Creating a Virtual Neighbor

Carina Corbin, Fabrizio Morbini, and David Traum

Abstract We present the first version of our Virtual Neighbor, who can talk with users about people employed in the same institution. The Virtual Neighbor can discuss information about employees in a medium sized company or institute with users. The system acquires information from three sources: a personnel directory database, public web pages, and through dialogue interaction. Users can interact through face to face spoken dialogue, using components from the ICT Virtual human toolkit, or via a chat interface.

1 Introduction

The Virtual Neighbor project strives to recreate the conversations workplace neighbors often share. When we first start a new job, we often want to learn more about those who work around us, and most people have a neighbor they can turn to for an update on the latest information. In this paper, we present first results in the creation of a Virtual Human that could do the same thing. A virtual neighbor shares much with a personal assistant dialogue system, such as Siri or Cortana, in terms of being able to answer a range of information questions, however the virtual human has a more persistent presence (beyond just answering an isolated query), and is aimed at finding out more about you, as well as just answering your questions.

Miki, the Virtual Human Neighbor, was designed to deal initially with topics related to co-worker information. She can answer questions that range from where someone is located to the last project he or she worked on. Miki can answer questions, prompt interlocutors with suggestions, and learn information from interlocu-

Carina Corbin · Fabrizio Morbini · David Traum
Institute for Creative Technologies
University of Southern California
Los Angeles, CA 90094, USA
e-mail: ccorbin17@amherst.edu, morbini@ict.usc.edu,traum@ict.usc.edu

tors. Miki does this with the help of three sources of information. She has a main relational database that contains the basic directory information about every employee. She connects to a web crawler that searches through a company or institute's website and gathers information such as who worked on which projects. Lastly, she learns from the user.

2 Related Work

There are a number of dialogue systems that provide directory assistance information, for example [8, 7, 2, 3]. These mostly focus on the challenges for ASR in dealing with a large database of names, and handle only contact information for the individuals. Our problem is focused a bit differently, in a medium-sized work environment (e.g. 100-200 people), where there is limited ambiguity between names, but where people also want to talk about other kinds of information related to the people at work, including who their supervisor is, and what they work on.

Our system is integrated within the ICT virtual human toolkit [4]. This provides connections to various speech recognition and synthesis options, as well as a visual environment and the virtual human body and non-verbal behavior.

For natural language understanding, generation, and dialogue management, we built on top of the FLoRes system [6]. This system allowed us to define the information state for employees and attributes, and a set of policies for prompting and responding to the user. In addition, related NLU and NLG routines allowed us to specify a set of speech acts, parameters and language templates for this domain.

3 System Overview

Figure 1 shows a screenshot of the Miki character. Users can talk to her face to face, using any toolkit-compatible speech recognizer [5] (we tested with apple dictation and pocketsphinx). Additionally, it is possible to text miki using a chat interface.

The system's knowledge is kept in the form of a database, with several types of information (people, locations, titles, projects), and several relations: for each individual, their first and last names, job title, their office location, and which projects they've worked on. Most of the information comes from an employee directory database. This database is used to power an internal web directory, but can also be saved as an excel spreadsheet. The virtual human neighbor uses the spreadsheet, since it can also easily be added to by hand. There are two other Excel sheets used to input training data to allow the Virtual Human Neighbor to understand and respond realistically. The first is the System Utterances sheet. This sheet gives the system different options for realizing speechacts with no parameters, such as greetings and commonly asked questions about the system itself. The sheet gives the system several different ways to say hello, goodbye, and several suggestions to help the user

Fig. 1: Virtual Human Neighbor.

interact with the system. Our other Excel sheet lists potential user utterance templates, and serves as training data for our natural language understanding module. The goal of the NLU module is to recognize the user's speech act and parameters. Training data can be given, including named entities (such as first name, or office location), and the NLU abstracts from this data, to allow recognition of other entities of the same type. Figure 2 shows an example of the System Utterances in real conversation.

System utterances with parameters that can be filled in from the information state use a natural language generator rather than the system utterances spreadsheet. For example, in Figure 3, utterance 3.1 is a query about a person with first name parameter filled in. The system has multiple people with that first name, and generates a list output in Utterance 3.2. When the last name is added in utterance 3.3, a unique

2.1 User: Hello?
2.2 Miki: Hi, I'm Miki, how may I help you?
2.3 User: What can I ask you?
2.4 Miki: You can ask me about your supervisor.

Fig. 2: A typical first interaction with the Virtual Neighbor

record is found, and the title is returned in utterance 3.4 (using the first and last name parameters). Utterance 3.5 illustrates an ability to do anaphora resolution to the person currently in focus. Utterance 3.6 uses only the first name parameter as well as the requested information about this person.

3.1 User: Who is David?
3.2 Miki: We have many people who share that name: David Traum, David Pynadeth, David DeVault,...
3.3 User Who is David Traum?
3.4 Miki: David Traum works at ICT as a Director of the Natural Language Dialogue Group
3.5 User: Where can I find him?
3.6 Miki: You can visit David at Rm 350

Fig. 3: Dialogue with named entity recognition and generation

In a true conversation, humans don't just ask questions and receive answers. We alternate between being the inquirer and the informant, something we want the Virtual Neighbor to do as well. The storing of information about the subject also allows the system to be updated through user interaction. Let's say the knowledge base does not have the office location of an employee. A person can approach the system and state where the person sits. The person in question can also introduce herself and then state where she sits using this same process. This makes interacting with the Virtual Neighbor a more well-rounded experience. Figure 4 illustrates this process, updating a location, which can be retrieved later.

4.1 User1: Where does Carina sit?
4.2 Miki: You can visit Carina off site
4.3 User1 I am Carina
4.4 Miki: I know you, Carina. You are a Summer Intern-R E U
4.5 User1: I now sit at 357.0
...
4.6 User2: Where is Carina's office?
4.7 Miki: You can visit Carina at 357

Fig. 4: Acquiring information from Users

To extend the topics of conversation, we also created a webcrawler that scrapes the organization's public webpages for more information about each employee. Our first example is acquiring information about which prototypes were worked on, us-

ing a page that lists institute prototypes
(`http://ict.usc.edu/prototypes/all/`). Each of the linked pages has a "team" section, with a list of employees who worked on the prototype. This enables dialogue like that in Figure 5.

5.1 User: What projects did Morbini work on?
5.2 Miki: Fabrizio worked on 2 prototypes including Combat Hunter Action and Observation Simulation (CHAOS).

Fig. 5: Acquiring information from Websites

4 Experimental Design

To evaluate the system, our main goal is to analyze how comfortable people feel interacting with the Virtual Neighbor: to what extent people actually could see themselves using the Virtual Neighbor as a common resource. We plan to test this, by having participants fill out a Pre-Test to see if the user commonly uses the employee directory, visits the company webpages and has an interest in working with the Virtual Neighbor. Next, we will give each participant a set of tasks, such as looking up information in the employee directory. Then we will ask participants to get the same information from the Virtual Neighbor by asking the system given questions. We will leave a few questions open ended, for example Participants will be invited to ask the location or title of another employee in the office of their choice. After the experiment, we will give each participant a Post-Test. This asks the participant to rank their experience with both the employee directory and the Virtual Neighbor. We will also query their favorite and least favorite parts about interacting with the system.

5 Conclusion

In this paper, we present the first version of our Virtual Human Neighbor, who can talk with users about people employed in the same institution. The Virtual Human neighbor can acquire information from three sources: a personnel directory database, public web pages, and through dialogue interaction. Future work involves completing the evaluation described in the previous section, as well as adding more kinds of information. We would also like to include other kinds of speech acts and dialogue games about this material, such as quizzes to get to know more about other employees, or gossip [1].

Acknowledgements The first author was supported by the National Science Foundation under grant 1263386, "REU Site: Research in Interactive Virtual Experiences" (PI: Evan Suma).

The effort described here has been sponsored by the U.S. Army. Any opinions, content or information presented does not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

We would like to thank Ed Fast and Eli Pincus for help with this work.

References

1. Brusk, J., Artstein, R., Traum, D.: Don't tell anyone! two experiments on gossip conversations. In: 11th SIGdial Workshop on Discourse and Dialogue, pp. 193–200. Tokyo, Japan (2010)
2. Buntschuh, B., Kamm, C., Fabbrizio, G.D., Abella, A., Mohri, M., Narayanan, S.S., Zeljkovic, I., Sharp, R.D., Wright, J.H., Marcus, S., Shaffer, J., Duncan, R., Wilpon, J.G.: Vpq: A spoken language interface to large scale directory information. In: Proceedings of InterSpeech, pp. 2863–2867. Sydney, Australia (1998)
3. de Córdoba, R., San Segundo, R., Montero, J.M., Colás, J., Ferreiros, J., Guarasa, J.M., Pardo, J.M.: An interactive directory assistance service for spanish with large-vocabulary recognition. In: INTERSPEECH, pp. 1279–1282 (2001)
4. Hartholt, A., Traum, D., Marsella, S.C., Shapiro, A., Stratou, G., Leuski, A., Morency, L.P., Gratch, J.: All together now: Introducing the virtual human toolkit. In: International Conference on Intelligent Virtual Humans. Edinburgh, UK (2013)
5. Morbini, F., Audhkhasi, K., Sagae, K., Artstein, R., Can, D., Georgiou, P., Narayanan, S., Leuski, A., Traum, D.: Which asr should i choose for my dialogue system? In: Proceedings of the SIGDIAL 2013 Conference, pp. 394–403. Association for Computational Linguistics, Metz, France (2013). URL http://www.aclweb.org/anthology/W/W13/W13-4064
6. Morbini, F., DeVault, D., Sagae, K., Gerten, J., Nazarian, A., Traum, D.: Flores: A forward looking, reward seeking, dialogue manager. In: In proceedings of International Workshop Series on Spoken Dialouge Systems Technologies(IWSDS-2012) (2012)
7. Seide, F., Kellner, A.: Towards an automated directory information system. In: in Proc. Eurospeech, pp. 1327–1330 (1997)
8. Yoshioka, O., Minami, Y., Shikano, K.: A multi-modal dialogue system for telephone directory assistance. In: ICSLP'94 (1994)