Conversation-Based Medication Management System for Older Adults Using a Companion Robot and Cloud

Zhidong Su, Fei Liang, Ha Manh Do, Alex Bishop, Barbara Carlson, and Weihua Sheng

Abstract—**Memory loss is a part of normative aging. Many older adults commonly forget to take prescribed medication, which can have an adverse effect on health. Therefore, it is important to provide older adults with a medication reminder service. There are several mobile apps and devices capable of reminding medication, but their user interfaces and operations are usually unfriendly to seniors. Considering the above demands and shortcomings, we proposed a conversation-based medication management system (CMMS). The CMMS uses a companion robot and the cloud to create medication reminders and check medication adherence. We implemented the CMMS in our ASCC (Advanced Sensing, Computation and Control) companion robot. To evaluate the CMMS, we tested our system through the mobile app end with 23 human subjects and the robot end with 15 human subjects. The feedback from the post-test survey shows that the convenience, usefulness and total rating of our system is 8.217, 8.696 and 8.478 out of 10 from the mobile app end and 9.000, 8.933 and 8.533 out of 10 from the robot end, respectively. The System Usability Scale (SUS) score of our system is 81.333 from the robot end users, which means the participants had a high satisfaction level when using the system.**

Index Terms—**Health care management, social HRI, companion robot, elderly care.**

I. Introduction

With an increase in population aging, more and more older adults need to be taken care of by the younger generation [1]. The elderly care problem is becoming more and more serious. Older adults tend to have a deteriorating memory and suffer from age-related diseases as they get older. Failed medical adherence can have an adverse effect on health, including premature death in old age [2]. Several mobile apps [3] provide the medication reminder function, but the user interface on a smartphone or tablet is unfriendly to seniors. Almao *et al.* [4] evaluated the usability and accessibility of 22 mobile apps. Those apps were designed for older adults and popular in the domains of health (including medication reminder apps), emergency, assistance and so on. Their experimental results show that out of the 22 evaluated apps, none of them has a high level of accessibility, especially for older adults. The main problem is the lack of consideration for the visual degradation in older adults with respect to text resizing and zooming. In addition, most of the apps lacked proper instructions and options to ease the data input process for older adults. Some devices like the smart medicine boxes [5] are used to store the pills, but those devices are bulky and unfriendly for older adults to use. While several existing companion robots can provide the functions of natural language chatting, home appliance control, taking photographs, collecting health data, etc., it would be desirable to add a function of medication reminder to the robots.

Companion robots are usually equipped with a microphone or a camera, sometimes even a RGB-depth camera, which means that the robots have the potential to interact with the older adult through robot vision or audition. Some of them are able to carry out tasks like cognitive orientation assessment [6], companionship [7], pain evaluation, mood detection and loneliness detection through verbal conversation [8], which improve the well-being of older adults [9]. We conducted a focus group in 2019. There were 40 participants in this focus group, including doctors, nurses, and caregivers from Stillwater and Oklahoma City areas. They provided various suggestions on the robot functionalities and features related to elderly care. One of their suggestions is to utilize verbal conversation to create medication reminders for older adults. Speech is most commonly used by humans to communicate with each other. If the robot can communicate with humans through verbal conversation in addition to a touch screen or buttons, the relationship between elderly people and social robots will be well formed. Instructions like taking medicines, exercising, cognitive training will be more effective through verbal communication. We believe that compared to mobile apps companion robots capable of talking

Manuscript received October 15, 2020; accepted February 3, 2021. Date of publication February 24, 2021; date of current version March 17, 2021. This letter was recommended for publication by Associate Editor X. Zhong and Editor J. Yi upon evaluation of the reviewers' comments. This work was supported in part by the National Science Foundation (NSF) under Grants CISE/IIS 1910993, EHR/DUE 1928711 and in part by the Open Research Project of the State Key Laboratory of Industrial Control Technology, Zhejiang University, China ICT20026. *(Corresponding author: Weihua Sheng.)*

Zhidong Su, Fei Liang, and Weihua Sheng are with the School of Electrical and Computer Engineering, Oklahoma State University, Stillwater, OK 74078 USA (e-mail: zhidong.su@okstate.edu; fei.liang@okstate.edu; weihua.sheng@okstate.edu).

Ha Manh Do is with the Louisville Automation & Robotics Research Institute (LARRI), University of Louisville, Louisville, KY 40292 USA (e-mail: ha.do@okstate.edu).

Alex Bishop is with the Human Development and Family Science Department, Oklahoma State University, Stillwater, OK 74078 USA (e-mail: alex.bishop@okstate.edu).

Barbara Carlson is with the Fran and Earl Ziegler College of Nursing, University of Oklahoma Health Sciences Center, Oklahoma City, OK 73117 USA (e-mail: Barbara-Carlson@ouhsc.edu).

Digital Object Identifier 10.1109/LRA.2021.3061996

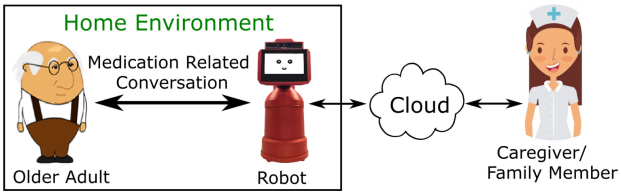

Fig. 1. Conversation-based medication management system for older adults using companion robot and cloud.

and listening can provide an easier and more elder-friendly way to create medication reminders and record medication history for older adults. In addition, the connection between the older adults and the caregivers is also important. If they can be connected through cloud, it will provide a convenient way for caregivers to help create reminders and monitor older adults' medication adherence.

In this letter, we proposed a conversation-based medication management system (CMMS) using a companion robot and cloud. Such a system is lacking in existing companion robot systems. As shown in Fig. 1, the older adult can launch a conversation with the companion robot to create a medication reminder and record the medication history. The companion robot can also start a conversation to check medication adherence. Additionally, caregivers and family members are connected to the companion robot through cloud to assist seniors to create a reminder and monitor the medication adherence status. The main contributions of this letter are three folds. First, the proposed CMMS combines mobile app and conversation based human robot interaction to deliver an elder-friendly medication reminder experience to the users. The CMMS connects older adults with their family members and caregivers through the cloud. This is one of the first in this kind of medication management system. Second, the proposed CMMS is able to close the loop by proactively asking the user regarding the medication compliance and history and recording the exact medication information. Third, we conducted human subject test and evaluated the performance of the CMMS, and the preliminary results are very promising with an overall rating of 8.53 out of 10.

The rest of this letter is organized as follows. Section II introduces the related work. Section III details the proposed method. Section IV gives the experimental results. Section V concludes our work and discusses the future work.

II. Related Work

Robotic technology has been developed to help older adults. Shishehgar *et al.* [10] conducted a literature review which examines research on robots for elderly care. In their studies, nine types of robots, including companion, manipulator service, telepresence, rehabilitation, health monitoring, reminder, entertainment, domestic, and fall detection/prevention robots, are identified to solve age-related problems. Do *et al.* [11] proposed a smart home architecture which integrates a companion robot, sensor networks, a mobile device and cloud servers to monitor older adults' activities and provide assistance. The work in [12] and [13] enabled a robot to detect emergency situations like falls. In order to improve older adults' experience when speech recognition fails during the dialogue, Lio *et al.* [14] developed a twin-robot dialogue system to take initiative in the dialogue by asking the user various questions.

Many older adults take medicine on a daily basis. When medicine is not taken as prescribed, it can put the older adult's life at risk. The academic community has done some research on medication reminder. Mohammed *et al.* [15] developed a mobile app for older adults to create medication reminders. However, the operation is not convenient for the older adults with vision problems or unfamiliar with smartphone operations. Baranyi *et al.* [16] proposed a system to remind older adults to take medicine and measure blood pressure. They used a mobile app to input reminder information, and the different colors of the light connected to the app to remind the older adult to take medicine or measure blood pressure. The drawback of this system is that it cannot provide exact medication information to remind the older adult. Bai *et al.* [17] built a medicine box and Mondol *et al.* [18] proposed a wearable wrist device to carry out medication reminder. However, those devices are not flexible and cannot provide the necessary functions. Zheng *et al.* [19] proposed a dialog system to collect medication information. Their dialog system obtains the necessary information in multiple rounds of conversation. As a result, the conversation system takes a long time to complete one medicine recording and it ignores the important function of reminding the older adult to take medicine. Florence [20] is a text-based chatbot providing medication reminder and health tracking functions. However, it does not offer audio input and cannot extract information from oral expression. Only users can create reminders while caregivers cannot.

Previous research has explored different methods to help older adults. However, when it comes to create reminders, those methods are not convenient for older adults and largely ignore the connection between the older adults and their families and caregivers. Therefore, we proposed a conversation-based medication management system. In this system, older adults can create reminders in an elder-friendly way while the robot is able to proactively check the medication compliance. The connections between the older adults and their families and caregivers are also realized through a mobile app and cloud. Therefore, they can help create reminders and monitor older adults' daily adherence.

III. Methodology

A. Software Architecture

In order to build the essential functions to manage medication in an efficient way, the robot software architecture is proposed as shown in Fig. 2. This system has two human-machine interfaces: 1) human-robot interface for older adults, 2) caregiver interface using a mobile app.

For the human-robot interface, the system provides a voice-based interaction. For the voice-based interaction, the voice captured by the microphone is processed by the Automatic Speech Recognizer (ASR) module to get the text output. The Natural Language Understanding (NLU) module is used to understand user's intent and acquire entities. Based on the text, intent

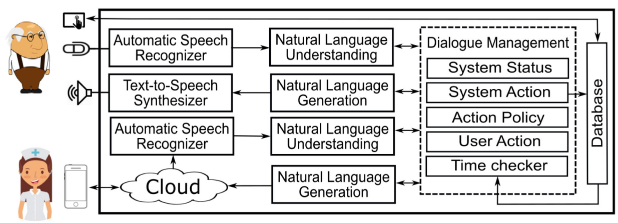

Fig. 2. The software architecture of the proposed system.

and acquired entities, the Dialog Management (DM) module is employed to manage system actions to control the dialog flow. Natural Language Generation (NLG) is used to generate utterance to respond to the user. The Text-to-Speech Synthesizer (TTS) unit synthesizes speech and plays it to older adults.

For the caregiver interface, a social networking app named Telegram [21] is employed to provide a communication platform for caregivers. This app is connected to the database through cloud. The conversational functions used in the human-robot interface is deployed to Telegram to implement a conversational chatbot. Therefore, caregivers can manage patients' medication through the conversation with Telegram chatbot.

B. Human-Robot Conversational Interface

A human-robot conversational interface is developed and deployed in the ASCC companion robot to assist older adults. In the human-robot conversational interface, both the older adult and the robot can initiate a conversation proactively to manage medication. Older adults can use the wake-up word "Hey Elsa" to initiate a conversation with the robot. We employ Google cloud services to implement the ASR and TTS modules. For the NLG module, we use a rule-based method which is more suitable than the deep learning-based method in our scenario due to the lack of domain training data. This section mainly focuses on the NLU and DM module.

The NLU module takes the output of the speech recognition module and generates the corresponding intent of the user utterance and necessary parameters. For example, when the user utterance is "Remind me to take penicillin for 3 ml," it can be parsed into *create a reminder* intent with *medicine* (*medicine* = "penicillin") and *dosage* (*dosage* = "3 ml") slots. Therefore, the NLU module should include two parts: intent recognition and slot filling. Intent recognition is used to distinguish intents such as creating a new reminder or recording a medicine history or launching other functions of the system. Slot filling is to extract necessary entities related to medication reminder or recording.

1) Natural Language Understanding:

a) $Intent\ Recognition$: For intent recognition, various techniques and approaches can be applied, such as Naive Bayes classifier [22], K-nearest neighbors (KNN) [23], support vector machines (SVM) [24], decision tree [25] and deep learning methods (Recurrent neural network and convolutional neural network). In this system, we utilize the fastText neural network [26] to implement intent recognition. It is a shallow neural network which achieves similar performance as deep neural networks but the training and inference speed is faster. It has 3 layers, input layer, hidden layer and output layer. The input is the combination of word embedding and n-gram embedding. The character embeddings make the model achieve better generalization ability. The output layer generates the confidence of each class. We defined a total of 10 intent classes, including the main intents like Create a reminder, Record medication history, Check medication, and other intents that are used in the Dialog Management and Caregiver Interface like Stop, Deny, Introduce, Help and so on. The number of training epoch is set to 25 and the word embedding dimension to 50. We set the threshold of the recognition confidence to 0.80 to improve the robustness of the system. We collected 450 utterances to train the intent recognition model and 50 utterances to test the model while achieving an accuracy of 98%.

b) $Slot\ Filling$: Slot filling, or Named Entity Recognition (NER), can be regarded as a sequence labeling task. Given a sequence of input utterance tokens, the slot filling algorithms output the corresponding labels of the tokens. The labels can be the person's names, product terminology, locations, etc. The state-of-the-art method like BiLSTM-CRF [27] needs a large amount of labeled training data. However, we do not have enough training data in our scenario. Therefore, we utilize the rule-based method combined with a predefined dictionary to extract the necessary entity information. In our system, we focus on five entities: "$dosage$", "$start\_time$", "end_time", "$remind\_time$" and "$medicine\_name$". If the utterance is classified as create a reminder or record a medication history, then the slot filling method will be used to extract entities for that utterance.

Set $sentence = \{w_1, w_2, \ldots, w_i, \ldots, w_n\}$ represents the user utterance tokens, where n is the total number of the tokens.

For the "$dosage$" entity, we firstly build a dosage dictionary, $dosage\_dic$ = ['pill,' 'gram,' 'milligram,' 'liter,' 'milliliter,' 'cup,' 'drop,' 'spoon,' 'pack,' 'tablet,' 'ml,' 'l,' 'mg,' 'g']. Based on the English expression patterns, if the Part-of-Speech (POS) feature of w_i is numeral and $w_{i+1} \in dosage\_dic$, then we can judge that (w_i, w_{i+1}) is the dosage information. According to the definition and analysis, we can use the first-order predicate logic theory to represent this reasoning process. The predicate logic reasoning formula (1) can be used to extract the dosage information.

$$\begin{aligned} &POS(w_i, Number) \cap Contains(w_{i+1}, dosage\_dic) \\ &\quad => Dosage(w_i, w_{i+1}) \end{aligned} \tag{1}$$

where $POS(a, b)$ checks if the POS feature of token a is b. If the value of $POS(a, b)$ is TRUE, it means the POS feature of token a is b. $Contains(a, A)$ checks if the set A contains token a. $Dosage(w_i, w_{i+1})$ is the logical consequence of predicate logic reasoning, and the True value indicates that (w_i, w_{i+1}) is the dosage information.

For the $time$, we built a $time\_extractor$ python package to extract information. For the $time\_extractor$, we firstly get the current time, then find the time information (year, month, day, week, time point) and normalize it to digital time, calculate with the current time and get the target time. If there are words like "start", "begin", "end" and "finish", we align the extracted time to the "$start\_time$" and "end_time". If not, we align the extracted time to "$remind\_time$".

For the "$medicine\_name$", we combined the dictionary-based approach with the rule-based approach. The dictionary $medicine\_dic$ is built by collecting common medicines. We collected 951 medicines from drug information websites [28]. If the dictionary does not contain the medicine, then a rule-based approach is used. For the rule-based approach, firstly a Python package NLTK is used to obtain the POS features of the utterance. The medicine name is extracted using the template "Noun", "Noun + Noun", "Proper Noun", and "Adjective + Noun" as a sliding window to traverse the POS result, which needs to be confirmed by the user. If confirmed, then this new medicine is added to the dictionary. The POS results of some dosage and time information are "Noun". Therefore we firstly extract and remove the dosage and time information in the utterance and then utilize the newly generated utterance to obtain the medicine. The newly confirmed medicine will be added to the dictionary to improve the recognition performance. The predicate logic reasoning formula (2), (3), (4), (5) and (6) can be used to extract the medicine information.

$$Contains(w_i, medicine\_dic) => Medicine(w_i) \tag{2}$$

$$\begin{aligned} &POS(w_i, Noun) \cap \neg POS(w_{i-1}, Noun) \\ &\cap \neg POS(w_{i+1}, Noun) => Medicine(w_i) \end{aligned} \tag{3}$$

$$\begin{aligned} &POS(w_i, Noun) \cap POS(w_{i+1}, Noun) \\ &\qquad => Medicine(w_i, w_{i+1}) \end{aligned} \tag{4}$$

$$POS(w_i, ProperNoun) => Medicine(w_i) \tag{5}$$

$$\begin{aligned} &POS(w_i, Adjective) \cap POS(w_{i+1}, Noun) \\ &\qquad => Medicine(w_i, w_{i+1}) \end{aligned} \tag{6}$$

where $Medicine(a)$ checks if the token a is the medicine information. If the value of $Medicine(a)$ is TRUE, it means the token a is the medicine information.

Sometimes one utterance can include more than one medicine names. Different medicines may share the time information. We measure the number of words N among the medicine names. If N is one or two between two medicines, then we can share the information for the two medicines. If not, we align the dosage according to the position of those information.

2) Dialog Management: The dialog management is used to manage the dialog flow. In our system, it has five parts: $system$ $action$, $system$ $status$, $action$ $policy$, $user$ $action$ and $time$ $checker$.

a) System Status: According to the task of the reminder system, we defined 5 system statuses, which are shown in the **System status** column of Table I. ss_1 and ss_2 are the intent recognition results. ss_3 means the robot asks the user about the medicine status when it's time to take medicine. If all slot values are filled, the current system status will include ss_4, if not, the current system status will include ss_5.

b) System Action: The main goal of the dialog system is to acquire all the slot values to store the information. We defined 7 system actions, which are shown in the **System action** column of Table I. To create a new reminder, the system will ask all slots information using action sa_1. To record a medication history, the system needs to get the missing information such as dosage and time information using action sa_2. Different intents store the slots information in different databases using action sa_3 and sa_4. When the user didn't take the medicine at the remind time, then the robot carries out action sa_6 to remind later.

TABLE I
STATUS AND ACTION

System status	System action	User action
ss$_1$: create a reminder	**sa$_1$**: ask missing slots for reminder	**ua$_1$**: response
ss$_2$: record a taken medicine	**sa$_2$**: ask missing slots for record	**ua$_2$**: no response
ss$_3$: ask user	**sa$_3$**: store slots in reminder database	**ua$_3$**: stop
ss$_4$: full slots	**sa$_4$**: store slots in record database	**ua$_4$**: deny
ss$_5$: not full slots	**sa$_5$**: stop	**ua$_5$**: confirm
	sa$_6$: remind T minutes later	
	sa$_7$: ask if taken medicine	

c) User Action: The user action can be classified into 5 classes, which are shown in the **User action** column of Table I. User action is obtained from the ASR recognition result. ua_1 means ASR module outputs text, ua_2 means ASR module outputs nothing, ua_3 means ASR module output has the meaning of stop. ua_4 means the user denies the question posed by the robot. ua_5 means the user confirms the question posed by the robot.

d) Time Checker: The time checker is used to check the reminder database. If the reminder time is between the start time and the end time of the reminder duration, then the robot synthesizes a speech using the medicine name and dosage information to remind the user.

e) Action Policy: The action policy controls the system action according to the system status and user action. Table II shows the designed action policy. The system status is initialized by the intent category or time checker. For example, if the user input is "I need to take flu medicine for 3 pills at 8 am and 1 pm", the intent is classified as creating a new reminder, while the system status is initialized to ss_1. The slot filling detects the missing slots which are "$start\_time$" and "end_time", then the system status becomes ss_1 and ss_5. According to the action policy, the next system action is sa_1, ask for the missing "$start\_time$" and "end_time". When the time information is complete, the system status will change to ss_1 and ss_4, according to the action policy, the next system action is sa_3 and sa_5, store the slot values and stop the conversation.

C. Caregiver Interface

Considering the situation that the older adults might forget to create medication reminders for their new prescriptions or forget to take medicine even if there is a reminder, which will seriously affect medication compliance and their health, we provide a cloud service as a complementary tool, namely virtual chatbot built in a mobile app, for caregivers or family members to assist older adults. This app has four functions: 1) create new reminders, 2) help patients record medication history, 3) check

TABLE II
ACTION POLICY

```
# Part 1: User talks to robot about medication information
If ss1 or ss2, Then
  Use formula (1) to (6) and time_extractor to
  extract information from user input;
  Update system state;
  While True
    If ss1 and ss5, Then
      sa1;
      If ua1, Then
        Use formula (1) to (6) and time_extractor to
        extract information from ua1;
        Update system state;
      If ua2, Then Continue;
      If ua3, Then sa5, Break;
    If ss1 and ss4, Then sa3, sa5, Break;
    If ss2 and ss5, Then
      sa2;
      If ua1, Then
        Use formula (1) to (6) and time_extractor to
        extract information from ua1;
        Update system state;
      If ua2, Then Continue;
      If ua3, Then sa5, Break;
    If ss2 and ss4, Then sa4, sa5, Break;

# Part 2: Robot asks user about medication information
If ss3, Then
  sa7
  If ua4 or ua2, Then sa6;
  If ua3, Then Return;
  If ua5, Then
    sa2;
    Use formula (1) to (6) and time_extractor to
    extract information from ua1;
    Update system state;
    While True
      If ss4, Then sa5, Break;
      If ss5, Then sa2;
        If ua4 or ua2, Then sa6;
        If ua2, Then Break;
        If ua1, Then
          Use formula (1) to (6) and time_extractor to
          extract information from ua1;
          Update system state;
```

daily medication adherence and 4) modify reminders and history records. We provide a mixed interface for caregivers to fully take advantage of the attribute of the chatbot and smart phones. Caregivers can use the chatbot to create reminders through voice and text. They can also use inline buttons and option buttons to control the dialog flow, check and modify records.

An open source mobile app Telegram is employed to build the connection between the caregivers and our companion robot through cloud. We registered a robot account, which is provided to caregivers to communicate with the robot. By using the $token\_id$ of the robot account generated by Telegram, the companion robot is authorized to obtain the messages sent by caregivers through Telegram. The message type can be text or audio. We build another human-robot conversational interface for caregivers, which is similar to the interface proposed in Section III(B). We employ the multi-threaded programming method to reconstruct the code to make it available for multiple users at the same time. Therefore, the caregivers can use the chatbot at the same time to help more than one patients. The patients can also acquire help from different caregivers. The wake-up function is removed because the user input can be regarded as the wakeup command. "Check" and "Modification" intents are built by utilizing the neural network mentioned in 1) of Section III(B). We also modify the robot response and conversation flow of the human-robot interface to enable the caregiver interface to fit the app-based conversation well.

IV. Experimental Evaluation

We implemented the medication management system in our companion robot. We also open sourced the code.[1] This section presents the experiments and evaluations of this system.

A. Reminder Creation Effectiveness Evaluation

In order to test the effectiveness of this system, we compared the time that is used to create reminders by existing apps and our CMMS.

1) Experimental Setup: We installed two mobile apps named Medisafe and Pill Reminder in the smart phone which rank in the top in the category of "Medication Reminder" in the app store. Nine male and one female volunteers are recruited to create reminders. Two of them are at the age between 31 and 35. Eight of them are at the age between 26 and 30. The subjects were firstly asked to familiarize themselves with the two apps and our chatbot through Telegram app. Then they used the three methods to create reminders. We asked the participants to create reminders through verbal conversation with our chatbot. Table III shows the reminders that we need to create in the experiments. We asked the users to create the 6 reminders in the order from Reminder 1 to Reminder 6 using CMMS, Medisafe and Pill Reminder sequentially. The length of medicine name, the difference of dosage units, the number of times required to take medicine in a day and the time to take medicine affect the time required to create a reminder. The six reminders have different complexities in order to cover different situations. Besides, the reminding time, start date and end date of Reminder 1 and 2 are the same, mainly because the patients need to take more than one medicine and the only difference is the medicine name and dosage.

2) Results and Analysis: The time used to create the 6 reminders one by one through the 3 methods is shown in Table V. Each time is obtained from the average of the ten users' time consumption. We also conducted the independent sample t-test to check if the averages of the ten users' time consumption between Medisafe and CMMS (M-C), and between Pill Reminder and CMMS (P-C) are significantly different when creating the 6 reminders. Table IV shows the independent sample t-test results. M-C means the independent sample t-test was conducted between Medisafe and CMMS. The results indicate the following:

- As can be seen from Table IV, all the p-values are smaller than 0.05; six of them are below 0.001 and two of them are below 0.01. This means that there is a significant difference in the averages of the ten users' time consumption when

[1] https://github.com/suzhidong/medication_reminder

TABLE III
REMINDER INFORMATION

Reminder	Medicine	Dosage	Reminder time 1	Reminder time 2	Reminder time 3	Start date	End date
1	Flu medicine	3 pills	08:00:00	13:00:00	19:00:00	2020/9/22	2020/9/30
2	Amoxicillin	5 pills	08:00:00	13:00:00	19:00:00	2020/9/22	2020/9/30
3	Cold medicine	4.5 milligrams	07:30:00	15:45:00	None	2020/9/22	2020/9/27
4	Penicilin	3.6 mg	06:00:00	None	None	2020/9/21	2020/10/5
5	Ibuprofen	6 ml	10:00:00	23:00:00	None	2020/9/21	2020/9/27
6	Aspirin	2 pills	07:00:00	14:00:00	20:00:00	2020/9/22	2020/9/29

TABLE IV
INDEPENDENT SAMPLE t-TEST RESULTS

Reminder	Categories	t	df	sig. (2-tailed)	MD
1	M-C***	5.542	9.546	0.000	32.549
1	P-C***	8.048	9.850	0.000	38.158
2	M-C***	4.619	10.636	0.001	26.745
2	P-C**	4.003	9.331	0.003	49.938
3	M-C**	4.116	10.389	0.002	26.587
3	P-C***	4.855	10.049	0.000	35.819
4	M-C*	2.137	18.000	0.047	15.019
4	P-C*	2.651	18.000	0.016	12.730
5	M-C*	2.580	9.121	0.029	34.160
5	P-C***	8.957	18.000	0.000	18.486
6	M-C*	2.633	9.105	0.027	33.241
6	P-C***	5.038	18.000	0.000	17.060

[1] $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$.
[2] df: degree of freedom, MD: Mean Difference.

TABLE V
THE AVERAGE TIME REQUIRED TO CREATE REMINDERS (SECOND)

Reminder	CMMS	Medisafe	Pill Reminder
1	25.962	58.511	64.120
2	28.421	55.166	78.359
3	34.850	61.437	70.669
4	32.506	47.525	45.236
5	31.247	65.407	49.733
6	32.481	65.722	49.541

creating the 6 reminders between our CMMS and another two apps.

- As can be seen from Table V, the mean time consumption to create Reminder 1 through CMMS is 25.962 seconds, which is 32.549 seconds less than Medisafe and 38.158 seconds less than Pill Reminder. Similarly, the mean time consumption to create Reminder 2 - Reminder 6 is also less than Medisafe and Pill Reminder. Those statistical results indicate that our proposed CMMS is faster to create reminders.

Another advantage of the CMMS is that it can create multiple reminders at one time and share the information like $remind\_time$, $start\_date$ and end_date. Therefore, the users do not need to repeat that information, which improves the efficiency. For Reminder 1 and 2, we also asked the participants to create them together by including the two medicines in one utterance so that the two reminders share the time information. The test result shows that the average time to create the two reminders together is 37.420 seconds, which is 16.963 seconds faster than creating them one by one.

TABLE VI
QUESTIONNAIRE

Q1: What is your professional background?
Q2: How convenient is it for you to manage patients' medication?
Not convenient at all: 1 — Very convenient: 10
Q3: How useful is it for you to manage your patients' medication?
Useless: 1 — Very useful: 10
Q4: Your overall rating for this reminder system.
Bad: 1 — Excellent: 10
Q5: Do you have any suggestions for this medication reminder?
Q6: Besides this function, are there any other important functions?
Q7: Do you have anything else to say regarding this reminder?

B. Mobile and Robot Interface Evaluation

We conducted two experiments to test the performance of the mobile and robot interface of the CMMS and asked users about their experience with the CMMS.

1) Experimental Setup: We recruited 31 users to test our system through Telegram. Considering that when testing the system, there would be different reminders from different users, we generated responses automatically for the reminders and send them to the corresponding users so that they could have a better understanding of how the entire system works. We also provided a demonstration video[2] to help them get started on the app. We asked the users to install Telegram in their phones to friend our robot and create at least 5 reminders for patients, check their daily adherence and modify records remotely. The users tested Telegram for 7 days to thoroughly experience the reminder function. In order to acquire users' feedback, we provided them a post-test questionnaire which is shown in Table VI.

We also recruited 15 users to test the performance of the robot end of the CMMS in our lab. We gave users a brief introduction about the overall system and how to use our robot. Two scenarios were tested. In Scenario 1, the participants tested the whole CMMS, including both the robot end and the caregiver end. A caregiver created reminders for the participants and showed the participants how he checked and managed users' medication. Our robot reminded the participants to take medicine and the participants told the robot about their medication adherence and record the medicine history. In Scenario 2, the participants tested part of the CMMS with the help only from the robot. The participants had to create reminders and record the medicine history by themselves. After the test, the participants were asked to provide a feedback based on their experience on the whole CMMS through the questionnaire shown in Table VI. In order to evaluate the usability of the CMMS, we used the System

[2] https://youtu.be/xM45QwP-ayw

TABLE VII
STATISTICAL RESULTS OF THE CONVENIENCE, USEFULNESS AND OVERALL RATING OF THE TELEGRAM END WITH DIFFERENT USERS

Metrics	Caregivers (n=8)		Non-caregivers (n=15)	
	Mean	Std. Dev.	Mean	Std. Dev.
Convenience	8.625	1.302	8.000	1.464
Usefulness	8.875	0.991	8.600	1.298
Overall rating	8.500	1.195	8.467	1.302

TABLE VIII
STATISTICAL RESULTS OF THE CONVENIENCE, USEFULNESS AND OVERALL RATING OF THE TELEGRAM END AND ROBOT END

Metrics	Telegram end (n=23)		Robot end (n=15)	
	Mean	Std. Dev.	Mean	Std. Dev.
Convenience	8.217	1.413	9.000	0.926
Usefulness	8.696	1.185	8.933	1.100
Overall rating	8.478	1.238	8.533	1.767

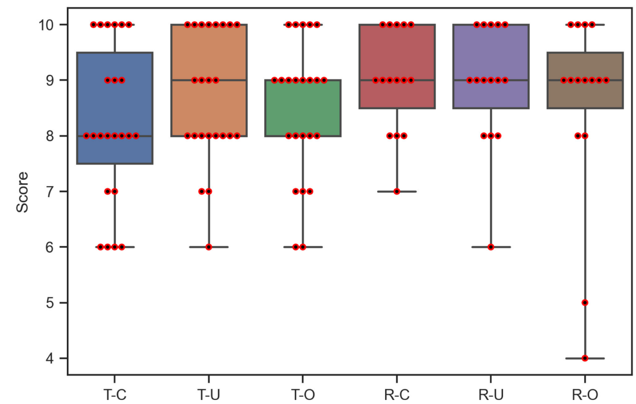

Fig. 3. The Telegram end and robot end box/swarm plot of convenience, usefulness and overall rating of CMMS.

Usability Scale (SUS) [29], including 10 items assessed with a 5-point Likert scale, to collect users' perception based on their experience on Scenario 1 and 2, respectively. We also asked the 15 users to install a pill reminder app from the App market on their cell phone and use it. After that, we asked them to give an anonymous feedback regarding which method they prefer to use, the app from the market or our CMMS.

2) Results and Analysis: For the Telegram end test, the experiment was conducted from October 1 to 14, 2020 and from December 20 to 30, 2020. We finally received 23 questionnaire responses out of 31 users while 8 users did not finish the questionnaire. Three of the 23 users are at the age between 41 and 55. Twenty of them are at the age between 26 and 35. Among the 23 responses, 8 responses are from formal caregivers and 15 responses are from non-caregivers. We checked the users' log file and found that some users prefer to use text input and some of them do not like typing. There are also some mixed input samples when it comes to the long sentences or complex words like "milligrams" and time information. About 91.30% users utilized all the functions of our system. They all created 2 or more reminders successfully. For the robot end test, the experiment was conducted from January 2 to 9, 2021. We received responses from all users. There are 3 females and 12 males. Three of them are at the age between 31 and 35. Twelve of them are at the age between 26 and 30. Due to the COVID-19 pandemic in U.S., we did not recruit older adults in our experiments.

Table VII shows the statistical results of convenience, usefulness and overall rating of the Telegram end with different users. Table VIII shows the same statistical metrics of the Telegram end and the robot end. Fig. 3 is the box plot with a swarm plot of user ratings for the three metrics from the Telegram end and robot end. T-C, T-U and T-O are the convenience, usefulness and overall rating metrics of the Telegram end. R-C, R-U and R-O are the three metrics of the robot end. The results indicate the following:

- For the Telegram end experiment, we observe from Table VII that the mean evaluation of convenience, usefulness and overall rating obtained from 8 caregivers and 15 non-caregivers are all above 8. The overall rating is 8.500 and 8.467 from caregivers and non-caregivers, respectively. It means from the perspective of professional caregivers and non-caregivers, the Telegram end of our system is acceptable. It is useful and convenient for them to manage and monitor older adults' medication.
- For the robot end experiment, we obtained a similarly good result as the Telegram end. As can be seen from Table VIII, the mean evaluation of convenience, usefulness and overall rating of the robot end is 9.000, 8.933 and 8.533, which is 0.783, 0.237 and 0.055 higher than that of the Telegram end, respectively. Besides, from Fig. 3 we can see that 93.3% (14 out of 15), 93.3% (14 out of 15) and 86.7% (13 out of 15) of the users graded the robot end with the convenience, usefulness and overall rating metrics above 8 and 73.9% (17 out of 23), 87.0% (20 out of 23) and 78.3% (18 out of 23) of the Telegram end ratings are above 8. More than half of the users graded both the Telegram and the robot end above 8, which means our CMMS is acceptable. In addition, the participants in the robot end experiment had a more thorough experience than the participants in the Telegram end experiment. The robot end users are the main target users of the CMMS. Therefore, they rated our system slightly higher and had better acceptance than the participants in the Telegram end experiment.
- We also acquired the feedback regarding which method they prefer to use. One out of 15 users chose the app. All the other 14 users chose our CMMS, which shows the user's preference for our system.

We received 15 SUS feedback from the robot end users regarding Scenario 1 and 2. For Scenario 1, the mean of SUS score is 81.333, the standard deviation is 13.980 and the median is 80.000. According to the SUS assessment criteria, it indicates that the participants had a high satisfaction level when using the whole system since it was perceived as easy to use. For Scenario 2, the mean is 75.500 and the standard deviation is 13.503. The mean of SUS score dropped by 5.833 because of the lack of help from caregivers, which also shows the importance of the CMMS's integrity. The score of 75.500 also means acceptable according to the SUS assessment criteria.

From the users' feedback, we learned that the user interface of the app is not accepted by all users. Some users suggest us to improve the Emojis. Another suggestion is that caregivers can provide the pictures of the medicines when creating reminders and show the picture in the robot face at medicine time. It would be useful for older adults to take the correct medicine. The effect of the medicine is another important issue for older adults, which prompts us to add medication effect monitoring into the robot in the future.

V. Conclusion and Future Work

In this letter, we proposed a conversation-based medication management system. In this system, older adults can talk to the companion robot to create reminders and record their medication history. Besides, caregivers and family members are connected to the older adults through cloud using a mobile app to manage their medication. We tested our system and the results show that our system has a good performance. However, there are still some limitations in our current evaluation. 1) The robot was evaluated in the lab environment only while real home environments are noisy. Therefore the interaction would be much messier and people might not hear a reminder; 2) Long-term evaluation is not considered in our study; 3) The robot side was not evaluated by the actual target of the system: older adults. In our future work, we will test our system in real settings and utilize computer vision methods to detect older adults' action of taking medicine and trigger the conversation to record medicine information. Besides, the effect of medicine is also important, we will use multi-modal data collected from microphones, camera and other sensors to detect older adults' physical and mental status after taking the medicine.

References

[1] J. Zhao and X. Li, "The status quo of and development strategies for healthcare towns against the background of aging population," *J. Landscape Res.*, vol. 10, no. 4, pp. 41–44, 2018.

[2] S. T. Charles, M. Mather, and L. L. Carstensen, "Aging and emotional memory: The forgettable nature of negative images for older adults," *J. Exp. Psychol.: Gen.*, vol. 132, no. 2, pp. 310–324, 2003.

[3] A. Alkandari, D. Aladem, S. Asaad, and S. Moein, "Medical care reminder for infants using android application," *J. Adv. Comput. Sci. and Technol. Res.*, vol. 7, no. 2, pp. 57–66, 2017.

[4] E. C. Almao and F. Golpayegani, "Are mobile apps usable and accessible for senior citizens in smart cities?" in *Human Aspects of IT for the Aged Population. Design for the Elderly and Technology Acceptance*, J. Zhou and G. Salvendy, Eds. Cham: Springer International Publishing, 2019, pp. 357–375.

[5] G. Gimpel, U. Varshney, and P. Ahluwalia, "Emerging it for medication adherence," *IT Professional*, vol. 18, no. 03, pp. 30–36, May 2016.

[6] F. Erivaldo Fernandes, H. M. Do, K. Muniraju, W. Sheng, and A. J. Bishop, "Cognitive orientation assessment for older adults using social robots," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, 2017, pp. 196–201.

[7] S. Saint-Aime, B. Le-Pevedic, D. Duhaut, and T. Shibata, "Emotirob: Companion robot project," in *Proc. RO-MAN 2007 - The 16th IEEE Int. Symp. Robot Hum. Interactive Commun.*, 2007, pp. 919–924.

[8] H. M. Do, W. Sheng, E. E. Harrington, and A. J. Bishop, "Clinical screening interview using a social robot for geriatric care," *IEEE Trans. Automat. Sci. and Eng.*, doi: 10.1109/TASE.2020.2999203.

[9] A. Scales, J. Loup, C. Juang, E. Sakai, F. Ma, and C. E. Gould, "Development and evaluation of educational materials regarding mental health mobile apps among older veterans," *Innov. Aging*, vol. 3, no. Suppl 1, pp. S329–S330, 2019.

[10] M. Shishehgar, D. Kerr, and J. Blake, "A systematic review of research into how robotic technology can help older people," *Smart Health*, vol. 7/8, pp. 1–18, 2018.

[11] H. M. Do, M. Pham, W. Sheng, D. Yang, and M. Liu, "Rish: A robot-integrated smart home for elderly care," *Robot. Auton. Syst.*, vol. 101, pp. 74–92, 2018.

[12] T. Sumiya, Y. Matsubara, M. Nakano, and M. Sugaya, "A mobile robot for fall detection for elderly-care," *Procedia Comput. Sci.*, vol. 60, pp. 870–880, 2015.

[13] P. Di *et al.*, "Fall detection and prevention control using walking-aid cane robot," *IEEE/ASME Trans. Mechatronics*, vol. 21, no. 2, pp. 625–637, Apr. 2016.

[14] T. Lio, Y. Yoshikawa, M. Chiba, T. Asami, Y. Isoda, and H. Ishiguro, "Twin-robot dialogue system with robustness against speech recognition failure in human-robot dialogue with elderly people," *Appl. Sci.*, vol. 10, no. 4, 2020.

[15] H. B. M. Mohammed, D. Ibrahim, and N. Cavus, "Mobile device based smart medication reminder for older people with disabilities," *Qual. Quantity*, vol. 2, no. 52, pp. 1329–1342, 2018.

[16] R. Baranyi, S. Rainer, S. Schlossarek, N. Lederer, and T. Grechenig, "Visual health reminder: A. reminder for medication intake and measuring blood pressure to support elderly people," in *Proc. IEEE Int. Conf. Healthcare Inform.*, 2016, pp. 432–438.

[17] Y. Bai and T. Kuo, "Medication adherence by using a hybrid automatic reminder machine," in *Proc. IEEE Int. Conf. Consum. Electron.*, 2016, pp. 573–574.

[18] A. S. Mondol, I. A. Emi, and J. A. Stankovic, "Medrem: An Interactive Medication Reminder and Tracking System on Wrist Devices," in *Proc. IEEE Wirel. Health*, 2016, pp. 1–8.

[19] J. Zheng, R. Finzel, S. Pakhomov, and M. Gini, *Spoken Dialogue Systems for Medication Management, Ser. Studies in Computational Intelligence*. Springer Verlag, Jan. 2020, pp. 119–127.

[20] Florence. Accessed: Jan. 12, 2021. [Online]. Available: https://florence.chat/

[21] Telegram. Accessed: Jan. 12, 2021. [Online]. Available: https://telegram.org/

[22] W. Dai, G.-R. Xue, Q. Yang, and Y. Yu, "Transferring Naive Bayes Classifiers for Text Classification," in *Proc. 22nd National Conf. Artif. Intell.*, 2007, pp. 540–545.

[23] O. Kramer, K-Nearest Neighbors. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 13–23.

[24] M. Mendoza and J. Zamora, "Identifying the Intent of a User Query Using Support Vector Machines," in *String Processing and Information Retrieval*, J. J. Karlgren Tarhio, and H. Hyyrö, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 131–142.

[25] A. J. Myles, R. N. Feudale, Y. Liu, N. A. Woody, and S. D. Brown, "An introduction to decision tree modeling," *J. Chemometrics*, vol. 18, no. 6, pp. 275–285, 2004.

[26] A. Joulin, E. Grave, P. Bojanowski, and T. Mikolov, "Bag of Tricks for Efficient Text Classification," 2016, *arXiv:1607.01759*.

[27] L. Luo *et al.*, "An attention-based BiLSTM-CRF approach to document-level chemical named entity recognition," *Bioinformatics*, vol. 34, no. 8, pp. 1381–1388, Nov. 2017.

[28] Drug Information Portal. Accessed: Jan. 12, 2021. [Online]. Available: https://druginfo.nlm.nih.gov/drugportal/

[29] A. Bangor, P. Kortum, and J. Miller, "Determining what individual sus scores mean: Adding an adjective rating scale," *J. Usability Stud.*, vol. 4, no. 3, pp. 114–123, May 2009.