

# Identity Models for Role-Play Dialogue Characters

**Patricia Chaffey**

Institute for Creative Technologies  
University of Southern California  
chaffey@ict.usc.edu

**David Traum**

Institute for Creative Technologies  
University of Southern California  
traum@ict.usc.edu

## Abstract

An ongoing challenge in dialogue systems is maintaining consistent personalities and attributes throughout a conversation. To this end, our ongoing work aims to address this obstacle by implementing an 'identity model' which references personality traits from the Big 5 model and personal attributes that appear frequently in conversation. This effort builds off of previous work conducted using a Wizard-of-Oz (woz) system, which set the dialogue agents in a wildfire search-and-rescue scenario. The data gathered in this study helps identify what attributes are most important for the specific context of our system, but the approach being taken for the dialogue agents may be favorable for more general approaches to the development of robust dialogue agents.

## 1 Introduction

Dialogue agents are capable of communicating a variety of topics, depending on their purpose and intended scope, and in recent years, there has been a focus on increasing their likeness in communication to how humans interact with one another. A related challenge to this focus is the difficulty in presenting a persistent personality and character attributes throughout an interaction. As such, an area of research within this domain seeks to incorporate elements of personality and character attributes as a means of developing more unique and approachable characters for people to converse with. To this end, we are developing an approach for modelling identities, and working towards implementing these modelled characters in a previously developed framework relating to a wildfire search and rescue simulation developed by [Chaffey et al. \(2019\)](#).

One area that we are currently expanding upon is the development of what we refer to as 'identity models'. These models track personality scores as well as character attributes (e.g. a character's age,

name, occupation, etc.). A character's personality and/or their attributes can be adjusted very quickly, either by changing the personality scores or by modifying the specific attributes.

## 2 Related Research

Models of personality and character attributes are not new concepts in the field of natural language processing. Some approaches have utilized large corpuses with aims to model consistent behavior in their chat agents ([Li et al., 2016](#); [Mazaré et al., 2018](#); [Zhang et al., 2018](#)). Others, such as [Mairesse and Walker \(2007\)](#), looked to demonstrate varying degrees of a personality trait across dialogue options. A challenge that is consistently brought up is the lack of persistence in personality and character attributes in conversations, which can be jarring to users.

[Fillwock and Traum \(2018\)](#) conducted an analysis on how personal attributes are shared in conversational settings. Their work identified a set of common topics that arise frequently in conversations, which provides a general set of characteristics that a dialogue agent should be able to contend with. In addition, a study conducted by [Mitsuda et al. \(2017\)](#) found that how a system responded to inclusions of personal attributes could be satisfactory or unsatisfactory depending on the nature of the attributes. Namely, expressing understanding towards attributes that were considered permanent was better received than attempting to refer to temporary states, which at times could feel unnatural.

There is also a wealth of information on Big 5 traits themselves, and how they relate to other facets of personality. [Roccas et al. \(2002\)](#) connect motivational values to traits in the Big 5, which include extraversion, agreeableness, openness to experience, conscientiousness, and neuroticism. The motivational values described by [Roccas et al. \(2002\)](#) help elaborate how the Big 5 traits come into play in an individual's personality.

### 3 The Identity Model

At this stage, we currently utilize the Big 5 personality model as discussed and expanded on by [Roccas et al. \(2002\)](#). Agents receive an array of five scores, indicating how closely they align (or don't align) with a trait from the Big 5 model. For example, an agent scoring higher in the neuroticism category may demonstrate more anxiety, and would select dialogue reflecting that anxiety. How high or low the agent is scored in other categories will influence whether or not they select dialogue reflecting their responsibility (or lack thereof), their introversion/extraversion, whether or not they are imaginative, and more. To go into more detail on the flow of the system, we utilize the sequential model via Keras API with 3 layers to rank all possible character responses to user input. Once we have this list of possible responses, we leverage the identity model to select responses that a.) fit the character's attributes (e.g. characters with an age attribute of under 16 would not suggest they drive themselves to safety), and b.) best match their personality scores (i.e. responses are tagged with associated Big 5 traits and scores, and can be used to select responses more in line with a character's personality scores).

In addition, [Fillwock and Traum \(2018\)](#)'s identified common attributes are used to develop the sustained information that an agent will reference when sharing details about themselves with the person interacting with them. Currently, the personality model is used to determine how a statement is phrased, and the list of attributes assigned to a character determines which features they share with the user.

Excerpts from two transcripts have been included below. The first one demonstrates an individual who has a higher score in conscientiousness, as well as a higher score in agreeableness. As can be seen below, this individual does not argue or question too much, and readily agrees with what the user is suggesting.

USER: Hi Ross, nice to meet you. We're doing a sweep through the town to let people know a fairly large fire is headed in this direction, about a day or two out. we're asking residents to evacuate ahead of time

RESIDENT: Wow. I'd heard about it on the news, but I didn't really think it'd make it this far

USER: Yeah, unfortunately it's looking like it will. So we need you, and anyone with you, to gather up the essential items and head south

RESIDENT: How long do you think we have to get everything together?

USER: I would say maybe a day or so, but I wouldn't recommend lingering

RESIDENT: Yeah, okay, I hear you. I'll grab what I need. Shouldn't take too long to go through the apartment

USER: That sounds great. Do you need any assistance evacuating?

RESIDENT: I think I'm all set, actually

This second excerpt is of an individual with a lower score in agreeableness, which leads them to resist the operator's suggestion they evacuate.

RESIDENT: I can't leave my condo! I've lived here for 3 years

USER: i understand how important your home is to you, but we're worried about this fire. we recommend you gather all of your most important items and get out of the area until the danger is past

RESIDENT: I hear what you're saying, but...I don't know, leaving my home is so hard

USER: I sympathize. We're going to do everything we can to control the fire, but it will help us a lot if we know you're safe

RESIDENT: If you're well-read like me, you don't just give up on the place you love. Surely there's something I can do to keep my home safe

USER: Here's what I recommend: take a little bit of time to clear out the yard of any dead plants or debris. Then gather important things, and head out.

RESIDENT: Okay, I'll try to gather the most important things

### 4 Future Work and Conclusions

Efforts towards automating agents capable of displaying distinct personalities and maintaining their own persistent set of attributes are ongoing. The data collected in previous experiments with the wildfire simulation is useful for honing in on the attributes necessary for the specific scenario our system currently supports, but we aim to expand our dialogue agent's capabilities using the observations discussed by [Fillwock and Traum \(2018\)](#). This, alongside with the added attention to modifiable personalities, currently sets us on a promising path for more robust dialogue agents capable of demonstrating a distinctive personality while maintaining persistent attributes in a conversation. Another avenue we intend to explore is applying our model to the users themselves during conversation, by taking note of when the user shares personal information.

### Acknowledgments

This research was sponsored by the Army Research Office and was accomplished under Cooperative Agreement Number W911NF-20-2-0053. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies of the Army Research Office or the U.S. Government.

## References

- Patricia Chaffey, Ron Artstein, Kallirroi Georgila, Kimberly A. Pollard, Setareh Nasihati Gilani, David M. Krum, David Nelson, Kevin Huynh, Alesia Gainer, Seyed Hossein Alavi, Rhys Yahata, and David Traum. 2019. [Developing a Virtual Reality Wildfire Simulation to Analyze Human Communication and Interaction with a Robotic Swarm During Emergencies](#). In *Proceedings of the 9th Language and Technology Conference*, Poznań, Poland. LTC.
- Sarah Fillwock and David Traum. 2018. [Identification of personal information shared in chat-oriented dialogue](#). In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)*, Miyazaki, Japan. European Language Resources Association (ELRA).
- Jiwei Li, Michel Galley, Chris Brockett, Georgios Spathourakis, Jianfeng Gao, and Bill Dolan. 2016. [A persona-based neural conversation model](#). In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 994–1003, Berlin, Germany. Association for Computational Linguistics.
- François Mairesse and Marilyn Walker. 2007. [PERSONAGE: Personality generation for dialogue](#). In *Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics*, pages 496–503, Prague, Czech Republic. Association for Computational Linguistics.
- Pierre-Emmanuel Mazaré, Samuel Humeau, Martin Raison, and Antoine Bordes. 2018. [Training millions of personalized dialogue agents](#).
- Koh Mitsuda, Ryuichiro Higashinaka, and Junji Tomita. 2017. [Investigating the effect of conveying understanding results in chat-oriented dialogue systems](#). In *Proceedings of the Eighth International Joint Conference on Natural Language Processing (Volume 2: Short Papers)*, pages 389–394, Taipei, Taiwan. Asian Federation of Natural Language Processing.
- Sonia Roccas, Lilach Sagiv, Shalom H. Schwartz, and Ariel Knafo. 2002. [The big five personality factors and personal values](#). *Personality and Social Psychology Bulletin*, 28(6):789–801.
- Saizheng Zhang, Emily Dinan, Jack Urbanek, Arthur Szlam, Douwe Kiela, and Jason Weston. 2018. [Personalizing dialogue agents: I have a dog, do you have pets too?](#) *CoRR*, abs/1801.07243.