The Making of ANGELA

The technologies that enable a Virtual Medical Clinic Receptionist

By Yilissa Tang

1.1 Virtual Receptionist: The market niche and the expectations

Several years ago, I was a medical mess and had to see doctors of varying specialties. Generally my experiences with medical clinics were positive, however I very much disliked being putting on hold or reaching a voice mail when I called in for small things such as trying to make an appointment.

Having already designed a few speech enabled applications with my previous job, it occurred to me that a speech-enabled virtual medical receptionist could be a win-win service to everyone: it would be a nice help to patients, doctors and staff. When I mentioned the idea to my friends and a few doctors, they not only agreed, but also suggested tasks that they would like a virtual clinic receptionist to perform.

- Be available 24/7,

- Answer calls when staff are away from their desk.

- Help patients make self-service appointments

- Make automated daily reminder calls for to patients for next-day appointments

- Automate after-hours answering service. Forward emergency calls to on-call doctors

- Collect and validate new patient insurance information over the first clinic call.

Deployed recently by **AngelSpeech** Inc, **ANGELA** is the virtual appointment manager and medical clinic receptionist capable of all except the last task listed above.

It is not coincidental that last few years have seen many new speech recognition applications such as **ANGELA** being developed and deployed to market. New technologies that are recently available have made the development and commercialization of these applications possible. In turn, the success of these applications has raised market awareness and market readiness to the new technologies.

To cite the two most critical achievements of speech industry:

1) Although the speech recognizers have been available for many years, it is only recently that their performance and reduced cost have reached maturity for commercialization.

2) **VoiceXML** (Voice Extensible Markup Language) as an industry standard has been finalized and by now widely adopted by the industry. This standard has provided a platform for implementing portable, complex speech applications that connect telephone callers to Internet.

1.2 De-constructing ANGELA: How she speaks, listens, thinks, and acts

ANGELA, as a typical speech application using **VoiceXML**, contains three logical modules as shown below.

ANGELA Software Modules

VoiceXML pages are technically very similar to HTML (web) pages in that both standards are about connecting users to Internet. The difference is that an **HTML** (or web) page interacts with user via text and graphic, while **VoiceXML** interacts with user via voice and audio channeled by telephone. Using **VoiceXML**, a telephone caller communicates with **ANGELA** by listening to an audio output that is either pre-recorded or computer synthesized and submitting request through the caller's natural speaking voice or via telephone keypad.

A typical **VoiceXML** page provides instructions to caller, asks a question, and then "listens" to an answer from the caller. This **VoiceXML** page controls WHEN **ANGELA** will "listen" to a caller, WHAT speech input to collect from the caller, and HOW she will "speak" to the caller.

In **VoiceXML**, the "listening" mechanism is controlled by speech recognition grammar. The so called grammar is a set of rules that specify the words and patterns of words to be listened for by a speech recognizer. For example, to implement **ANGELA**, we need to design grammar for "yes/no", "date and time" inputs. A grammar should in general include many variations of phrases; (i.e. A yes/no grammar should include "yes, yeah, right, correct, no, nope, wrong", etc…)

The controlling logic module, implemented using JSP (Java Server Pages), is the "brain" of **ANGELA**. It makes many decisions based on information available in the backend database as well as the voice requests received from the caller. Following each decision, it generates a new **VoiceXML** page for the next part of the dialog with the caller.

The backend database is where **ANGELA's** "actions" take place. For example, once an appointment is chosen by caller, it will be reserved in a database.

1.3 Humanizing ANGELA: Dialog Design with Personality and Purpose

Dialog design is critical to the success of any speech application and the task is never as obvious as it seems. In designing **ANGELA's** dialog with a patient who needs an appointment, two factors have posed major challenges. First of all, she can not "see" the available appointment slots nor can she help the caller "pick-and-choose" a slot as intelligently as a human receptionist can. Secondly, even if she is able to find many available slots quickly as a computer, she MUST NOT read too many time slots to caller. Doing so would make the dialog unbearably long, boring and ineffective.

In general, a successful dialog design must achieve two goals:

1) To be efficient; in other word to help the caller get the job done as quickly as possible.

1) To make the caller experience as natural and pleasant as possible; for any self-service application, user friendliness is at the heart of quality-of-service.

To develop a efficient voice-search for appointment, we first selected a list of questions that ask a caller's preferences on time and date, then arrange them in such an order that each question will bring the caller much closer to a final time slot selection. By controlling the conversation, and asking the caller questions in an optimized interrogative flow, the duration of the call is minimized.

To put callers at ease with **ANGELA**, we focused our efforts in creating her personality, and to make her sound just like the girl next door. We chose a language style that is slightly casual yet professional and friendly. To add human personality to her, we have programmed her to say things such as "Well, the office did not leave me a phone number for emergency contact…", or, "I am having trouble with my telephone line" (When in fact the speech-recognizer fails to understand the caller!).

Her voice should be sweet and clear. However, this does not necessarily mean she needs the polished vocal quality of a singer. In fact, a polished voice tends to sound formal. A formal voice that lacks distinct characterizations tends to be forgotten quickly. A clear, pleasant voice with slightly distinct accent is what we have chosen for her. In addition, our voice talents behind **ANGELA** have delivered the last touch to help bring **ANGELA** to life; their acting on the voice-over scripts has created the receptionist exactly as we wanted.

1.4 "Hello, this is ANGELA, I am here to help you…"

We are pleased that **ANGELA** has finally arrived to help out everyone around busy, under-staffed medical clinics. Here I emphasis the word "help" as I think it is important to point out that she is not designed to replace medical staff whose primary responsibility is to focus their valuable time on patients care.

ANGELA will greatly increase medical practice efficiency and productivity by doing what a computer program does the best: she performs trivial, repetitive tasks at very low cost. She works 24/7 and she accomplishes her tasks with consistency and reliability.

1.5 About the Author
Mrs. Yilissa Tang is the founder and President of **AngelSpeech** Inc, A Plano, Texas based medical practice software company. (www.angelspeech.com) She can be contacted at: Yilissa@angelspeech.com