MSc in Computing (Information Technology)

An Investigation of the Multi-Modal Agent Based PDA and Developing a Framework for It

TAO WANG

School of Computing

Dublin Institute of Technology

Kevin St. Dublin 8, Ireland.

A Dissertation Submitted in Partial Fulfilment of the Requirements of Dublin Institute of Technology for the Degree of M.Sc. in Computing Science (Information Technology)

November 2007

MSc in Computing (Information Technology)

Declaration

I herby certify that this dissertation which I now submit for assessment by the School of Computing, Dublin Institute of Technology on the programme of study leading to the award of The Award Goes Here is entirely my own work and has not been submitted for assessment for any academic purpose other than in particular fulfillment for the stated above.

Signed:

Date: 12th/Nov/2007

ABSTRACT

This project combines multi-modal interfaces, agent technology, web services and speech server to create a next generation intelligent PDA system. The system would use intelligent agent technology to aggregate information important to the users' daily life from web services and present it intelligently using a speech and pen interface. For example, a user employs a speech and pen interface to communicate with the PDA for downloading or uploading data from a web service, if the PDA detects there is voice command during an interaction between presenting information and receiving information, the system will invoke speech server to convert this voice command in order to enable it working with the web service. Information could include calendar and email information, weather and traffic reports, entertainment listings, information about friends and colleagues.

Keywords:

PDA, Web Service, Agent Technology, Multi-Modal Interface, Speech Server,

ACKNOWLEDGEMENTS

I would like to express my sincere thanks to my supervisor Mr. Damian Gordon and my course coordinator Mr. Bryan Duggan and other lecturers who supervised my work for their valuable support and guidance. Without their kind and helpful participation I would find it impossible to do this project. This dissertation could not have been written without Mr. Damian Gordon who not only served as my supervisor but also encouraged and assisted me throughout the research programme. And also million thanks to Mr. Damian Gordon again for assisting me in gathering data, correcting my poor spelling and giving valuable information for the dissertation.

Also I want to thank all my lectures in the MSc in computing science course, Mr. Ronan Bradley, Dr. Svetlana Hensman, Dr. Brian MacNamee, Mr. Richard Lawlor and Dr. Fred Mtenzi, who patiently lectured me throughout the entire academic program, never accepting less than my best efforts.

Finally, I cherish the both moral and financial support from my family members during the period of my studies abroad. There is no doubt that my success is based on all of your encouragement, love and support. Furthermore I would thank for my two classmates Mr. Ronan Carty and Mr. Seo Cathán as well, to appreciate all advisers and assistants on both English and thoughts. In addition I also would like to thank all of you who painstakingly provided comments and offered advice to my PDA online survey, my grateful thanks

TABLE OF CONTENT

TABLE OF FIGURES

TABLE OF TABLES

GLOSSARY OF ACRONYMS AND TERMS

Acronyms

PDA	Personal Digital Assistant
PC	Personal Computer
GUI	Graphic User Interface
LCD	Liquid Crystal Display
GPS	Global Positioning System
OS	Operating System
PIM	Personal Information Manager
SD	Secure Digital
CPU	Central Processing Unit
RAM	Random Access Memory
CF	Compact Flash
GIS	Geographic Information System
RS	Remote Sensing
LBS	Location Based Service
MMA	Multi Modal Agent
MAS	Multi Agent Simulation
MONA	Mobile MultimOdal Next-generation Applications
RPC	Remote Procedure Call
XML	Extensible Markup Language
XSD	XML Schema Definition
SOAP	Simple Object Access Protocol
UDDI	Universal Description, Discovery, and Integration
WSDL	Web Services Description Language
FIPA	Foundation for Intelligent Physical Agents
ACL	Agent Communication Language
DF	Directory Facilitator
JADE	Java Agent Development Framework
API	Application Programming Interface
GUI	Graphic User Interface
IVR	Interactive Voice Response

WSIG	Web Service Integration Gateway
JADE	Java Agent DEvelopment Framework
FIPA	Foundation for Intelligent Physical Agents
TTS	Text-to-Speech
SD	Service Description
OWL	Ontology Web Language
SALT	Speech Application Language Tag
SR	Speech Recognition
HTML	HyperText Markup Language
XHTML	eXtensible HyperText Markup Language
URL	Uniform Resource Locator
SSML	Speech Synthesis Markup Language
SES	Speech Engine Service
MSS	Microsoft Speech Server
IP	Internet Protocol
ASP	Active Server Pages

Chapter 1 Introduction

This dissertation details the background of the project which is about the usage of Personal Digital Assistant (PDA)'s with multi-modal agent; PDAs have become extremely popular for both business and personal uses. Among the main reasons for their popularity is their portability with the advantage of having the functionality of a PC, for example, a recent report has stated that Global shipments of handheld computers rose 18.4% to 17.7 million in 2006 *[TelecomAsia, 2007]*.

A PDA is a mobile handheld device that provides services tailored to users with mobility, it has very familiar applications that most users need for daily life, such as personal organizer, calculator, alarm, and notepad. More recently, PDAs feature special challenges for technology and software, which have expanded the functionality envelope of PDAs without sacrificing portability. Consequently, word processing, spreadsheets and even more complicated applications have started appearing on PDAs. The portability of PDAs has become as a catalyst for current development of journey tracking systems, for example, Global Position System (GPS) based on PDAs has well applied in many fields *[earthmate 2007]*. Therefore an excellent prospect for developing the next generation PDA application has been already grown.

1.1 The Multi-Modal Agent System

The multi-modal agent system is next generation PDA software kits, it enables users interacting with PDAs to use speech and pen interface rather than the conventional way, and also the agent allows users to gather information with less work rather than manual manipulation. In the process of these operations, this system requires different technologies of the multi-modal interface; intelligent agent, web services and speech server are described as follows:

- The Multi-Modal interface allows the users to check their daily information rather than only one method displayed on screen, it also allows the PDA talk to the users, for example, when the PDA finishes updating, it will talk to that users presenting them with the information or data about.

- The Intelligent agent can assist users in order to accomplish their needs with a lot less pen tapping, for example, it could read newspapers and magazines to find articles significant to us, even better, once users set preference to the intelligent agent (e.g. one could ask an agent to find new friends), the agent would aggregate information based on users' preference (such as gender), using this knowledge that the agent would then search the Web.
- When once the intelligent agent has gathered information or data which matches to the users' previous preferences, it will invoke the Web service to present the information or data using a friendly interface, which means there is a webpage to finally demonstrate the information or data via a Multi-Modal interface.
- Furthermore the speech server operates behind the scenes and carries out the functions of speech recognition and text-to-speech conversion, also it communicates with web services to request information that is required to process speech recognition and text-to-speech conversion.

These technologies cooperate and affect each other in the overall architecture of the multi-modal agent system. The web service is integrated with an intelligent agent know as the Web Service Integration Gateway (WSIG) structure. The multi-modal interface resides on a web browser on a PDA. And the speech server works at converting voice commands that should be on an intranet communicated with a PDA using either WiFi or USB connection.

1.2 Research Problem

The main purpose of this dissertation is to propose an investigation of the development of a multi-modal agent based PDA and developing a framework for it. The dissertation proposes an architecture of preferred activities participated in during structuring the multi-modal agent system. While numerous papers and researchers have investigated the importance of marginality of different functions in the multi-modal agent system, most of them have not been fully discussed in identifying the entire system that how each of the functions work together. Additionally, two primary sub-problems will be addressed. First what are the significant challenges in building a next generation PDA application? And second to propose a framework for the next generation multi-modal agent based on PDA system. To date, efforts have focused

primarily on research to investigate the impact of PDAs, and the roles within the multi-modal agent system for structuring a guideline.

1.3 Research Aim and Objectives

The research aim is to evaluate the impact of PDAs on users, research the particular area of PDAs applications on multi-modal agents on use, and analyse each component or function in the multi-modal agent application, Finally to develop a framework that implements each component or function into one application, and then design the possible elements of this application to test and validate them by use of questionnaires.

The research objectives are analysis and evaluation of multi-model agent system based on PDA that carries out the functions of an intelligent agent technology, the multi-modal interfaces.

- **The impact of PDAs**: to investigate the impact of the PDA on people's daily life, and various fields the PDA has been used for.
- **Analysis of each function of multi-modal agent system**: to discuss the key roles of intelligent agent, multimodal strategy, and web services employed in the multi-modal agent system, and the proposal of the integration of intelligent agent and web services.
- **A Framework for creating a multi modal agent system**: to develop a framework for creating a multi modal agent system from both functional and technical perspectives, in order to evaluate this system.

1.4 Organization of the Dissertation

Along the progress in investigating the multi-modal agent based PDA and developing a framework for it, this dissertation is structured into the following six chapters. A research into the impact of PDAs is conducted on chapter 2 which evaluates the PDA's impact in the user environment and their functionality. This chapter focuses on the validation of technical impact of PDAs, the social impact and future research on users, each perspective has been particularly discussed in the area of introducing the utility of PDAs, analyzing the functionality of PDAs and paradigms that PDAs have been applied successfully.

Chapter 3 is investigating existing research and an analysis of the key roles of the multi-modal agent systems which are; agent technology, multi-modal interface and web service. These components has been specified in detail as being the foundational evidences to support following research and framework of the multi-modal agent system, moreover the interaction of each role also is presented in this chapter, and proposed a standard for the seamless connection of web services and agent technology.

Before a framework of the multi-modal agent system will be structured. Chapter 4 demonstrates an overview of the functional perspective, and describes the design goals and functional requirements in order to create a preparation for developing a framework of the multi modal agent system. As well as the system scenario diagrams are illustrated in this chapter for presenting the functional specification.

Chapter 5 is a technical document that investigates each of roles of the multi modal agent system from the technical perspective, and finally creates a system framework according to the functional specification. In this chapter these components are completely analyzed in detail as to how they work with each other, even some parts of the system have examples of codes for implementing and testing.

During the research of the multi-modal agent system, an online survey was undertaken and some interviews as well, in order to collect valid information from people who are experts, experienced and non-experienced users to evaluate this investigation. Chapter 6 is based on analyzing the data, which was gathered from the survey and interviews, to create some charts and introduce methodologies used in this dissertation.

Finally Chapter 7 concludes the entire research work that has investigated this multi-modal agent system and brings issues for future research.

Chapter 2 The Impact of PDAs

Over the course of this chapter, the impact of personal digital assistant (PDAs) are described focus on three main points which are technical, social and the future impacts in order to present an overall perspective on PDAs.

Firstly this section begins with a general discussion of how the PDA getting popular, after this basic introduction a review is presented of the technical impacts of PDAs, which include the evolution of PDAs, a comparison of traditional PCs and PDAs based on aspects such as capability, the development of hardware and software and some limitations of PDAs. Secondly this review moves to the social impacts on people using PDAs and covers communication which contains knowledge attainment and information exchange, comparison of traditional PCs to people and PDAs to people, practical implementation of PDAs and also the shortcomings in the social impacts. Finally this chapter does which a look at the future of PDAs.

2.1. The Popularity of the PDA

A Personal Digital Assistant (PDA) is a handheld computer of palm size that assists people in managing, storing, displaying, manipulating and organizing personal information and data, for example, it works as a clock and calendar that can help users to keep track of scheduled events or appointments, descriptions and dates or other information, as well as some have games. It generally has similar facilities to the traditional PC, that means a microprocessor, data storage memory and an input/output mechanism (such as a virtual keyboard in a touch/pressure sensitive screen which it can be activated by a attached stylus used to input information or data) compose a handheld device in order to provider a graphic user interface (GUI) environment, that allows users to direct commands and invoke applications by inputting information using a simulated keypad displayed on a liquid crystal display (LCD) *[about.com 2007]*. A PDA offers computing processes and storing information and retrieving data capabilities. In common, it can store considerable amounts of data, for example schedule data, to-do list data, and a note pad data, and meanwhile a data communication refers to the transmission of digital data to PDAs. It can transport inside data to a traditional PC and also receive data from the traditional PC by using a portable information terminal known as *synchronization*. Moreover the PDA is not

only available for managing personal information, but it includes most of the functionality of a computer, and has been already extended to offer computing functionality, faxing, browsing, networking, emailing and can even make a telephone call, but the most important thing is that they all can be implemented using a wireless function. In general PDAs keep several extensible connecting ports for extending functionality and connecting other peripheral devices, at least one for each extension and cradle port, meanwhile the extension port is posited on the top of PDA for add-on card in order to extend additional functions of PDA, such as extensible and plug-in Global Positioning System (GPS) card based on the Compact Flash (CF) port, also there is another cradle port which is used for fixing the PDA on a cradle and connecting the PDA to the cradle with other sources, such as the PC can connect this cradle port for controlling processes, updating data, accessing the Internet and can even charge PDA's battery *[US patent 2004]*.

Today's PDAs are multi-media products having new functionalities such as telephone call and facsimile transmitter which has recently become the most significant function of the PDA, the PDA can be used for browsing the web, running data sharing applications through the internet, internal and external networks and sending or receiving electronic mails, alternatively they are be able to run a variety of applications, for example, nowadays PDAs provide the capability of wirelessly sending and receiving email, browser sharing information, e-books, games and databases. The use of PDAs has evolved to turn into an effective tool for managing personal and business daily processes due to the convenience of its portability and data access via the internet, in term of accessing the Internet the PDA can either use a USB cable connect to Internet through a PC or incorporate a wireless connection, for example, a PDA handles facsimile service or Internet access may use additional connector during a synchronization process with an external database located in a PC, also can communicate over the Internet through a mobile phone connected to a wireless communications network to conduct information transmission, moreover the information the user needed can as well as be downloaded into a PDA by ways of both connections mentioned above *[Mahd 2006]*.

Recently the uses of personal digital assistants (PDAs) have become an extremely popular apparatus with consumers for many social and technical uses, because of the

multi-function data processing and large storing capacity. According to a report summarized in the first quarter of 2007, PDAs using Windows CE *[Windows embedded community 2005]* grew by 64% and drove the growth in the market overall by 39.7% *[Padilla 2007]*. Among the main reasons for their popularity is their size which is roughly the size of a deck of cards. The PDA is a compact computer that can be carried easily in the pocket, for the last few years PDAs have become a necessary device for personal and business uses, one of the reasons which the PDA is a portable and expandable item, easy to start, and has a thoroughly complete accessories to allow users setting a variety of application for their preferences *[About.com 2006]*. During the designing of the size, their systems such as Windows CE, Palm OS, Symbian and Linux are as well as lightweight and extraordinarily convenient and portable. As to their compact size, PDAs are easily carried about which is very convenient for today's uses.

2.2. PDAs: A Technical Perspective

A Personal Digital Assistant device consists of a variety of versatile technologies consisting of hardware and software, such as touch screen and microprocessor, as well as video recording, typewriting and word processing. It has been applied in many fields, combined with other different technologies in order to make more features, such as GPS and electronic facsimile. This section will describe the technical uses of PDAs technologies.

2.2.1. The Evolution of PDA

According to Evan Koblentz *[Koblentz 2005]*, in the early 1980 Toshiba was the first firm to sell an actual integrated organizer known as the LC-836 Memo Note 30, which was an electronic digital memorandum, its features emphasised the following functions; calculator, phone numbers, addresses and simple memos. However the first true PDA device was developed in 1993 by Apple Computers Inc which launched the Newton Message Pad; it was the first handheld device to use a touch-sensitive screen and a handwriting recognition software *[Commentary 1993]*. John Sculley, the former chairman of Apple Computer Inc predicted the PDA device would become an omnipresent item *[Kirvin 2001]*, including address list, note pad, calendar and data exchange utilities wireless functions, but at the time the Newton was not able to deliver those features entirely while it was released. after three years in March 1996

the Palm Computing Inc was incorporated into US Robotics Inc, after that it delivered the first commercial handheld computer called Palm Pilots (Pilot 5000, 1000), this device was developed using new generation Palm OS v1.0 with a Motorola 68328 processor at 16MHz, and had 128k memory *[Kairer 2007]*, it was extremely technical handheld computer of that time especially when the Personal Computer was still in infancy. It was a smaller portable device in size and more functionality such as holding 500 addresses that assisted people to manage and organize their personal and professional information, also it has a graphic user interface device and the ability to synchronise data with a traditional PC. Until 1999 Palm Inc still kept developing better and better devices, for example, Palm V has the size about 115mm x 77mm x 10mm, weight about 115g and pels about 160 x 160 with features about daily book, address list, personal information manager, notes and games *[Palm.com 1999]*. But in October 1999, Palm Inc delivered the wireless Internet market with the introduction of the Palm VII organizer which allows users a wireless means of quickly, securely and easily obtaining information from the Internet, such as flight schedules and news headlines, and processing online transactions.

PDAs are currently able to run powerful applications and software that would have been on a fully-functional PC only a few years ago, for instance word processing, data analysis, information exchange, communications and multimedia applications, furthermore the microprocessor of the PDA can perform high- capability tasks, and electronic memory storage make it possible to store plenty of documents and applications on a PDA, eventually new protocols for effective wireless communications extend the control from the PDA to other useful accessory devices.

2.2.2. The Capability of PDAs

PDA technology has emerged as a viable technology for accessing and updating information on personal and business use, it is in a sense a handheld computer which has a significant advantage in size and portability. The most important capabilities of PDAs are discussed below.

- **Connectivity**

Wi-Fi is a major function of the PDA that provides a fast Internet connection to a network or PC through wherever there is a wireless signal, although some pervious

PDAs do not have built-in Wi-Fi, but there is expandable memory in forms of Secure Digital (SD) and Mini-SD card. Also PDAs have the ability to access networks wirelessly via built-in wireless capability or other optional wireless cards *[MobileTechReview.com]*.

- **Synchronization**

The synchronization of a PDA module enables the synchronization of contact data being held on the PDA and the data held by complementary applications on the computer, which means the PDA can be synchronized with a PC to update data such as list of contacts, memo pads, to-do lists, e-mail clients, schedule planners and off-line internet and web information. Another function of the synchronization is the periodic synchronization used as a backup mechanism for preventing the loss of data stored on the device, information in this mechanism is stored independently on the PDA or either PC side. Alternatively the data input can be done on PC side via the synchronization, because transferring data to a PDA via the computer is thus a lot faster than having to text input all data via a touch screen which is not optimal.

On the PDA the implementation of synchronization with the PC is generally based on software, such as Microsoft ActiveSync *[Citrix.com 2007]* for the Windows CE system on mobile handhelds and HotSync Manager *[Citrix.com 2005]* which comes with the Palm OS system, These applications enable the PDA device to be synchronized with the Personal Information Manager (PIM) whose purpose is to facilitate the recording, tracking and managing personal data, the PC however allows PIM synchronizations with software such as Microsoft Outlook along with such as Internet "favorites", files, tasks and other data types. Other synchronization software developed by third parties is also competent for the ability of synchronization such as CompanionLink and Intellisync *[Gilliland & Roach 2006]*.

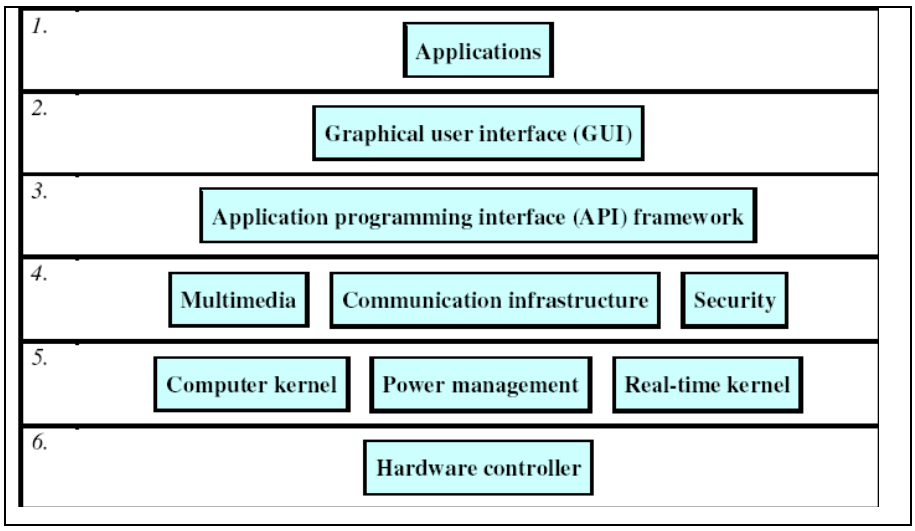
- **Customization**

Comparing the PDA to the traditional PC, the PDA is similar that it has a customization function, for example, it is possible to install additional applications on most PDAs for applying more uses in which applications can be bought or downloaded from the Internet or other sources. It allows people to set their own preferences to personalize, such as changing display theme, background picture and

alarm's ringtone for the PDA. From the hardware perspective PDAs always allow users to extend some form of hardware. An example of this would be a memory card slot, which enables the users to get extra and exchangeable storage space on their handheld devices; furthermore there are some headsets, mice and miniature keyboards that can be linked to most PDAs for quicker text input through wireless connection such as Bluetooth technology *[Cecil 2005]*.

2.2.3. The Software on PDAs

A PDA system is one of embedded systems designed for a particular kind of application device, unlike a general purpose workstation, desktop or laptop computer, its system usually uses microprocessors, or customised chips or both to perform predefined or very specific tasks, In embedded systems, because the system is designed to specific tasks, so it can be optimized for reducing the size and cost of the devices. The figure below presents a generalized structure of the PDAs' Operating System (OS) that can be expressed as a six-layer stack.



A generalized structure of PDA OS [source: hu, yeh & chung 2005]

Figure 2. 1

Figure 2.1 illustrates a generalized PDA operating system structure, which can be visualized as a six-layer stack. The very bottom layer is hardware controller which controls the management of hardware; the fifth layer describes high real-time performance and power management; the fourth layer supports the basic infrastructure for secure multimedia communications; the Application Programming Interface (AIP) framework enables multimedia be added to applications; the second layer is a user

interface that allows people to interact with a PDA; finally the applications is PDA programs.

In this case, there are four most representative Operating System environments for PDAs are Palm OS, Symbian, Windows CE and Linux.

- **Palm OS:** was first introduced for the PDA operating system, it was developed initially to be a easy-to-use and similar to desktop OS for the Palm Pilots and is a very compact OS and a single purposing environment that allows users to access Microsoft Word, Excel and Power Point files. The Palm OS was first time released in 1996; the most current version is called "Garnet" which is Palm OS v5.4 *[Operating-System.org 2004]*.
- **Symbian:** is structured in a similar way to the Windows system with pre-emptive multitasking, multithreading and memory protection to allow users to create files such as addresses, copy and paste files and view properties. It is a proprietary operating system that restricts the software and applications, which can run only on their operating system *[Sulaiman 2005]*.
- **Windows CE:** is high modular embedded system that users can request the customization of it for satisfying their requirements that means only some needed modules will be kept and other unnecessary modules will be removed. Its functionality and graphic User Interfaces (UI) are both the same as the standard version of Windows. A good example of this Operating System is a Smartphone that is fully functional cellular phones with integrated many of the features of a traditional PC, it has equipped along with a wide range of software applications specially designed for mobile devices. They are also capable of accessing the Internet, usually over a cellular wireless connection (GPRS) and can be used to access the Web for checking email on anywhere *[Sulaiman 2005]*.
- **Linux:** is an open source operating system which has migrated to the PDA and a first Linux-based PDA was released by the end of March 2000 called Yopa 3700 PDA *[Tuxmobil.org 2006]*, which brings the flexibility and power that Linux can add to the PDA. It is multimedia PDA designed closely resembles the Sony Clie's clamshell design *[Story 2000]*.

2.2.4. Software Limitations

The most significant limitation on the software restrictions of the PDAs is the instability of system. Although it is not an extreme or critical error, it is a source of frustration, because the PDA units are limited by their size; portability and cost. Furthermore the crashes may result in the loss of data, because all data is always kept in RAM disk that will not be erased by a "soft" reset. In addition if the crashes require a "hard" reset on the PDA, this means all the data will be erased from the PDA *[Carroll, Saluja, Tarczy-Hornoch 2002]*.

In addition, a multimodal agent is quite a big application, which requires high quality of both PDA processors and memory, because the agent technology has to invoke different functions such as Web services using multi modal methods in order to implement all of the features of this application. Therefore the support of good hardware is necessary.

2.2.5. Hardware on PDAs

The PDA is a palm-sized device with a display screen; it takes up much of the device's front surface. It has no hardware keyboard, but software one used by a stylus that is a vital accessory for the PDA. Moreover the term handhelds usually refers to devices with an integrated keyboard, in the PDA case there is external folding keyboards available, but they are peripherals that attach to the unit.

- **PDA Processors**

Today's PDAs enable users to do a lot of operations and transactions, and run some big applications, for example, playing MP3 music files, accessing the Web, running complicated games, and doing much of what a full-fledged PC can do, with current high-speed processors they have over 400 MHz, for example the Dell Axim X51v PDA 624 MHz Processor.

Slower Central Processing Unit (CPUs) are generally used in the early entry level Palm-based devices that are sufficient for the basic uses of the PDA such as managing tasks, editing contact lists, jotting down to-do lists, and working with simple spreadsheets. Otherwise faster CPUs are implemented in Pocket PC (Windows CE system) and other advanced Palm devices, they are essential for rapid performance in

multimedia applications and for dealing with more complicated transactions or processes such as Word and Excel files comparing to older CPUs handhelds *[Fox, Felkey 2003]*.

- **PDA Memory**

All PDAs definitely have a certain amount of internal Random Access Memory (RAM) used to support the applications running and temporary data storing while users use PDAs to handle processes. In the early period those slower PDAs had internal memory around 32 MB that is sufficient to deal with, for example, storing contact lists, calendars, simple documents and worksheets. Users have to have more space for holding larger documents and storing multimedia applications such as photos, mp3 files, and video, in this situation the PDAs have to have either an SD and/or a CF slot to plug in expansion slot *[Pdaonlinecenter.com 2006]*.

- **Touch Screen**

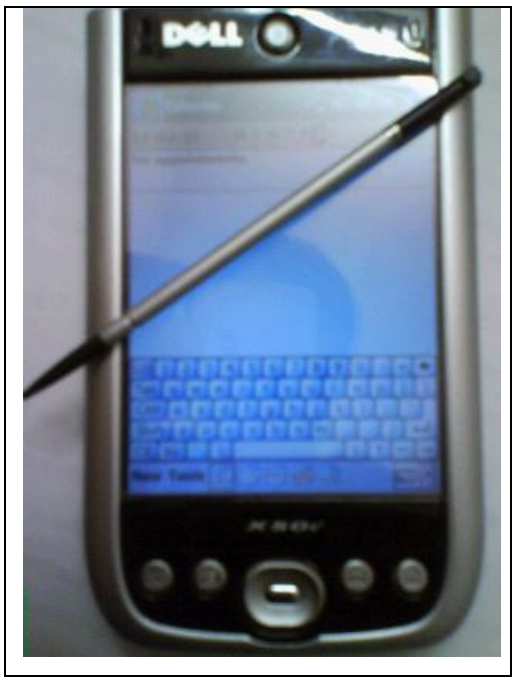
A touch screen is used together with a stylus for taking the ability of the mouse for application control. is PDAs feature touch screens for user interaction, the touch screens have a high resolution display in PDA's to ensure that the display is colour and displays 265k colours, which is always better to enable users more effective and straight to communicate with PDAs while navigating with a detachable stylus which allows users to tap the screen to activate buttons or menu choices, and drag the stylus to move files. All current PDAs come with in colour backlit LCD screens, which allow users to take full advantage of multimedia and entertainment options, also another innovation is to offer more viewing area and landscape mode *[Geiger 2006]*.

- **Expansion Slots**

The expansion slots are ports which support the expandable memory for PDAs to obtain extra memory instead of the built in memory which is not always satisfaction to the users' requirements. They may take different kinds of memory, such as SD cards which is an outstanding expandable memory for PDA's, it is cheap and samll, currently with up to 4GB of memory and more add-on features, for example TV tuner SD cards allow users to watch TV on PDAs, Wi-Fi and Webcams SD cards enable wireless internet and video chatting on PDAs. But if other facilities are added on, the slot will not always be available for adding memory, furthermore some PDAs also have a USB port for USB flash drives mainly *[Pdaonlinecenter.com 2006]*.

- **Keyboards**

The PDAs system for input of data is usually handwriting recognition software, it is quite easy to use a stylus to enter each character via PDA's touch pressure screens. But this is slow for inputting therefore some users prefer to use a real keyboard, the option is to set a peripheral keyboard via the PDA's expansion slot.



An Internal Keyboard and a Stylus of PDA [source: the author]

Figure 2. 2

As illustrated in Figure 2.2 above, the standard way to input data into PDA is that a virtual keyboard inside a PDA and a peripheral stylus attached with the PDA. In the following chapters this document describes a multi modal interface for voice signals inputting instead of the slow and small keyboard.

2.2.6. Hardware Limitations

The key limitation of PDA technology is the small display/screen which is factor to affect the visible effectiveness of sight, for example, in Palm OS PDAs have a typical display with eleven lines of possible text on it along with a maximum of thirty characters for each line, but this is obviously not sufficient space to contain information for any uses of large content in the current day, such as reading large amounts of text on a palm sized display will be difficult, otherwise users are not simply aware to all information that may have not displayed properly but it was intended to be communicated with users on the PDA, while the presence of information is restricted.

Another limitation of the PDA is the speed to input data. Recently the PDA system have developed a new data entry tool that is superior to pervious handwriting recognition, called "graffiti", but it is not convenient and fast as actual handwriting (pen and paper). Therefore there are a few additional keyboards units for overcoming this problem, but again it will cause another problem which the peripheral keyboard would annoy people installing/unstalling it, also it may cause thumb strain over time and is rather fragile *[Carroll, Saluja, Tarczy-Hornoch 2002]*.

2.3. PDAs: A Social Perspective

The use of PDA devices is increasing at a dramatic rate. A PDA offers the user a convenient environment to be able to manage things such as meeting schedules, diaries and addresses lists; they also enable the user to directly add new entries such as meeting appointments in a form of a pocket, portable and transportable unit. In this section the social impact of using PDAs is presented, additionally how PDAs are affecting people's communication on the social impacts and some paradigms of PDAs uses, finally it is the shortcoming of the social impacts on PDAs.

2.3.1. Social History of PDAs

Computers and information systems are becoming more and more ubiquitous nowadays, as well as a large increasing number of inexperienced users wanting to use high technologies. Thus it is essential to build a lightweight system in such a way that allows general users who are unfamiliar with high technology also are able to use them. This is of extreme significance when the lightweight system could be able to be developed for the involvement in public social life. Nevertheless, PDA manufacturers are doing innovations of evolving with new technology, and offer users a gradually more efficient and functional experience.

The first PDAs were introduced to the market in the mid 90s that were little more than glorified notepads, making it possible to record notes, contact lists, calendars, and up to date meetings/appointments on an form of electronic units *[Haskin 2007]*. Progressively these electronic units have evolved other the last few years, such as The Palm Pilot devices gained a loyalty of reputation and a following of both general-level and enterprise-level users. Eventually new features containing 3D games, multimedia applications, and WI-FI uses have become commonplace on modern day PDAs, that

are performed by PDAs are being slowly but unquestionably integrated into the many advanced cellular phones (Pocket PC phones are embedded with the Windows CE system) and mobile email devices (Blackberry based on open standards platform) phone now on the current market. Although the cost of PDAs is a prohibitive barrier for the majority of consumers, but the convenience, portability and functionality have been already exceeded the only expenses on PDAs while compare to their features, such as the SmartPhone which is a converged device and integrates mobile computing technology with wireless communication, meanwhile the wireless features provide a connectivity to the Internet capability, organizational features, entertainment and other PDAs with add-on hardware and software. Ultimately the SmartPhone has become an institution among enterprise users, and is now a necessity for many mobile businesspeople during their social life users *[Koblentz 2005]*.

In the last few years since the PDA has become more omnipresent, a increasing number of businesspeople, ordinary and professional users have been purchasing Personal Digital Assistants, a electronic unit designed to organize and simplify social life. For instance, according to an online diary is called "A Day in the Life of a PDA User" *[Killick 2007]* which describes the author's one day daily life about using PDA. In detail this diary shows how the author used a PDA to be an Advice Worker during all of events in a general day he has been through, the following schedules are his one-day life referred to use the PDA at each certain period of time.

"0800: *Wake up by the clock on my PDA.*
0830: *Plug my PDA into computer and synchronise the podcast I downloaded last night on issues regarding social housing. I might get a chance to listen to it later!*
0900: *On the way to work I decide to listen to the podcast.*
1000: *A reminder goes off on my PDA to tell me that I have a meeting in half an hour. Better get going!*
1005: *I can't seem to remember the address of that meeting! Never mind, from my calendar I find the contact of the person I am going to meet along with the address.*
1300: *Fresh out of the meeting, I have arranged to a second meeting again*

next month. I add the appointment to my PDA and synchronise it with my office server. This way, my office know that I am busy and can't double-book me. This also brings down my latest email and I check these quickly.
1400: *I'm at my next appointment now and the client has asked for an update on booking a place for them on the Carers' holiday we arrange every year. Our database is web-based so I open a browser on the PDA and log in. I find the details and confirm that two places have been booked.*
I record on the system that our Administrator needs to send a confirmation letter regarding the holiday so transport can be arranged. Now we have a record of this the Administrator will see that this is done this afternoon.
1530: *On the way back to the office I get a phone call on the PDA. I need to visit a client who is unhappy about their condition of the housing.*
1555: *I drop in to look at the property. I take some photographs of it using the camera on the PDA so that I can show colleagues later.*
1625: *Back in the office and time to get some work done. I plug the PDA in to my computer so that it can recharge.*
While this is happening I download the photographs I took earlier. I then email this them to a colleague so that we can discuss the best way to resolve the issue.
1800: *Leaving for home I do one final sync. This puts in all the appointments that have been recorded for me whilst I have been out. Lucky I did as my first meeting tomorrow is cancelled!*
2300: *I reflect on this little marvel. Without it, I would have to have organisational skills of the highest order."*

A Day in the Life of a PDA User *[Source: Killick 2007]*

Table 2. 1

In Table 2.1 that the PDA is the determinant item for this user's life, each period of time the PDA has been used as a crucial device that helped the user a lot, for example the user might forget some important appointments or meetings during each eventful day that could cause significant losses for the user.

2.3.2. Communication with a PDA

PDAs have been used for a great deal of purposes, such as enterprise, education and entertainment etc, but the most important usage of the PDA is that of

communication. PDAs may function as intermediates of communication between people and offer people new interaction channels for social communication; there are many ways of the social interaction that can be implemented on PDAs, for example, it can take the form of texts between PDAs users, making voice connection; it can have knowledge attainments for updating data, reading news, email and browsing the Internet; and it can make information exchanges to connect people.

For example PDAs have been used as a portable device for communicating between people with aphasia *[Davies, Marcella & McGrenere 2004]*, this communication based on PDAs which offers people with aphasia, when communicating with limited language skills, it has a form factor and feature set for being an effective communication tool while people with aphasia. A PDA can be incorporated into many strategies in order to support communication between people with aphasia, like a graphic application can provide the users an interface when they try to communicate with other people; and they can use PDAs to exchange any information to communicate each other.

A PDA has a range of uses in the interactive communication for social interaction; it is an effective tool to establish the technology affects people's experience and emotions for diverse social settings. Another good paradigm is the business card exchange, whenever people ask for business card exchange, it can be beamed from one PDA to another PDA, and thus the information of business card is instantly recorded in the receiving PDA for eliminating the possibility that the normal business card might be lost. Also the information can be immediately updated to a database over a network, so people do not have to go back office with the normal business cards, and inform everyone the database has been updated *[Brown, Brown 2001]*.

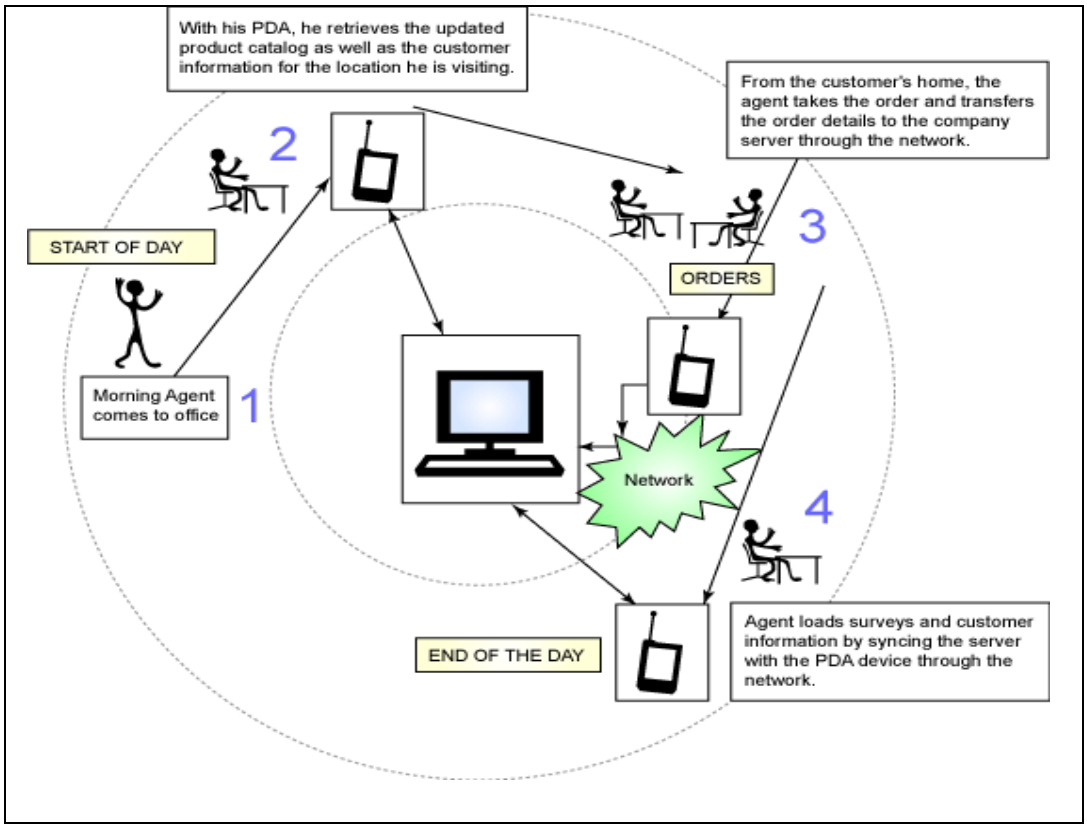
2.3.3. Practical Implementations of PDAs

The Personal Digital Assistant has developed from a task organizer for business, ordinary and professional users to a powerful, multi-featured, and multi-purpose device; it is being used in wide domains with different features, such as business, education, medicine and GPS, for example, students use of PDAs have moved well beyond the interesting, intelligent and technological gadget to become a useful and helpful academic tool *[Carroll 2002]*.

2.3.3.1 Business Use of PDAs

PDA allows users to conduct business without being tied to their desk, which is like an electronic planner and helps the users to organize their work and communicate with one handy device. The following example shows the importance of PDA for a user using in business.

According to a report from the IBM website called as *"Pervasive computing: e-business anywhere, anytime" [Jhoney, Kuchhal & Jaiswal 2003]* which depicts a pervasive computing-enabled e-business, also known as wireless e-business solution, it is integrated into the existing e-business infrastructure to give people convenient access to relevant information and the ability to act on that information anytime, anywhere, for example, the efficiency of the computing-enabled e-business can bring the immediately benefits of increasing sales, better customer service, and lower transaction costs. A PDA based wireless e-business solution can be the same to reduce paperwork and processing time, but also being a lightweight device, the PDA has the advantage of portability and a cheaper alternative to laptops. Therefore it can be easy and fast to transport and use without considering the limitation of weight and size. Figure 2.3 presents a convenient PDA-based sales application with an agent's daily routine.



One day life of using an agent based PDA to handle e-business. *[Source: Jhoney, Kuchhal & Jaiswal 2003]*

Figure 2. 3

This paradigm presents a business user using an agent based PDA, in the morning the agent authenticates itself to connect office network and update catalog of products and availability status, and information about customers for visiting them; and then the agent based PDA stores customer's orders, feedbacks and survey data, also shows customers any information required, all those operations are from the company server through the network; finally the agent will update all stored information to main server while the user get home.

Otherwise for selecting a good PDA for business uses is very important with promises, guarantees, evolving equipment and changing features, an excellent example there are devices such as, Dell Axim X50v PDA is good choice that has the ability to handle any business applications and/or documents, as well as the integrated Bluetooth and Wi-Fi for wireless connectivity makes users have a mobile office *[Brian 2005]*.

2.3.3.2 Clinical Environment

Using personal digital assistants in the medical environment, specifically in the social context of interaction between doctor and patient is important, which the uses of PDAs in clinical environment have been divided as the following aspects.

- **Advice system**

The application combing a database on a PDA can be used to record patients' relative information/data, such as the features of basic profile, which doctor in charge of the case, previous medical record, and current medical record, checking report, laboratory data, and diagnosis of diseases, treatment plan, discharge notes and so on, in fact this system is sort of Electronic medical record. Recently there are applications such as Patient Keeper, Ward Watch and Patient Tracker. The PDA can transfer prescribed instructions to the main system via the functions of WI-FI connection or synchronous data exchange in the first time *[Chignell & Lottridge 2006]*.

- **Medical clerk system**

This system can be more interactive with patients in which doctors treat the patients, after treatment the doctors can use an application running on PDAs to complete the synchronizing, capturing and storing needed data on a database through wireless networks or hotsync server, such as iScribe and ePad software *[Luo, Hales & Servis 2002]*.

- **Clinical computing**

The PDA can be used for computing clinical formulas; there still have a lot of programs that they merely have to input correlative data to compute the results automatically as doctors required *[Chignell & Lottridge 2006]*.

- **Clinical electronic document**

The PDA can be used as a storage media for storing a variety of medical document while doctors can access medical data at any moment. Generally Microsoft office can be useful for creating clinical documents *[Chignell & Lottridge 2006]*.

- **Clinical project tracking**

A PDA allows doctors to be able to supervise project tracking. Such as some project software enables management of many researches program, it can be integrated with Memo for the control of time tracking *[Chignell & Lottridge 2006]*.

- **Clinical Internet access**

Doctors can browse the medical Web *[MedicWeb 2007]* to obtain any useful or helpful data through a PDA device, which has the very powerful advantage as portability, for example, doctors can carry a PDA anywhere for accessing the articles on medical journals, medical news and the medical website content [Chignell & Lottridge 2006].

2.3.3.3 Navigation System

Mobile computing systems based on personal digital assistant products have been used widely. The navigation system function provides vast benefits for mobile users to operate computers when moving and connecting to the network anytime and anywhere in different ways and combine themselves and immense digital space seamlessly. The system is exploited on such as mobile computing, embedded developing, computer graphics and wireless networks, and the science of surveying and mapping, such as the principle and data organization methods of geographic information system, the visualization of digital maps, location and navigation based GPS and the technology of "3S" integration, which is the integration of Remote Sensing (RS) plus GPS and Geographic Information System (GIS) *[Zhang, Wu, Chen & Sun 2004]*.

The developing navigation function embedded into PDA devices actually is a spatial information visualization system based on PDA, which this system integrates the mobile computing technology and mobile map technology and it has satisfied the visualization, query and analysis of spatial information. Besides it is more important on the social impact for that on the one hand the system is mobile and enables users to enjoy Location-based Service *[Bennett & Capella 2002]* anytime and anywhere, on the other hand users can customize information at will so as to meet the requirements in different professional fields.

A Photo of Earthmate Application. *[Source: earthmate 2007]*

Figure 2. 4

The above Figure 2.4 is Delorme's second generation GPS that can work with a PDA device, it enables the user to download maps and create trip directions that can then be downloaded into a PDA device in order to have the newest information of maps *[earthmate 2007]*.

2.3.4. Shortcomings in the Social Impacts

The security of personal information on PDAs is among biggest shortcomings of using PDA devices, among the most popular uses in business, the users need to access enterprise data, like corporate e-mail, contacts documents, business messages and some meeting schedules. All this data have to be ensured as confidential privacy, but the unmanaged and unsecured results are increasingly growing to put users' devices in damage, two examples that can be clearly presented as follows;

- **Virus impacts**: the mobile malicious software and viruses are rising in frequency and impact. This exposes an extremely fatal vulnerability whenever these compromised PDAs connect to company data networks, wirelessly or through cradled synchronization, to access company data. Moreover the unsecured public network might be a source of any infections while also an unsecured PDA tries to request access *[Strauch 2004]*.
- **Accidentally loss**: for any reasons people could accidentally leave a phone or PDA in a public area like hotel, taxi or restaurant. This situation happens

commonly and sometimes it can not be avoided, finally it can expose any personal or business privacy *[Strauch 2004]*.

2.4. The Future of PDAs

The development of wireless technology has combined wireless communication and the internet, among the connection devices, the mobile device is the main connection of interface between a wireless network and the major end-users. Again as the main device of mobile device the PDAs are becoming more a part of daily life for business and personal use, they still suffer from the same problem that they always have had in previous incarnation of PDAs. There is a common drawback which the input of typing information into a PDA, particularly a touch screen model that remains very slow. Most people can type and write faster than they can jot into a PDA, in this case previously suggested a future product which has currently became reality, a speech-to-text interface, great example is Microsoft Voice Command *[microsoft.com 2006]*, it uses speech recognition technology to record the voice command in order to replace inconvenient pen and touch screen interface.

PDAs have been a subject of concern; its market is quite higher than the actual market demand. Currently PDA trend is predicted generally towards four directions:

- Smart electronic products: smart appliances, but more unlike form of the PDA.
- Multi-featured handheld computer: satisfying personal mobile office, communication requirements and becoming into one handheld-computer equipment with information processing and data communicating, such as high-end smart phones, tablet PCs.
- Object Oriented applications: the applications of PDAs are the focus in the nearly future market which would concentrate on professional uses only, such as the GPS usage is current the big part of PDAs application market.
- The connectivity of the next generation of PDAs will focus on wireless networks, for example, using Wi-Fi and Bluetooth is set to become cheaper and much more common over the coming year

2.5. Conclusion

This chapter introduced the impact of PDAs dependent on the technical impact of PDAs which is based on the technical perspective of view, introduced the history and evolution of PDAs with how would they innovate PDAs from the originally simple assistants to the currently powerful toolkits, as well as another important social impact which is based on the view of influence on people's general life with a range of different examples, eventually to depict the future of the PDA on a variety of prospects. In the next chapter the research will be focused on the details of technologies for constructing a multi-modal agent system based PDA.

Chapter 3 Discussion of Multi Modal Agent System

In this chapter, the focus is on knowledge attainments of researches and investigation based on three key components in the Multi Modal Agent (MMA) system based PDA, and also this chapter introduces those technologies in the perspectives of introduction and exploitation and/or application. In particular, this chapter finally explores the influence on interaction of each technology from such the deployment of agent technology and Web Service for outputting information using a Multi Modal Interface according to users' preference.

3.1. Introduction

The multi-modal agent system consists of Agent Technology, Web Service and Multi Modal Interface; each of these roles has different features in the MMA system. An agent technology based on multi modal environments to form a new technology, known as multi modal agent, which combines agent technology within multi modal environments to handle operations using the benefits of them. The following terms are specifically needed to discuss, such as "mode" "multimodal," "interface," "agent", "web service".

- **Mode**: generally used to refer different and separate methods or mechanisms carrying different kinds of data for the purpose of communication with other technologies, such as the way of speech, intonation and display on screen.
- **"Multimodal"** refers thus this collection of many such mechanisms.
- **"Interface"** conventionally means the place where two different systems meet each other, in here it is the machine meet human by using "multimodal interface".
- The term **"agent"** has served a variety of meanings, but can be referred here for assisting users and acting on their behalf to carry out non repetitive tasks.
- **Web service** provides the interoperable interaction between different Web applications over a network.

A multimodal agent in other words, an agent can be invocated and communicated with two or more different modes of operation in which the communication build through different modalities and modes, such as for modalities there are speech, gesture, screen and tapping, whereas the modes determines the context in which the

data is interpreted, refer to numerous functions and conditions in which two or more different methods, processes or forms of delivery are used. For example a multimodal agent handles an operation in multiple ways through intonation and speech, facial expression and gesture, posture and body actions. When this agent obtains instructions, it has to communicate with master systems in all modalities and to control the distribution of information across modalities, selecting which one is most suitable for each content and how they can combine with each other.

The following Sections describe the different functions of agent technology, Multi modal and Web service. The Web Services technologies provide the basis for interoperability and dynamic discovery and integration of services; their benefits exist not only on the Internet but also in the any integrated usages *[W3C 2002]*. In this research finally Web Services will be implemented on mobile devices combining with agent technology by using multi modal interface.

3.2. Agent Technology

Agent technology is becoming increasingly used within information technology to depict a wide scope of computational entities which these entities can be relatively simple systems such as that the TIP WIZARD assistant in Microsoft office *[Maes, 1994]*, which helps users with advices working in EXCEL 5 *[Maes, 1994]*, or even professional interoperability systems or databases such as Oracle *[orafaq.com 2005]*. In the level of agent technology, it can be seen as performing three types of behaviour, first level which is simplest level called *"gopher"* agent, it performs comparatively easy tasks based on well-defined, pre-specified rules and assumptions such as alarms which can simply reminder users with per-setting; second level *"service performing"* agents which is more sophisticated to carry out well-defined task at the request of user such as find information users needed with setting preference; the final level which is capable of flexible autonomous behaviour known as predictive/proactive agent such as robots *[Jennings & Wooldridge, 1996]*.

3.2.1. Benefits of Agent Technology

Agent technology is a rapidly developing interdisciplinary field in the computer science; it focuses on the use of autonomous software entities with the interoperation of abilities to with other such software entities in a dynamic environment, which

performs intelligent and sophisticated actions. An agent normally has a variety of functions for dealing with different type of circumstances, in term of functions they are explored as follows.

- **Reactivity:** the ability to perceive its environment and respond with correct action in a adaptive behaviour, for example an agent can observe its environment to make correct decisions while any new changes occurred
- **Autonomy:** an agent is proactive, goal-directed and act in response to the states in its environment, it can also perform tasks without external initiation, confirmation, and notification.
- **Mobility**: an agent can be able to move some where in a self-directed way following an itinerary from one host to another.
- **Adaptive**: an agent can dynamically learn and adapt its behaviours from previous experience.
- **Personality**: an agent has the capability to manifest the traits such as emotion.
- **Collaboration**: an agent can cooperate with other agents in achieving a common aim that they are self-organized to coordinate and negotiate
- **Interaction**: an agent has the interoperability to communicate with such other agents, humans and legacy systems that the language between them is similar as human language rather than typical computer programming languages and protocols *[Bradshaw 1997]*.

3.2.2. Adoption of Agent Technology

The number of agent based adoptions being deployed and developed in real world settings is rapidly growing; the adoption of agent technology is the procedure that the developing computer software use agent technology, for example, developing of agent based systems contains building sophisticated, self-contained components *[Jennings 1999]*. Which can continually active in order to reach their goals and can flexibly interact with other a number of independently developed similar components.

3.2.3.1. Enterprise

There is a paradigm of a simple model of the lifecycle that companies using agent technologies for improving effectiveness and achievement showed in table 3.1.

Step 1	Lack of knowledge on agent technologies
Step 2	Understanding to potential and existing of agent technologies for application software development
Step 3	Test of agent technologies for the evaluation on individual company
Step 4	Completion of the executing agent technologies

A Lifecycle of Companies Adopting Agent Technologies [*Source:Marık & McFarlane 2005]*.

Table 3. 1

3.2.3.2. Entertainment

Agent technology based systems in the social simulations that are increasing importance in entertainment applications, this special sort of agent system called *"real life agent"* that means to simulate its real world counterpart by means of animated visualisation. These applications are a range of single-player computer games to multi-player computer games, where the other players may be both humans and agents. For example there is a popular social simulation game in point, Maxis SimCity *[Kogler 2003]* that human players build their own virtual societies, also players may have their own avatar to participate in the virtual societies and interact within factional characters *[Jiang 2000]*. However those entertainment applications are mainly based on Multi Agent Simulation (MAS), which is an agent system with multiple agents, used for modelling and simulation of complex social systems, the MAS supports an efficient development and execution for agent technology based computational models.

3.2.3. Challenge

Nowadays the information society is becoming ever increasingly sophisticated; agent based technology is rapidly evolving to meet the demands of this new information time. Numbers of broad agent technology based systems have now been applied, also that a number of superior prototypes for real world problems have been generated as well, many issues and challenges still remain. In term of deployment there are few major challenges.

- Lack of sophisticated software tools, techniques and methodologies to design, develop, verify and manage agent systems. The components and features in agent

technology is difficult to be integrated; which agent systems have to be specified, designed, implemented and managed, before integrating different theories and infrastructures, for example, the agent-oriented methodologies still need improvement of seamless integration with other mainstream software engineering methodologies *[Mckean Shorter Luck McBurney Willmott 2007]*.

- Agent systems are not really in the automation of specification, development and management, there are still some features that may have to be handled manually, for example, the creation and management of the agent technology coalitions and virtual organizations *[Norman, Preece, Chalmers, Jennings, Luck, Dang, Nguyen, Deora, Shao, Gray, Fiddian 2004]*.

Challenges of agent systems are continuous to vary making the advance and mature of the field of agent based computing technologies, any appearance of the challenges are actually efforts that contribute improvements of agent technologies. However the standards of agent technologies will inevitably continue to be critical whether they come from within the agent community or emerge from other general computing infrastructure progress *[Bradshaw 1997]*.

3.2.3.1. Security Issues of Agent Technology

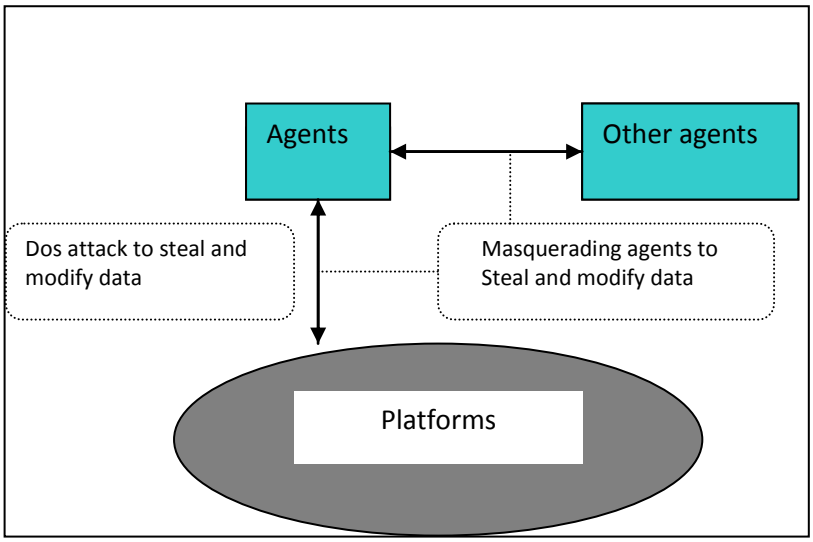
In term of security issues of agent technology, the mobile agent system has four potential security problems as follows.

- **Two agents**: security of agents against other agents and malicious agents, attacks in this problems that may contain masquerading, (e.g. masquerading as another agent to obtain private information) unauthorized access, (e.g. Attempt buffer overflow, reset to initial state) and Denial of Service (e.g. keep sending and spamming other agents messages that may place undue burden on the message handling routines of the recipient). For the solution suggested on those issues are authentication of agents (e.g. by using digital signatures) and employment of a service contract mechanism *[FIPA Security SIG 2002]*.
- **Agents and platform**: this issue can be separated in two sub areas as followed
  - Agents against platforms

The platforms can be compromised from malicious agents for stealing and modifying data on the platform, such as committing Denial of Service (DoS) attacks *[Zhang, Karmouch, Impey 2001]*.

- Platforms against agents

  The agent can be attacked from malicious platforms to steal and modify its own data, such as masquerading agents. The figure 4 shows the possible attacks of mobile agent technology *[Zhang, Karmouch, Impey 2001]*.



Samples of Security Issues on Agent Technology. *[Source: the Author]*

Figure 3. 1

- **Platforms**: the possibility of attacks between platforms is including cheating, masking and denial of service, the common mechanisms can be applied in this circumstance without modification for authenticating, accounting and interaction control while platforms are stationary entities *[Jansen & Karygianis 2000]*.
- **Platforms and unauthorized third parties**: the problem can be occurred in this area between platforms and unauthorized third parties that is the communication over an insecure network *[Jansen & Karygianis 2000]*.

3.2.3.2. The Future

The future of agent technology based computing science is challenging, and will lead to many important discoveries. Agent technology is autonomous software entities with a degree of reasoning ability, which has the potential for solving problems in new application areas. The future trend of agent technology will be indisputably applied and evolved for ubiquity in most areas while the use of agent technology is

becoming more and more common and important, For example, the future agent technology with enterprise process; with the increasing of globalization and competition between enterprises, the business process will require the advanced helps from agent technology, because the agent technology has the ability and coordination for real-time problem solving *[Lange 1998]*.

3.3. Multi Modal Strategies

Multimodal simply defines a number of operations or processes handling with more than one mode of communication, it provides users with interaction of multiple modes of interfacing a system beyond the traditional keyboard and mouse input or output, for example Multimodal browsers allow users to interact via a combination of modalities, such as, speech recognition and synthesis, displays, keypads and pointing devices.

Generally a system based on multimodal strategies which support communication within interaction of different modalities, such as voice and gesture, in term of "multi" and "modal", the "multi" means more than one something; the "modal" refers to functions and conditions. In this case the term "modal" which means two or more types of communication way used for transferring and receiving information/data, and the way or manner an idea is perceived and performed.

In the current mobile software field there are still only a few applications for personal digital assist (PDA) which refer to several modalities in a flexible manner, even though the generally small screen and frequent context-changes suggest that a multimodal interface yields great use benefits. For example, MONA (Mobile Multimodal Next-generation Applications) is the development of a generic platform for multimodal services for mobile devices such as PDA; it allows a new class of mobile applications featuring rich multimodal user interfaces based on a variety of mobile devices using through mobile phone networks and wireless LAN *[Niklfeld, Anegg, Gassner 2005]*.

MONA Login Screen and Example MONA Voice Messaging Application on PDA and WAP Device Emulators *[Source: Simon & Jank 2004]*.

Figure 3. 2

3.3.1. Paradigms of Multi Modal System

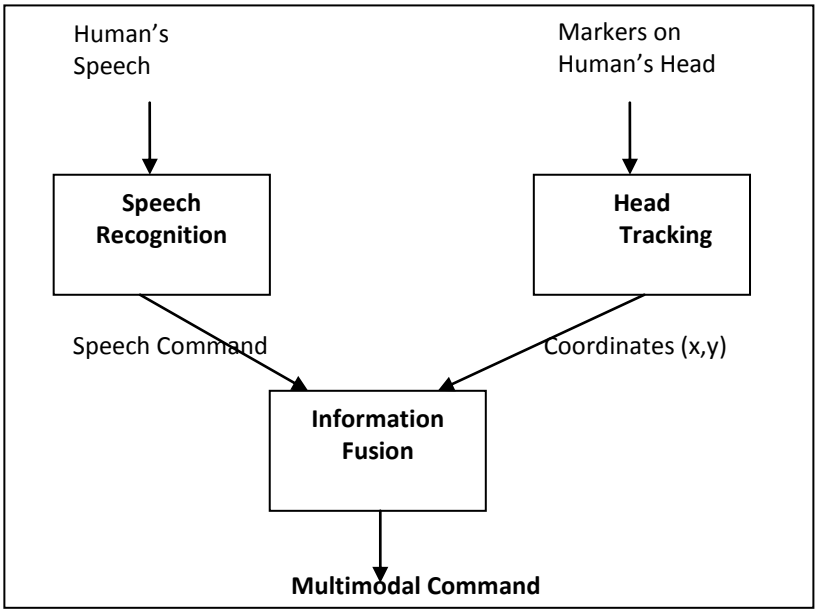
Recently computer technology allows us to build interactive multimodal system for some specific requirements; however multimodal systems enable users to achieve a certain aim based on multiple way and different functions/processes, such as for disabled people using PCs the multimodal system permits users the modalities of head and speech movements. Another example in voice browser the multimodal system enable users access web applications using the ways of emotion recognition, facial moves, eye detection instead of only conventional mouse, keyboard and screen *[Vallés, Manso, Arredondo Pozo 1996]*.

3.3.1.1 Multi Modal System Based on Speech Recognition and Head Tracking

Ronzhin and Karpov(2005) describes an assistive multimodal system for disabled people based on the multiple methods of voice and head movements, such as head movements to control the cursor across the computer screen, voice movement is using voice commands. This multimodal system processes the input information and controls the PC devices, the output of information are implemented by screen for graphical information and speakers for audio information.

Figure 3.3 below demonstrates the general architecture of developing this system; in detail there must be no any stop during the input by controlling of the system while

both the modalities of speech and head movements are active, as this is a multimodal system each modality is independent to process their own operations/functions. Meanwhile the head tracking specifies coordinates of a marker (cursor) in the current time moment, and the speech recognition transfers commands about the action from human's speech, this modality must be performed with an object selected by the cursor (or irrespectively to the cursor). Information fusion is used in this frame, when the needed data fills the fields of some structure and the command execution of signal on completion is given.



General Structure of Multimodal System *[Source: Ronzhin & Karpov 2005]*

Figure 3. 3

3.3.1.2 Multi Modal Being Addressed into the Voice Browser

A voice browser is an ordinary web browser that permits users to use their voice; in general it is a software component separate from the visual browser to allow users to navigate the Web through the voice content using the microphone and speaker instead of a conventional keyboard, mouse, and monitor, and processing pages of VoiceXML instead of the visually-oriented HTML. The multimodal system is used in the voice browser based on the techniques, which provide multiple methods/ways and different functions/processes to allow various forms for outputting the same results, for example, a typical paradigm for voice browser is multimodal browser known as Opera, which has being developed by IBM and Opera, this multimodal browser is based on the XHTML+Voice (X+V) specification, and includes the multimodal

system to provide the interchangeable use with different modalities of input/output methods, for example voice command mechanism, keypads with stylus in the same interaction *[Barnett 2005]*.

3.4. Web Service

A Web service is a set of standards which is a new platform of building interoperable distrusted applications; it defines how applications can be interacted on the Web, it can be compiled by any Web language on any platforms, only need to test whether it can be inquired and accessed by Web service standards *[Orchard, Newcomer, Ferris, Champion 2002]*. Eventually there must be a method to remotely invoke the Web service, which the method actually is a protocol of Remote Procedure Call (RPC), for achieving interoperability the RPC protocol must have no relative with Web languages.

3.4.1. The Components of Web Service

To design and implement a Web service that has to require the key components, each component has different function and interacts with each other as follows;

3.4.1.1 XML and XSD

- XML is used for building and analyzing which represents the form of data on a platform; the main advantage of XML is that it is independent of platform. XML solves representation of data, but it does not define a set of standard data types, such as which type of integer, is it 16bit, 32bit or 64 bit? These details are very important to implement interoperability, so W3C instituted XML Schema Definition (XSD) in order to solve this problem.
- XSD defines a set of standard data types, and a language to extend these types. The platform of Web services uses XSD to be its data type system, when a Web service is built by some languages (such as VB.NET or C#), all data types must have to change to XSD type for according with Web service standard *[Salchner 2005]*.

3.4.1.2 SOAP

Simple Object Access Protocol (SOAP) is an XML Web service communication protocol used by Web service for the exchange of information, this protocol must be platform-independent flexible, and based on standard, ubiquitous technologies. It

provides a way to communicate between applications (application-to-application) running on different platforms with different programming languages, this permits it to be used in a variety of systems ranging from messaging systems to remote procedure call. Meanwhile the SOAP message is an XML document instance usually carried as the payload of some other network protocol. For example the common methods for exchanging SOAP messages are through decentralized, distributed environment over HTTP, FTP and SMTP. The SOAP consists of three parts as followed *[Box, Ehnebuske, Kakivaya, Layman, Mendelsohn, Nielsen, Thatte, Winer 2000]*:

Envelope Construct	An entire framework to present the contents of messages, purpose and whether it is optional or mandatory.
Encoding Rules	A serialization mechanism used for exchanging instances of application defined data types in communicating between applications (A2A)
RPC Representation	A convention used for representing feedback and RPC

Three Parts of SOAP, *[based on Box, Ehnebuske, Kakivaya, Layman, Mendelsohn, Nielsen, Thatte, Winer 2000]*

Table 3. 2

3.4.1.3 WSDL

Web Services Description Language (WSDL) is an XML language that describes documentation of Web services and specifies communications of Web service. It describes three critical pieces:

- What: includes elements of type, message and port type;
  a) Type: defines structure of data in Web service (XML Schema)
  b) Message: one message is an essential communication element of SOAP, each message can have one or more parts, and each part represents one parameter.
  c) PortType: includes a supported set of operations. Each operation includes the input and the output messages of the operation, furthermore each PortType represent an interface, the Web service can have multiple interfaces that all use PortType to represent.

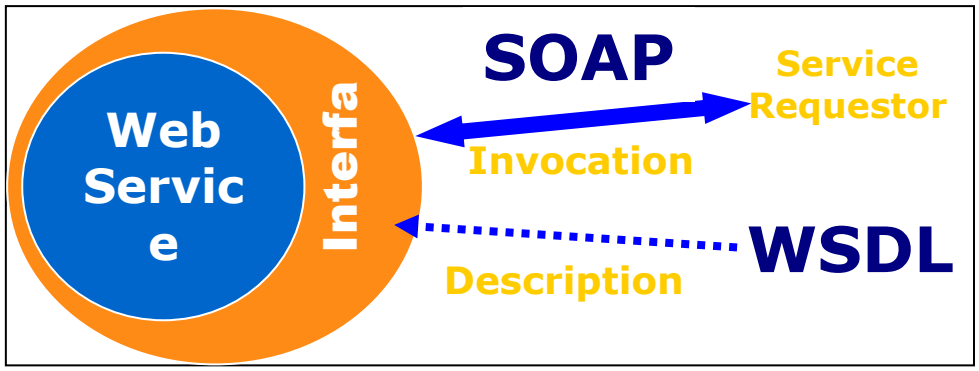
- How: includes binding element which the Binding information is about the transport protocol to be used.
- Where: consists of service elements which describe the position (URL) of PortType, Binding and Web service together *[Christensen, Curbera, Meredith, Weerawarana 2001]*.

3.4.1.4 UDDI

Another standard of Web service is UDDI know as Universal Description, Discovery, and Integration, it is an open framework and platform independent, and a registry based on XML for discovering and integrating businesses and describing Web service using the Internet. The UDDI is a directory for storing data about the Web service and providing access to WSDL documents, and it is communicating via SOAP. The purpose of UDDI is to streamline transactions through the Internet by enabling enterprises to make their systems interoperable for e-commerce with other enterprises on the Web *[Jewell, Chappell 2002]*.

3.4.1.5 The Interaction of Each Component

These techniques discussed above present the major roles to compose a Web service, each of the techniques has different functions applied in the Web service. Figure 3.4 shows a very general structure of a Web service, the Web service can be known as an interface to describe a set of accessible operations that refers to clients and servers to be communicated by using XML messages, which follows the approach of SOAP standard, the SOAP is a simple XML based protocol to allows information can be exchanged through such as network between services, and the WSDL is a specification for describing web services in a common XML grammar *[Gottschalk, Graham, Kreger and Snell 2002]*.



The very basic overview of Web service. *[Source: the author]*

Figure 3. 4

However Web services are interfaces which explore APIs that can be invoked via the Web, which means the integrating of Web based applications dynamically interact with other Web applications by using open standards. For example if a user want to create a web service whose function is simply to return the weather of current region. So the user can build an Active Server Pages (ASP) page which accepts that zipcode to be the query in order to send a HTTP GET request, and then the server response a separate string by comma including temperature and weather. This ASP page should be a Web service, because it bases on HTTP GET, explores an API that can be invoked through the Web *[Orchard, Newcomer, Ferris, Champion 2002]*.

3.4.2. Architecture of Web Service

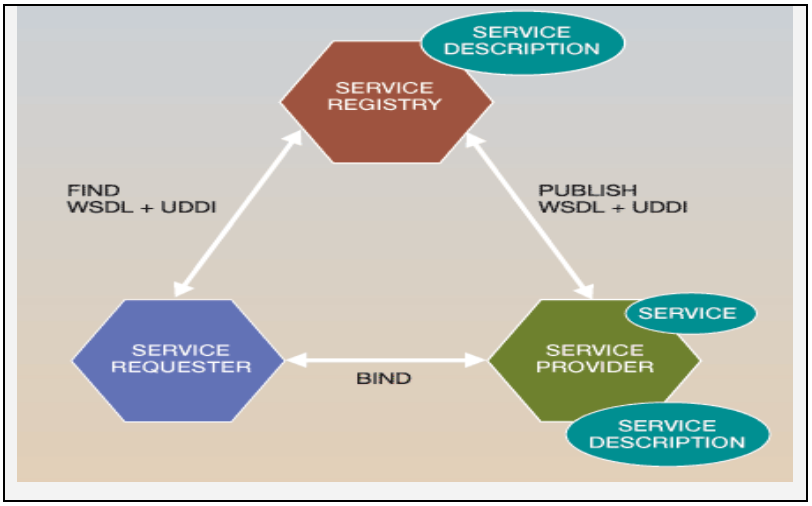
Web Services are independent and modular applications interoperated with an interactive environment through networks, it describes three actors (service registry, service provider and service requester) and operations (publish, find and bind). In the Web service world these three operations include different technologies:

- Publish using UDDI,
- Find using UDDI and WSDL,
- And finally bind using WSDL and SOAP.

In these three operations, the bind operation is most important which includes the practical use of service, as well as this is easy where to produce interoperation problem. But as a result of service providers and service requesters cooperate with

each other to support the SOAP standard in order to achieve seamless interoperability; this problem has been solved properly *[Gottschalk, Graham, Kreger and Snell 2002]*.

For an overall process to develop new applications, this may use UDDI operator or UDDI search engine to look for a required Web service from UDDI registry through Web interface. Afterward finding this Web service invoke-standard in the UDDI registry or by the connection of UDDI registry, that this invoke-standard usually is described using WSDL. Developers add the definition of this invoke-standard into their own applications using any developing tools or manual manipulation to invoke this standard; therefore these applications can then invoke specific Web services by using SOAP. Otherwise those relevant auto-integrated applications and services, users access UDDI operator or UDDI registry through SOAP standard to find needed Web service in which the UDDI operator and UDDI registry both response the Web service of invoke-standard or connection of invoke-standard, once applications have the text of invoke-standard described by WSDL, applications will parse this text description to automatically generate local invoke-interface for binding required invoke-parameters, and complete this invoke. These ideas are shown in the following figure



Roles of Three Operations in Web Service. *[Source: Gottschalk, Graham, Kreger and Snell 2002]*

Figure 3. 5

Figure 3.5 depicts the interaction of each role of three operations which performed by service providers publishing services to a service registry, and Service requesters finding required services using a service registry and binding to them.

In addition to above, the Web service usually has three types of clients as followed:

- Business Partners: includes distribution, retail and large consumers, this type of clients connect with Web service using XML technology (such as SOAP, WSDL, ebXML, UDDI) *[Holdsworth,Venkatapathy 2002]*.
- Thin clients: includes Web browsers, PDA and some WiFi devices, this type of clients connect with Web service using "thin" protocol (such as HTTP) *[Nichol 1997]*.
- Fat (thick) clients: includes Applet, systems and even PC, this type of clients connect with Web service using "thick" protocol (such as IIOP) *[Nichol 1997]*.

3.4.3. Security Issues of Web Service

Web Services are the next stage of evolution for E-Enterprises which from the perspective of a business is a service, dynamically discovered and orchestrated, using messaging via networks. Thus the security issues would be major problem while deploying a Web service to open up access to the service for everyone. For example when contemplating the impact of this issue, protecting data against people who is seeking compromise to services. According to Web service security issues *[Edgar & Baruah 2002]*, the main consideration of security issue can be described as three major aspects as follows;

- *" **Integrity Control**: Protection against malicious or accidental attempts to alter data*
- ***Confidentiality**: Protection from unauthorized attempts to read data*
- ***Availability**: Protection against unauthorized deletion or otherwise cause a denial of access to the data or service"*

Integrity Control: means to ensure the information can not be modified, deleted, fabricated, added, destroyed and lost without authorised access and must keep the information as original integrity while it is being accessed and transmitted, the information, the integrity of information reveals the reliability, validity, availability and consistency of information.

Confidentiality: protects information, can not be accessed without authorisation, it is usually implemented by information encryption, ID authentication, access control

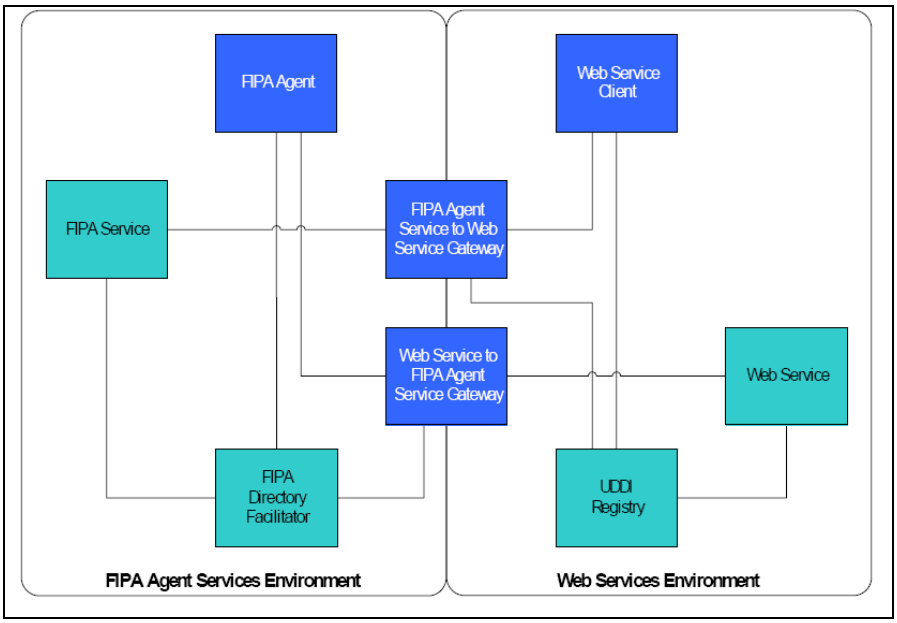
and communication protocol etc, and this is common idea when people talk about the network security

Availability: the information and resources are accessible by authorized parties, the Web services used to process the information dependent on the data availability which provide all available and functioning correctly when the information is required, the attack on availability is denial of service which means denying resources to a legitimate users or processes to access any resources.

3.5. Agent Integrating with Web Services

In the current information society, a large amount of information has become accessible at anytime and anywhere, as a result people can be rapid and effective in locating information in a convenient form. But the information can not be necessarily useful all the time; it has to be presented at the right time, in the right place and the right way for becoming welcome. The Web service technology has gained popularity to be able to widespread interoperability among applications, in one of many scenarios, enabling agent technology to access and control with Web Service is promising and recently becoming attractive attention for people in both the agent and Web services communities.

To seamlessly integrate the autonomous and problem solving capabilities of agent technology and Web services, and using agent technology to manage and access the universal Web services infrastructure and interoperability. Therefore a new environment for integrating these two technologies with separately different standards and specifications is promising, the issue of providing access to Web services from agent platforms and vice versa has been addressed in the AgentCities project *[Agentcities Web services Working Group 2003]*. This working group has offered a proposal that provide a proxy approach as a most suitable solution for Web services and Foundation for Intelligent Physical Agents (FIPA) compliant agent platforms, to solve the interoperation between accessing Web Services via an agent based service gateway and accessing agent services via a Web Service based service gateway.



Integration Architecture of Web Service and Agent Technology *[Source: Agentcities Web services Working Group 2003]*

Figure 3. 6

In Figure 3.6, the FIPA agent service environment and web service environment comprises the logical components of agents and web service clients and servers. With the FIPA agent querying a FIPA Directory Facilitator (DF) and the web service clients querying with a UDDI Registry Server and then invoking the appropriate SOAP method, between the two environments, the FIPA agent service to web service gateway by translating Agent Communication Language (ACL) messages to web service invocations, and the Web Service to FIPA agent gateway allow web service client to access a FIPA service seamlessly by registering them in UDDI registry.

In another paradigm, there is enhanced solution for Web services and FIPA compliant agent platforms based on WS2JADE system, which is a toolkit for dynamically translating and controlling Web services into Java Agent Development Framework (JADE) agent services at run-time for deployment flexibility and active service discovery, meanwhile the JADE agent is a middleware for supporting the development of applications based on the Peer-to-Peer intelligent agent approach. As it is an enhanced solution, it provides a greater level of automation in Web services discovery and the use of Web services *[Nguyen and Kowalczyk 2005]*

A DF is storage that provides registration to agents such as yellow pages. Agents may register their services with the DF or query the DF to find out what services are offered by other agents *[Shen, Norrie, Barth & 2000]*.

A FIPA ACL message contains a set of one or more message parameters. Precisely which parameters are needed for effective agent communication will vary according to the situation; the only parameter that is mandatory in all ACL messages is the performative, although it is expected that most ACL messages will also contain sender, receiver and content parameters *[ACL specification 2002]*.

3.6. Conclusions

The goal of this chapter is to introduce the required technologies on multi-modal agent system based PDA, that they are "mode" "multimodal," "interface," "agent", "web service", each of these technologies has been discussed within the multi-modal agent system.

- Mode is the way to perform operations or processes.
- Multimodal refers different mechanisms to achieve the same goals.
- Interface defines the communication boundary between two different systems.
- Agent is used to carry out tasks on users' behalf.
- Web service is a information container to enable clients to access the information via a network

In the end it has briefly presented the interaction and integration of those technologies, for example, the integration of web service and agent technology In next chapter the emphasis will be on developing the framework of multimodal agent based PDA from both functional and technical perspective.

Chapter 4. Functional Specification of Multi-Modal Agent System

The Multi-Modal Agent System delivers a wireless and knowledgeable environment to PDA users without the need to consider time, place and physical constraints. In this chapter the Multi-Modal Agent System will be introduced from a functional perspective, and it will be also analyzed in detail of design goals, components, architecture and user interface.

4.1. Introduction

This chapter outlines a functional specification of the Multi-Modal Agent System; it depicts a theoretical introduction and requirement of the overall system based on each component, and the interaction between those components as follows:

- The Web Service: an online system designed for the access and interaction between the web software over a network, it provides the information for the requirement of either a speech server or a PDA client.
- The Agent Technology: refers to one or more services, or behaviors, which can be processed at the same time and activated by specifying goals rather than methods to be invoked, it brings the functionality to correspond between a web service and a PDA client in an intelligent way.
- The Speech Server: a flexible and integrated platform for speech deployments, it supplies the services to PDA client, such as speech recognition and text to speech processing.
- The PDA Client: a handheld device designed as Personal Digital Assistant that carries out functionalities similar to a PC, it acts the role as a user interface to communicate with users in the dynamic and interactive way.
- The User: who are an information demander and receiver outside this system.

Each of the roles above have their own responsibilities and representing functionalities in the multimodal agent system, which those roles permit the cooperation with each other and conducting a variety of tasks, such as the communication between a PDA and a speech server has to be required when the PDA does speech recognition feature *[multimodal speech group 2001]*.

Section 3.2 describes the overall design goals of the system. The functional architecture of this system with the detailing of operations and responsibilities of each component will be presented in Section 3.3. A use case diagram and user interface design of the multi modal system is both showed in Section 3.4. Conclusions are drawn in the end of this chapter.

4.2. Overall Design Goals

From a theoretical perspective, this project provides multimodal and intelligent system based on innovative concepts that will consistently appeal to a range of people. The depth of contents of this system must be designed to challenge and to satisfy the basic requirements the level of a system user. The execution of the general method to design will consist of both establishment and innovative concepts through incorporation of each component.

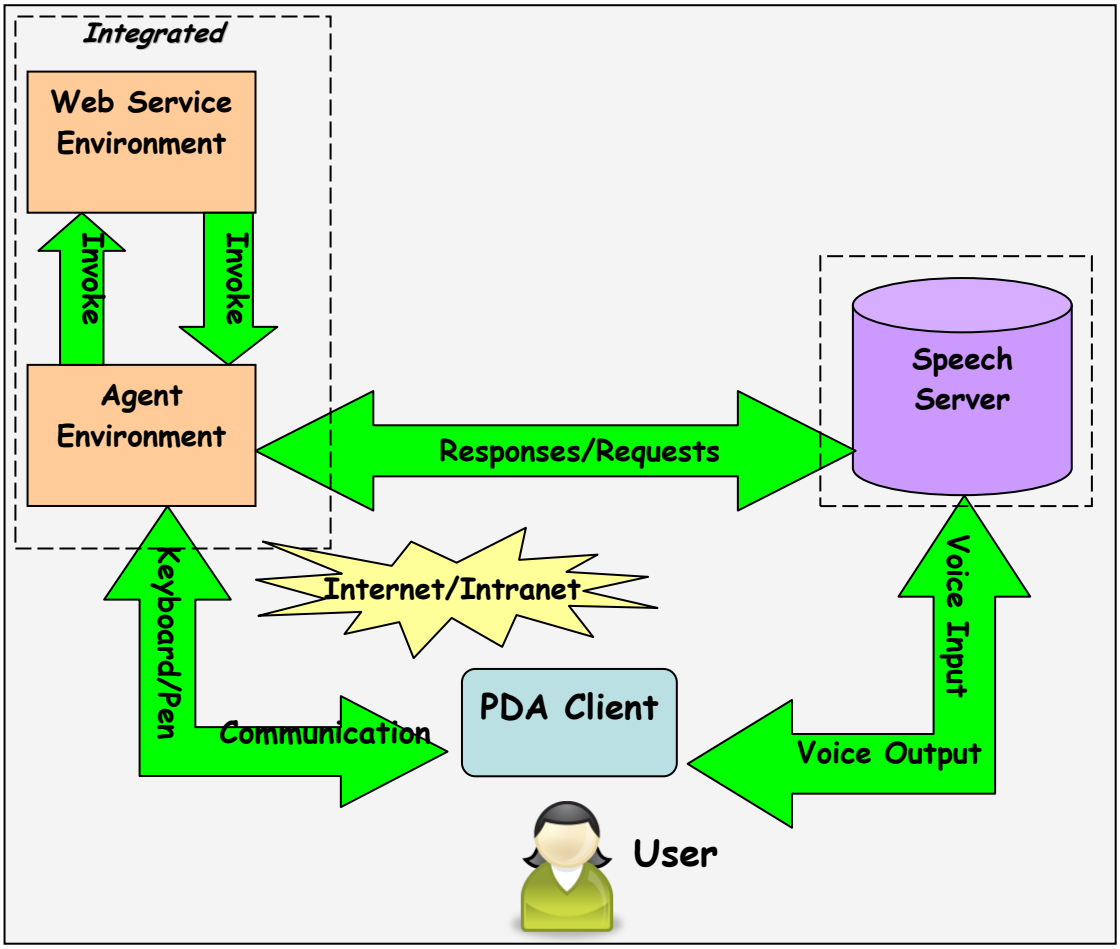
The Multi-Modal Agent System is designed to build a reliable system in order to implement a multimodal interface with agent technology and web services, which represents a next generation system based on PDAs for multifunctional usages. The design goals intend to provide desired functions for general users through such aspects as:

- **Interaction**: Input and Output of information or the interactive play system will enhance the participation of the users, and allow users to have multiple options with a dynamic environment, for example, a user could have voice input and voice output instead of normal keyboard interaction.
- **Innovation**: Use of and Consist of each innovative technique and method in building the system and creating the environment to enhance innovative value with innovation, such as the system can be used by disabled users.
- **Excitement**: Dynamic and unexpected presentation methods that increase the awareness and enjoyment level of the users, for example the users could choose voice commands to control the browsing on the web.

However the Multi-Modal Agent System has a wide range of future development in order to exploit more features in advance, it is the new trailblazer using a PDA to implement the next generation system for facilitating people's lives.

4.3. Components of this system

In Figure 4.1 the functional scenario for the multimodal agent system is depicted, which clearly describes the interaction and integration of each component. In general, a user holds a PDA client for sending and receiving information using multi-modal interface. If the user chooses the keyboard/pen to communicate with the web service which is integrated with agent technology, the connection will be built directly between integrated web service and PDA browser through the Internet or Intranet. If the user chooses voice input/output to communicate with an integrated web service, the connection will go through to speech server firstly in order to process voice signals through the Intranet. Once the connection is built with the integrated web service, the internal agent technology will automatically search the web for updating or downloading new information, such as calendar and email information, weather and traffic reports, entertainment listings, information about friends and colleagues, and finally represent the information in an opposing way *[Vo, Wood 1996]*.



The Functional Architecture of the Multi-Modal Agent System *[Source: the Author]*

Figure 4. 1

In this diagram there are four main functions interacting with each other, each function carries out different operations as showed below.

4.3.1. Web Service

The Web Service provides the environment for users to access information or applications over the web; it also offers the environment for users to publish their knowledge sharing with others. Within the Multi Modal Agent System the role of web service is like storage, which allows the users to have abilities to access any information as they want, such as checking email, updating blogs and making friends via the web.

4.3.2. Agent Technology

The Agent Technology is an efficient mechanism built within this system to solve specific problems on the behalf of the users and according to their directions, which allows an incremental development of complex architectures to make a more robust intelligent web server system for accessing information. It is integrated with web server in order to completely exhibit a novel environment for providing more capable services.

4.3.3. Speech Server

The Speech Server is speech-enabled Interactive Voice Response (IVR) and multimodal application solutions platform, which is designed for the translation of text-to-speech and speech-to-text, in order to provide the features of automatically speech recognition and text to speech, which particularly enable speech input and output for implementing the multi modal interfaces. It is a speech server that merely handles the speech input and output from PDA client, to be able to allow users to talk to PDAs *[Song 2007]*.

4.3.4. PDA Client

The PDA Client is a major part of the multi-modal agent system, to enable communication with bi-directional transformations between information sources and users. It can be seen as metaphor of a physical “user interface” to deal with users which the PDA carries out the functions of directly presenting any information by the requirements of users. Moreover because the portability of its physical advantage, the

PDA can be used at any time and place irrespective of additional demands, with this potential benefit the application can be very well expected.

4.4. User Interface and Use Case Diagram

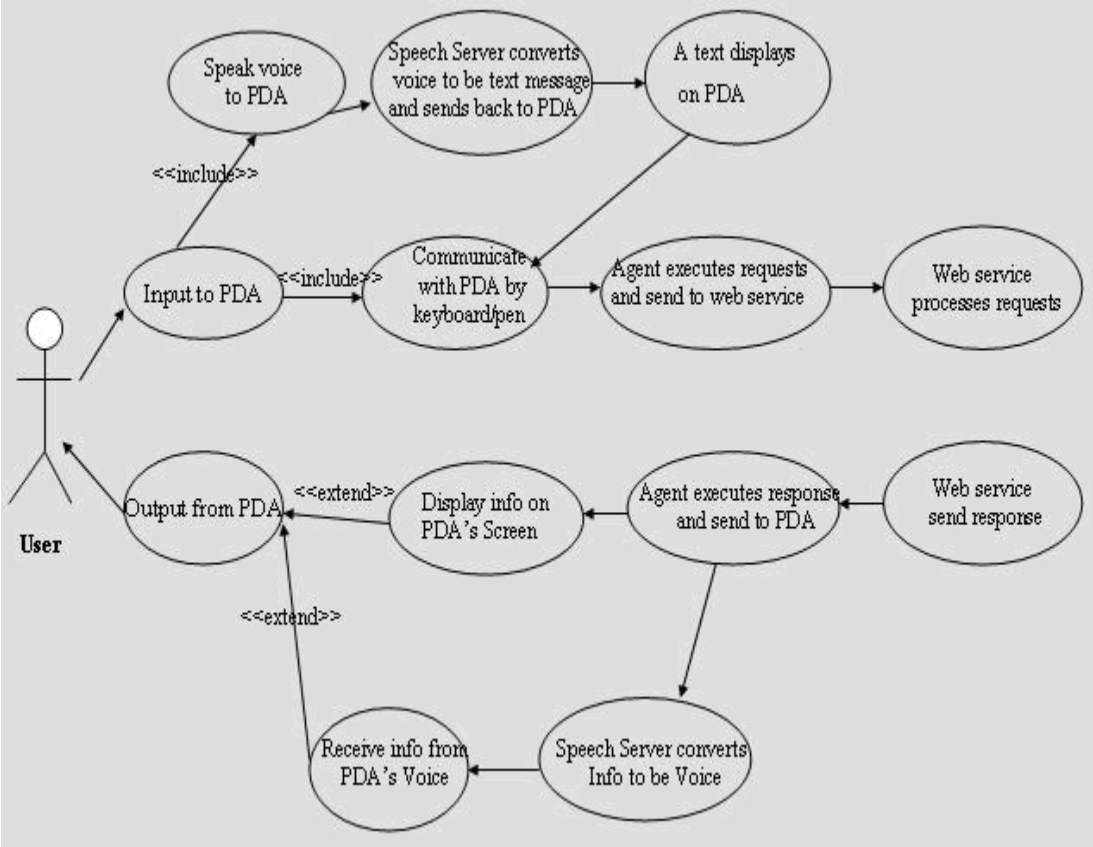
In Figure 4.2 below a diagram presents a compact framework user interface for the users interacting with the PDA devices, which is actually a basic browser with a multi modal interface. In this browser, its major function is to perform the features of keyboard/pen and voice signals. From this picture, the User Interface contains three internal sections; the top one is the function bar including file, zoom, tools and help functions, the middle one is control panel for managing the features of keyboard/pen and voice signals, and the last one is display part for showing any sort of information the users require.

Meanwhile a very important section in the middle section which carries out the foremost function of this user interface, the very first text area on the top with little button “go”, it is the keyboard/pen interface for the users typing keyboard or using a stylus to enter information. The second text area with gray button “speak” is what implements text-to-speech, which means the PDA speaks to people in which if there is any information displayed and the “speak” button activated. The last three buttons, the “listen” button receives voice signal from users, and the “retrieve” button gives users a second chance to speak again *[Larson 2003]*.

A Compact Framework for User Interface *[Source: the Author]*

Figure 4. 2

Figure 4.3 showed below is the general Use Case Diagram for the entire multi-modal agent system. The PDA has a multimodal interface to allow a user to communicate with the PDA dynamically and interactively, which the user can speak and enter texts into the PDA, and contrarily the PDA talks and presents information back to the user. During the operating process of the PDA, a speech server carries out the speech recognition from the user's voice signal and the text to speech from information having to present. But if there is no speech during the entire process, it will be the same as the traditional communication approach to connect with a web service directly. Meanwhile the agent technology handles the executions which are requests from the PDA to the web service and the responses from the web service to the PDA. Alternatively the web service is an online server on the net; it contains the required information and data for which the PDA wants to access. During the diagram, the case of input to PDA actually includes two different processes voice input and keyboard/pen input, also the case of output from PDA extends from voice output and screen display.



Use Case Diagram of Multi-Modal Agent System *[Source: the Author]*

Figure 4. 3

4.5. Conclusion

This document describes the functional specification of multi modal agent by presenting four theoretical components in a typical multi modal agent system; furthermore the document also described the interaction and consequence of these four components within the multi modal agent system. In the functional specification the entire architecture and a use case diagram of this system has been developed with design goals as follows;

- Provide a multi-modal interface to allow users to have multiple choices in their interest.
- Enhance the use of PDA in an interactive and dynamic environment.
- Allow innovative techniques and methods to be built within this system.
- Although the desktop PC can carry out similar features to the PDA, the portability of the PDA is a huge advantage for those who require mobility and constant access to information.

The functional specification introduced here will be described from a technical perspective in detail in the next chapter; it will be presented by using specific techniques and methods, and in order to create a framework of the multi modal agent system.

Chapter 5. Technical Specification for the Multi-Modal Agent System

The Multi Modal Agent System will provide a meaningful and knowledgeable platform to permit users having an intelligent and multi modal access based on PDA through a network. This chapter will technically focus on the analysis of each component and the integration and interaction of each component, as well as some diagrams and screenshots will be delivered in order to increase the visually understanding of development.

5.1. Introduction

The document outlines a technical specification of the overall multi modal agent system in a technical perspective; it should not be measured as a complete technical specification in which there has to require initial prototypes of system elements to be foundation, before the deploying of the multi modal agent system. But it provides an entire framework for the intuitive analysis of further development.

However this technical specification describes and specifies a new topic of next generation system for the less of interaction and integration of information interoperability between a PDA and a web service, it has proposed a new approach to overcome this issue in constructing a novel architecture by four roles as follows:

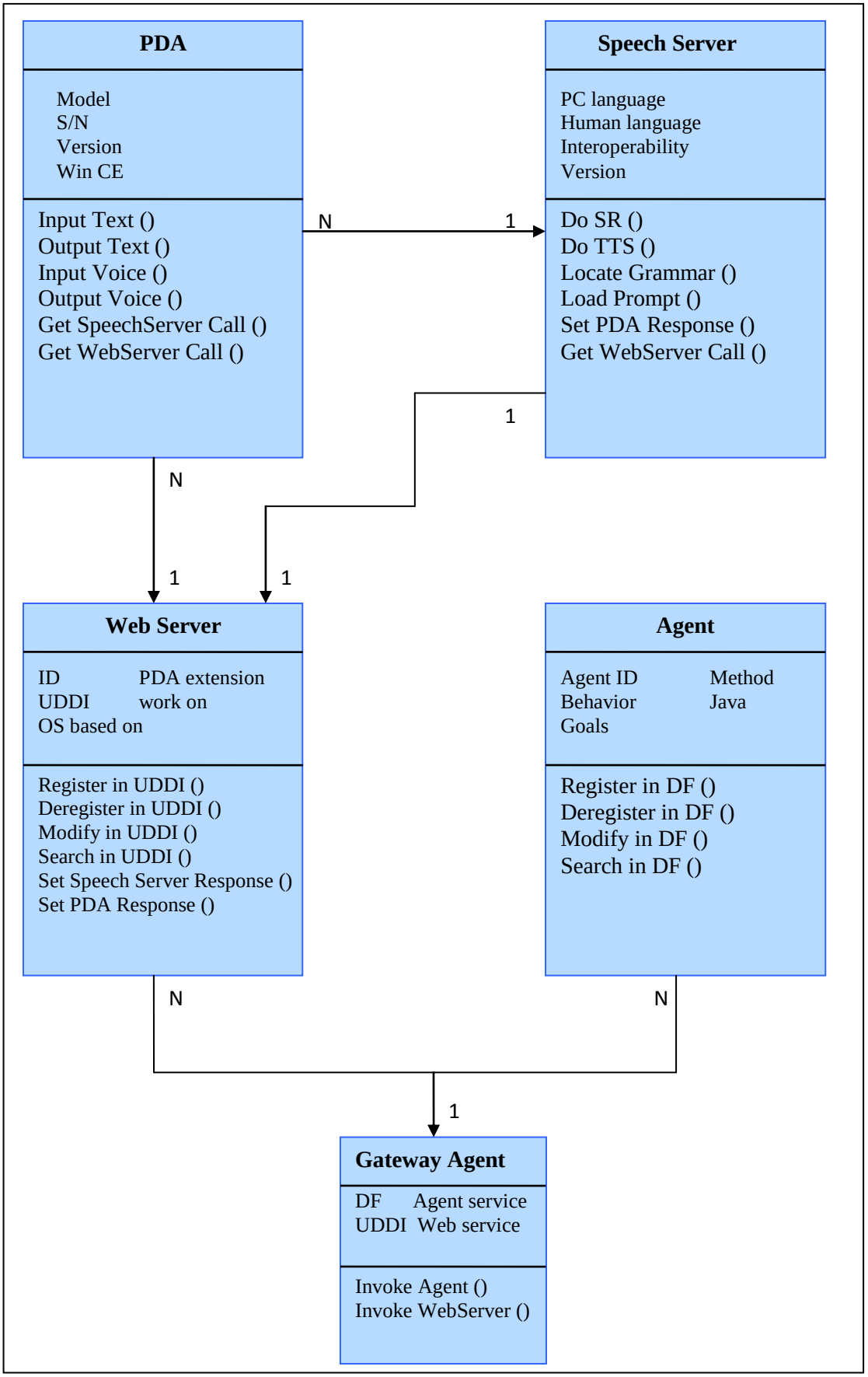
- **The Web Service**: enables the services can be accessed by PDA clients, and generates the required data to permit normal performance of speech server can be carried out. It is registered in UDDI and described by WSDL and its messages of transmission are based on SOAP, also it integrated with agent service by Web Service Integration Gateway (WSIG) technology.
- **The Agent Technology**: provides an efficient manipulation to allow the PDA client has more intelligent access with a web service, it is also integrated with a web service by WSIG technology, but in this system it will be described by Java Agent DEvelopment Framework (JADE).
- **The Speech Server**: coverts voice signals to text and vice versa converts text to audio speech; in this system it does the feature to provide a multimodal interaction between a user and a PDA for implementing a multi modal interface.
- **The PDA Client**: allows users to have a visual interface for directly

communicating and browsing web information, it has to have a speech add-in control binding within its browsers in order to be capable doing Speech Recognition (SR) and Text-to-Speech (TTS).

This technical specification will be used for a future proposal of the system, and some future work is still required on more definitive details. Nevertheless the next Section 5.2 outlines a general Class Diagram to show the methods and attributes of each roles of the system. Section 5.3 describes a detailed structure of a WSIG mechanism with the integration of a web service and agent technology. And the multi-modal interface is described in the speech server Section 5.4 with the examples of speech recognition and text-to-speech. Thus the interaction between speech server and PDA client is also introduced in Section 5.5. Finally a conclusion summarizes this chapter.

5.2. Architecture of the Multi-Modal Agent System

Figure 5.1 shows a class diagram of a multi-modal agent system from an analysis point of view, as illustrated in this diagram this system has been divided three main parts. The PDA part is a client used by a user who wants to input and output text and voice, and the PDA itself does the sending request messages to web server and speech server for getting required service. The Speech server is designed to convert the voice-to-text and contrarily convert text-to-voice, but in the process of converting, the speech server has to request additional resource based on web server to accomplish its operations. Furthermore the web server and agent are actually interactive by gateway agent in order to establish the integration between both of them. This whole structure is called Web Service Integration Gateway (WSIG) which will be specified in the following sections. However the gateway agent provides a platform to allow the invocation from agent to web server and vice versa.



The Class Diagram of Multi Modal Agent System *[Source: the Author]*

Figure 5. 1

5.3. Agent Technologies Integrated with Web Services

Allowing software agents to access and control Web Services is important and hence the Integration of Software agents and Web services can create immediate benefits in order to connect application domains hosting one or the other technology. An agent service should be able to invoke a web service and vice versa, which is identified by the Web Services Architecture specification of the W3C;

> "*... software agents are the running programs that drive Web services - both to implement them and to access them as computational resources that act on behalf of a person or organization." [W3C 2004]*

Nevertheless if the interconnection between the software agent and web service is established, the software agent technologies and concepts should be able to allow new operational and usage modalities of Web services. Therefore an approach to identify an intermediary service entity for the connection between software agents and web services which is the motivation of this work, meanwhile the apparent technology mismatches between them containing representational encodings and strong vs. loose coupling *[Greenwood, Calisti 2004]* is the difficulty that requires overcoming. The Web Service Integration Gateway (WSIG) state that;

> *"The objective of WSIG is to expose services provided by agents and published in the JADE DF as web services with no or minimal additional effort, though giving developers enough flexibility to meet specific requirements then may have." [JADE Board 2007]*

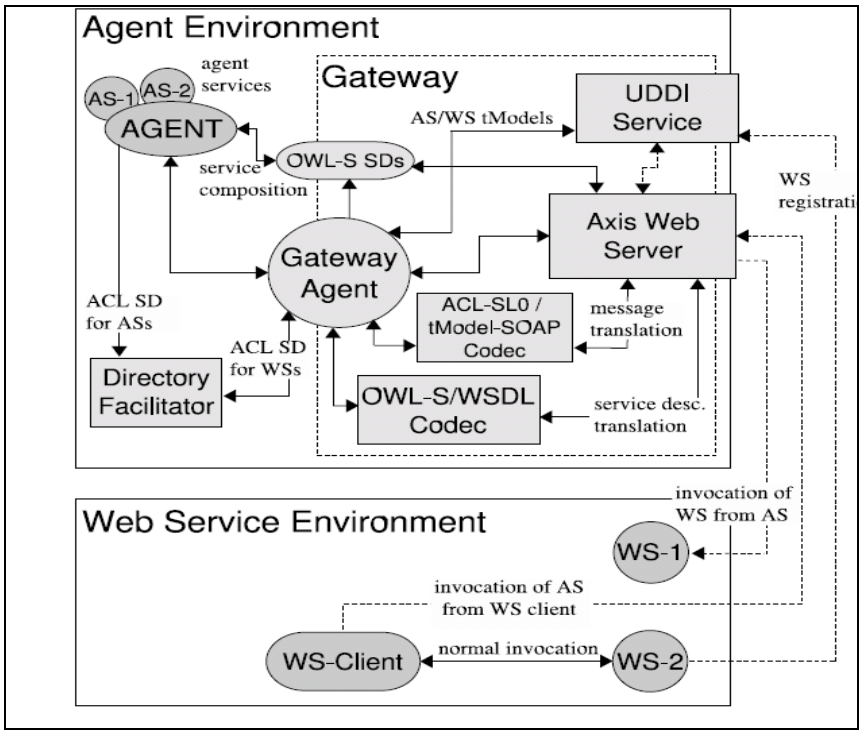
The WSIG is a recent proposal of web services and agents integration architecture and is designed to encapsulate the functionality required connecting the two domains, and also reducing the human intervention and service interruption. There are three key features of WSIG claimed by the developers as followed:

- **Transparency**: The WSIG is like an interface platform for operationally transparent to invoking agents or Web services. For example if there is an agent trying to invoke a web service, it will send a invocation message including the

name and address of the web service and the message content to WSIG, which the WSIG will seamlessly transform the corresponding message and issue it, once any response return to the invoking agent, the WSIG will transcribe it.

- **Automation**: During the operation of WSIG, it will not require either manual intervention or configuration; it is like an autonomous system without any manual operation.
- **Integration**: The WSIG will not require any additional external resources to deliver all features it has, it encapsulates and integrates the functionality in an assumed environment.

The WSIG permits both FIPA compliant agents registering with DF and Web services registering with UDDI in order to issue their service documentations to clients' external their operational domain. Thus the calls of transforming message encodings and creating service access endpoints within registered services in the WSIG environment can be intercepted during agents invoking web services and vice versa. Figure 5.2 below is a standard model of WSIG architectural.



Standard WSIG Architectural Model *[source: Greenwood, Nagy, Calisti 2005]*

Figure 5. 2

This illustration above is a gateway that consists of several components, at the heart of the WSIG the FIPA compliant DF and Stack compliant UDDI repository are both linked either directly or indirectly from other components. The DF is storage for agents descriptions used by all agents within the FIPA compliant agent platform as invisible form from outside, but the UDDI repository is visible from both external web services and web clients. The descriptions of agents and web services are both registered with the DF and UDDI, which is automatically translated into an entry for the other; this is because of making sure any registered web service is visible to agents via the DF and agent service visible to web service clients via the UDDI. These registries expose the four standard operations *[Greenwood 2005]*:

- **Registration**: registering a new service description
- **Deregistration**: deregistering an existing service description
- **Modification**: modifying an existing service description
- **Discovery**: searching a existing registered service description

Each agent registries for agents and web services will be described further in sections 5.3.1 and 5.3.2.

In this standard WSIG architectural model, the kernel is the logical component Gateway agent which receives and processes operations as independent threads to a number of other logical components such as codec and web server etc. It is responsible for a set of tasks as follows:

- Receiving notifications from DF that the agent has registered an Agent Communication Language Service (ACL) Description (SD), managing the mapping of newly registered ACL SDs into tModels for sending to the outside web services through the gateway UDDI which wishes to invoke. And vice versa receiving notifications from UDDI that a web service has registered a WSDL SD, managing the mapping of newly registered WSDL SD into OWL-S service ontologies and ACL SDs for sending to the inside agent through the agent platform DF which wishes to invoke. The detail of both web service and agent invocation will be described in following Sections 5.3.1 and 5.3.2.

Meanwhile, the tModel is a data structure used to describe taxonomies in UDDI, which is used to define a service's technical interface, for example, there are two companies want to register their service in UDDI in order to cooperate with each other, so the tModel is used to solve the compatibility problem. Each tModel has a name, identifier and explanatory description, the name means the name of a service, the description provides more information about the service and the identifier is a key to identify the service which consists of a series of alphanumeric characters *[Yu 2005]*.

- Receiving request messages from the agent for invoking an outside web service, which these messages will be translated into the form of FIPA messages, whereas the response messages returned to the agent, will be again translated into the form of FIPA messages. And vice versa receiving requests messages from web service for invoking an inside agent which these message will be translated into the form of FIPA messages, whereas the response message returned to the web service will be also translated into the form of FIPA messages *[Greenwood, Calisti 2004]*.

The component of ACL-SL0/tModel SOAP Codec in the WSIG is a very complex module which translates agent ACL/SL0 SD into a web service into WSDL, tModels and SOAP according to the specific context; it also has a socket connection to an Axis web server from where agent services can be exposed as if they were web services, the Axis server also receives calls for invocation from external Web service clients onto agent services *[JADE Board 2006]*.

The component of Axis web server is like an interface between web services with their clients and agent service ontologies. It is issued by web services in order to process invocation with agents, making agents externally like endpoints by those web service clients; it also internally issues OWL-S schema of agent service ontologies, which is from WSDL SD of web services registered with UDDI. Meanwhile the OWL-S is Ontology Web Language (OWL)-based Web service ontology *[Greenwood, Nagy, Calisti 2005]*.

The component of OWL-S/WSDL is a pure codec which translates OWL-S based service ontology into WSDL based service descriptions, which used the mapping

guidelines described in then its own language for publishing and sharing ontologies *[W3C 2004]*.

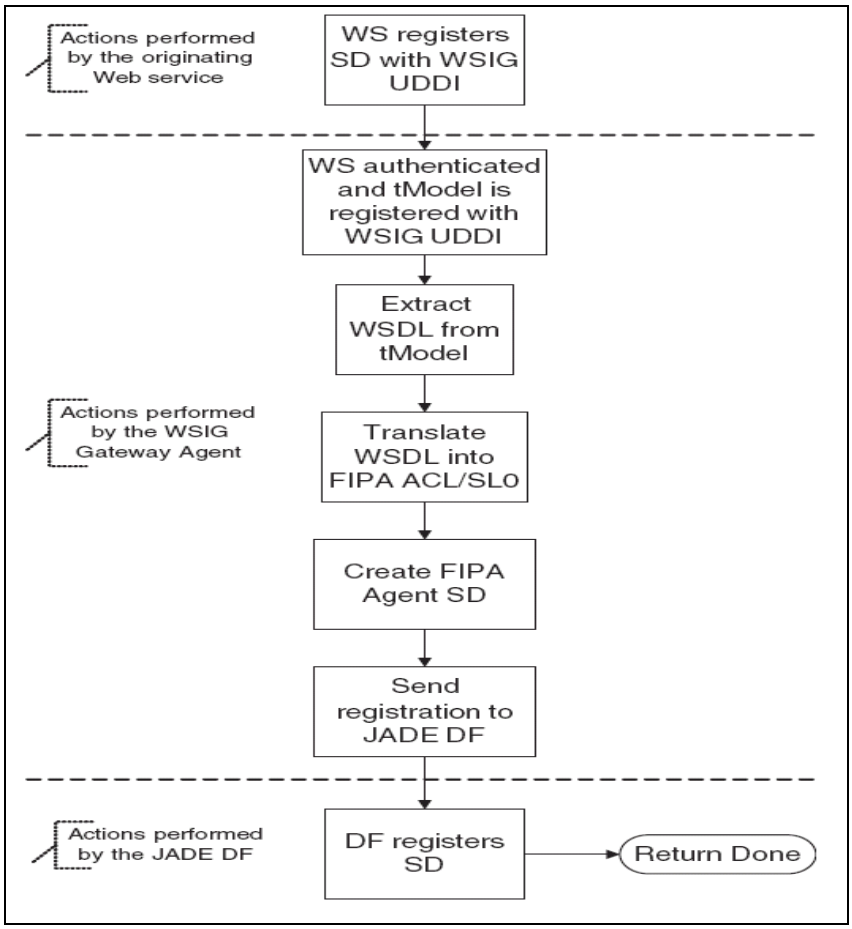
However in the entire platform of multi modal agent, the web service client would be a PDA handheld device with or without speech server which requests and responses messages from this WSIG environment including web service and agent service. The detail of invocation from web service client to agent service and agent service to web service client will be presented in the following paragraphs.

5.3.1. Web Service Environment in WSIG

The web service is normally registered, deregistered, modified and discovered in WSIG UDDI for publishing web service descriptions, and also the agent service is registered and deregistered with the WSIG UDDI by a gateway agent in order to be discovered and invoked by external web services. Therefore the web service has first to look for agent service descriptions in the UDDI registered by gateway agent, before sending an invocation request message to the gateway agent through the Axis web server component *[JADE Board 2005]*.

5.3.1.1. Web Service Registration in WSIG

Figure 5.3 describes a process which the outside web services publishes a SD can be permitted by the WSIG as an agent SD publishes with the gateway agent DF. However in this diagram the web service SD is registered with the WSIG UDDI as a normal web service registration. Once the new registration of the web service within UDDI in WSIG is detected, the gateway agent will extract WSDL from the tModel in order to translate into a FIPA ACL/SLO, and create a FIPA agent service description registered into platform DF which performed in JADE. Furthermore if the WSDL of a web service includes several operations, the gateway agent will process them sequentially with translated into separate agent SD for the DF *[JADE Board 2005]*.



Web Service Registration with WSIG Gateway *[source: JADE Board 2005]*

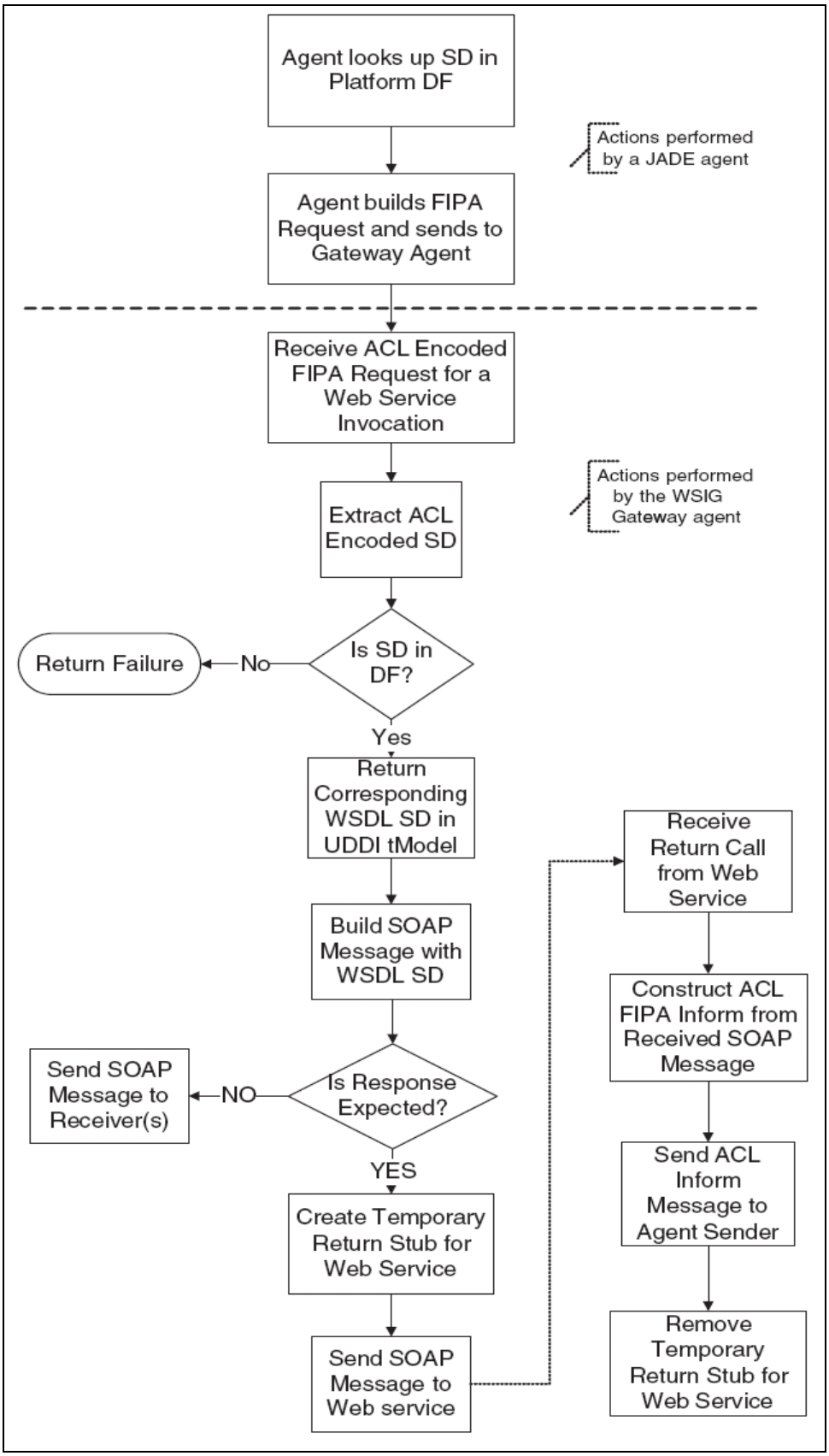
Figure 5. 3

5.3.1.2. The Invocation of Web Service in WSIG by an Agent

A process of an agent invoking web services is described in Figure 5.4. Before the invocation, there is a constraint that the WSIG UDDI repository must have the registration of a web service. Firstly the JADE agent has to look up ACL encoded web service description in platform DF which the web SD was translated from the original WSDL by the gateway agent. Once the JADE agent has discovered this encoded SD, it will then send a FIPA request message for invoking, but the FIPA request message has to include any necessary parameters and the identity of the service for being invoked.

In the WSIG gateway agent section, the gateway agent will map the ACL encoded SD into the corresponding WSDL, and stored in the UDDI repository. When the UDDI repository has the WSDL SD which will be transformed into the form of SOAP encoded message from ACL form (this is an ACL to SOAP bi-directional codec).

Then this message is used for invoking a web service with a temporary return stub which receives any expected response, if any message received by the stub will be constructed back into ACL and returned to the original sender agent using a FIPA inform message. This temporary stub is removed once a response is received. In addition, the web service stub provides replacement implementations for an invocation is received from a Web service client *[JADE Board 2005]*.



The Web Service Invocation *[source: Greenwood 2005]*

Figure 5. 4

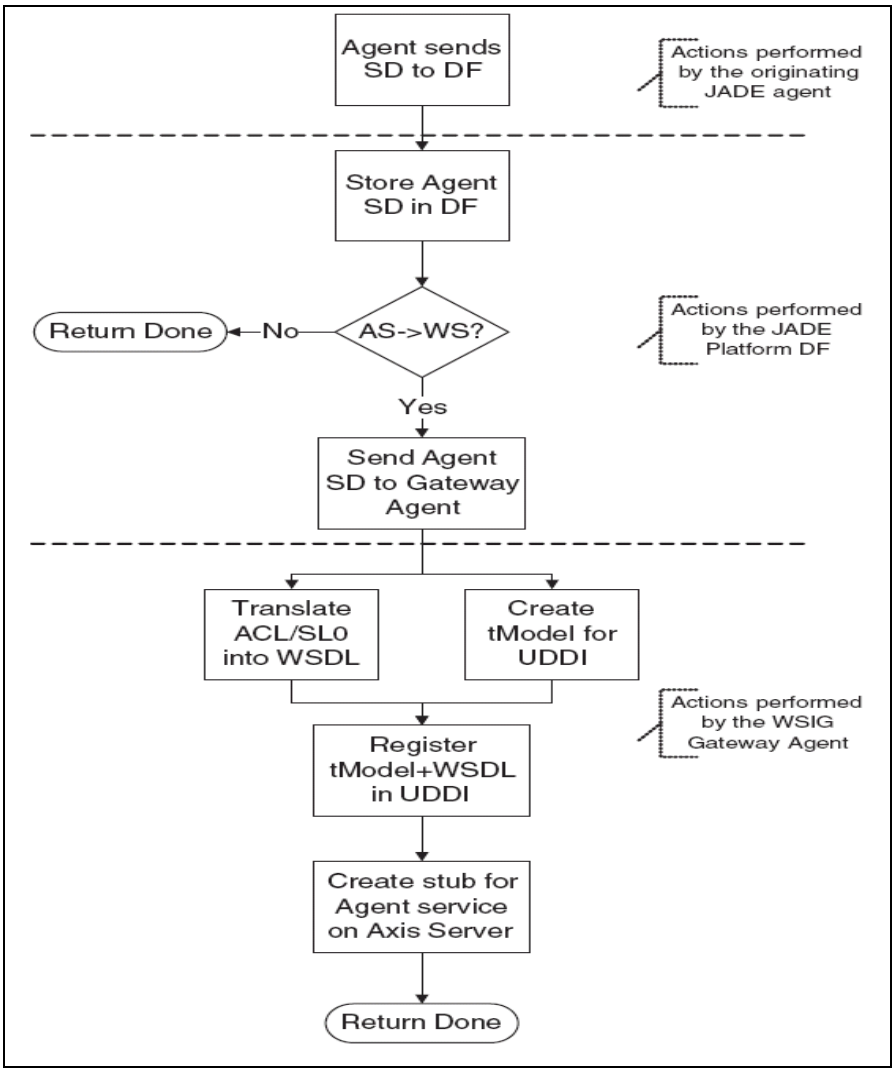
5.3.2. Agent Environment in WSIG

The agent service description is normally published by the gateway agent, and the web service can be also modified, discovered, registered and deregistered with the DF via the translation of gateway agent, that the web service can be directly discovered and invoked by the JADE agent, but it can be only deregistered and modified by the gateway agent. Therefore the JADE agent has first to look for the web service description in the DF in order to send an invocation request message to it through the gateway agent.

5.3.2.1. Agent Registration in WSIG

Figure 5.5 below describes the process which the internal JADE agent publishes a service description can be allowed by the WSIG as a web service description publishes with the WISG UDDI. If the JADE agent tries to provide the accessible to outside web service that it has to set the type parameter of the agent service description to "web-service": sd.setType("web-service") *[Bellifemine, Caire, Greenwood 2007]*.

However the initiation has been set the JADE agent just simple requests the registration (sending service description) to the platform DF without any transforming. Once the DF has the registration, it will automatically forward to the gateway agent which will perform the translation of the ACL/SL0 into the WSDL tModel, and register the WSDL tModel into UDDI for the directly invoking from external web service clients, also create a stub for JADE agent on Axis server *[Shafiql, Ali, Suguri, Ahmad 2006]*.



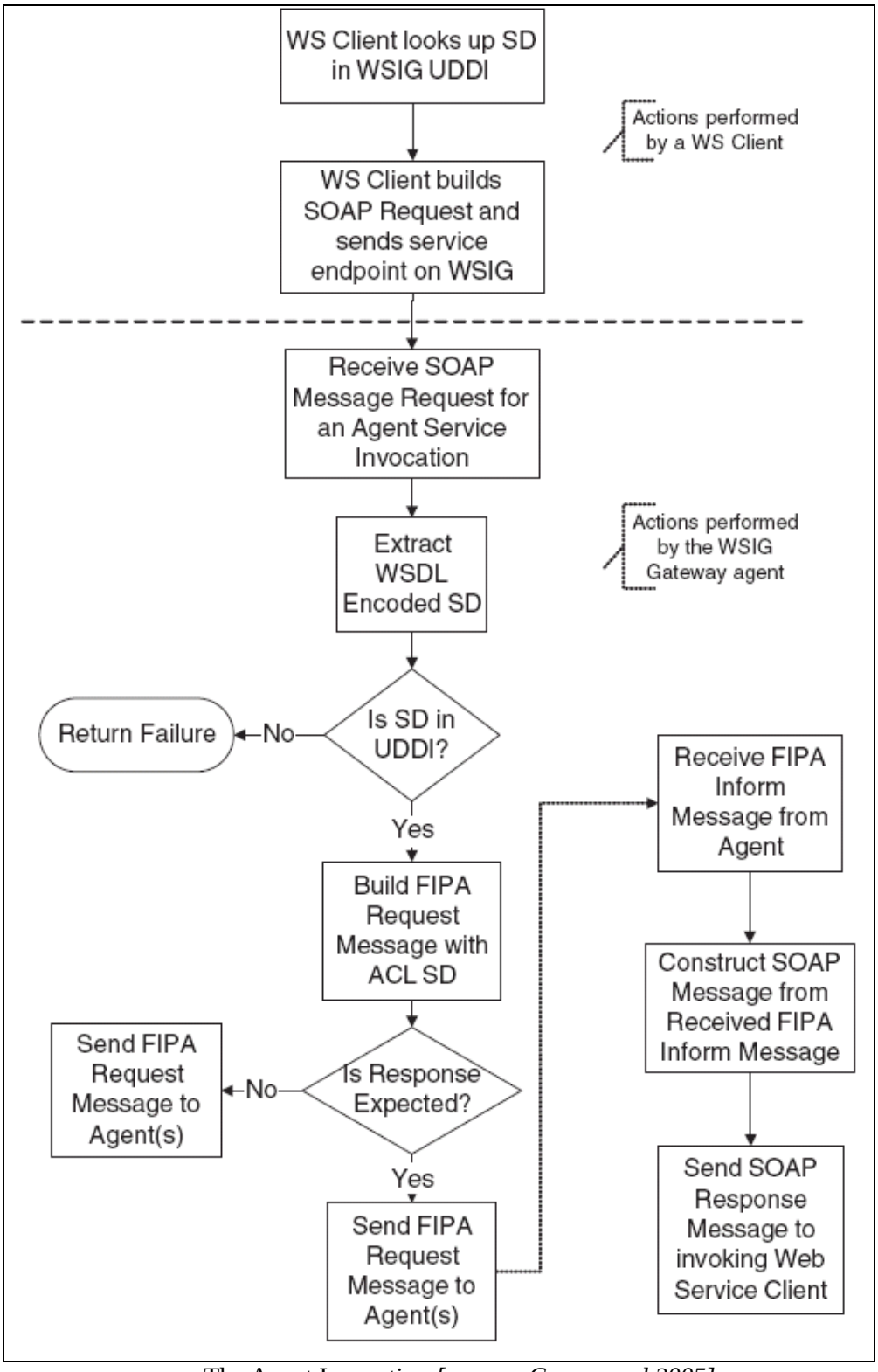
Agent Service Registration with WSIG Gateway *[source: JADE Board 2005]*

Figure 5. 5

5.3.2.2. The Invocation of Agent in WSIG by Web Service

A process of a web service invoking JADE agent is described by Figure 5.6 below. Before the invocation, there is a limitation that the WSIG DF repository must have to have the registration of a JADE agent. Firstly whether web services or web service clients have to look up the service description in the WSIG UDDI repository, and be able to build SOAP encapsulated messages onto service endpoints exposed Axis server, meanwhile the SOAP request message has to include any necessary parameters and the identity of the service for being invoked, those actions are performed by the web services or web service clients *[Shafiql, Ali, Suguri, Ahmad 2006]*.

When the gateway agent is received by the SOAP request message, it will be parsed in order to translate the identity to corresponding DF service description and the parameters value to a new created FIPA request message. This message is then sent to the agent exposing the agent service and if a response is expected, the gateway agent will expect to receive a FIPA Inform from the target agent. Finally a response is received it is constructed into a SOAP message and returned to the invoking Web service client back *[Greenwood, Calisti 2004]*.



The Agent Invocation *[source: Greenwood 2005]*

Figure 5. 6

5.4. Speech Server

Accessing information on the web by using multimodal interface is currently popular technology; especially using voice to control application is a hot topic. Nevertheless the speech server is a Speech Application Language Tag (SALT)-compatible server which enables the capabilities of doing SR and TTS.

"Speech Recognition recognizes caller utterances by means of one or more application specific grammars and converts spoken audio into text.

Text-To-Speech produces a speech audio stream from text provided by the speech application or stored in databases for playback over a telephone [IBM 2001]."

SALT is a web programming specification extending existing mark-up languages such as HyperText Markup Language (HTML), eXtensible HyperText Markup Language (XHTML) and XML; it helps in the creation of speech interfaces to allow speech to be used as an approach to communicate between users and web applications *[Slatforum 2007]*.

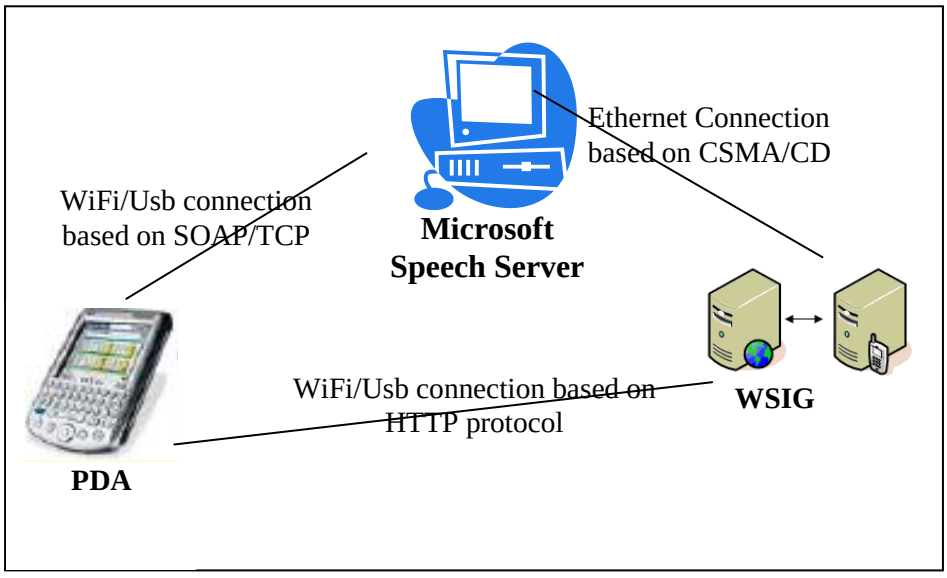
Speech Server is used when a device does not have enough abilities to do SR or TTS on its own. Therefore the two most prominent functions for Speech Server would be as followed:

- **Voice-Only Telephone Applications**: This function is deal with telephone as a terminal; it allows users to use voice commands to interact with servers by using telephones. For example, when a user calls up his automated phone bank to check his account balance, his inputting commands can be done by using a keyboard or through his voice on the phone, simply the output is either a pre recorded voice or a computer synthesized voice. This system will not cover this function of speech server, because the user client is a PDA to communicate with web service using internal a web browser.
- **Multi-Modal PDA Application**s: the multimodal applications are a new subject that are currently speech-enable web applications, they permit users to interact with web applications which are not only limited on traditional input and output (such as keyboard type, mouse click and screen display), users can also use SR

and TTS instead of. For example, when a user is online to check his bank balance, he can use the speech to speak instruction and receive information instead of conventional keyboard typing and mouse clicking. This system will focus on this function of speech server, and use Microsoft speech server as a sample of speech server *[Chesnut 2005]*.

5.4.1. Microsoft Speech Server

The Microsoft Speech Server (MSS) is a multi component and distributed server environment running on windows Operating System, such as Pocket PC and Internet Explore, its purpose is to allow the speech can be enabled on web application in order to assist some devices such as PDA, which do not have sufficient memory, capability and performance to carry out the implementation of speech engine service. In Microsoft Speech Server, the SALT mark-up tag is embedded with a HTML document which is automatically generated by use of ASP.NET dialog controls and is processed by a browser with add-in for Pocket Internet Explorer that directly communicates with the speech engine service to do recognition and synthesis for SR and TTS. Figure 5.7 is showed below for describing the use of speech server in the multimodal agent system [Microsoft 2005].



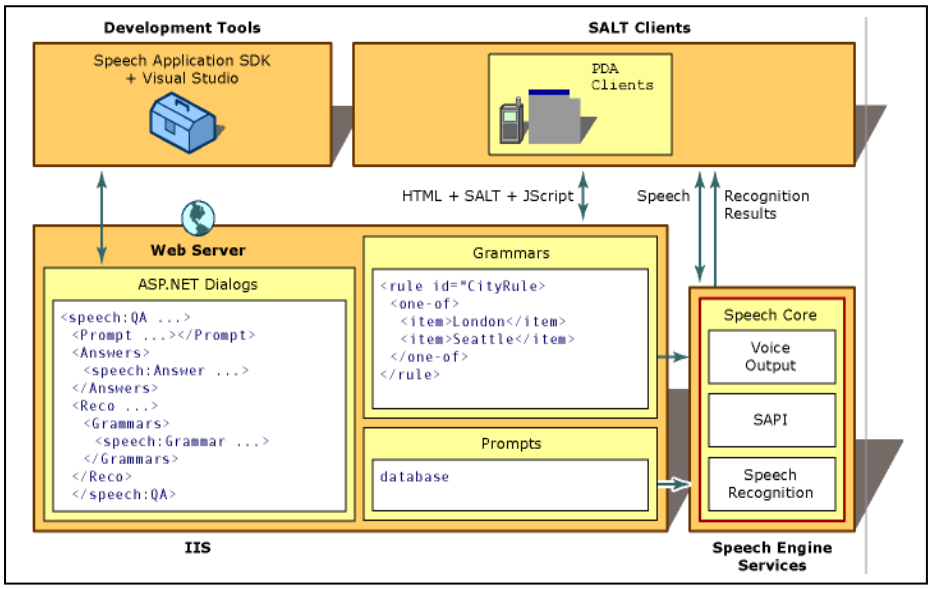
Work Principle of Speech Server in Multimodal Agent System *[Source: the Author]*

Figure 5. 7

As Figure 5.7 illustrates, a multimodal device in here is a PDA, a user employs a speech add-in browser (such as pocket internet explorer) to surf a web page which is

generated with HTML, SALT and script mark-up language from a web service, that the web service hosts the speech application. Primarily the browser initiates a session with Microsoft speech server by using SOAP message, and then the web page on this browser could require either speech input from the user or TTS output. If the web page requires a speech input, the pocket internet explorer does the processing and compressing of the voice commands in order to transmit that to speech server over TCP/IP network connection. On the speech server side, it then performs and recognizes the voice commands according to a pre-determined "list" of things that users will be expected to say, this "list" is called a "grammar", and returns the results as an XML message to the speech add-in pocket Internet Explorer then displays it on the web page. The experience of the user conducting this speech recognition and text-to-speech is that the speech server is always behind scenes to build those connections with web server and PDA browser; it is invisible for the user to think that the connection is directly established to web server.

Note a PDA has to establish a "trust" connection with speech server called Access Control List (ACL). Which means before the PDA does either speech recognition or text to speech. That means during each initiation time, a user has to configure speech server to trust my web server for communicating with a new device.



Technical Architecture of Speech Server *[Source: Microsoft 2005]*

Figure 5. 8

Figure 5.8 illustrated is a speech server scenario, the PDA clients have to have the speech add-in browser (Pocket Internet Explorer), and the web server is a ASP.NET

speech enabled web service developed by Speech Application Software Development Kits and Visual Studio 2003/2005, which can generate speech enabled web application pages, as well as this server contains some pre-recorded grammars and a prompts database for TTS. Importantly it should also have a configuration file containing the Uniform Resource Locator (URL) to the speech server that implements speech operations.

Initially the user enters a URL in address box on PDA browser to connect web server, when the web server has received this request, it sends the URL pointing to speech server to allow user's PDA can find the specific speech server. Once the user requests a web page including SALT speech mark-up language to talk to PDA and listen from the PDA. Thus there is either Speech Recognition or Text to Speech will be carried out.

Speech Recognition: the PDA sends either an inline recognition grammar or a location address of external stored recognition grammar (speech elements enable recognition of users' voice). If it is an inline grammar, the speech server simply loads the grammar and carries out recognition. Whereas the grammar is an external grammar, the speech server has to download it first, and then loads the grammar and carries out recognition.

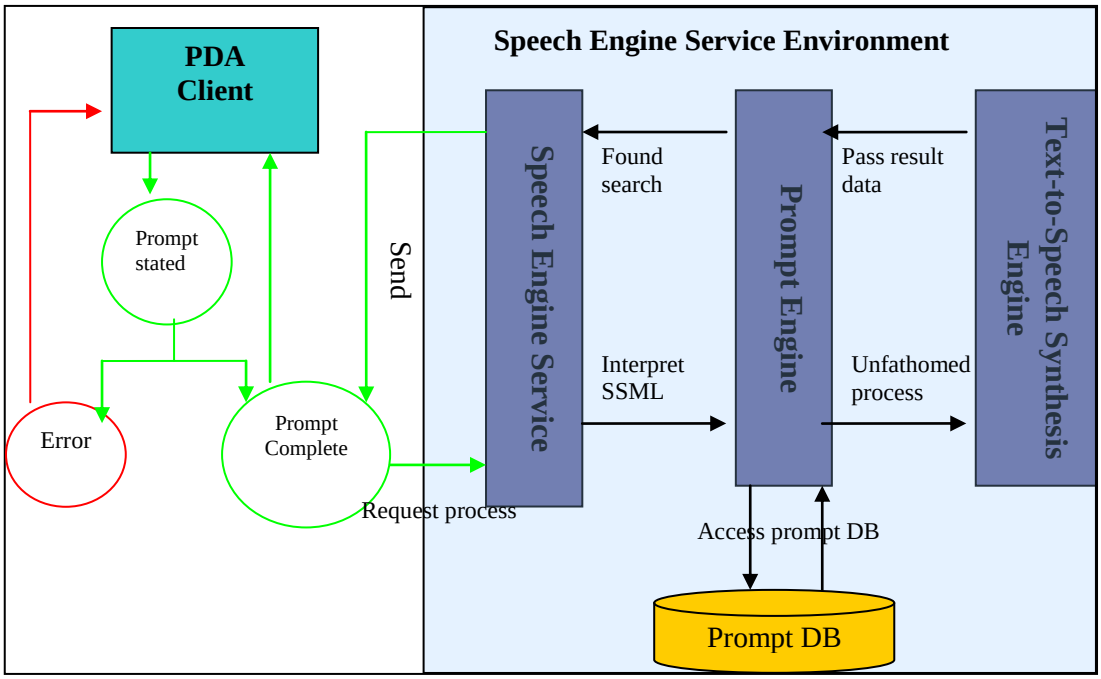
Text to Speech: When the speech recognition is finished, the speech server sends Semantic Markup Language output to the PDA through the process of Speech Engine Service showed in this scenario, if the output requires the web application page to play an audio file for prompts, that the PDA client parses the SML output and plays any prompts that speech server sends.

The further details of Text-to-Speech, Speech Recognition and Speech Engine Service are all discussed in the following Sections 5.4.1.1 and 5.4.1.2.

5.4.1.1 Text-to-Speech

Text-to-Speech capabilities includes the ability to play back printed text as spoken words, the synthesized speech can be generated from a concatenating pieces of pre-recorded audio files which are stored in a pre-defined database, these pre-recorded

storage of entire words or sentences allows for high-quality and clarity output. For example the utility of an intelligible TTS with quality of a speech synthesizer may assist people with physical disabilities to listen and written works on a speech enable device. The detail of TTS engine in multimodal agent system is drawn below at Figure 5.9 *[Microsoft 2005]*.

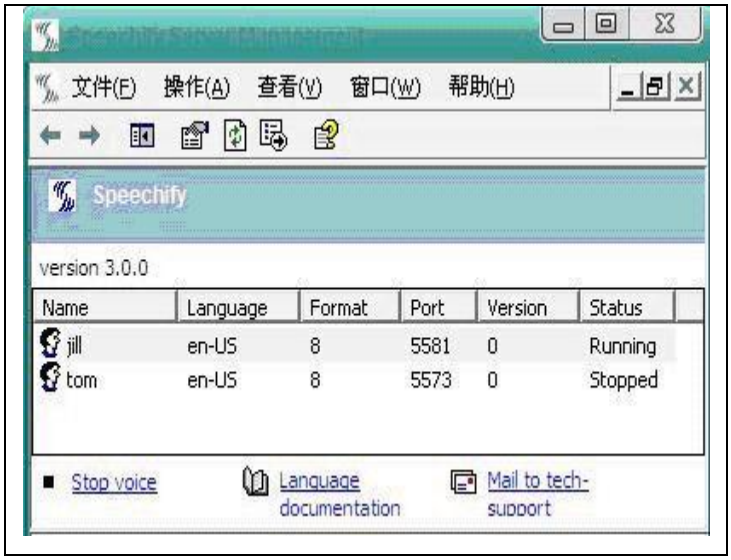


Work Flow of Text to Speech *[Source:the Author]*

Figure 5. 9

Figure 5.9 describes the theoretic work principle of Text-to-Speech; basically a user starts with controlling a prompt command, which has to be carried out in a Speech Engine Service Environment. In this environment Speech Engine Service (SES) loads a request from a PDA client with speech application, which the user wants to play a text-to-speech. Therefore the SES interprets Speech Synthesis Markup Language (SSML) from that request, and then passes to prompt engine which compares this SSML to the prompt database (they can be defined in advance or in-line as a component of the environment), whether or not there is any pre-recorded audio files or prompts that match to this SSML according to the text. But if the prompt engine can not find any part of the text to a pre-recorded audio files or prompts, it sends the text to the text-to-speech synthesis engine handling it *[W3C 2002]*.

The purpose of a text-to-speech synthesis engine in this speech engine service environment is to approximate the audio stream in order to allow a human voice reading the text which can only be able to read an approximate speech. Figure 5.10 below shows a interface for speechify server management control panel, usually the voice "Jill" (female) is a default audio speech, and users can use "stop" and "run" to configure whether choice "Jill" or "Tom" to speak. In addition users can still install any other language packets (such as Chinese language packet) for permitting different languages speaking *[Microsoft.com 2007]*.

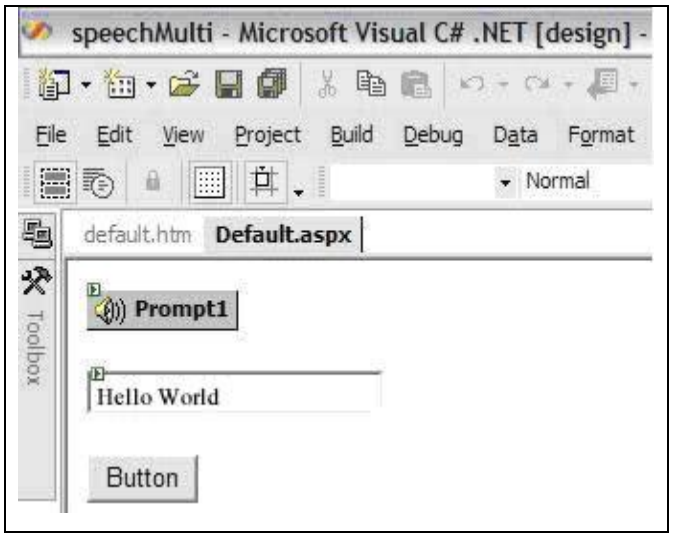


The Control Panel of Speechify Server Management *[Source: the Author]*.

Figure 5. 10

An example of Text-to-Speech based on SALT mark-up language.

The TTS can be developed by different tools, but in this example it would be programmed using Visual Studio 2003/2005 with Speech Software Development Kits 1.1/5.1. A simple variation on the "Hello World" program allows a PDA to speak text entered by the user into an input text area as the development picture below.



The development picture of TTS program *[Source: Casey 2004]*

Figure 5. 11

Figure 5.11 is a Prompt control to the page along with an HtmlTextBox and an HtmlButton which the HtmlTextBox allows users to enter speak text, and the HtmlButton allows users to start the speech from the text entered in the HtmlTextBox. At this point, the code of this picture is showed as followed:

Note the little "Prompt" button means there is a "prompt" element in this page.

```
1:<html
xmlns:salt="http://www.saltforum.org/2002/SALT">
2:<object id="speech-add-in" CLASSID="clsid:33cbfc53-
a7de-491a-90f3-0e782a7e347a">
</object>
3:<?import namespace="salt" implementation="#speech-
add-in"/>
4:<salt:prompt id="Prompt1">
5:<InlineContent>Hello World</InlineContent>
</salt:prompt>
7:<input type="text" id="iptext" name="iptext">
8:<input type="button" name="speak" value="button"
onClick="startSpeak()">
</body>
9:<script>
10:function startSpeak()
{
  11:var pfield=document.getElementById("iptext");
  12:var pprompt=document.getElementById("Prompt1");
  13:pprompt.Start(pfield.value);
}
</script>
</html>
```

The comments for each line of this program are also followed:

Line 1:

Any SALT program is just an HTML compiled page with the SALT mark-up tags embedded, thus the HTML page refers to a URL with the defined namespace of SALT which means this definition states there will be a SALT embedded tag that is going to be used.

Line 2:

Once the HTML defines the SALT, it has to refer specific performance at the beginning of following implementation (tags), therefore the performance is an Active-

X object, it loads the speech add-in active with the specific CLASSID identifier, which the CLASSID means this HTML page is the normal end user version as a part of speech add – in redistributable files. Moreover there actually is another one is comes with the visual studio debugger.

CLASSID="clsid:DCF68E5B-84A1-4047-98A4-0A72276D19CC".

Line 3:

This is an XML process instruction to the Internet Explorer which connects the SALT namespace to the Speech add-in object.

Line 4:

It means the following tags are contained from here and closed at "</salt:prompt>", which is a SALT tag with "prompt" element, and the "id="prompter"" provides a unique object name for the prompt element.

Line 5:

It is a property which includes one or more basic speech controls, they could be "content" and "value" type.

Line 6:

Simply the body of HTML begins here

Line 7:

A text area allows users to enter any text for speaking called "iptext"

Line 8:

A little button with "startSpeak" function that provides the speech feature if users click on it called "speak"

Line 9:

The following codes are all script language and end by "</script>", here the script language is JavaScript.

Line10:

A JavaScript function method named as “startSpeak”

Line11 and 12:

These lines use two variables to define the text area and prompt element to obtain their values

Line 13:

To get object references for both the text area and the prompt element. Then pass the value of the text area into the Start( ) (pprompt.Start(pfield.value);) method of the prompt element in order to execute the speech function.

Furthermore those codes can be only executed on speech add-in Internet Explorer on PC; they are not tied to speech server yet. In order to interact with speech server on PDA, there also have some additional codes to be added.

```
Private void Page_Load(object sender, System.EventArgs
e)
{
1:          string  speechServer =
ConfigurationSettings.AppSettings[“SpeechServer”];

2:          Param  param = new Param();
3:          param.Name = “server”;
4:          param.Value = speechServer;
5:          prompt1.Params.Add(param);
}
```

Line1:

To get the variable “speechServer” reference for the “trust” connection, for example the “SpeechServer” address can be configured as “http://localhost/speechTest”

Line2:

To use Microsoft.Speech.Web.UI object, and define a new variable to refer the “Param” element

Line3 and 4:

To specify the location and server name of a remote speech recognition server for distributed architectures

Line5:

To use this "param" method *[Casey 2004]*

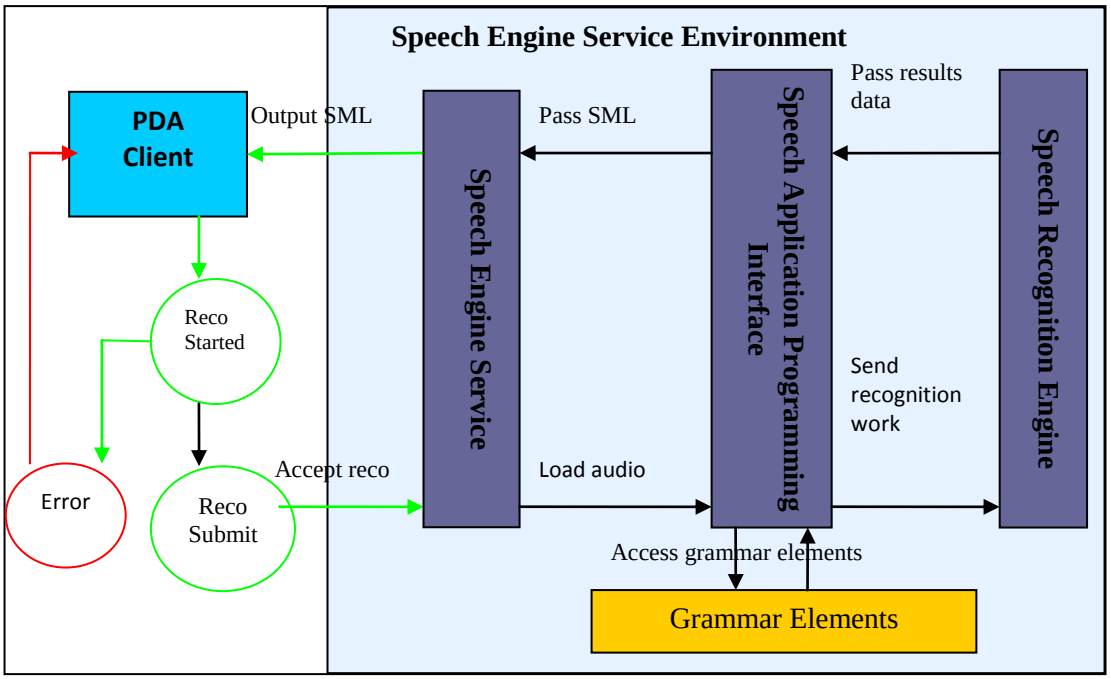
The User Interface of TTS *[Source: Casey 2004]*

Figure 5. 12

Finally a user can have a user interface similar Figure 5.12 below for a simple text to speech web page based on speech add-in browser on a PDA, The user could connect the "trust" site to link with speech server, and use button to start the speech from the words entered in the text area. Note the initial speech is "Hello World" when the user refreshes this web page.

5.4.1.2 Speech Recognition

In the MSS, the speech recognition is to process the transforming of a speech signal to a sequence of words in the form of digital data. It records audio input from the user, and interprets to digitized audio data implemented in a speech recognition engine which is a heart of performance, thus a speech application can understand it and response. From technical perspective, this speech recognition actually contains many standard procedures as illustrated in Figure 5.13.



Work Flow of Speech Recognition *[Source: the Author.]*

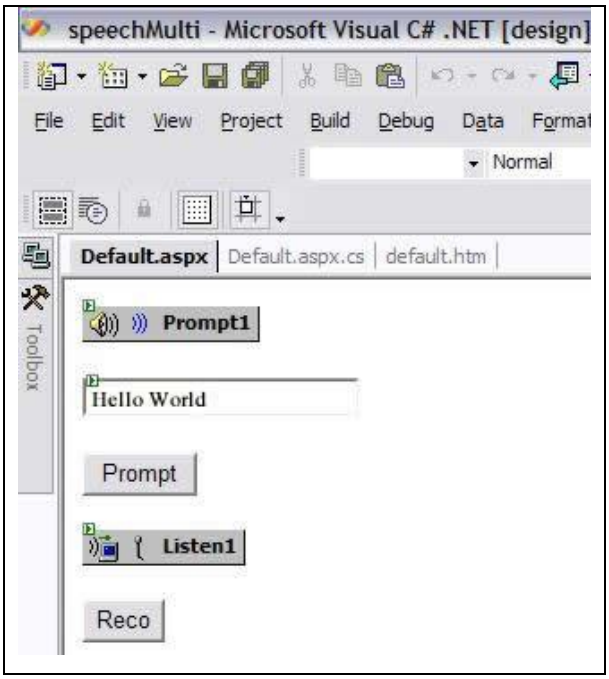
Figure 5. 13

When a user has a PDA device and talk to it, the PDA should firstly start a record function in order to obtain the audio stream for recognizing. Inside the speech recognition service, once the record audio is submitted to it which will load the grammar elements (they can be defined in advance or in-line as a component of the environment) in order to find the per- determined "list" of things for matching what the user said, if the words of user saying is out of grammar elements the speech server will require user to speak again. Meanwhile the grammar elements have to be either in W3C-complaint XML format or context-free grammar format.

However the speech engine service loads corresponding audio and grammar elements to SAPI, which parses the grammar elements into appropriate properties and syntaxes to dispatch it with the corresponding audio to a speech recognition engine. And then the speech recognition engine processes the real recognition performance, the result date will be sent back to SAPI after this implementation. Therefore the SAPI uses the grammar elements again to format the recognition result date into a speech mark-up language in order to pass it back to speech engine service, which finally will output this SML results back to the PDA client *[Microsoft 2005]*.

An example of Speech Recognition based on SALT mark-up language.

With the example of Speech Recognition, the design page should be like Figure 5.14. From this development picture, there added a Listen control and a new HtmlButton, which the Listen control represents record feature for obtaining users voice, and the HtmlButton is simply a control button to trigger the record feature.



The development picture of SR program *[Source: Casey 2004]*

Figure 5. 14

However in this design page, there should also have an in-line grammar to per-record some elements that expect users to say, and finally display on a PDA screen.

```
<html xmlns:salt="http://www.saltforum.org/2002/SALT">
<?import namespace="salt" implementation="#speech-add-in"/>
<!-- SALT: Recognise 7 colors -->
1: <salt:listen id="Listen1">
2: <salt:grammar>
3: <grammar version="1.0" xml:lang="en-US"
   xmlns="http://www.w3.org/2001/06/grammar" root="Rule1">
4: <rule id="Rule1" scope="public">
5:  <one-of>
6:   <item>Red</item>
     <item>Orange</item>
     <item>Yellow</item>
     <item>Green</item>
     <item>Blue</item>
     <item>Cyan</item>
     <item>Purple</item>
    </one-of>
   </rule>
  </grammar>
 </salt:grammar>
7: <salt:bind targetelement="iptext"
targetattribute=”value” value="/SML" />
</salt:listen>
<body>
8: <input type="button" name="Reco" value=”Reco”
onclick="Listen1.Start()" />
</body>
</html>
```

Line1:

The SALT mark-up language sets up a recognition object called “Listen1” which is used to do the recognition of the seven colors.

Line2:

The following tags are part of grammar elements, which is it defines what words sequences may be recognized by the “Listen1” object.

Line3:

This is a standard format of SALT grammars by the World Wide Web Consortium (W3C), which the language could be recognized will be US English, and refers to the grammar namespace at http://www.w3.org/2001/06/grammar.

Line4:

This is a “rule” element specifies one of grammar rules according to the W3C standard with the ID”Rule1”, furthermore as mentioned before, a grammar can be also referred from other applications, so the property of this grammar can be attributed as “public”.

Line5:

To define this rule is a single set of phrase alternatives demarcated by the 'one-of' element.

Line6:

To set those phrases within these “item” tags, which those phrases will be expected to be recognized

Line7:

To return the recognition results from the “Listen1” object, and bind with the “iptext” text area which belongs the TTS application, that means this SR application can recognize the voice from users speaking, and display the voice as a text message in “iptext” text area to allow the TTS application to read it as an audio. Alternatively the “value” attribute specifies what aspect of the recognition is to be returned, and the argument is Speech Markup Language.

Furthermore those codes can be only executed on speech add-in Internet Explorer on PC; they are not tied to speech server yet. In order to interact with speech server on PDA, there also have some additional codes to be added.

```
Private void Page_Load(object sender, System.EventArgs
e)
{
1:          string  speechServer =
ConfigurationSettings.AppSettings["SpeechServer"];

            Param  param = new Param();
            param.Name = "server";
            param.Value = speechServer;
            Prompt1.Params.Add(param);

// this block of codes are similar to the TTS which
configures // Listen1 feature on a PDA web page
            Param  param1 = new Param();
            Param1.Name = "server";
            Param1.Value = speechServer;
            Listen1.Params.Add(param1);
}
```

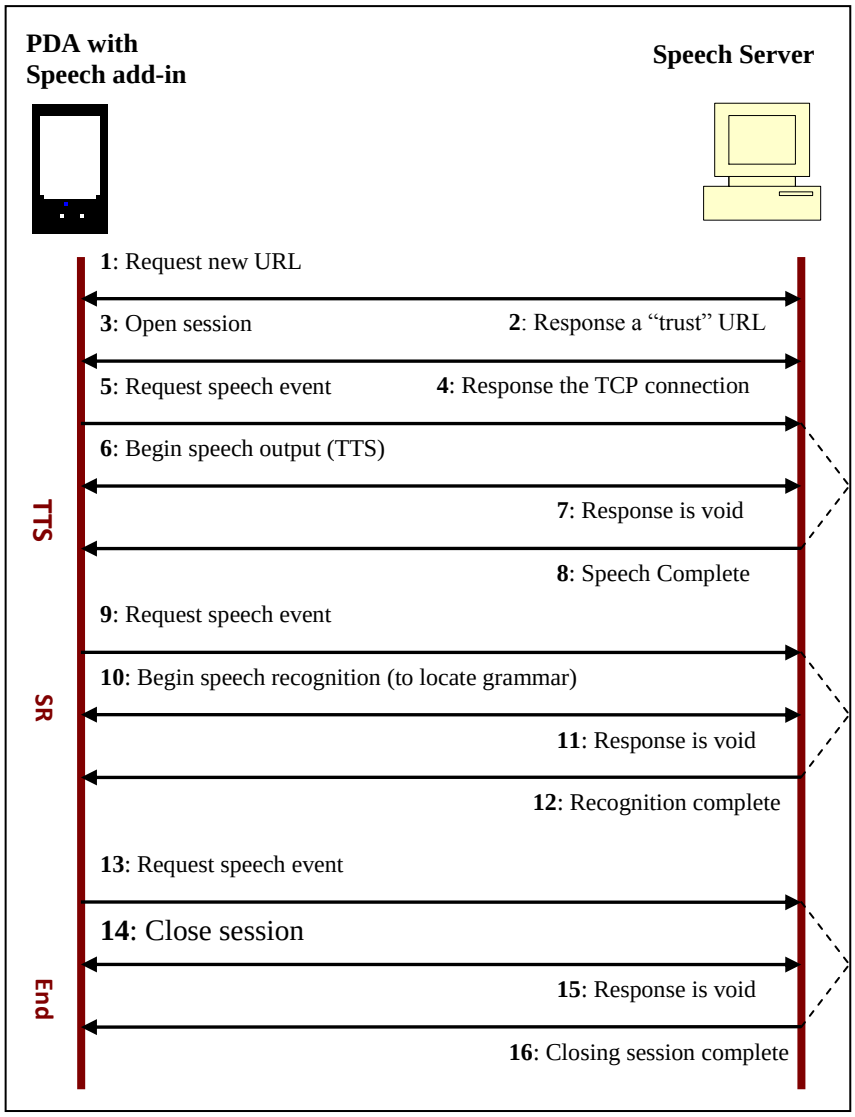
The User Interface of SR *[Source: Casey 2004]*

Figure 5. 15

As Figure 5.15 illustrated above, the user could click the Reco button to start SR function, once the record started, there is an audio meter in the bottom right which is a display of the input voice. Unfortunately this SR application does not allow users to speak whatever they want; it only recognizes the voice according to grammar elements which are seven colors. In addition, the user could also click the Prompt button to read the text from previous recognition displayed in the text filed.

5.5. The Interaction of Speech Server and PDA Client

After introducing the specification of speech server and some examples, in this section the interaction between a speech server and a PDA client is presented by using a sequence diagram as below in figure 5.16.



The Interaction of Speech Server and PDA *[Source: Casey 2004]*

Figure 5. 16

This sequence diagram is generally demonstrated the flow of the conversation between a PDA client and a speech server. Originally a PDA client has to send a request to make sure a speech server is there, such as using "ping" to detect the Internet Protocol (IP) address, once the speech server has discovered this request, it sends a response back with the "trust" site address, in the previous example the site address is configured as http://localhost/speechTest/. After this connection has been built, again the PDA client has to send a request to open the session, like to tell speech server it will do either TTS or SR, and the speech server has to return a response back with the transfer control protocol, for example TCP://192.168.1.1:8080.

So far the preparation has been carried out, thus the PDA is ready to process TTS and SR, firstly in the TTS operation, the PDA has again to request speech server some necessary preparation which is going to handle the text-to-speech, and then sends a text message to speech server to let it convert this text and finally output the prompt (audio). As well as it is similar to TTS process, when the PDA request a speech recognition event, that the PDA has to record voice and transfer to speech server to allow it convert this voice as a text, in addition during this convert in speech server, the speech server has to locate a grammar file either inline or pre-stored, finally the speech server outputs the text message back and display on the PDA client. In the end of this interaction, the PDA client sends a request to inform the speech server that this conversation is accurately finished, and then the speech server close the session.

5.6. Conclusions

The present document has laid the foundations for any further work with completely specifying a multi modal agent system, to date, the achievement of a consistently reliable development quality has entailed high testing costs which has impacted on the result of improvement. In this technical specification it has described the common view of system architecture of four roles; each role has been discussed in a separate and clear consequent structure. This work is focused in two principal areas:

- A Web Service Integration Gateway platform --- the invocation between a web service and an agent service in this platform has been discussed in the processes of registration, deregistration, modification and search.
- An interaction between speech server and PDA client --- the operations of speech recognition and text-to-speech has been discussed by using specific examples and interaction.

However this document has stated the general framework of a multi-modal agent system, it still has to be progressed in the future research, implementation and testing.

Chapter 6. Research Evaluation

In this chapter a detailed discussion of research of the multi-modal agent project is presented, which focuses on the objectives of the research methodologies, data collection and data analysis, each objective is stated as follows;

- Research Methodologies define the activity of researching, this research is based on the ways of proceeding, measuring progress and constituting success.
- Data Collection refers to the ways of collecting data in order to evaluate the research.
- Data Analysis provides the results of analysis for collected data.

Finally a conclusion will be delivered in the end to summarize the overall work of this chapter.

6.1. Research Methodology

The research methodologies that employed to gather the information in this research to address the evaluation objectives were as follows:

- Primary research: experiments survey and interviews
- Secondary research: review of literature
- Critical analysis
- Data collection

These methodologies are introduced in next Sections.

6.1.1. Primary Research

The primary goals of the research are to arrive at valid evaluation through scientific enquiry, and produce information precisely tailored for direct and hands-on access to this document. The primary research in this document was comprised of a range of different methods that they are described below.

- **Exploratory Experiments**

The test hypotheses in this document were to explore how a system would be implemented from both functional and technical perspectives, a use case diagram and architecture of this multi-modal system were structured in the functional perspective,

and also there were experiments of speech recognition and text-to-speech which were coded and implemented from technical perspective.

- **Online Survey**

An online survey has been undertaken to acquire both facts and opinions and attitudes that most accurately identify personal experiences and preferred mode of content delivery, in order to gather information to measure the valid evaluation of this document. Table 6.1 has listed a range of people who has been invited from email and the message boards and forums invitation.

PDA Forum and Board	Dell Community Forum	PDAPhoneHome.com
	Pocket PC magazine Forum	PDA Essentials forum
	BellaOnline Forums	PDA-EDU@JISCMAIL.AC.UK
	WiFi-Forum	Mobiledia Forums
	Total PDA Forums	PDA Street forums
Users and Non Users	Experience Users	Non-Experience Users
	Developers	Strangers
	Expert users	Novice users

A List of Survey Sources *[Source: the Author]*

Table 6. 1

- **Interview**

Interviews are semi-structured and in-depth to collect specific knowledge from expert and experienced persons who have direct insights on PDA, Web service, Agent technology and Multimodal interface. These persons are from a variety of different fields, such as lectures, programmers, professors, doctors and experts.

6.1.2. Secondary Research

The secondary goals of the research are to collect hard data that already exists about the multi-modal agent system. In this dissertation the method that has been adopted for secondary research was a literature review, it gives a good general overview of the multi-modal agent system subject and provide citations to primary research, which was research relating to multi-modal agent system using materials, such as search capabilities for recognised papers, text books, journals, conference articles, work paper, internet resources and CD-ROM resources. It was performed to

better identify how a multi-modal agent system is measured and compared and contrasted to existing PDA systems. As part of the literature review, some review was given to multi-modal agent system well sampling methodologies and direct push technologies in investigations of multi-modal agent system. Furthermore these literature review topics were all related to the underlying theme of the analysis and design of multi-modal agent system.

6.1.3. Critical Analysis

The critical analysis is employed in this dissertation to review the secondary research in order to extract reasons and evidence to support the arguments of multi-modal agent system, and create the critical success factors and the framework of the dissertation. At this point the critical analysis was used for identifying the focus of the multi-modal agent system, which has been critically discussed about the requirements of the system, and also identifying author's perspective of the multi-modal agent system has been clearly and directly discussed about the design of the system. In the dissertation there is a well-reasoned argument (such as coding examples and diagrams) that leads to a conclusion based on evidence to persuade readers of author's point of view.

6.1.4. Data Collection

During the process of the project, the data was collected to evaluate the project, which was both quantitative and qualitative data from academically rigorous sources, assessments and industry experts, as well as an online survey and interview have been undertaken in order to gather primary data.

- In the quantitative research the data was researched from an online survey and other work as simply a number through statistical manipulations to represent the sample of large population of interests.
- In qualitative research the data was collected from which contains virtually varied data as non numerical information in nature through group discussions, individual depth interviews, direct observation and written documents.

Nevertheless the data was collected from the online survey which recorded by the general population, who could be either PDAs users or non-PDA users. Alternatively

all of the data in interviews was collected by sending email to experts, lectures, programmers, professors and doctors.

6.2. Data Analysis

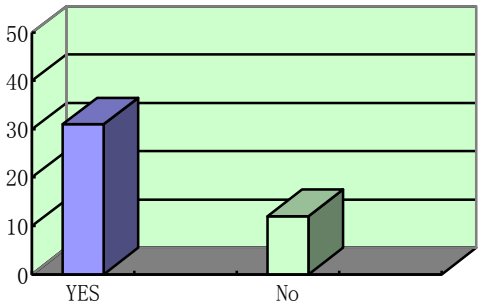
In this dissertation, interviews and questionnaires were the major sources of data to provide quantitative and qualitative data. As expected results and conclusions varied greatly from experts and users, those who have contributed the textural and numerical data according to their experience.

In the questionnaires and interviews they also provided the opportunity to allow experts and users to comment on how they felt the multi-modal agent system went for them, and reactions in preconceived notions. Experts and experience users with prior experience of using PDA have recorded significant and powerful comments, whereas non-experienced users with no prior experience of using PDA have immersed themselves in the questionnaires and interviews and thus recorded less significant comments.

A breakdown of questions answered in questionnaires and interviews is illustrated below.

Question 1: Do you currently have a PDA (Personal Digital Assistant)?

This question was designed to determine which of the people answering the questionnaire already own a PDA. The alternatives offered were 'yes' or 'no'.



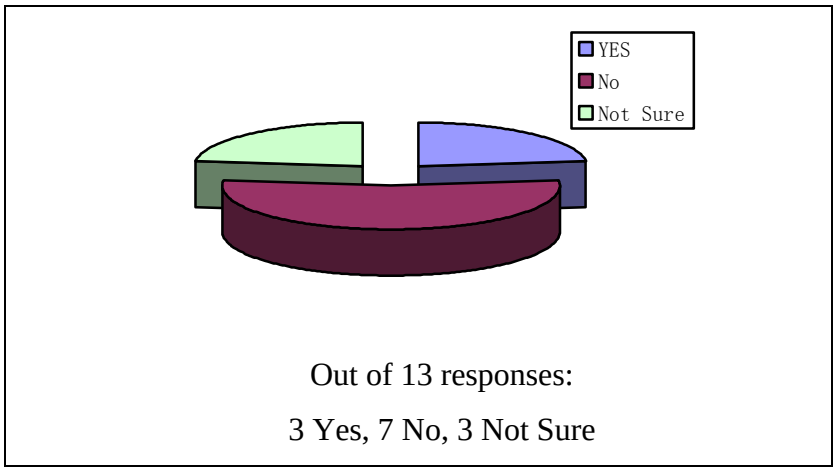
Out of 43 responses:

31 Yes, 12 No

As can be seen above, 31 out of 43, or 72% of respondents to this question already own a PDA. Thus the audiences of this questionnaire are mainly experienced or expert users of PDAs.

Question 2: Do you plan to buy a PDA?

This question was designed to see if those people who do not own a PDA plan to buy one. The alternatives offered were 'yes', 'not sure' or 'no'.



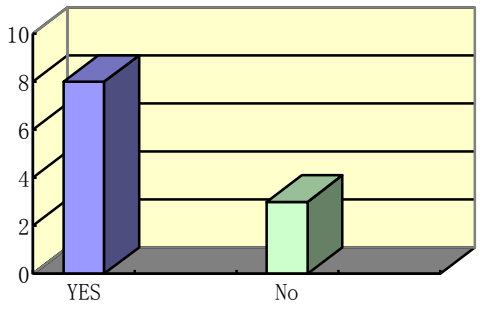
Out of 13 responses:

3 Yes, 7 No, 3 Not Sure

3 out of 13 respondents (or 23%) do intend to buy a PDA, 7 out of 13 respondents (or 54%) do not want a PDA, and 3 out of 13 respondents (or 23%) are not sure whether or not to buy a PDA.

Question 3: Will you use PDA, if you have the opportunity?

This question was designed to know whether or not non-experienced people are keen to use PDA, once they have the opportunity to use a PDA. The alternatives offered were 'yes' or 'no'.



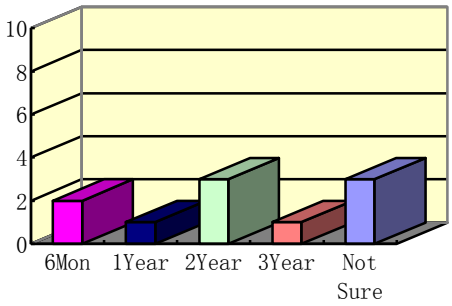
Out of 11 responses:

8 Yes, 3 No

8 out of 11 responses (or 73%) will do use PDA once they have an opportunity, and 3 out of 11 responses (or 27%) will not use PDA. Therefore a high potential market for non-experienced people has interests on PDA.

Question 4: When will you buy a PDA?

This question was designed to investigate people has no PDA, but who has interests to buy a PDA during a certain time period, it has presented a few different period to allow responders to ticket which the alternatives offered were '6 months', '1 year', '2 years', '3 years', or 'not sure'.



Out of 10 responses:

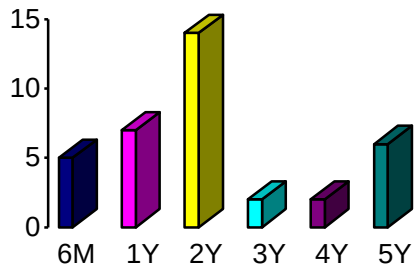
2 6M, 1 1Y, 3 2Y, 1 3Y, 3 Not Sure.

2 out of 10 responses (or 20%) will intent to buy a PDA during 6 months, 1 out of 10 responses (or 10%) will intent to buy a PDA during 1 year, 3 out of 10 responses (or 30%) will intent to buy a PDA during 2 years, 1 out of 10 responses (or 10%) will intent to buy a PDA during 3 years, 3 out of 10 responses (or 30%) do not know that

whether or not they will buy a PDA. Thus most of responders are quite interested to buy a PDA.

Question 5: If you currently own a PDA, how long have had it?

This question was designed to research the loyalty of PDA users. The alternatives offered were '6 months', '1 year', '2 years', '3 years', '4 years', or '5years'.



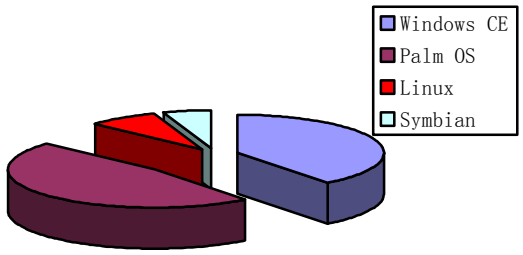
Out of 38 responses:

5 6M, 7 1Y, 14 2Y, 2 3Y, 2 4Y, 6 5Y.

5 out of 38 responses (or 13%) have had a PDA for 6 months, 7 out of 38 responses (or 18%) have had a PDA for 1 year, 14 out of 38 responses (or 37%) have had a PDA for 2 years, 2 out of 38 responses (or 5%) have had a PDA for 3 years, 2 out of 38 responses (or 5%) have had a PDA for 4 years, 7 out of 38 responses (or 18%) have had a PDA for 4 years. Therefore most of PDA users have owned PDAs during the recent 2 years.

Question 6: What is the Operating System of your current PDA?

This question was designed to see what the most popular operating system is on the PDA device. The alternatives offered were ' Windows CE ', 'Palm', 'Linux', or 'Symbian'.



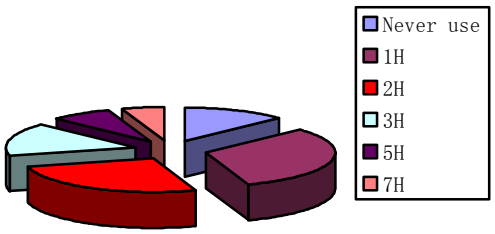
Out of 38 responses:

15 Windows CE, 18 Palm, 7 Linux, 3 Symbian

15 out of 38 responses (or 393%) currently are using Windows CE operating system, 18 out of 38 responses (or 47%) currently are using Palm operating system, 7 out of 38 responses (or 18%) currently are using Linux operating system, 3 out of 38 responses (or 8%) currently are using Symbian operating system. Thus the currently popular operating systems based on the PDA are both Window CE and Palm.

Question 7: How many hours per days do you use your PDA?

This question was designed to investigate the frequent that how often a PDA user uses the PDA per days. The alternatives offered were ' never use ', '1 hour', '2 hours', '3 hours', '5 hours', or '7 hours'.



Out of 38 responses:

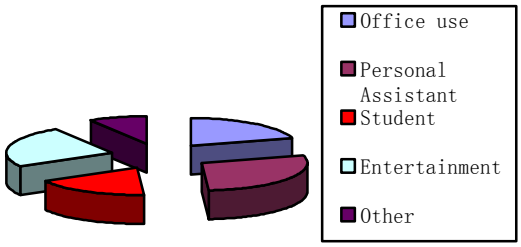
5 Never, 12 1h, 10 2h, 6 3h, 3 5h, 2 7h

Most of users only use their PDAs for couple hours, which 5 out of 38 responses (or 13%) do not use their PDAs per days, 12 out of 38 responses (or 32%) use their PDAs for 1 hour per days, 10 out of 38 responses (or 26%) use their PDAs for 2 hours

per days, 6 out of 38 (or 16%) responses user their PDAs for 3 hours per days, 3 out of 38 responses (or 7%) use their PDAs for 5 hours per days, 2 out of 38 responses (or 5%) use their PDAs for 7 hours per days. From these responses that most of users can be benefited from the impact of PDA.

Question 8: What do you use your PDA for?

This question was designed to identify some of the range of uses to which people employ their PDAs. In this question four alternatives were suggested; Office, Personal, Student, Entertainment and Other.



Out of 38 responses:

21 office uses, 29 personal assistants, 18 studies, 25 entertainments, 9 others.

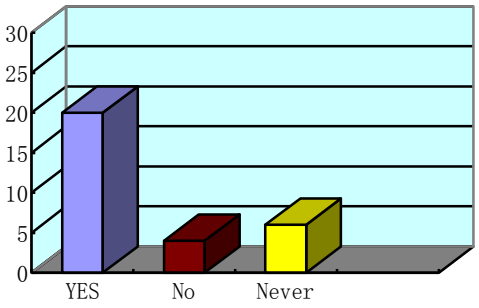
Most respondents choose more than one option as to what these use their PDA for; this shows the multifunctional nature of the PDA. 21 out of 38 indicated office use (or 55%), 29 out of 38 indicated personal use (or 76%), 18 out of 38 indicated studies (or 47%), 25 of 38 indicated entertainments (or 66%), 9 out of 38 indicated others use (or 24%) which the other category in this case gave a number of very interesting responses, some examples of the responses received include; for travel ("*Off road navigation*" and "*GPS navigation*"), for working with pictures ("*Image processing - using Pocket Artist and Resco*" and " *Photo Pro to edit images from my digital camera*"), for the internet ("*Internet* " and " *I would use it while on a break or for checking email "),* and also "*Address book*", " *Database of tunes, music*" and " *For debug my software*".

Question 9: Is there an alternative technology you favour over a PDA, and what?

This question was designed to know is there an alternative technology that users may favour over a PDA, these responses have chosen 'Yes' that give different technologies as follows; *"UMPC, Smartphones, Mobile phone, Iphone, Laptop";* and these responses have chosen 'No' that have very strong persuasion to think there is no an alternative technology to replace the PDA, for example, *"nothing around at the moment does everything in one box. It would mean separate gps/satnav, ipod, psp etc"*, otherwise these responses have chosen 'Maybe' that they are not sure there is an alternative technology to replace the PDA, for example, *"smartphone"*.

Question 10: Have you tried new software on your PDA?

This question was designed to investigate the potential PDA market for the next generation application. The alternatives offered were ' YES', 'No' or 'Never'.



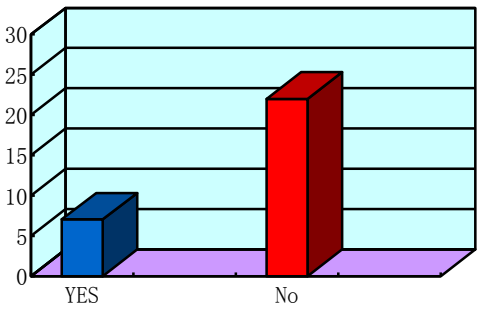
Out of 30 responses:

20 Yes, 4 No, 6 Never

20 out of 30 responses (or 67%) have tired new software in their PDAs; 4 out of 30 responses (or 13%) have not tired new software in their PDAs; furthermore 6 out of 30 responses (or 20%) have never installed and used any new software in their PDAs. Thus the potential PDA market is positive for developing the next generation application.

Question 11: Have you used speech recognition or text to speech software on your PDA?

This question was designed to see the popularity of speech recognition or text to speech software based on PDA. The alternatives offered were ' YES' or 'No'.



Out of 29 responses:

7 Yes, 22 No

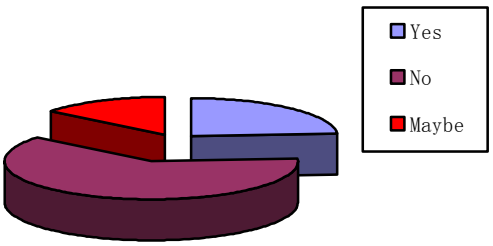
Most of responses have no experience with the using of speech recognition or text to speech software, 22 out of 29 responses (or 76%) have not used either speech recognition or text to speech software based on PDA, 7 out of 29 responses (or 24%) have used either speech recognition or text to speech software based on PDA.

Question 12: Please list the speech recognition/text to speech software you have used before?

This question was designed to know the existing voice software based on PDA, for investigating the potential conflict from other software to the multi-modal agent system. These responses have listed few voice software, for example, *"stadard WM6 voice regonition, Microsoft Voice Command, adobe voice reading software, VoiceIt, Microsoft speech"*. In addition these responses also have commented for each voice software they have used, for example, *"Microsoft Voice Command, but It wasnt that useful, so I uninstalled it; Microsoft speech - did not like"*.

Question 13: Will you try this multi modal agent system above on your PDA without considering cost, and why?

This question was designed to research whether or not users would like the current proposal of multi-modal agent system, and why. The alternatives offered were 'YES', 'No' or 'Maybe'.



Out of 21 responses

5Yes, 13 No, 3 Maybe.

5 out of 21 responses (or 24%) are happy with the current proposal of multi-modal agent system, and give some reasons, for example, *"easy for upgrade".* 12 out of 21 responses (or 62%) have decided to not consider the current proposal of multi-modal agent system, and give some reasons, for example, *"complicated, no interests, Keyboard/pen is faster & more accurate, No always-on connection, I just need an easy to use organizer".* 3 out of 21 responses (or 14%) are not sure to try the multi-modal agent system based on such as *"don't think voice regonition is the best way forward, Would try if I could see a benefit".*

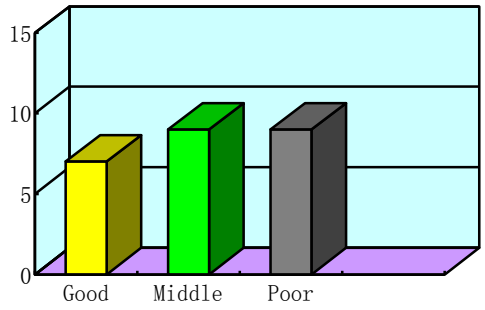
Question 14: What are the system most important features in your opinion?

This question was designed to see the most important advantages of this multi-modal agent system based on PDA. The following citation expresses the most impressive features of the multi-modal agent system.

- *"I think voice recoginition nice*
- *Clear speech recognition, the ability to switch between speech, pen, and keyboard input on the fly.*
- *Accuracy*
- *Cheap instant wireless connectivity*
- *Applications and rss feeds of my most popular websites..Igoogle is going down the direction with the gmail user homepage*
- *Voice input and recognition*
- *As a user I would like to gain access to useful content Also it should be easy to use and choose new content It should be reliable"*

Question 15: How is the interface?

This question was designed to exam the quality of the user interface proposed. The alternatives offered were 'Good', 'Middle' or 'Poor'.



Out of 25 responses:

7 Good, 9 Middle, 9 Poor

7 out of 25 responses (or 28%) have had good impression from the user interface, 9 out of 25 responses (or 36%) have had middle impression from the user interface, 9 out of 25 responses (or 36%) think the user interface is poor. Thus the following question was going to ask suggestions for improving this user interface.

Question 16: Have you any comments on how the interface could be improved? (For example: rich enough graphic, too rich, needs more memory..etc)

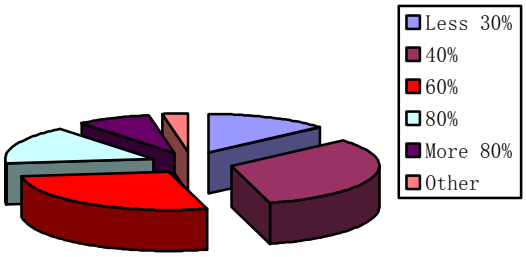
This question was designed to get recommendations for enhancing the user interface of the multi-modal agent system as follows;

- *"Rich*
- *Hard to know without usng it*
- *Buttons/controls too large. Wasted space in middle.*
- *It looks like you need tap the screen to switch between voice and pen, this is bad, at worst, and it should be a hard button on the chassis of the unit. Also, the screen didn't have a lot on it, and had to be scrolled. The options should fit on to one screen of the handheld, or at least the bulk of them should.*
- *Simple is best. I don't like the WinCE modal.*
- *Not sure what the UI is trying to accomplish. Button text extremely confusing.*
- *Smaller icons, prettier graphics*

- *Needs more memory*
- *I think its good, but maybe the user would not understand the functionality of the buttons. "Go" for example could mean anything! You could use picture buttons which might help the user to understand the functionality*
- *Less scroll bar"*

Question 17: Finally, how many marks you can give the PDA in your daily life?

This question was designed to summarize the impact of the PDA in the users' daily life. In this question 6 alternatives were suggested; less than 30%, 40%, 60%, 80%, more than 80% and other.



Out of 26 responses:

3 less 30%, 2 40%, 4 60%, 9 80%, 7 more 80%, 1 other.

Most of responses have given high marks for the impact of the PDA in their daily life, 3 out of 26 responses (or 12%) have marked less than 30%, 2 out of 26 responses (or 8%) have marked 40%, 4 out of 26 responses (or 16%) have marked 60%, 9 out of 26 responses (or 35%) have marked 80%, 7 out of 26 responses (or 27%) have marked more than 80%, 1 out 26 response (or 4%) have indicated other which in this case the response is *"depend on day".*

Furthermore many experts and users did not fell that they will be benefited from this multi-modal agent system, therefore they gave very negative responses and suggestions to this system, which will improve the next iteration of further research.

6.3. Conclusion

This chapter has provided a discussion of the data collection from expert domain users and general users, and has demonstrated the process for collecting data to support the experimental work described in pervious chapters. For the purpose of extracting appropriate information and knowledge about multi-modal agent system, this chapter has presented a suitable strategy for gathering data.

Research Methodology discusses scientific measurements for data collection.

Data Collection identifies quantitative and qualitative data.

Data Analysis analyses the data that has been collected from both questionnaires and interviews.

Chapter 7. Conclusion

In this conclusion chapter the document has explored an overview of this dissertation, gives summary of findings from previous chapters, lists the problems encountered in the research and investigation, and suggestions for the dissertation in further research, ultimately closing remarks defines remarks in the project.

7.1. Introduction

This dissertation set out to provide a clearer research of an investigation of multi-modal agent based PDAs and the development of a framework for it. The main point in this dissertation is to gather information and review literature to investigate this multi-modal agent system based PDA. Thus this project selected a discrete step to achieve the research gap.

- first explored was the impact of PDAs on users to research the popularity of PDAs,
- second analysed was each of key roles and interaction between these roles in the multi-modal agent system,
- third was proposed a functional specification and a technical specification to develop the framework of this multi-modal agent system,
- finally there was listed research methodologies used in this research, and an analysis of the data gathered from both an online surveys and interviews.

However the important contribution of this dissertation is that it provides guidelines to implement the framework of the multi-modal agent system based on PDA. Also it proposes a set of methods to analyse and design the performance of each function in the system, especially given the experimentation of some roles to test the execution in the system.

7.2. Summary of Findings

An introduction was presented in chapter 1 to introduce the dissertation in general, it started at a background on the research that PDA has been regarded so far for importing users' life, such as the increasing numbers of using PDA on the market. The multi-modal agent system was introduced for an understanding overview; each role in the system are listed and explained. The research problem was to suggest the

main purpose of this dissertation in identifying a research gap, which was again divided as two sub-problems to build the next generation PDA application and a framework for it. The research aim and objectives were both the roadmap to recommend a path to achieve the outcome of this dissertation. Furthermore a brief introduction for each chapter was provided to give readers a preview on the structure of this dissertation.

An investigation of the importance of PDAs on both personal and commercial usage was discussed in chapter 2, which separated discussions of the impact of PDA by dividing it into three sub-sections. First part provided a technical impact on hardware, software of the evolution of the PDA, discussed both the physical and functional restrictions; second part listed many paradigms as to why the PDA is important from a social perspective, such as a user could employ the PDA on business, study, clinical and navigation system, this section showed how in one day a PDA can be used in many ways, also the social limitation was expressed; the third section predicted the future trends of development of the PDA either from both technical and social views, some directions for the prospect also expected, for example the application of PDAs using Object Oriented applications and the next generation application of multi-modal agent system.

Before developing a framework for the multi-modal agent system based on PDA, the components within this system were thoroughly analyzed in chapter 3 which is a discussion of multi-modal agent system. The key components within this system are concluded as follows;

- **Agent technology:** A definition of agent technology with a variety of the benefits of its exploitation was depicted, in order to illustrate the adoption of the agent technology, for example, the adoption in entertainment to simulate its real world counterpart by means of animated visualisation. Furthermore the challenge and security issues were also discussed for looking forward to the future.
- **Multimodal strategy**: to explain the strategy of the multimodal interface that supports communication through interaction of different modalities, two examples that showed multimodal strategy have been used on an assistive multimodal system for disabled people based on the multiple methods of voice

and head movements, and one browser to allow users to navigate the Web through the voice content instead of only keyboard and mouse.

- **Web services**: to present the organisation of a web service based on XML, SOAP, WSDL and UDDI, an architecture of the interaction of each role within a web service was analyzed from three operations performed by service providers publishing services to a service registry, and service requesters finding required services using a service registry and binding to them, and the three security issues of the web service are integrity control, confidentiality and availability.

Alternatively, the difficulty of an agent integrating with a web service can be solved by the FIPA standard, which employs agent technology to manage and access the universal web services infrastructure and interoperability.

The functional specification of the multi-modal agent system was presented in the chapter 5; it was produced for a perspective that enables readers to have understanding of the functions of a multi-modal agent system, and to know the principles of how it works. In this chapter, the overall design goals of the multi-modal agent system were outlined, and provided a functional architecture diagram of the system was provided which mentioned the interaction between each key role within the system. In additional a use case diagram and an interface screenshot were illustrated in this chapter as well. The use case depicted all actions during the process of the multi-modal agent system, and the interface screenshot gives a picture of the multimodal interface that users should be interacting with.

The most important point in this dissertation is to develop a framework for the multi-modal agent system which is introduced in the chapter 6; this research also looked at multi-modal agent based PDAs from a technical point of view to propose guidelines for this system. This technical specification was discussed from four main aspects as follows;

- A class diagram of the multi-modal agent system used to explain the entire system within the roles, attributes and operations.
- A standard Web Service Integration Gateway (WSIG) was parsed to integrate agent technology and web service in order to handle the interoperability between them, and analysed the invocation from agent to web service and vice versa.

- Another key role within this system is speech server which was instanced as the Microsoft Speech Server (MSS); it coverts voice commands to text and vice versa. Furthermore there were developed two demonstrations to reveal how the speech recognition and text-to-speech conversion work based on the PDA.
- Ultimately a sequence diagram was employed to present the interaction between a PDA client and the MSS.

The validation of this dissertation was demonstrated in the chapter 6 with the research methodologies, data collection and data analysis. First the research methodologies expressed the methods used for primary research, secondary research and critical analysis, second the data collection presented the information has been gathered by quantitative and qualitative research, and finally the data analysis parsed the gathered data into charts and tables.

7.3. Problem Encountered in the Research

There are many contradictory and ambiguous papers and books in circulation on multi-modal agents related to this topic. One of the difficulties associated with this research was the access to thoroughly related sources on this topic; a majority of materials are only marginally associated with this topic which caused a considerable amount of confusion at the early phases of the research. However this problem was addressed by reading a variety of literature sources of that refer to all aspects of this dissertation, and also by engaging in discussions with experienced users, expert and supervisor, in order to acquire directions and assistance on this dissertation. Alternatively critical analysis of all documentation and derivations from this research was thus essential to ensure authority.

Another difficulty encountered during the course of this research was to get access costly software, such as the Microsoft Speech Server (MSS), an expensive software product required to support the speech synthesis for speech recognition and text-to-speech on PDAs, and also the web service should be deployed using Microsoft Visual Studio 2005 to develop the Speech Application Language Tags (SALT) embedded web pages, which again is another amount of expense.

7.4. Suggestions for Further Research

Recommendations for further research in this dissertation are proposed as the following different aspects;

- **A Speech database** that stores a range of complicated grammars that can predict what users are saying which is already recorded in the grammars, but actually in the real world grammars are impossible to know in advance, therefore a high quality speech database can only be deployed via iterative design and evaluation.

- **Standards for different platforms** enable the multi-modal agent system to interoperate with other technologies, for example, the multimodal interface should not only work on the SALT based web pages; in the speech systems vendor market there has to have a standard to allow the multimodal interface that also works on voiceXML based web pages. Adopting standards for different platforms helps reduce risk from volatility.

- **Business opportunities** for the multi-modal agent system based on PDA in Ireland provide a chance to catalyse the PDA commercial market which is merely limited by a certain amount speech products. The possibility of the next generation application based on PDA to boom the Irish PDA market is very high; it requires the further research to propose proposals.

- **Advanced web services** have the features that merely service the multi-modal agent system based on PDA, a further research of the advanced web services can be proposed for the functions that to produce the PDA size web pages in order to suit their screen; to gather the information and/or data from the web for PDAs only; to embed a speech server to reduce processes that PDA clients input and output voice command directly from advance web services.

7.5. Closing Remarks

This research set out to identify the key characteristics necessary to enable a multi-modal agent system to work on a PDA. A great deal of desk-based research was undertaken to identify some of these characteristics. To help clarify the relationship between different aspects of the system both a functional and technical specification of the system was developed. From this several further key features and requirements were identified, using the work a questionnaire was created which asked users questions about their use of PDAs. These results help confirm that the framework developed in this research represents an important step forward in building the next generation of PDAs.

Appendix A:

PDA Online Survey

1. Do you currently have a PDA (Personal Digital Assistant)?

- Yes (if Yes, please go question 5)
- No (if No, please go question 2)
- Other (Please Specify)

2. Do you plan to buy a PDA?

- Yes (if Yes, please go question 4)
- No (if No, please go question 3)
- Not Sure (if No Sure, please go question 4)

3. Will you use PDA, if you have the opportunity?

- Yes (if Yes, please go question 5)
- No (if No, now you can colse this page)

4. When will you buy a PDA?

- Within 6 months
- Within 1 year
- Within 2 year
- Within 3 year
- Other (Please Specify)

5. If you currently own a pda, how long have had it?

- Less than 6 months
- About 1 year
- 2 years
- 5 years
- Other (Please Specify)

6. What is the Operating System of your current PDA?

- Windows CE
- Palm OS

- Linux
- Symbian

7. How many hours per days do you used your PDA?

- Never think about it
- Less than 1 hour
- About 2 hours
- 3 hours
- 5 hours
- Other (Please Specify)

8. What do you use your PDA for?

- Office use
- Personal assistant
- Study
- Entertainment
- Other (Please Specify)

9. Is there an alternative technology you favour over a PDA, and what?

- Yes:
- No:
- Maybe:

10. Have you tried new software on your PDA?

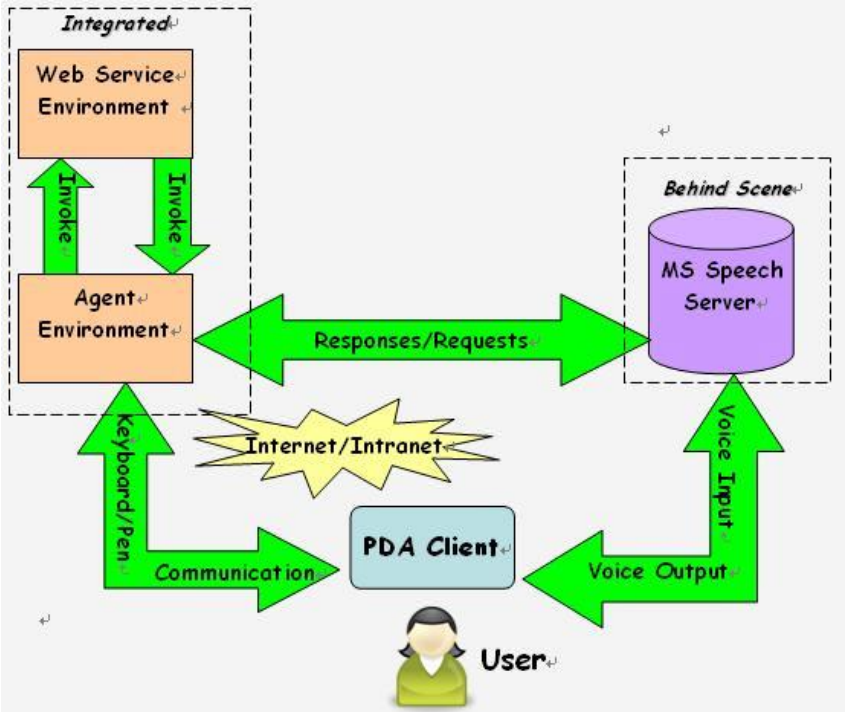
- Yes
- Sometime
- Never (if Never, please go last question)

11. Have you used speech recognition or text to speech software on your PDA?

- Yes (if Yes, please go question 12)
- No

12. Please list the speech recognition/text to speech software you have used before?

Answer:



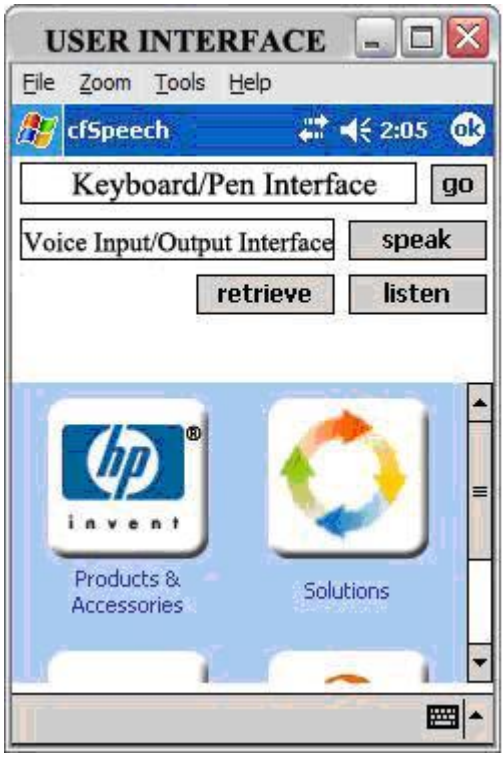
This picture depicts the functional scenario for the multimodal agent system, which clearly describes the interaction and integration of each component. In general, a user holds a PDA client for sending and receiving information using multi-modal interface. If the user chooses a keyboard/pen to communicate with the web service which is integrated with agent technology, the connection will be built directly between integrated web service and PDA browser through the Internet or Intranet. If the user chooses voice input/output to communicate with the integrated web service, the connection will go through to speech server firstly in order to process voice commands through the Intranet. Then the speech server (behind scene) will request and response with integrated web service. Once the connection is built with the integrated web service, internally the agent technology will automatically search sources such as the Web for updating or downloading new information such as calendar and email information, weather and traffic reports, entertainment listings, information about friends and colleagues, and finally represent the information in an alternative way.

- **13. Will you try this multi modal agent system above on your PDA without** considering cost, and why?
- Yes (if Yes, please go question 14):

- No:
- Maybe (if Maybe, please go question 14):

14. What are this system most important features in your opinion?

Answer:



This picture describes a compact framework user interface for the users interacting with the PDA devices, which is an actually basic browser with multi modal interface. In this browser, its major function is to perform the features of keyboard/pen and voice command. From the view of the picture, the User Interface contains three internal sections; top one is the function bar including file, zoom, tools and help functions, middle one is control panel for managing the features of keyboard/pen and voice command, and the last one is display part for showing any sort of information the users require. Meanwhile the very important section is the middle section which carries out the foremost function of this user interface, the very first text area on the top with little small button “go”, it is the keyboard/pen interface for the users typing keyboard or using stylus to enter information. The second text area with gray button ”speak” implements text-to-speech, that means the PDA talks to people in here if there is any information displayed and the “speak” button is activated. Eventually the last three buttons, the “listen” button does the feature for receiving commands

from users, and the ”retrieve” button does to give users the second chance for speaking again.

15. How is the interface?

- Good
- Middle
- Poor
- Other (Please Specify)

16. Have you any comments on how the interface could be improved? (For example: rich enough graphic, too rich, needs more memory..etc)

Answer:

17. Finally, how many marks you can give the PDA in your daily life?

- Less than 30%
- 40%
- 60%
- 80%
- More than 80%
- Other (Please Specify)

Appendix B:

Multi Modal Agent System Based on PDA Interview Questions

Name: _____
Organisation: _____
Date: _____

Multi Modal Agent system based on PDA Interview Questions:

1. **How many months/years you have use a PDA?**
   Answer:

2. **What operating system you have used on a PDA?**
   Answer:

3. **Which operating system you prefer to?**
   Answer:

4. **How many hours do you use a PDA per day?**
   Answer:

5. **What do you generally use your PDA for? (e.g. office, study and entertainment uses)**
   Answer:

6. **Have you developed any application for your PDA? If YES, what was its functionality?**
   Answer:

7. **Do you ever use speech recognition/text-to-speech products based on your PDA? If YES, what are these products?**
   Answer:

8. **From the figure 1 below what do you think, will the design of this multimodal agent system is structured well or not? and why?**
   Answer:

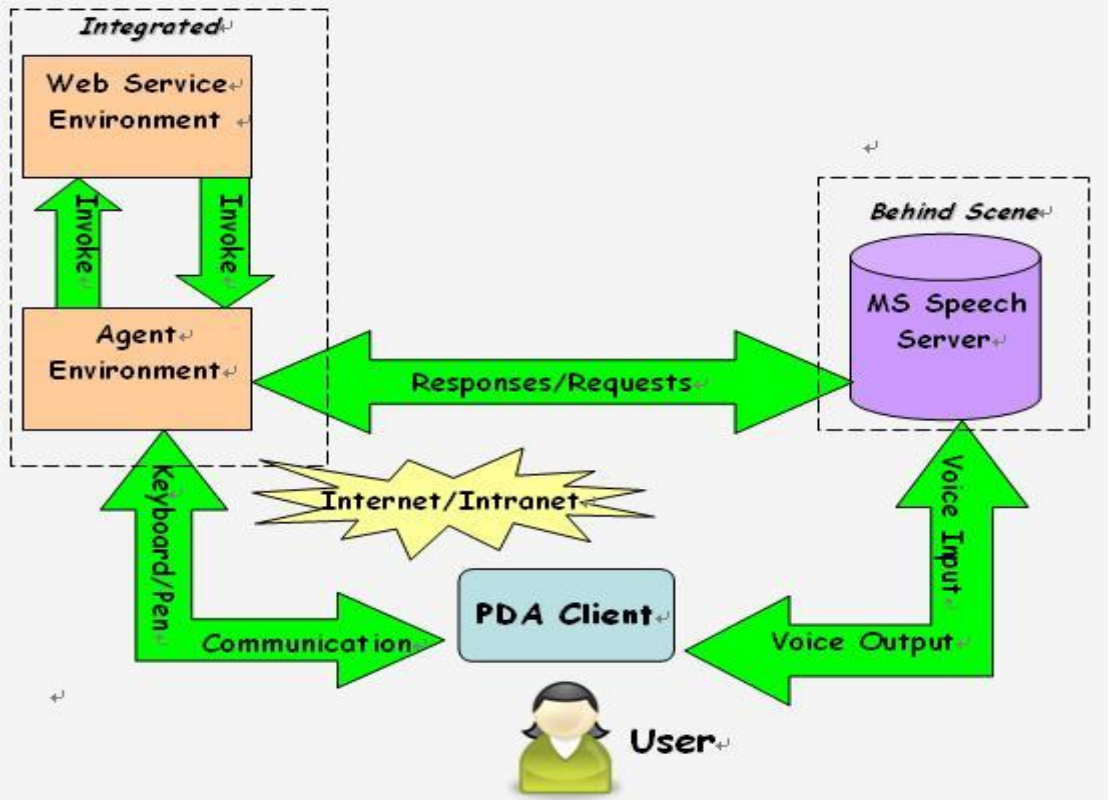


Figure 1

Figure 1 is a Functional Structure of the Multi-Modal Agent System which includes a PDA device, speech server, and the integration of web service and software agent. In general, a user holds a PDA client for sending and receiving information using multi-modal interface. If the user chooses a keyboard/pen to communicate with the web service which is integrated with agent technology, the connection will be built directly between integrated web service and PDA browser through the Internet or Intranet. If the user chooses voice input/output to communicate with the integrated web service, the connection will go through to speech server firstly in order to process voice commands through the Intranet. Then the speech server (behind scene) will request and response with integrated web service. Once the connection is built with the integrated web service, internally the agent technology will automatically search sources such as the Web for updating or downloading new information such as calendar and email information, weather and traffic reports, entertainment listings, information about friends and colleagues, and finally represent the information in an alternative way.

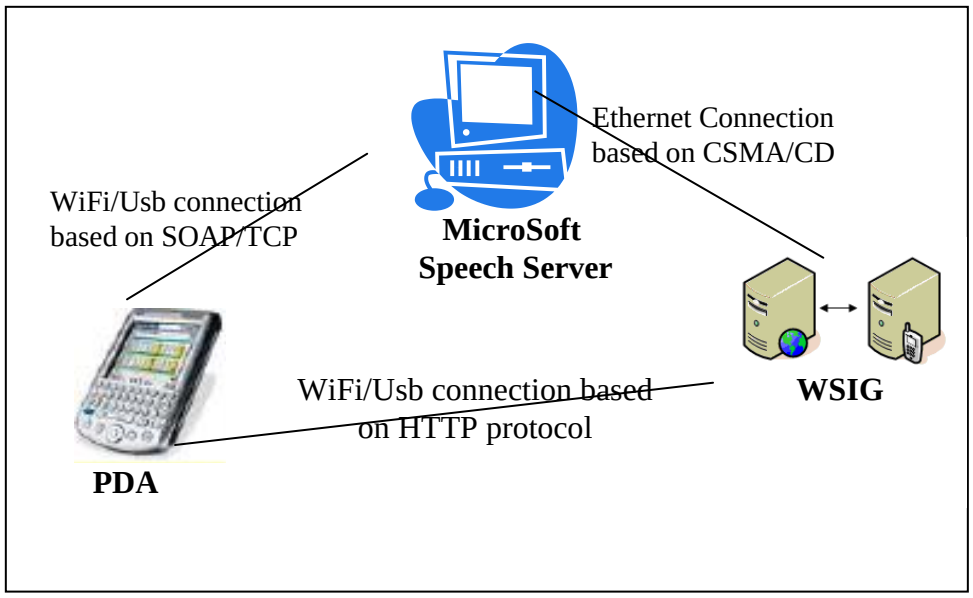


Figure 2

Figure 2 above is a general architecture of the multimodal agent system, there are basic three parts, and each part is specified as follows:

PDA:

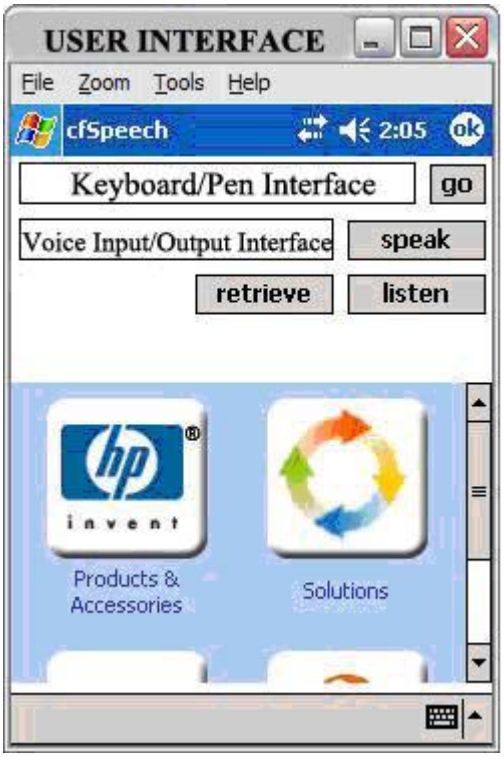


Figure 3

This is a PDA user interface which enables a multimodal interaction; a user will be permitted to do speech recognition and text-to-speech, also allowed to use the normal stylus and internal keyboard. In this browser, its major function is to perform the features of keyboard/pen and voice command. From the view of the picture, the User Interface contains three internal sections; top one is the function bar including file, zoom, tools and help functions, middle one is control panel for managing the features of keyboard/pen and voice command, and the last one is display part for showing any sort of information the users require. Meanwhile the very important section is the middle section which carries out the foremost function of this user interface, the very first text area on the top with little small button “go”, it is the keyboard/pen interface for the users typing keyboard or using stylus to enter information. The second text area with gray button ”speak” implements text-to-speech, that means the PDA talks to people in here if there is any information displayed and the “speak” button is activated. Eventually the last three buttons, the “listen” button does the feature for receiving commands from users, and the”retrieve” button does to give users the second chance for speaking again.

Speech Server

To coverts voice signals to text and vice versa converts text to audio speech; in this system it does the feature to provide a multimodal interaction between a user and a PDA for implementing a multi modal interface.

Web Service Integration Gateway

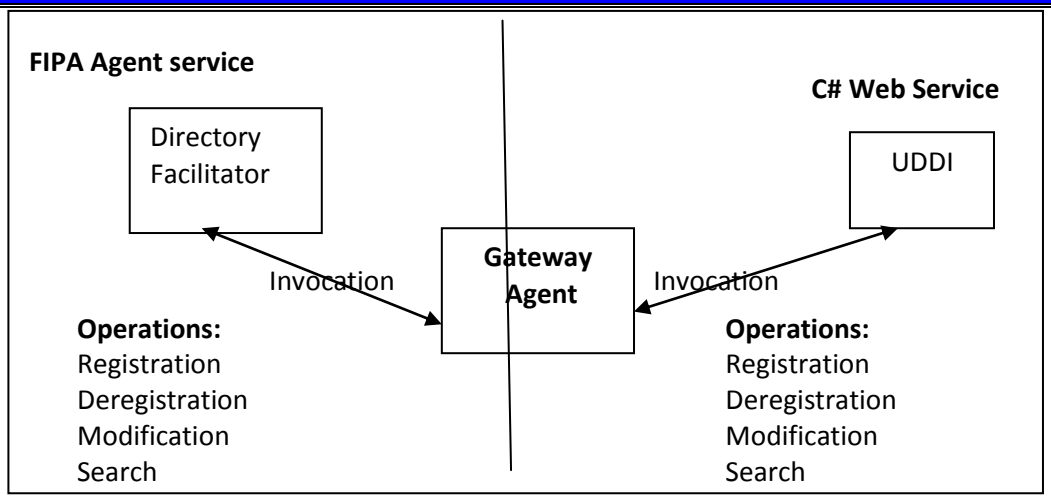


Figure 4

Figure 4 is the integration of web service and JADE agent using the standards of FIPA, gateway agent is a core in this architecture which does the invocation between UDDI and DF, and also does the operations of registration, deregistration, modification and search in order to provide the bi-directional requests and responses from UDDI and DF.

9. **What do you think the user interface in figure 3? and how could it be improved?**
Answer:

10. **What do you think the entire architecture of the multimodal agent from a technical perspective in figure 2? and why?**
Answer:

11. **Are there any potential pitfalls to this approach? and how may they be avoided? (Problems during implementing, testing, etc.)**
Answer:

12. **In what ways would extend the functionality proposed?**
Answer:

BIBLIOGRAPHY

[earthmate 2007] earthmat http://www.earthmate.com/ last accessed at July 2007

[about.com 2007] about.com "PDA Hardware Reviews." http://palmtops.about.com/od/pdahardwarereviews/PDA_Hardware_Reviews.htm 2007

[US patent 2004] US Patent. "Personal digital assistant system" http://www.patentstorm.us/patents/6810450-description.html October 26, 2004

[Padilla 2007] Alfredo Padilla "Gartner Releases Global PDA Report" May 23, 2007. http://www.wirelessinfo.com/content/Gartner-Releases-Global-PDA-Report.htm last accessed on July 2007

[Mahd 2006] F Mahd. "Top 5 features for a PDA and my feelings as to why" http://www.phonedog.com/cell-phone-research/blog/top-5-features-for-a-pda-and-my-feelings-as-to-why.aspx October 18, 2006

[About.com 2006] About.com. "Using a PDA to Simplify Your Life." http://add.about.com/od/adultadd/a/PDAs.htm July 11, 2006

[Windows embedded community 2005] Windows embedded community. "PDA market robust, Windows CE tops, analyst finds." http://www.windowsfordevices.com/news/NS8639381498.html May 04, 2005

[Commentary 1993] G Commentary. "Newton MessagePad Models: original (OMP) and 100 (MP100)." http://www.msu.edu/~luckie/gallery/mp100.htm Released 8/93

[MobileTechReview.com] MobileTechReview.com "WiFi (802.11b/g) Networking for your PDA and Smartphone." http://www.mobiletechreview.com/tips/wifi.htm

[shroff, Vanessa 2007] Vanessa, Sunil Shroff "History of PDA" Apr 22 2007 http://www.medindia.net/pda/pda_history.htm last accessed on July 2007

[Flaminio 2000] Michael Flaminio "PDA to GPS:Using a Handheld Computer as a Navigation Tool" 27th/Nov/00 http://www.insanely-great.com/reviews/pdagps.html last accessed on July 2007

[Kairer 2007] Ryan Kairer "Palm Pilot 1000 Retrospective" 13 Jul 2007 http://www.palminfocenter.com/news/8493/pilot-1000-retrospective/ last accessed July 2007

[Kirvin 2001] J Kirvin "Dance With The Girl That Brung Ya." http://www.writingonyourpalm.net/column011112.htm 12 November 2001

[Palm.com 1999] Palm.com "Press Release Archive: 1999" http://www.palm.com/us/company/pr/archive_1999.html

[Koblentz 2005] Evan Koblentz "THE EVOLUTION OF THE PDA" May 2005 http://www.snarc.net/pda/pda-treatise.htm last accessed July 2007

[Killick 2007] Morgan Killick "A Day in the Life of a PDA User" Jul 2007 http://www.icthubknowledgebase.org.uk/pdadiary last accessed July 2007

[Citrix.com 2007] Citrix.com "How to Configure ActiveSync for PDA Synchronization." http://support.citrix.com/article/CTX821115 Aug 10, 2007

[Citrix.com 2005] Citrix.com "How to get Palm's Hotsync Manager to Work with Published Applications." http://support.citrix.com/article/CTX103125 Jan 21, 2005

[Gilliland & Roach 2006] C Gilliland .S Roach "How to sync ACT! with any PDA device." Power of 3 Consulting. www.po3inc.com. November 2006

[Cecil 2005] P Cecil "Palm Tungsten C – Customizing" http://www.permuted.org.uk/mypda2.html 5/02/2005

[Sulaiman 2005] H Sulaiman "Highlights: Knowing the differences in PDA operating systems." http://www.bytebot.net/media/pda-os-01032005.html 2nd/Mar 2005

[Operating-System.org 2004] Operating-System.org. "Palm Computing." http://www.operating-system.org/betriebssystem/_english/bs-palmos.htm 2004-04-03

[Tuxmobil.org 2006] Tuxmobil.org. "YOPY Linux PDA."http://tuxmobil.org/pda_yopy.html Dec 23 2006

[Fox, Felkey 2003] Fox BI, Felkey BG. "PDA interface: Expanding your hardware capabilities." Hosp Pharm. 2003 Jan;38(1):90-2, 96

[Pdaonlinecenter.com 2006] Pdaonlinecenter.com. "PDA Memory Cards - A Cheap And Effective Way To Expand Your PDA's Abilities." http://www.pdaonlinecenter.com/PDA-Memory-Cards.html Copyright © 2006

[Geiger 2006] G Geiger. "Using the Touch Screen as a Controller for Portable Computer Music Instruments." 2006 International Conference on New Interfaces for Musical Expression (NIME06), Paris, France

[Carroll, Saluja, Tarczy-Hornoch 2002] Carroll AE, Saluja S, Tarczy-Hornoch P. The implementation of a Personal Digital Assistant (PDA) based patient record and charting system: lessons learned. Proc AMIA Symp. 2002::111-5

[Haskin 2007] D. Haskin "Hey, Apple: Remember the Newton before releasing iPhone." http://www.computerworld.com/blogs/node/5066 02/26/2007

[Brown, Brown 2001] B.Brown M.Brown "Quick PDA Data Exchange." http://www.pcmag.com/article2/0,4149,1962,00.asp 05.22.01

[Carroll 2002] B, T Carroll "A Case for PDA Student Computing." http://www.electronetwork.org/articles/psc/ 6 Nov 2002

[Brian 2005] Brian "Dell Axim X50v - User Reviews." http://www.brighthand.com/?newsID=2479 April 09, 2005

[Luo, Hales & Servis 2002] Luo J, Hales RE, Servis M, et al. Clinical computing: use of personal digital assistants in consultation psychiatry. Psychiatric Serv. 2002 Mar;53(3):271-2, 279

[MedicWeb 2007] MedicWeb, Inc http://www.medicalweb.com/ ©2007

[Zhang, Wu, Chen & Sun 2004] F Z Zhang; H Wu; X W Chen; M Sun; YH Luo; "Spatial information integration modes and its key technologies." Geoscience and Remote Sensing Symposium, 2004. IGARSS '04. Proceedings. 2004 IEEE International Volume 4, 20-24 Sept. 2004 Page(s):2571 - 2574 vol.4 Digital Object Identifier 10.1109/IGARSS.2004.1369822

[Bennett & Capella 2002] V Bennett. A Capella. "Location-based services Wherever you are, wherever you go, get the information you want to know." http://www.ibm.com/developerworks/ibm/library/i-lbs/ 01 Mar 2002

[Strauch 2004] J Strauch "PDA virus found in the wild." http://www.computerworld.com/printthis/2004/0,4814,95090,00.html August, 2004

[microsoft.com 2006] microsoft.com. "Voice Command 1.6." http://www.microsoft.com/windowsmobile/voicecommand/default.mspx Nov 2006

[Chignell & Lottridge 2006] Chignell Mark. Lottridge Danielle. 2006. "Social impacts of handheld computer information retrieval during physician-patient communication" Human Factors of Telecommunications Conference (HFT).

[Hu, Yeh & Chung 2005] Wen-chen Hu, Jyh-haw Yeh and Chung-wei Lee, "Handheld Devices and Computing and Payment Methods for Mobile Commerce,"

Proceedings of the 38th Annual Midwest Instruction and Computing Symposium, Eau Claire, Wisconsin, April, 2005

[Story 2000]D Story. "All Linux PDA: Fact or Fiction?" http://www.oreillynet.com/pub/a/network/2000/05/05/magazine/LinuxPDA.html 05/05/2000

[Jhoney, Kuchhal & Jaiswal 2003] Albee Jhoney, Manu Kuchhal, Rohit Jaiswal . "Pervasive computing: e-business anywhere, anytime". 01 Jul 2003. http://www.ibm.com/developerworks/wireless/library/wi-pvcapps/ last access at July/2007

[Davies, Marcella & McGrenere 2004] Davies, R., Marcella, S., McGrenere, J., and Purves, B.The ethnographically informed participatory design of a PDA application to support communication. Proc.SIGACCESS. ACM Press (2004), 153-160.

[W3C 2002] W3C. "Web Services Activity." http://www.w3.org/2002/ws/ 2002

[Padilla 2007] Alfredo Padilla "Gartner Releases Global PDA Report" May 23, 2007. http://www.wirelessinfo.com/content/Gartner-Releases-Global-PDA-Report.htm last accessed on July 2007

[shroff, Vanessa 2007] Vanessa, Sunil Shroff "History of PDA" Apr 22 2007 http://www.medindia.net/pda/pda_history.htm last accessed on July 2007

[Flaminio 2000] Michael Flaminio "PDA to GPS:Using a Handheld Computer as a Navigation Tool" 27th/Nov/00 http://www.insanely-great.com/reviews/pdagps.html last accessed on July 2007

[Maes, 1994] P. Maes. 1994. Agents that reduce work and information overload. Comms. of the ACM 37 (7) pp 31-40.

[orafaq.com 2005] orafaq.com "What is the Oracle Intelligent Agent." http://www.orafaq.com/faq/what_is_the_oracle_intelligent_agent 2005-11-10

[Jennings 1999] N. R. Jennings (1999) "Agent-Oriented Software Engineering" Proc. 12th Int Conf on Industrial and Engineering Applications of AI, Cairo, Egypt, 4-10. (Invited paper) [Also appearing in Proc. 9th European Workshop on Modelling Autonomous Agents in a Multi-Agent World (MAAMAW-99), Valencia, Spain 1-7 (Invited paper)]

[Kogler 2003] M Kogler Simulation and Visualization of Agents in 3D Environments Diploma Thesis February 26, 2003

[Norman, Preece, Chalmers, Jennings, Luck, Dang, Nguyen, Deora, Shao, Gray, Fiddian 2004] T. J. Norman, A. Preece, S. Chalmers, N. R. Jennings, M. Luck, V. D. Dang, T. D. Nguyen, V. Deora, J. Shao, W. A. Gray, and N. J. Fiddian. "Agent-based formation of virtual organisations." Knowledge-Based Systems, 17:103–111, 2004.

[Mckean Shorter Luck McBurney Willmott 2007] J. Mckean, H. Shorter, M. Luck, P. McBurney and S. Willmott. "Technology Diffusion: the Case of Agent Technologies." University of Liverpool Department of Computer Science ULCS-07-014: August 7, 2007

[FIPA Security SIG 2002] FIPA Security SIG. "FIPA Security SIG Request For Information." http://www.fipa.org/docs/output/f-out-00065/f-out-00065.html 2001/02/02

[Zhang, Karmouch, Impey 2001] M. Zhang, A. Karmouch, R. Impey "Towards a Secure Agent Platform Based on FIPA." Mobile Agents for Telecommunication Applications: Third International Workshop, MATA 2001, Montreal, Canada, August 14-16, 2001. Proceedings

[Lange 1998] D. B. Lange. "Present and Future Trends of Mobile Agent Technology." Lecture Notes In Computer Science; Vol. 1477. Proceedings of the Second International Workshop on Mobile Agents. Year of Publication: 1998 ISBN:3-540-64959-X

[Niklfeld, Anegg, Gassner 2005] G. Niklfeld, H. Anegg and A. Gassner, "Device independent mobile multimodal user interfaces with the MONA multimodal Presentation Server," Eurescom Summit 2005 Ubiquitous Services and Applications Exploiting the Potential, Heidelberg, Germany, April 2005.

[Vallés, Manso, Arredondo Pozo 1996] M. Vallés, F. Manso, M. T. Arredondo, and F. Del Pozo, "Multimodal envoronmental control system for elderly and disabled people.," presented at 18th Annual Conference of the IEEE Engineering in Medicine and Biology Society., Amsterdam, 1996.

[Barnett 2005] J. Barnett "Multimodal Architecture and Interfaces." W3C Working Draft http://www.w3.org/TR/2005/WD-mmi-arch-20050422/ 22 April 2005

[Orchard, Newcomer, Ferris, Champion 2002] D.Orchard, E.Newcomer, C.Ferris, and M.Champion. "Web Services Architecture." http://www.w3.org/TR/2002/WD-ws-arch-20021114/ 14 November 2002

[Salchner 2005] K Salchner. ".NET and XML: XSD Schemas." http://www.developer.com/xml/article.php/3388311 2005

[Box, Ehnebuske, Kakivaya, Layman, Mendelsohn, Nielsen, Thatte, Winer 2000] D.

Box, D. Ehnebuske, G. Kakivaya, A. Layman, N. Mendelsohn, H. F. Nielsen, S. Thatte, and D. Winer "Simple Object Access Protocol (SOAP) 1.1." http://www.w3.org/TR/2000/NOTE-SOAP-20000508/ W3C Note 08 May 2000

[Christensen, Curbera, Meredith, Weerawarana 2001] E. Christensen, F. Curbera, G. Meredith and S. Weerawarana,."Web Services Description Language (WSDL) 1.1."http://www.w3.org/TR/wsdl W3C Note 15 March 2001

[Jewell, Chappell 2002] T. Jewell, D. Chappell Chapter 6 UDDI: Universal Description, Discovery,and Integration." ISBN: 0596002696 March 2002

[Jennings & Wooldridge, 1996] N. Jennings, M. Wooldridge. Software Agents. IEE Review, January 1996, pp17-2

[Bradshaw 1997] Jeffrey M. Bradshaw (Editor), "Software Agents",AAAI Press/The MIT Press, April 18 1997

[Marık & McFarlane 2005] Marık, V. - McFarlane, D.: Industrial Adoption of Agent-Based Technologies. IEEE Intelligent Systems. 2005, vol. 20, no. 1, s. 27-35. ISSN 1094-7167.

[Jiang 2000] Jiang B. (2000), "Agent-based approach to modelling urban and environmental systems within GIS". Proceedings of 9th International Symposium on Spatial Data Handling, August 10-12, Beijing

[Jansen & Karygianis 2000] W. Jansen & T. Karygiannis, NIST special publication 800-19:Mobile agent security, NIST, Gaithersburg, MD, 2000.

[Simon & Jank 2004] R. Simon, M. Jank, and F. Wegscheider, "A Generic UIML Vocabulary for Device- and Modality Independent User Interfaces",in Proc. of the 13th International World Wide Web Conference, New York, USA, May 17-22, 2004, ACM Press.

[Stormer & Knorr 1998] Stormer H. & Knorr K. (1998). PDA- and Agent-based Execution of Workflow Tasks, Proc. Informatik 2001 Conference, Vieanna, Austria.

[Chiu, Cheung, Leung 2005] Chiu, D.K.W.; Cheung, S.C.; Ho-fung Leung "A Multi-Agent Infrastructure for Mobile Workforce Management in a Service Oriented Enterprise" System Sciences, 2005. HICSS apos;05. Proceedings of the 38th Annual Hawaii International Conference on Volume , Issue , 03-06 Jan. 2005

[Cheyer & Julia 1998] Cheyer, A. and Julia, L. (1998) Multimodal Maps: An agent-based approach. In: H. Bunt, Beun, Bor-ghuis (Eds) Multimodal Human-computer commu-nication, Springer Verlag, pp. 111-121.

[Ronzhin & Karpov 2005] A.L. Ronzhin, A.A. Karpov. Assistive multimodal system based on speech recognition and head tracking. In Proc. of 13-th European Signal Processing Conference EUSIPCO-2005, Antalya, Turkey, 2005.

[Gottschalk, Graham, Kreger and Snell 2002] K. Gottschalk, S. Graham, H. Kreger, and J. Snell "Introduction to Web services architecture" Number 2, 2002.

[Edgar & Baruah 2002] Gerald Edgar & Pranab Baruah "Web Services Security Issues" March 5, 2002

[Agentcities Web services Working Group 2003]. Agentcities Web services Working Group. "Integrating Web services into Agentcities Technical Recommendation". *28 November, 2003* http://www.agentcities.org/rec/00006/

[Nguyen and Kowalczyk 2005] Xuan Thang Nguyen, Ryszard Kowalczyk: "WS2JADE: Integrating Web Service with Jade Agents". 30 July 2005 Technical Report: SUTICT-TR2005.03

[Shen, Norrie, Barthès 2000] W. Shen, D.H. Norrie, and J.-P.A. Barthès, Multi-Agent Systems for Concurrent Intelligent Design and Manufacturing, Taylor & Francis, 2000.

[ACL specification 2002] Foundation for Intelligent Physical Agents. Acl message structure specification, December 2002. http://www.fipa.org/specs/fipa00061/

[Nichol 1997] S. Nichol. "Distinguishing the Mulitple Meanings of Thin Client and Fat Client." http://scottnichol.com/thincli2.htm 27-Feb-97

[Holdsworth,Venkatapathy 2002] S.Holdsworth,C.Venkatapathy."An introduction to Web Services Gateway." http://www.ibm.com/developerworks/webservices/library/ws-gateway/ 01 May 2002

[Song 2007]X. Song "Comparing Microsoft Speech Server 2004 and IBM WebSphere Voice Server V4.2"http://www.developer.com/java/other/article.php/3381851Oct2007

[Vo, Wood 1996] Vo, M. T. and Wood, C. (1996). Building an Application Framework for Speech and Pen Input Integration in Multimodal Learning Interfaces. In Proceedings of IEEE International Conference of Acoustic, Speech and Signal Processing, volume 6, pages 3545–3548, Atlanta, USA.

[multimodal speech group 2001] multimodal speech technology group "MULTIMODAL SPEECH SYNTHESIS" http://www.speech.kth.se/multimodal/ September 24, 2001

[Larson 2003] J, A. Larson Commonsense Guidelines for Developing Multimodal User Interfaces http://www.larson-tech.com/MMGuide.html April 3, 2003

[Niles 2007] R.Niles "Data Analysis" http://www.robertniles.com/stats/dataanly.shtml

[W3C 2004] W3C Web Service Architecture Working Group. "Web Service Architecture Recommendation." http://www.w3.org/TR/2004/NOTE-ws-arch-20040211/ last accessed 26th/Sep/2007

[JADE Board 2007] Jade Board, "JADE WSIG Add-On Guide JADE Web Services Integration Gateway (WSIG) Guide", last update 18-June-2007 http://jade.tilab.com/doc/tutorials/WSIG_Guide.pdf

[JADE Board 2006] Jade Board, "JADE WSIG Add-On Guide JADE Web Services Integration Gateway (WSIG) Guide", last update 15-November-2006. JADE 3.4.1 http://sharon.cselt.it/projects/jade/doc/tutorials/JADE%20WSIG%20Guide.pdf

[JADE Board 2005] Jade Board, "JADE WSIG Add-On Guide JADE Web Services Integration Gateway (WSIG) Guide", last update 03-March-2005. JADE 3.3 http://jade.cselt.it/doc/tutorials/JADE_WSIG_Guide.pdf

[Greenwood, Nagy, Calisti 2005] D. Greenwood, J. Nagy, M. Calisti "Semantic Enhancement of a Web Service Integration Gateway". 2005 http://www.whitestein.com/resources/papers/SOCABE05.pdf

[Greenwood 2005] D. Greenwood, WHITESTEIN Technologies, "JADE Web Service Integration Gateway (WSIG)", JADE Tutorial, AAMAS 2005, http://jade.tilab.com/papers/2005/JADEWorkshopAAMAS/AAMAS05_JADE-

Tutorial_WSIG-Slides.pdf

[Bellifemine, Caire, Greenwood 2007] F, Bellifemine. G, Caire .D, Greenwood. Agent Technology. "Developing Multi-Agent Systems with JADE". Wiley. April 20, 2007. ISBN-10: 0470057475

[Greenwood, Calisti 2004] D. Greenwood, M. Calisti. "Engineering Web Service – Agent Integration" in Proceedings of the IEEE Conference on Systems, Man and Cybernetics, The Hague, The Netherlands, 2004.

[W3C 2004] W3C "OWL Web Ontology Language Reference". W3C Recommendation 10 February 2004 http://www.w3.org/TR/owl-ref/

[Shafiql, Ali, Suguri, Ahmad 2006] M. Omair Shafiq1, A. Ali, H. Suguri, H. Farooq Ahmad. (Presented by Hiroki Suguri). "AgentWeb Gateway integrating FIPA agents and Web Services". http://www.semanticgrid.org/OGF/ggf16/slides/ 19-Feb-2006

[Shafiql, Ali, Suguri, Ahmad 2006] M. Omair Shafiq1, A. Ali, H. Suguri, H. Farooq Ahmad. "AgentWeb Gateway integration of FIPA Multi Agent System and W3C Web Service System". http://www.semanticgrid.org/OGF/ggf16/slides/ 19-Feb-2006

[Slatforum 2007] SaltForum. "SALT Forum" http://www.saltforum.org. last accessed 1st/Oct/2007

[IBM 2001] IBM "IBM WebSphere Voice Server gives voice to e-business applications". ftp://ftp.software.ibm.com/software/pervasive/info/products/WSVS.pdf Printed in the United States of America 11-2001

[Chesnut 2005] C. Chesnut. An Introduction to Speech Server For voice-only and multimodal applications. http://dotnet.sys-con.com/read/84454.htm. May.11.2005

[Microsoft 2005] "Microsoft Speech Server Architecture Overview"

http://technet.microsoft.com/en-us/library/bb684748.aspx June 20, 2005

[W3C 2002] W3C. "Speech Synthesis Markup Language Specification ".http://www.w3.org/TR/2002/WD-speech-synthesis-20020405/ W3C Working Draft 5 April 2002

[Xxxcyy 2007] Xxxcyy. "The study of TTS in Microsoft Speech SDK" http://blog.csdn.net/xxxcyy/archive/2007/07/14/1690535.aspx 14/July/2007

[Microsoft.com 2007] Microsoft.com. "How To Configure and Use Text-to-Speech in Windows XP." http://support.microsoft.com/kb/306902 May 7, 2007

[Casey 2004] Casey. "Multimodal Speech on a Pocket PC". http://www.mperfect.net/speechMulti/ 8/1/2004

[Microsoft 2005] Microsoft "Speech Server 2004: Processing Speech Recognition". http://technet.microsoft.com/en-us/library/bb684754.aspx June 20, 2005

[Microsoft 2005] Microsoft "Speech Server 2004: Processing Speech Output". http://technet.microsoft.com/en-us/library/bb684753.aspx June 20, 2005

[Microsoft 2005] Microsoft "ASP.NET Distributed Speech Application Scenarios". http://msdn2.microsoft.com/en-us/library/ms869987.aspx June 20, 2005

[Yu 2005] liyang yu "Understanding UDDI's tModel". http://www.codeproject.com/soap/understandingTModels.asp 11 Nov 2005