**KTH Computer Science
and Communication**

Lorem ipsum dolor sit amet, sed diam nonummy nibh eui mod tincidunt ut laoreet dol

Duis autem vel eum iruire dolor in hendrerit in vulputate velit esse molestie consequat, vel illum dolore eu feugiat null

ANGELINA VON GEGERFELT
KASHMIR KLINGESTEDT

Abstract

This is a skeleton for KTH theses. More documentation regarding the KTH thesis class file can be found in the package documentation.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Mauris purus. Fusce tempor. Nulla facilisi. Sed at turpis. Phasellus eu ipsum. Nam porttitor laoreet nulla. Phasellus massa massa, auctor rutrum, vehicula ut, porttitor a, massa. Pellentesque fringilla. Duis nibh risus, venenatis ac, tempor sed, vestibulum at, tellus. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos.

Referat

Titel på rapporten på svenska

Denna fil ger ett avhandlingsskelett. Mer information om LaTeX-mallen finns i dokumentationen till paketet.

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Mauris purus. Fusce tempor. Nulla facilisi. Sed at turpis. Phasellus eu ipsum. Nam porttitor laoreet nulla. Phasellus massa massa, auctor rutrum, vehicula ut, porttitor a, massa. Pellentesque fringilla. Duis nibh risus, venenatis ac, tempor sed, vestibulum at, tellus. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos.

Contents

Chapter 1

Introduction

Natural language is the language that is intuitive and natural to the user. The aim of this project is to use gamification in order to acquire better knowledge of how to build a good Natural Language Interface (NLI), which is a field in human-computer interaction. We have implemented a game with different implementations of NLIs, one is a typing based interface where the user give commands using a keyboard and the other is a speech based interface where the user give commands by speaking into a microphone. The game is inspired by other text based adventure-games such as Zork. We are going to compare the different NLI implementations to find which provides the best user experience.

1.1 Problem Statement

Which Natural Language Interface is most user friendly, one using text input or one using speech input?

1.2 Scope

The area of natural languages include various types of communication, such as gestures and touch, but in this project the focus has solely been on text and speech. There also exist different kinds of NLIs for both of these natural languages. For example a text based NLI could use strictly grammatical processing and another could be less strict in this matter. We will however focus on one kind of NLI for both text and speech, since the main goal is to compare the usability of these two input methods.

1.3 Purpose

Natural language is flexible and people need little to no training in how to use it. If it can be successfully translated to a system, it would be possible to create interfaces readily usable by a wider array of people and with a lower entry-level for adaptors. The goal of this project is therefore to gather information about the difference in usability between speech and text based NLIs in order to be able to use the results to build more user friendly NLIs in the future.

1.4 Disposition

1.5 Terminology

Word	Deffinition
Natural Language	Any language that develops naturally in humans through use and repetition
Natural Language Interface	A way for the user to interact with a system or program by the use of human natural language
Natural Language Processing	Derives meaning from natural language input and converts it into something the computer can understand and vice versa.
SUS	System Usability Scale

Table 1.1. Caption for the table.

Chapter 2

Background

2.1 Natural Language

A natural language is any language that develops naturally in humans through use and repetition without any conscious planning or premeditation of their own. These are the languages human beings use to communicate with each other, whether by speech, signing, touch or writing. They are distinguished from constructed and formal languages such as those used to program computers or to study logic.

2.2 Natural Language Interface

A Natural Language Interface (NLI) is a way for the user to interact with a system or program in a more natural and intuitive way. NLI have many advantages over other systems, such as that it is flexible, people need little training to use it and it can be allowed to do multiple things at once due to uses of pronouns, quantification and context. In general , NLI's primary function is that they support and deal with the user's view of the system and translates it into those actually used by the system (Hendrix, 1982). A few examples of working consumer NLI are Wolfram Alpha, Siri and Google Search.

2.3 Natural Language Processing

Natural Language Processing (NLP) explores how computers can be used to understand and manipulate natural language (Chowdhury, 2003). The input might be text, speech or other. NLP can be used for translation into another

language, to comprehend and represent the content of text, to build/search a database or to maintain a dialogue with a user as part of an interface for database/information retrieval. (Allen, 2003) NLP can be seen as a backend to an NLI.

2.4 Speech Recognition

Speech Recognition (SR) systems have been researched and developed as a worldwide activity because of the potential this brings for applications, such as to have voice-interactive management, voice dictation and spoken language translation. Although many successes has been had in the development of practical and useful SR systems, there are still limitations to what can be done. The speech signal is one of the most complex signals that humans try to work with. There is also the fact that human's vocal systems differ between humans and things can be said in different ways. However, various SR systems have been integrated into consumer-technology today (Siri and Google Now, for example), so we are still making advances with the technology. (Lee et al., 1996)

2.5 Zork and Text-Based Adventure Games

The first version of Zork was written in 1977–1979 and it is one of the earliest interactive fiction computer games. Zork distinguished itself in its genre as an especially rich game, in terms of both the quality of the storytelling and the sophistication of its text parser, which was not limited to simple verb-noun commands (e.g. "hit troll"), but recognized some prepositions and conjunctions (e.g. "hit the troll with the Elvish sword"). (Webpage-name, 1997)

2.6 Previous Research

Chapter 3

Method

3.1 The Game

The game is inspired by the previously mentioned Zork. It consists of a few rooms and tasks to be performed before reaching a victory scenario. The game differs in environments and plot between the different implementations of NLIs, which requires the user to input different commands in order to win. One version uses typing to control your character's actions and the other uses speech. The reason we decided to create two implementations is so that a user who has played one control-scheme could still play the other without having the benefit of knowing what is required to win.

3.1.1 Plot

You play as a bunny that has escaped its cage and is on the hunt for food. You need to eat three crackers in each game in order to ease your hunger and win the game. In the speech version you are a house-pet and are in an apartment and has the ability to be in the kitchen, the livingroom and the bedroom. In the text version you are a class-pet in a school and can visit the classroom, the hallway and the cafeteria.

3.2 Evaluation

"[...] the usability of any tool or system has to be viewed in terms of the context in which it is used, and its appropriateness to that context" (John Brooks, 1996). In 1986 John Brooks created a way to test a user interface and its usability. The idea is that a user tries out a system after which they answer a specific questionnaire, developed by Brooks, concerning the usability

of the system. The user should not think for a long period of time or discuss their opinion with anyone before or while filling out the questionnaire. It is important that the user's initial thought and own experience is recorded.

The questionnaire consists of 10 statements and the user must rank each statement by a scale of 1-5, where 1 is "strongly disagree" and 5 is "strongly agree". Some examples of the statements are "I thought the system was easy to use" and "I thought there was too much inconsistency in this system". All the odd questions are "positive" while all the even questions are "negative" in nature. In order to make them work together, we convert the points as follows:

$$f_i = \begin{cases} 5 - Q_i, & \text{when i is even} \\ Q_i - 1, & \text{when i is odd} \end{cases}$$

Where Q_i is the answer to question numbered i. The total point is then:

$$2.5 * \sum_{i=1}^{10} f_i$$

The score lies in the range 0-100. A high score means that the system is easy to use and liked by the users, while a low score means that it should be improved before publishing.

3.2.1 Testing

Users tested both versions of the game in order to compare them. First we explain to them the basics of a text-adventure: there is a description of the room you are in, you type or say simple commands in natural language in order to do things and sometimes must interact with your surroundings in order to progress.

The user gets to play one of the game-versions and upon completion they will answer a short form that will be expanded upon in the following section: "System Usability Scale"(SECTION x.y). We record various data, such as amount of commands, completion time and if there were anything specific that the user failed on multiple times. The last check is in case we require the users to say a certain word or phrase at one part but the game has a hard time parsing the command for each user. This is done in an attempt to filter out bad programming and game design on our part.

After the user played one version and filled in the questionnaire the user then played the other version and filled in the questionnaire again but with regard to the new control scheme. We repeated this test with multiple users and we alternated between which game version was to be played first.

Chapter 4

Results

4.1 Effectiveness

4.1.1 Time Requirement

4.1.2 Effective Commands

4.2 Ease of use

User level Ease to learn User requirements User skills bloblbobnklafnmdsdkma

4.2.1 English Confidence

4.3 Satisfaction (SUS)

Speech	Text
54.17	78.06

Table 4.1. Caption for the table.

satisfaction

Chapter 5

Discussion

5.1 Known Problems

5.2 Future Research

Chapter 6

Conclusion

Bibliography

Allen, J. F. (2003). *Natural language processing.* Published in Encyclopedia of Computer Science.

Chowdhury, G. G. (2003). *Natural language processing.* American Society for Information Science and Technology.

Hendrix, G. G. (1982). *Natural-Language Interface.* SRI International, California.

Lee, C.-H., Soong, F. K., and Paliwal, K. K. (1996). *Automatic Speech and Speaker Recognition, Advanced Topics.* Kluwer Academic Publishers.

Webpage-name (1997). *Internet Archive: Wayback Machine.*

Appendix A

RDF

And here is a figure

Figure A.1. Several statements describing the same resource.

that we refer to here: A.1