Sudeep R. Gandhe
Ph.D., Computer Science
srgandhe@gmail.com

Research Interests

Dialogue Management, Spoken Dialogue Systems, Natural Language Processing, Machine Learning, Human Computer/Robot Interaction, Virtual Humans, Mobile Speech Applications.

Employment

Google LLC. Mountain View, CA
Senior Software Engineer *Jan 2018 - Current*

Essential Products Inc. Palo Alto, CA
Software Engineer *May 2017 - Jan 2018*

Google LLC. Mountain View, CA
Software Engineer *Jul 2012 - May 2017*

Institute for Creative Technologies, University of Southern California Los Angeles, CA
Graduate Research Assistant *Summer 2003 - Jul 2012*

Microsoft Research Redmond, WA
Summer Research Intern *Summer 2006*

Education

University of Southern California Los Angeles, CA
Ph.D. Computer Science (GPA: 3.86) *May 2014*

Thesis title: **Rapid Prototyping and Evaluation of Dialogue Systems for Virtual Humans.**

Advisor: Prof. David Traum

University of Southern California Los Angeles, CA
M.S. Computer Science (GPA: 3.93) *May 2004*

Veermata Jijabai Technological Institute, University of Mumbai Mumbai, India
B.Eng. Computer Engineering (GPA: First Class) *June 2002*

Experience

Google LLC. Mountain View, CA
Senior Software Engineer *Jan 2018 - Current*

- Developing/Prototyping Conversational AI features for Google Assistant specifically for Recommendations and Shopping.
  Technologies/Tools: C++, Python, Tensorflow, Large Language Models, Dialogue Management, Natural language processing.

Essential Products Inc. Palo Alto, CA
Software Engineer *May 2017 - Jan 2018*

- Developing the multi-modal dialogue system for Essential Products.
  Technologies/Tools: Multi-Modal Dialogue Management, Production quality software development.

Google LLC. Mountain View, CA
Software Engineer *Jul 2012 - May 2017*

- Developing the dialogue system for Google Assistant. Launched features such as Local search dialogs, conversations for Recipes and delightful features such as Lucky Trivia.
  Technologies/Tools: C++, Google.com search stack, Production quality software development, Dialogue Management, Natural language processing.

- Adding Natural language processing based features to a Computer Vision based production backend powering Google Photo Search, Google goggles.
  Technologies/Tools: C++, MapReduce, Flume, Google Protocol buffers, Production-quality backend server development, Natural language processing.

Institute for Creative Technologies, Los Angeles, CA
University of Southern California
Graduate Research Assistant *Summer 2003 - Jul 2012*

- Conducted research on rapid prototyping of spoken dialogue systems for Virtual Humans. Build a dialogue model that bootstraps from human-human dialogue corpus in the form of role-plays and Wizard-of-Oz studies. The model primarily works at the surface text level but also allows for additional deeper information state annotations.
  Technologies/Tools: Information Retrieval approaches, Text classification, Natural Language Processing, Java, Swing.
  Publications: Gandhe & Traum 2007a, 2007b.

- Conducted research on automatic evaluation of dialogue coherence models for virtual humans.
  Technologies/Tools: Tomcat web server, Java servlets, JSP, JavaScript, AJAX, PostgreSQL, R.
  Publications: Gandhe & Traum, 2008.

- Developed Integrated Authoring
  Tool that allows non-experts to author advanced QuestionAnswering Virtual Human characters.
  Technologies/Tools: Java, Swing, XML, XSLT.
  Publications: Gandhe et al., 2011, Gandhe et al., 2009.

- Developed Dialogue Manager for Advanced Question-Answering Virtual Humans. These virtual humans can engage in bargaining and negotiation dialogue with humans. These characters are used for simulation training.
  Technologies/Tools: Spoken Dialogue System, Java, SCXML, XML.
  Publications: Gandhe et al., 2008.

- Set up web interfaces for gathering coherence ratings from human judges for evaluating dialogue systems.
  Technologies/Tools: Tomcat web server, Java servlets, JSP, JavaScript, AJAX, PostgreSQL.
  Publications: Artstein et al., 2009, Artstein et al., 2008.

- Developed Natural Language Generation Module based on phrase expansion grammar that converts a semantic frame into surface text for an utterance.
  Technologies/Tools: C++, Perl.
  Publications: Traum et al., 2005.

- Conducted research on coherent interactions. This research explored the idea of introducing short linking dialogues between a question and its answer which is in the form of a video retrieved from a large collection. These linking dialogues increased the coherence of interactions.
  Technologies/Tools: Text Classification, Support Vector Machines, Perl.
  Publications: Gandhe et al., 2006, Gandhe et al., 2004.

- Build a Dialogue Manager and User Interface for English to Farsi Speech-to-Speech Translation in medical domain.
  Technologies/Tools: Dialogue Management, Text Classification, Java, Swing
  Publications: Belvin et al., 2005, Narayanan et al., 2004.

Microsoft Research Redmond, WA
Summer Research Intern *Summer 2006*

- Conducted research on improving Speech Recognition performance for out of grammar utterances in command-and-control speech application, VoiceCommand, on Windows Mobile devices. Investigated use of active learning in developing recognition grammars.
  Technologies/Tools: Grammar based Speech Recognition, Active Learning, C, MS SQLserver, MS Visual Studio, Windows Mobile 6.5
  Publications: Paek et al., 2007, 2008

USC Med-COR program Los Angeles, CA
Tutor *Aug 2002 - May 2003*

- Teaching Math to 9th and 12th grade students.

Academic Honors

- The **National Deans list** 2002-2003.
- One of the top 25 students selected for Indian National Mathematics Olympiad exam, from **Regional Mathematics Olympiad**, Mumbai 1997.
- Secured All India Rank 1427 in IIT (Indian Institute of Technology) JEE 1998 **(Joint Entrance Exam) among 100,000 students. (98.57 percentile)**
- Secured All India Rank 196 in IIT (Indian Institute of Technology) GATE 2002 **(Graduate Aptitude test in Engineering (Computer Science)) (98.96 percentile)**

Skills

Languages	C++, Java, C, C, Perl, Javascript, AJAX, XML, XSLT, SCXML
Platforms	Linux, Windows, Mac OS X
Tools & Technology	MySQL, PostgreSQL, Tomcat Web server, JSP, lex & yacc, LaTeX, soar, R, NetBeans IDE, Eclipse, Microsoft Visual Studio

Relevant Coursework

Computational approaches to natural language dialogue modeling, Empirical methods in natural language processing, Natural Language processing, Lexical Semantics, Advanced AI, Software Multi-Agent systems, Analysis of Algorithms, Database systems, Information integration on the web.

Professional Activities

- Review committee member for:
  - NAACL-HLT 2013, 2015, 2021,
  - ACL 2014, 2020, 2021,
  - ACL Rolling Review 2021, 2022,
  - Coling 2014, 2020,
  - EMNLP 2020,
  - EACL 2021.
- Member of Association for Computational Linguistics (ACL), International Speech Communication Association (ISCA), SIGdial - ACL Special Interest Group on Discourse and Dialogue.
- Organizing Committee member for Young Researchers Roundtable on Spoken Dialog Systems, 2007.
- Advisory Committee member for Young Researchers Roundtable on Spoken Dialog Systems, 2008.
- Organizing Committee member for SemDial 2011 (Los Angelogue) Workshop on the Semantics and Pragmatics of Dialogue.

Publications

1. Gurmeet Manku, James Lee-Thorp, Bhargav Kanagal, Joshua Ainslie, Jingchen Feng, Zach Pearson, Ebenezer Anjorin, Sudeep Gandhe, Ilya Eckstein, Jim Rosswog, Sumit Sanghai, Michael Pohl, Larry Adams and D. Sivakumar. **ShopTalk: A System for Conversational Faceted Search.** arXiv preprint arXiv:2109.00702, 2021

2. Yury Zemlyanskiy, Sudeep Gandhe, Ruining He, Bhargav Kanagal, Anirudh Ravula, Juraj Gottweis, Fei Sha and Ilya Eckstein. **DOCENT: Learning Self-Supervised Entity Representations from Large Document Collections.** Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics, 2021.

3. Sudeep Gandhe and David Traum. **SAWDUST: a Semi-Automated Wizard Dialogue Utterance Selection Tool for domain-independent large-domain dialogue.** Demo presented at 15th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL), Philadelphia. 2014.

4. Sudeep Gandhe and David Traum. **A semi-automated evaluation metric for dialogue model coherence.** IWSDS2014 Workshop on Spoken Dialog Systems, Napa, USA. 2014.

5. Sudeep Gandhe and David Traum. **Surface Text based Dialogue Models for Virtual Humans.** 14th Annual Meeting of the Special Interest Group on Discourse and Dialogue (SIGDIAL), Metz, France. 2013.

6. Peter Khooshabeh, Sudeep Gandhe, Cade McCall, Jonathan Gratch, James Blascovich and David Traum. **The effects of virtual agent humor and gaze behavior on human-virtual agent proxemics.** Proceedings of 11th International Conference on Intelligent Virtual Agents, IVA 2011, Reykjavik, Iceland. 2011.

7. Sudeep Gandhe, Michael Rushforth, Priti Aggarwal and David Traum. **Evaluation of an Integrated Authoring Tool for Building Advanced Question-Answering Characters.** Proceedings of Interspeech-11, Florence, Italy. 2011.

8. Ron Artstein, Michael Rushforth, Sudeep Gandhe, David Traum and Aram Donigian. **Limits of Simple Dialogue Acts for Tactical Questioning Dialogues.** Proceedings of 7th IJCAI workshop on Knowledge and Reasoning in Practical Dialogue Systems, Barcelona, Catalonia (Spain). 2011.

9. Sudeep Gandhe, Alysa Taylor, Jillian Gerten and David Traum. **Rapid Development of Advanced QuestionAnswering Characters by Non-experts.** Demo presented at SIGdial 2011, Portland, Oregon, June 2011.

10. Peter Khooshabeh, Cade McCall, Sudeep Gandhe, Jonathan Gratch and James Blascovich. **Does it matter if a computer jokes?** Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems, CHI EA 11, Vancouver, BC. 2011.

11. Sudeep Gandhe and David Traum. **Ive said it before, and Ill say it again: An empirical investigation of the upper bound of the selection approach to dialogue.** Proceedings of the SIGdial 2010 Conference, Tokyo, Japan, 2010.

12. Michael Rushforth, Sudeep Gandhe, Ron Artstein , Antonio Roque, Sarrah Ali, Nicolle Whitman and David Traum. **Varying Personality in Spoken Dialogue with a Virtual Human.** Intelligent Virtual Humans Conference (IVA-09), Lecture Notes in Artificial Intelligence Vol. 5733, Amsterdam. 2009. An extended version is available as USC-ICT Technical Report, ICT-TR-03-2009.

13. Ron Artstein, Sudeep Gandhe, Jillian Gerten, Anton Leuski and David Traum. **Semi-formal evaluation of conversational characters. In Languages: From Formal to Natural.** Essays Dedicated to Nissim Francez on the Occasion of His 65th Birthday (Lecture Notes in Computer Science 5533). 2009.

14. Sudeep Gandhe, Nicolle Whitman, David Traum, Ron Artstein. **An Integrated Authoring Tool for Tactical Questioning Dialogue Systems.** 6th Workshop on Knowledge and Reasoning in Practical dialogue Systems. Pasadena, California, July 2009.

15. Ron Artstein, Sudeep Gandhe, Michael Rushforth, David Traum. **Viability of a Simple Dialogue Act Scheme for a Tactical Questioning Dialogue System.** DiaHolmia 13th Workshop on the Semantics and Pragmatics of Dialogue. Stockholm, Sweden, June, 2009.

16. Ron Artstein, Jacob Cannon, Sudeep Gandhe, Jillian Gerten, Joseph Henderer, Anton Leuski and David Traum. **Coherence of Off-Topic Responses for a Virtual Character.** 26th Army Science Conference, Florida. 2008.

17. David Traum, Anton Leuksi, Antonio Roque, Sudeep Gandhe, David DeVault, Jillian Gerten, Susan Robinson and Bilyana Martinovski. **Natural Language Dialogue Architectures for Tactical Questioning Characters.** 26th Army Science Conference, Florida. 2008.

18. Sudeep Gandhe, David DeVault, Antonio Roque, Bilyana Martinovski, Ron Artstein, Anton Leuski, Jillian Gerten and David Traum. **From Domain Specification to Virtual Humans: An**

integrated approach to authoring tactical questioning characters. Proceedings of Interspeech-08, Brisbane, Australia, 2008.

19. Sudeep Gandhe and David Traum. **Evaluation Understudy for Dialogue Coherence Models.** 9th SIGdial Workshop on Discourse and Dialogue, Ohio, 2008.

20. Tim Paek, Sudeep Gandhe, David Maxwell Chickering. **Rapidly Deploying Grammar-Based Speech Applications with Active Learning and Back-off Grammars.** 9th SIGdial Workshop on Discourse and Dialogue, Ohio, 2008.

21. Ron Artstein, Sudeep Gandhe, Anton Leuski, and David Traum. **Field Testing of an Interactive QuestionAnswering Character.** Proceedings of ELRA Workshop on Evaluation, LREC, Marrakech, Morocco. 2008.

22. Sudeep Gandhe and David Traum. **Creating Spoken Dialogue Characters from Corpora without Annotations.** Proceedings of Interspeech-07, Antwerp, Belgium, 2007.

23. Tim Paek, Sudeep Gandhe, David Maxwell Chickering and Yun Cheng Ju. **Handling out-of-grammar commands in mobile speech interaction using backoff filler models.** Proceedings of the ACL Workshop on Grammar-Based Approaches to Spoken Language Processing (SPEECHGRAM), Prague, Czech Replublic, June 29, 2007.

24. Sudeep Gandhe and David Traum. **First Steps towards Dialogue Modeling from an Un-annotated Human- Human Corpus.** 5th Workshop on knowledge and reasoning in practical dialogue systems, Hyderabad, India, Jan 8, 2007.

25. Sudeep Gandhe, Andrew S Gordon and David Traum. **Improving Question-Answering with Linking Dialogues.** 2006 International Conference on Intelligent User Interfaces, Sydney, Australia, Jan 29 - Feb 1, 2006.

26. Randall Hill, Julia Kim, Michelle Zbylut, Andrew S Gordon, David Traum, Sudeep Gandhe, Stewart King, Salvo Lavis and Scott Rocher. **AXL.Net: Web-enabled Case Method Instruction for Accelerating Tacit Knowledge Acquisition in Leaders.** 25th Army Science Conference, Orlando, Florida. 2006.

27. Sudeep Gandhe, Andrew Gordon, Anton Leuski, David R Traum, and Douglas W. Oard. **First steps toward Linking Dialogues: mediating between free-text questions and pre-recorded video answers.** presented at the Army Science Conference, 2004.

28. Robert Belvin; Emil Ettelaie; Sudeep Gandhe; Panayiotis Georgiou; Kevin Knight; Daniel Marcu; Scott Millward; Shrikanth Narayanan; Howard Neely; David Traum. **Transonics: A Practical Speech-to-Speech Translator for English-Farsi Medical Dialogs.** Demo presented at ACL Interactive Poster and Demonstration Sessions, June 2005.

29. S. Narayanan, S. Ananthakrishnan, R. Belvin, E. Ettelaie, S. Gandhe, S. Ganjavi, P. G. Georgiou, C. M. Hein, S. Kadambe, K. Knight, D. Marcu, H. E. Neely, N. Srinivasamurthy, D. Traum and D. Wang. **The Transonics Spoken Dialogue Translator: An aid for English-Persian Doctor-Patient interviews.** In working notes of the AAAI Fall symposium on Dialogue Systems for Health Communication. 2004

30. David Traum, William Swartout, Jonathan Gratch, Stacy Marsella, Patrick Kenny, Eduard Hovy, Shri Narayanan, Ed Fast, Bilyana Martinovski, Rahul Baghat, Susan Robinson, Andrew Marshall, Dagen Wang, Sudeep Gandhe and Anton Leuski. **Dealing with Doctors: A Virtual Human for Non-team Interaction,** Demo presented at SIGdial 05, September 2005.

31. Paul Scerri, David Pynadath, Nathan Schurr, Alessandro Farinelli, Sudeep Gandhe and Milind Tambe. **Team-oriented programming and proxy agents: the next generation.** Proceedings of the workshop on Programming multiagent systems, Lecture notes in computer science. 2004.

Patents

1. Johnny Chen, Thomas L. Dean, Qiangfeng Peter Lau, Sudeep Gandhe, Gabriel Schine. **Conversational agents.** WO 2016018763 A1.
2. Johnny Chen, Thomas L. Dean, Gabriel Schine, Qiangfeng Peter Lau, Sudeep Gandhe. Transfer of configuration agent between devices. WO 2016007339 A1.
3. Johnny Chen, Thomas L. Dean, Gabriel Schine, Peter Lau, Sudeep Gandhe. **Event scheduling.** US 20160014222 A1.
4. Johnny Chen, Thomas L. Dean, Gabriel Schine, Sudeep Gandhe, Qiangfeng Peter Lau. **Conversational music agent.** US 20160092159 A1.

Research Summary

My research interests include spoken dialogue management, coherent interaction, natural language processing. Specifically I'm interested in applying machine learning techniques toward bootstrapping dialogue models from human-human corpus with minimal annotations.

Virtual Human Dialogue Systems

At Institute for Creative Technologies, I have been building several Virtual Human dialogue systems. The goal for such dialogue systems is to be as human-like as possible and their performance is generally evaluated using turn-by-turn appropriateness ratings.

There are multiple approaches towards building dialogue systems. The specific architecture chosen is dictated by the specific goals of the dialogue system and its evaluation criteria. This architecture, in turn, determines the specific types of resources required.

One of the contributions of my thesis towards practical goal of rapid prototyping of virtual human dialogue systems is **Reducing the cost of collecting required resources**. Our approach is to allow designers with limited or no expertise in computational linguistics to quickly build consistent resources with the help of an integrated authoring tool (Gandhe et al., 2009). Reducing the level of expertise required allows us to collect the resources at a reduced cost.

We have adopted a top-down approach for building virtual characters that can take part in tactical questioning simulations (Gandhe et al., 2008). Scenario designers start by authoring the domain knowledge for the virtual characters with the help of an integrated authoring tool. The authoring tool automatically generates all relevant dialogue acts following a genre specific dialogue act scheme and ensures consistency and completeness. Scenario designers can author appropriate surface text for all the dialogue acts. The tool also allows designers to specify simple dialogue policies for question-answering. The accompanying dialogue manager uses an information-state based approach (Traum Larsson, 2003). The information-state is in part based on the conversational game theory (Lewin, 2000) and implemented as State Chart XML (SCXML), a W3C working draft (Barnett et al., 2008). To date, seven virtual human characters have been authored by designers who had no experience in building dialogue systems.

The other contributions of my thesis work are,

- **Developing flexible dialogue architectures that allow novel combinations of different types of resources.**

  Each dialogue system architecture imposes its own constraints on what resources can be combined. Traditional dialogue systems such as information state based models (Traum Larsson, 2003) operate at a dialogue act level. For a dialogue system to be operational corpus annotation with dialogue acts and rule authoring for updating the information state are prerequisites. On the other hand, there have been approaches that operate at surface text level alone. (e.g., Wallace (2003), Abu Shawar Atwell (2005)). But the output of such systems frequently is ungrammatical.

  We present a model that primarily works at the surface text level but also allows for additional deeper information state annotations. Our initial model has been implemented and tested in the context of a Virtual Human dialogue system and has shown 82(Gandhe Traum, 2007b; Gandhe Traum, 2007a). The model uses information state annotations for tracking the topic in order to avoid violations of presupposition. It is significantly better than a model that takes only surface text dependencies into account as judged by the dialogue participants. Our model bootstraps from roleplay and WoZ dialogue corpora and allows for information state annotations in an incremental fashion. It also allows us to compare relative performance gains from different types of annotations over modeling the surface text dependencies alone.

- **Reducing the cost of estimating the performance for a specific dialogue system**
  One of the ways to calculate the performance is to actually evaluate the resulting dialogue system with human users. But this is a time consuming process and can be prohibitively costly for multiple evaluations. The cost of estimating the performance can be reduced by reducing the human involvement from the evaluation procedure.

  Our approach focuses on evaluating dialogue coherence, which is a suitable criterion for Virtual Human Dialogue Systems. There is a need for automated evaluation or methods that can estimate the performance of the dialogue system (evaluation understudy).

  We have developed an evaluation understudy for evaluating local models of dialogue coherence. A local dialogue coherence model gives a coherence score for an utterance given its context (the dialogue history). We propose to use the information ordering task (Lapata, 2006) which repeatedly applies the local model of dialogue coherence to construct the most coherent ordering from a set of dialogue utterances. We have found an objective measure for scoring the success in information ordering task that correlates well with human judgments for dialogue coherence (Pearsons r = 0.75) (Gandhe Traum, 2008). This objective evaluation measure can be used to compare different dialogue models using different resources.

Here are some of the other projects Ive worked on,

Coherent Interactions

Free text questions and pre-recorded video as a response is a recurrent theme in NL interfaces. It has proved very useful in various applications for training and entertainment as well. Users are allowed to input a free-text question which in turn elicits a pre-recorded video response. Although the video response tends to have very good value in terms of immersive experience, the very design of the system allows for a lack of coherence. It is due to the case when there is no video response that directly answers the question or when the response is not phrased in a desired manner. I tried to address this issue by introducing short linking dialogue between question and answer to bridge the gap. Experimental results showed that human-generated linking dialogs can significantly increase the coherence of interaction (Gandhe et. al., 2004). Further analysis of human-generated linking dialogs reveals that these carry more information than present in the answer or the question. I have implemented first techniques for creating such computer generated linking dialogues (Gandhe et. al., 2006).

Speech-to-Speech Translation

I helped develop a speech to speech translation system for medical domain. (Narayanan et. al., 2004, Belvin et. al., 2005) Using this system, an English speaking doctor can communicate with a Farsi speaking patient and carry out the medical diagnosis. My work focused on the dialogue manager and a java based graphical user interface that facilitates the communication between patient and the doctor. Only one participant, the doctor, can control the interaction. The dialogue manager component in this system is different from most of the dialogue systems, in the sense that it has no active participation in carrying out the dialogue. It can only assist the communication process.

Speech Applications on Mobile Devices

During my summer internship at Microsoft Research, I worked on improving the command and control speech application called VoiceCommand designed for Windows Mobile platform. VoiceCommand uses SAPI CFG grammar for speech recognition. The work focused on improving recognition results for out of grammar utterances (Paek et. al., 2007). I also investigated methods for rapid prototyping of grammars used in speech recognition for such command and control systems (Paek et. al., 2008).