Tell Me About Your Day: Designing a Conversational Agent for Time and Stress Management

Libby Ferland, Jude Sauve, Michael Lucke, Runpeng Nie, Malik Khadar, Serguei Pakhomov, and Maria Gini

Abstract Growing interest in applications of AI in healthcare has led to a similarly elevated interest in fully integrated smart systems in which disparate technologies, such as biometric sensors and conversational agents, are combined to address health problems like medical event detection and response. Here we describe an ongoing project to develop a supportive health technology for stress detection and intervention and discuss a pilot application for one component, the conversational agent.

1 Introduction

Stress is a universal constant in modern life, for both young and old alike. Stress has heavy societal impact – since stress plays a major role in physical and mental well-being, people repeatedly experiencing stress without appropriate coping mechanisms cannot achieve their full potential in any number of situations [15, 17]. A variety of mitigation and coping techniques have been proposed to address stress, notably effective time management and planning skills, but these can be difficult to learn without proper tools [10, 17].

Stress management is a key component of lifestyle modification in the treatment of many conditions among elderly people [4, 16]. The increased availability and ease of use of intelligent technologies such as conversational agents (CAs) gives them great potential as supplements for human caregivers in supporting older populations and for more effective stress management in general.

L. Ferland, J. Sauve, M. Lucke, R. Nie, M. Khadar, M. Gini
Department of Computer Science & Engineering, University of Minnesota, Minneapolis, MN, 55455, e-mail: ferla006@umn.edu

S. Pakhomov
Department of Pharmaceutical Care and Health Systems, University of Minnesota, Minneapolis, MN, 55455

Here we describe the first phases of an ongoing project, “SMARTHUGS,” which is a multidisciplinary effort to create a system designed to detect stress in users and deliver a multimodal touch- and speech-based intervention when users experience high stress. The end product is a prototype of a compression garment, such as a vest, with an integrated unobtrusive physiological signal monitoring using a wrist-worn device (Empatica E4) and a CA interface. The monitor and the CA in tandem will be able to detect potential stressors in real time and flag them, activating the garment to provide therapeutic compression similar to a hug. The CA component is the focus of this paper, specifically the design of its architecture and the domain as initial steps towards addressing open challenges in the field as informed by real user data.

2 Background and Related Work

The key motivating concepts to this work include the effects of stress on health and the usefulness of CAs as support mechanisms.

Stress and Health. Many studies have shown the effects of stress on conditions like heart disease, depression, and sleep disorders, to name a few [11]. It has also been shown that stress has negative effects on disease progression and quality of life, and managing stress can significantly affect disease management [16]. Stress also impacts mental health, which appropriate interventions can mitigate or reverse [4]. This is notably true for cognitive functioning, which is salient for age-related cognitive decline [7, 11]. Stress management and coping skills are important aspects of treatment plans and can be used as cognitive support [6].

Time Management and Coping. Time management is closely related to stress. The development of effective time management skills has been linked to stress mediation and anxiety reduction in many situations and age groups [10, 15, 17]. Time management skills lead to a perception of control over time, which in turn increases reported satisfaction levels in measures such as academic/job performance and overall life satisfaction [15]. Time management skills are also important for older adults, especially those experiencing cognitive decline [6].

Conversational AI and Support. CA systems have been used in many applications, such as changing habits, patient education, and as social support for older users [8]. Health applications are sensitive [14] and so arguably benefit most from data-driven ground-up design. However, the cycle of prototyping and user testing is costly [3], making the use of user-centered design in many support domains as difficult and uncommon as it is necessary, even despite its desirability [12].

3 Proposed Approach

Our prototype system incorporates three separate applications. The tasks done by the system consist of two separate yet similar workflows. In a *prospective* use case, the CA talks with a user about their upcoming schedule and stress levels for the day. In the *retrospective* use case, the CA acts like a diary and talks with a user about the events of their day and the stress they actually experienced.

The key challenges reflect open challenges with CAs and NLP in general: (1) interpreting ambiguous or complex temporal expressions [18], (2) appropriately responding to stress-related disclosures, and (3) doing the same for other affective states and self disclosure [14]. These challenges informed our architecture choices (Sect. 3.1) and were directly drawn from observations in human experiments (Sect. 3.2).

3.1 System Components

MindMeld. MindMeld is an open-source conversational AI platform recently released by Cisco [2], which neatly packages all the NLP modules necessary to create a fully-integrated dialogue system. We chose MindMeld for this pilot due to its robustness and flexibility, which will be needed to further address challenges 2 and 3. Unlike other frameworks which require a complex set of dialogue rules, MindMeld fully exploits machine learning methods to train an agent given only a set of domains, entities, and intents within the domain.

SUTime. SUTime, a library developed by the Stanford NLP group, is a robust deterministic temporal expression parser [1]. SUTime uses an extensive rule base to parse temporal expressions into the TIMEX3 format. We chose SUTime as an alternate pipeline for temporal information because Mindmeld does not handle some types of temporal expressions (eg., discussion of recurring or conditional events) natively, so more support is needed to effectively address challenge 2. SUTime also allows us to reduce the amount of training data required; the application only needs to recognize an entity, not fully understand its contents.

Google Calendar API. We included an external calendar service in our system as a way to increase system performance. With an external API, we can store, retrieve, and modify events without slowing the application down. We chose the Google Calendar API for ease of use and familiarity by developers and users [5].

3.2 Implementation

Data-Driven Dialogue Development. The dialogue flow for our CA is based on scripted interactions with participants in a Wizard of Oz protocol (WOZ) [3]. We sought to answer two questions: When faced with a machine, how do users actually talk about the events in their day? How do they talk about stress and time, and how comfortable are they when discussing these topics? To answer those questions, participants talked to a realistic “prototype” about their daily schedule and their stress levels and mood. Using real user data as a basis for our CA allows us to create a more natural dialogue flow and provides us examples of possible inputs. User data are also reserved as a test set for the CA. Fig.1 shows two dialogue segments from this dataset with rich features.

(1): Response to schedule prompt, with complex temporal expressions and habit disclosure

System: What’s next on your schedule after reading and studying?
Participant: Um...Around that time...once I’m done with that it should be about maybe 4:00 or something. I usually just wait until {PERSON} gets home or when {PERSON} needs to get picked up and then I’ll go do that but it’ll be around maybe 6:00 or 8:00 depending on what days of the week it is.

(2): Response to stress prompt with indirect voluntary preference disclosure

System: How stressful do you think that will be?
Participant: Um, probably around a two, because it’s cold outside.

Fig. 1 Sample real-world dialogue fragments

Domain Modeling. Our implementation began with developing a training dataset of calendar entries provided by other researchers. We then developed our ontology of calendar events and annotation scheme. The key entity in our application is an event, which contains these entities: (1) event name, (2) location, (3) start date/time, (4) end date/time, (5) duration, (6) frequency/frequency duration, and (7) stress level. Of these, only 3, 4, and 5 were predefined in the Mindmeld platform. In the first training set, 300 calendar entries were annotated by two independent raters.

User: I have a meeting with Paul and Ryan at one.
System: Where will that be?
User: in my office
System: How long do you think it will take?
User: Hopefully about an hour
System: How stressful do you think it will be?
User: I’m not sure.
System: How would you rate that on a scale from one to seven?
User: Maybe a 5
System: That sounds stressful. Please tell me how it goes later.
User: Okay.
System: Thank you. I’ll add an event called meeting with Paul and Ryan in your office to your calendar for today. I’ll schedule it to start at 1:00 PM and end at 2:00 PM. Does that sound good?
User: yep
System: Great. What’s the next item on your schedule?
User: That’s it
System: Goodbye.

Fig. 2 Sample dialogue for event entry during a prospective interaction, with features such as automated slot filling and ability to switch between intents and domains.

Conversational Agent. Our pilot CA is designed to exploit learned domain models rather than relying on a rule base. Users need not fill in event information in any particular order, and can mostly move freely between different domains and intents. The only rule enforced is that users must provide all necessary details for each event. However, we built some fail-safes into the application to minimize repetitiveness and user frustration. First, the application fills missing slots automatically using information already provided by the user. Second, in the case of entity resolution failure, the system will re-prompt for the desired information twice before switching to a rule base. Third, in the case of major failure, such as a user attempting an unknown workflow, the agent will try a redirect and then switch to a rule base.

Figure 2 shows a sample dialogue with the system when adding a calendar event.

4 Conclusions and Future Work

This work is in early stages and presents many opportunities for improvement. In the short term, we will increase data privacy and flexibility by both connecting this CA to an in-house speech recognition server and developing a simple calendar graphical interface to be hosted internally. We will also expand our dataset and validate our annotation scheme with naïve raters. Long-term questions to explore include:

Designing Conversational Interventions for Stress. CAs in mental health applications face a number of unique challenges and pitfalls. How does an agent appropriately detect affective state from conversation? What constitutes an appropriate reply to self-disclosure and different user states? The answers to these questions are vital and difficult, considering the sensitivity of mental health issues [14]. CAs used to address mental health issues must be carefully designed to be appropriate and sensitive to user needs, while not being artificial or repetitive.

Improved Handling of Complex Temporal Expressions. Humans talk about time in daily conversation in a way which is often coarse-grained and ambiguous. Fully understanding temporal expressions encompasses problems such as commonsense reasoning and temporal resolution [18, 19]. For instance, our prototype currently struggles to interpret expressions with long tails and conditional expressions ("Remind me to take the garbage out every Wednesday, except when it's after a holiday."), which it should understand to converse naturally.

Enhancing User Modeling with Self Disclosure. The need for better user modeling in CAs is not a new problem, but it is both complex and a notable source of dissatisfaction in users [9, 13]. Though user-dependent, our WOZ testing has shown that users display at least a moderate degree of self-disclosure. A history of a user's schedule, including patterns, could be used as part of a more complex user model.

Acknowledgements Work supported in part by the University of Minnesota Grand Challenges Research, NSF I/UCRC (IIP-1439728) and NSF EAGER (IIS 1927190) grants. The authors would like to thank Anja Wiesner and Sarah Schmoller for their assistance in model and data development.

References

1. Chang, A.X., Manning, C.D.: SUTime: A library for recognizing and normalizing time expressions. In: Proc. 8th Int. Conf. on Language Resources and Evaluation (LREC), pp. 3735–3740. European Language Resources Association (2012)
2. Cisco: The Conversational AI Playbook (2019). https://www.mindmeld.com/docs/ Cited 19 Nov 2019
3. Dahlbäck, N., Jönsson, A., Ahrenberg, L.: Wizard of Oz studies—Why and how. Knowl.-Based Sys. **6**(4), 258–266 (1993)
4. George, L.K.: Social factors, depression, and aging. In: Handbook of Aging and the Social Sciences, pp. 149–162. Elsevier (2011)
5. Google: Calendar API (2019). https://developers.google.com/calendar Cited 17 Nov 2019)
6. Jean, L., Bergeron, M., Thivierge, S., Simard, M.: Cognitive intervention programs for individuals with mild cognitive impairment: Systematic review of the literature. Am. J. Geriatr. Psychiatry **18**(4), 281–296 (2010)
7. Justice, N.J.: The relationship between stress and Alzheimer's disease. Neurobiol. Stress **8**, 127–133 (2018)
8. Laranjo, L., Dunn, A.G., Tong, H.L., Kocaballi, A.B., Chen, J., Bashir, R., Surian, D., Gallego, B., Magrabi, F., Lau, A.Y., et al.: Conversational agents in healthcare: A systematic review. J. Am. Med. Inform. Assoc. **25**(9), 1248–1258 (2018)
9. Luger, E., Sellen, A.: Like having a really bad PA: The gulf between user expectation and experience of conversational agents. In: Proc. CHI Conf. on Human Factors in Computing Systems (CHI), pp. 5286–5297 (2016)
10. Macan, T.H., Shahani, C., Dipboye, R.L., Phillips, A.P.: College students' time management: Correlations with academic performance and stress. J. Educ. Psychol. **82**(4), 760–768 (1990)
11. McEwen, B.S.: Central effects of stress hormones in health and disease: Understanding the protective and damaging effects of stress and stress mediators. Eur. J. Pharmacol. **583**(2-3), 174–185 (2008)
12. Meiland, F., Innes, A., Mountain, G., Robinson, L., van der Roest, H., García-Casal, J.A., Gove, D., Thyrian, J.R., Evans, S., Dröes, R.M., et al.: Technologies to support community-dwelling persons with dementia: A position paper on issues regarding development, usability, effectiveness and cost-effectiveness, deployment, and ethics. JMIR Rehabil. Assist. Technol. **4**(1) (2017)
13. Milhorat, P., Schlogl, S., Chollet, G., Boudy, J., Esposito, A., Pelosi, G.: Building the next generation of personal digital assistants. In: Proc. 1st Int. Conf. on Advanced Technologies for Signal and Image Processing (ATSIP), pp. 458–463. IEEE (2014)
14. Miner, A., Chow, A., Adler, S., Zaitsev, I., Tero, P., Darcy, A., Paepcke, A.: Conversational agents and mental health: Theory-informed assessment of language and affect. In: Proc. 4th Int. Conf. on Human-Agent Interaction (HAI), pp. 123–130. ACM (2016)
15. Monat, A., Lazarus, R.S.: Stress and Coping: An Anthology. Columbia University Press (1991)
16. Newman, S., Steed, L., Mulligan, K.: Self-management interventions for chronic illness. The Lancet **364**(9444), 1523–1537 (2004)
17. Prenda, K.M., Lachman, M.E.: Planning for the future: A life management strategy for increasing control and life satisfaction in adulthood. Psychol. Aging **16**(2), 206–216 (2001)
18. Rong, X., Fourney, A., Brewer, R.N., Morris, M.R., Bennett, P.N.: Managing uncertainty in time expressions for virtual assistants. In: Proc. CHI Conf. on Human Factors in Computing Systems (CHI), pp. 568–579. ACM (2017)
19. Zhou, B., Khashabi, D., Ning, Q., Roth, D.: "Going on a vacation" takes longer than "Going for a walk": A study of temporal commonsense understanding. In: Proc. of the 2019 Conf. on Empirical Methods in Natural Language Processing and the 9th Int. Jt. Conf. on Natural Language Processing (EMNLP-IJCNLP), pp. 3354–3360. ACL (2019)