Towards a Virtual Assistant Health Coach: Corpus Collection and Annotations

Itika Gupta
Department of Computer Science
University of Illinois at Chicago
igupta5@uic.edu

Barbara Di Eugenio
Department of Computer Science
University of Illinois at Chicago
bdieugen@uic.edu

Brian Ziebart
Department of Computer Science
University of Illinois at Chicago
bziebart@uic.edu

Bing Liu
Department of Computer Science
University of Illinois at Chicago
liub@uic.edu

Ben Gerber
Institute for Health Research and Policy
University of Illinois at Chicago
bgerber@uic.edu

Lisa Sharp
Institute for Health Research and Policy
University of Illinois at Chicago
sharpl@uic.edu

1. INTRODUCTION

Many health problems faced by individuals can be mitigated with changes in health behavior. However, successfully implementing healthy behaviors in one's daily life requires significant motivation that most people, individually, find difficult to initiate and maintain. Health coaching has been identified as a successful method for motivating and maintaining health behavior changes by having a peer or professional convey relevant medical information, help to set realizable, yet challenging goals tied to health behavior change, and provide encouragement in adhering to those goals [4]. But unfortunately, personal health coaching is time-intensive, uneconomical for the low-income patients, and has limited accessibility.

Therefore, we aim to create a dialogue-based virtual assistant health coach that will converse with the patient via text messages and will help them increase physical activity by setting Specific, Measurable, Attainable, Realistic and Time-bound (SMART) goals [3]. Even though an influential strand of work on conversational agents has been conducted by Bickmore and his group, their systems rely on predefined set of utterances from the patients [1]. We plan to build an autonomous system that learns from observed communications between human health coach and participants to increasingly automate the generation of dialogue [5]. The ability to recognize proposed health goals in dialogues is a key initial capability needed for this type of learning from demonstration.

In this paper, we talk about our data collection process, two annotation schemas, agreement results and future work on extracting goals from patient-coach dialogue.

2. DATA COLLECTION

ACM ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

We recruited 28 patients between the age of 21 to 65 years who were interested in increasing their physical activity. A health coach, trained in SMART goal setting, conversed with these patients to set goals on weekly basis for a month via SMS. The patients were given Fitbit Alta to monitor their progress. The coach also monitored patients' progress using Fitbit application. The conversation involved setting a specific, measurable and realistic goal, and establishing any barriers to goal attainment. The coach also sent reminders based on patient's preference and provided motivational feedback on their progress. Only one patient didn't complete the study. We have a corpus of 2857 messages. Among these messages, approximately 53% were sent by the coach and 47% by the patients. This tells us that both the coach and the patients were equally involved in setting or negotiating the goal.

3. DATA ANNOTATION

In this section we will look into two types of annotation schemas designed by us to help automate the extraction of health goal set by the user.

3.1 SMART Goal Annotation

3.1.1 *Annotation Schema*

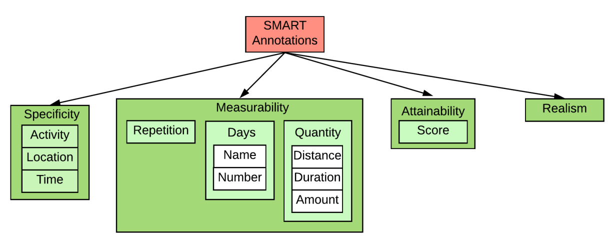

Figure 1: SMART goal annotation schema

15 patient-coach conversations were used to design the SMART goal annotation schema shown in Figure 1 [2]. We didn't annotate Timeliness as a new goal was set every week, and hence by default its value is one week. Each of the annotations can either be categorized as a *slot value* or an

Table 1: Kappa statistics for SMART annotations

Level	S	M	A	R
Message	0.967	0.965	**0.907**	**0.694**
Word	**0.878**	**0.895**	0.515	0.549

Table 2: Kappa statistics for stages and phases annotations.

gs_I	gs_R	gs_A	gs_S	ga_A	ga_S	ga_F
0.978	0.978	0.982	0.795	0	1.0	0.952

intention. A slot value is a word or group of words that captures a particular piece of information such as 'walk' is a slot value for *specific activity*. Where as intention is an utterance that tries to gain information about a slot. Each of the SMART annotation category can have other optional tags such as *previous* to annotate an attribute related to previous week's goal, *accomplished* or *remaining* to annotate the progress of the patient, *update* to add another slot value to an existing one, and *other* for anything which doesn't belong to previous or current week.

3.1.2 Annotation Results

Two annotators annotated four previously unseen patients' data for SMART goals. Inter annotator agreement (IAA) was measured using kappa statistics on individual SMART categories and the results are shown in Table 1.

We measured kappa on two levels: message and word. In message level, we consider an agreement if both the annotators labeled at least one word in the message with the given tag (not necessarily the same word). In word level, we consider an agreement if both the annotators labeled the same word with the given tag.

In total, 447 messages were annotated for IAA. There were 128 messages with *Specificity* tag, 120 with *Measurability* tag, 45 with *Attainability* tag and 13 with *Realism* tag. We observe approximately 90% reliability for *Specificity* and *Measurability* and only 50% reliability for *Attainability* and *Realism*. This is because {S, M} tags are easy to annotate and have high number of occurrences in the data as compared to {A, R} which are hard to distinguish from each other and have very few occurrences. It should also be noted that for {S, M} word level annotation is more important where as for {A, R} message level annotation makes more sense.

3.2 Stages and Phases Annotation

3.2.1 Annotation Schema

The aim of this annotation is to understand how the conversation unfolds in a health coaching dialogue. Stages and phases help to capture the coaching tasks and sub tasks being performed throughout the communication dialogue respectively. The higher tier is composed of stages; *goal setting* and *goal action*. Stages are composed of phases. Each message can be annotated as *goal setting*, *goal action*, or none. Phases are optional.

3.2.2 Annotation Results

Similar to SMART annotation, two annotators annotated four previously unseen patients for stages and phases. The kappa for each individual category is shown in Table 2, where 'gs' stands for *goal setting*, 'ga' stands for *goal action*, 'I' stands for *goal identification*, 'R' for *goal refining*, 'A' for *anticipate barrier*, 'S' for *solve barrier*, and 'F' for *follow up*. Other categories were not present in the data we annotated for agreement. *Goal action* stage with *anticipate barrier* phase has kappa value of zero as one of the annotators marked one message under this category, whereas the other annotator did not. The kappa statistics for all the categories together came out to be 0.93.

In total, 398 messages were annotated for IAA. There were 33 messages with 'gs_I' tag, 33 with 'gs_R', 49 with 'gs_A', 3 with 'gs_S', 1 with 'ga_A', 2 with 'gs_S' and 45 with 'ga_F'. The high value of kappa can be attributed to the fact that while following SMART criteria for goal setting, the stages/phases are bound to occur in a particular order and therefore, are easy to annotate.

4. DISCUSSION AND FUTURE WORK

In this paper we discussed our data collection process, SMART goal annotation schema, and stages and phases annotation schema. We plan to use these annotations to train a classifier that can extract and summarize the goal set by the patient. We are currently working on a pipeline that will first detect all the possible SMART goal features in the conversation, then detect the stages/phases in the conversation and then finally combine these two to extract the correct SMART goal. For example, if a goal was set during the goal identification phase, but then due to some barrier the goal changed in the goal negotiation phase, the summarization model should summarize the updated goal.

5. ADDITIONAL AUTHORS

Aiswarya Baiju (Department of Computer Science, University of Illinois at Chicago, email: `abaiju2@uic.edu`) and Rafe Davis (Institute for Health Research and Policy, University of Illinois at Chicago, email: `rdavis5@uic.edu`).

6. REFERENCES

[1] T. Bickmore, R. Asadi, A. Ehyaei, H. Fell, L. Henault, S. Intille, L. Quintiliani, A. Shamekhi, H. Trinh, K. Waite, et al. Context-awareness in a persistent hospital companion agent. In *International Conference on Intelligent Virtual Agents*, pages 332–342. Springer, 2015.

[2] T. J. Bovend'Eerdt, R. E. Botell, and D. T. Wade. Writing smart rehabilitation goals and achieving goal attainment scaling: a practical guide. *Clinical rehabilitation*, 23(4):352–361, 2009.

[3] G. T. Doran. ThereâĂŹsa smart way to write managementâĂŹs goals and objectives. *Management review*, 70(11):35–36, 1981.

[4] K. Kivelä, S. Elo, H. Kyngäs, and M. Kääriäinen. The effects of health coaching on adult patients with chronic diseases: a systematic review. *Patient education and counseling*, 97(2):147–157, 2014.

[5] G. Rizzolatti and L. Craighero. The mirror-neuron system. *Annu. Rev. Neurosci.*, 27:169–192, 2004.