

AVARI – Voice Recognition

Morris LeBlanc, Vicky Fowler, Toni Bloodworth, Lauren Cairco, Louis Fletcher, Dr. Larry F. Hodges, Dr. Dale-Marie Wilson, and Amy Ulinski
{msleblan, vdfowler, lfhodges, dwilso1, aculinsk}@mail.uncc.edu
{bloodworth2, cairco2}@winthrop.edu and lfletcher2@elon.edu



Morris LeBlanc,
UNCC
Vicky Fowler,
UNCC
Toni Bloodworth,
Winthrop
Lauren Cairco,
Winthrop
Louis Fletcher,
Elon

Introduction

During this REU program we designed an animated agent named AVARI, which stands for **A**nimated **V**irtual **A**gent **R**etrieving **I**nformation. My portion of the project consisted of the voice recognition.

Most humans interact with one another in natural conversation, but what if you could apply this concept to a human and computers interaction. So we took this idea and applied it to AVARI.

Background

I noticed that other types of animated agents have been created that use speech to interact with the user. Here at UNCC we have quite a few, from Charlotte, Diane, and Marve.

The real big thing that sets AVARI apart from them is that she uses a database and also SALT, which is uses grammars and tags to interact with the user.

In order to create the natural conversation, it was necessary to learn what spoken language really consists of.

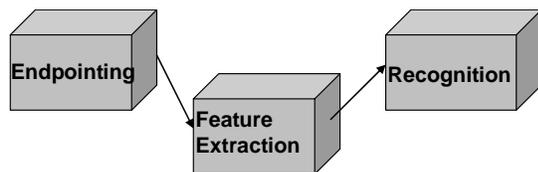


Figure 1: Elements of Spoken Language

- **Endpointing:** Detects the beginning and end of speech.
- **Feature Extraction:** Transforms endpoint utterance into sequence of feature vectors
 - Feature vector – list of numbers that represent measurable characteristics of speech.
- **Recognition:** Determines spoken words using feature vectors.

All these elements help create what is known today as voice recognition.



Figure 2: AVARI

Research

- **Voice Recognition**
 - In order to make the idea of natural language work, we decided to use SALT with JavaScript for AVARI to work.
 - JavaScript allows SALT to be used as a component.
 - Tags and grammar files were used so words could be recognized and also certain things can happen when the user may be silent, say something that is not part of the grammar file.

- While creating the grammar file for AVARI, I had to be aware of using words twice and words that could sound similar.
- I also had problems with some of the last names that had to be changed to sound the same as a human would say it.
- The issue came about because these words are compared to words in the English dictionary.



Figure 3: Working on AVARI's desk (just starting)

Impact

- AVARI is able to recognize any voice and is not uniquely trained on just on voice.
- She also has very good recognition accuracy of what someone is saying.
- AVARI is now capable of having a natural conversation with a human.
- Another reason AVARI is so unique is what was stated before with her using a database to answer questions which the user says.

Conclusions

- What I have learned from this project is:
 - I have a better understanding JavaScript, MySQL, and SALT.
 - How to keep a good research journal.
 - How to work with others on a big project.
 - Databases are tedious and not very fun.
 - That I have accomplished a task that has been put in front of me.



Figure 3: AVARI's builders

Future Work

Sometime in the future we plan on placing AVARI in front of the general lab on the 3rd floor in Woodward Hall. We would want to see how people interact with her from beginning to end.

Also, we could create more grammar files to make the accuracy for the recognition much better.