CHAD: Chat-Oriented Dialog Systems

Alexander Rudnicky
Carnegie Mellon University, e-mail: air@cs.cmu.edu

Abstract Historically, conversational systems have focused on goal-directed interaction and this focus defined much of the work in the field of spoken dialog systems. More recently researchers have started to focus on non- goal-oriented dialog systems often referred to as ”chat” systems. We can refer to these as Chat-oriented Dialog (CHAD)systems. CHAD systems are not task-oriented and focus on what can be described as social conversation where the goal is to interact while maintaining an appropriate level of engagement with a human interlocutor. Work to date has identified a number of techniques that can be used to implement working CHADs but it has also highlighted important limitations. This note describes CHAD characteristics and proposes a research agenda.

1 Introduction

Historically, conversational systems have focused on goal-directed interaction (e.g. [3]) and this focus defined much of the work in the field of spoken dialog systems. More recently researchers have started to focus on non- goal-oriented dialog systems [2, 9, 13, 11] often referred to as ”chat” systems. We can refer to these as Chat-oriented Dialog (CHAD) systems. CHADs differ from ”chatbots”, systems that have recently gained significant attention. The latter have evolved from on-line, text-based, chat systems, originally supporting human-human conversations and more recently finding acceptance in as a supplement to call centers. Chatbots have begun to resemble dialog systems in their use of natural language understanding and generation, as well as dialog management. Speech-based implementations have begun to appear in commercial applications (for example, see the recent Conversational Interaction Conference [1]).

By contrast, CHAD systems are not task-oriented and focus on what can be described as social conversation where the goal is to interact while maintaining an appropriate level of engagement with a human interlocutor. One consequence is that many of the metrics that have been develop for task-oriented system are no longer appropriate. For example, the purpose of a task dialog is to achieve a satisfactory goal (e.g. provide some unit of information) and to do so as rapidly and accurately as practical. The goal in CHAD on the other hand is almost the opposite: keep the human engaged in the conversation for an extended period of time, with no concrete goal in mind other than producing a pleasant experience. In this note we will use the term “chat” to refer to this latter class of system.

2 Contemporary chat systems

Chat systems have been of interest for many years. An early, and influential, example is ELIZA[12], more recent examples include A.L.I.C.E.[10] and Tank the Roboceptionist [6, 8]. These systems are rule-based, in that responses are generated based on hand-crafted pattern matching with limited use of context and history. More recent attempts have take a corpus-based approach [2, 13], where some source of conversations (say movie dialogues) is indexed and features of user inputs are used as retrieval keys. Deep Learning approaches, e.g. [9, 13, 11] have been successfully implemented for this task. A persistent limitation of these approaches is that they tend to reduce to an equivalent of question answering: the immediately preceding user input and perhaps the previous system turn are part of the retrieval key. This narrow window necessarily reduces continuity and users struggle to follow a conversational thread.

A key research question is how to create conversations that initiate and develop in ways similar to those found in human-human conversation. Human (social) conversations exhibit structure that guides their evolution over time. This structure includes elements familiar to participating interlocutors such as conventions for engagement and disengagement, a succession of topics (with heuristics for identifying promising successors), as well as monitoring engagement level and picking strategies for managing topic transitions. Good conversational management also implies sophisticated use of memory, both within the conversation to support back references and general world knowledge (or at the very least domain-specific knowledge), to generate proper turn succession. The implementation of such capabilities is not well understood at this time.

3 Research Challenges

Work to date has identified a number of techniques that can be used to implement working CHADs but it has also highlighted important limitations. These include the following:

- **Continuity**: Response generation techniques have focused primarily on database retrieval; even deep-learning approaches have that character. In this approach a response is generated based on a very short history, typically the previous human input. The unfortunate consequence is that system behavior begins to resemble questions-answering, for example exhibiting sudden topic changes when appropriate responses cannot be retrieved. Missing is some sense of continuity and history that would allow the system to build momentum for a topic, recognize topic re-introductions, decide when and how to change the topic or otherwise communicate that it is paying attention to the conversation.
- **Dialogue Strategies:** Humans are quite adept at managing their conversational behavior, by monitoring interlocutor engagement [14] and by maintaining higher-level goals [7, 5]. Current approaches are still at an early stage and do not yet exhibit the level of flow one is accustomed to observe in human conversations. A related challenge is how to effectively interleave social and task goals in a single conversation and be able to coordinate strategies at both levels to guide behavior into a fluent that is consistent with human expectations.
- **Evaluation**: Current evaluation techniques for chat rely on human judgments. This indicates that we do not as yet understand how to measure chat success. It is a serious problem on several levels. The lack of a clear, automatically computable metric hinders development of corpus-based approaches (the dominant scheme in current research) which lack suitable objective (i.e. loss) functions to guide learning. But the problem extends to human judgment, as currently proposed rating schemes exhibit wide cross-judge variation [4], making it difficult to understand underlying phenomena.

Irrespective of these challenges CHAD continues to attract a significant amount of interest and is generating an increasing body of research. The problem of creating fluent chat provides the opportunity to address a broader range of phenomena than ones that have been the focus in task-oriented systems and develop a richer understanding of human communication.

References

1. AVIOS: Conversational interaction conference. San Jose, CA (2017). URL `http://www.conversationalinteraction.com/`
2. Banchs, R.E., Li, H.: IRIS: a chat-oriented dialogue system based on the vector space model. In: Proceedings of the ACL 2012 System Demonstrations, pp. 37–42. Association for Computational Linguistics (2012)
3. Bohus, D., Rudnicky, A.I.: The RavenClaw dialog management framework: Architecture and systems. Computer Speech & Language **23**(3), 332–361 (2009)
4. Charras, F., Dubuisson Duplessis, G., Letard, V., Ligozat, A.L., Rosset, S.: Comparing system-response retrieval models for open-domain and casual conversational agent. In: Second Workshop on Chatbots and Conversational Agent Technologies (2016)
5. Foster, M.E., Gaschler, A., Giuliani, M., Isard, A., Pateraki, M., Petrick, R.P.: Two people walk into a bar: Dynamic multi-party social interaction with a robot agent. In: Proceedings of the 14th ACM International Conference on Multimodal Interaction, ICMI '12, pp. 3–10. ACM, New York, NY, USA (2012). DOI 10.1145/2388676.2388680. URL `http://doi.acm.org/10.1145/2388676.2388680`
6. Gockley, R., Bruce, A., Forlizzi, J., Michalowski, M., Mundell, A., Rosenthal, S., Sellner, B., Simmons, R., Snipes, K., Schultz, A.C., et al.: Designing robots for long-term social interaction. In: Intelligent Robots and Systems, 2005.(IROS 2005). 2005 IEEE/RSJ International Conference on, pp. 1338–1343. IEEE (2005)
7. Hiraoka, T., Georgila, K., Nouri, E., Traum, D.R., Nakamura, S.: Reinforcement learning in multi-party trading dialog. In: SIGDIAL Conference (2015)
8. Makatchev, M., Fanaswala, I., Abdulsalam, A., Browning, B., Ghazzawi, W., Sakr, M.F., Simmons, R.G.: Dialogue patterns of an arabic robot receptionist. In: HRI (2010)
9. Ritter, A., Cherry, C., Dolan, W.B.: Data-driven response generation in social media. In: Proceedings of the conference on empirical methods in natural language processing, pp. 583–593. Association for Computational Linguistics (2011)
10. Shawar, B.A., Atwell, E.: Using dialogue corpora to train a chatbot. In: Proceedings of the Corpus Linguistics 2003 conference, pp. 681–690 (2003)
11. Vinyals, O., Le, Q.V.: A neural conversational model. In: ICML Deep Learning Workshop (2015). URL `http://arxiv.org/pdf/1506.05869v3.pdf`
12. Weizenbaum, J.: ELIZA—a computer program for the study of natural language communication between man and machine. Commun. ACM **9**(1), 36–45 (1966). DOI 10.1145/365153.365168. URL `http://doi.acm.org/10.1145/365153.365168`
13. Yu, Z., Papangelis, A., Rudnicky, A.: TickTock: A non-goal-oriented multimodal dialog system with engagement awareness. In: Proceedings of the AAAI Spring Symposium (2015)
14. Yu, Z., Xu, Z., Black, A.W., Rudnicky, A.I.: Strategy and policy learning for non-task-oriented conversational systems. In: 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue, vol. 2, p. 7 (2016)