

1.1 Background of the Study

An artificial-intelligence system is often built from natural-language specifications, either oral or written. It will be involved in interpreting and extracting information from written text, such as simple commands to a computer.

"Understanding" language means, among other things, knowing what concepts a word or phrase stands for and knowing how to link those concepts together in a meaningful way. It would be ironic that natural language, the symbol system that is easiest for humans to learn and use, is hardest for a computer to master. Long after machines have proven capable of inverting large matrices with speed and grace, they still fail to master the basics of our spoken and written languages.

If this type of understanding could be achieved by a system, it would be much easier to interact with it. Users often get tired and confused on querying information from databases using computer jargons. After all, users are not concerned on the significance and meaning of those commands but rather on the result. Thus, issuing commands to a computer using conversational English

would be the best alternative. Users might be able to learn something new about the computer without memorizing commands that means absolutely nothing to them. They can now tell the computer to query information from the database in conversational English command.

1.2 Statement of the Problem

The study aims to answer the problem on How can Natural Language Processing Concepts be used to develop an Interface that recognizes conversational English input for a Banking System.

Specifically, the study aims to find the answers to the following questions:

- What Natural Language Processing concepts can be applied to develop an interface for relational databases?
    - What are the steps in Natural Language Understanding?
    - What will be the core components of the Interface?
- How to integrate the Natural Language Processing concepts to retrieve data from a database?

1.3 Objective of the Study

The general objective of the study is to apply Natural Language Processing Concepts to develop an Interface for a Banking System that recognizes conversational English input.

The specific objectives are:

- To consider the concepts of Natural Language Processing that can be applied in developing an interface for relational databases
    - To identify the steps in Natural Language Understanding
    - To identify the components needed to develop the NLP Interface
- To integrate the Natural Language Processing concepts to retrieve data from a database

1.4 Scope and Limitation of the Study

This study will be about a Natural Language Interface for a Banking System. Instead of using the most common data retrieval commands for databases, users can issue conversational English commands in order to query information from the database. The system is capable enough to convert the plain English command into its equivalent SQL command. It will be using the key aspects of natural language processing concepts such as Parsing, Lexical

Analysis, Syntactic and Semantic Analysis and other natural language rules and understanding.

However, only a part of the SQL-Data Retrieval commands are to be included. These are the commands and methods you use to retrieve data in an existing database. This includes the facilities for querying the database and navigating through its tables.

On the other hand, the proponents will integrate the said interface into a Bank Database Application for the proof of concept. However, It is only limited to gathering information and producing outputs based on the customer's information as well the transactions with regards to their deposits and withdrawals. Moreover, the proponents will be using Microsoft Visual Basic 6.0 for the front-end Programming Language and Microsoft Access as the back-end support for the database.

1.5 Significance of the Study

This study is significant for Database Management System users such as Universities, Banks, Institutions, Business Corporations and Organizations. In this case, the bank officers and employees that have access to the database for reports and bank statistics can query information by just typing conversational English command. The workload then is lightened and it is easier and faster for

the bank employees to access the database since the focus is more on "what the results are" rather than "how to get the results." Also, queries are not limited to commands that will be executed when the user clicks a button, selects an item from a list box and the like.