State Machine-Based Multimodal Dialogue System for the Elderly Care Service*

Takatsugu Suzaki[1,*,†], Masayuki Numao[1,†]

[1]The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu, Tokyo, Japan

Abstract

Dialogue is a useful communication way between human and robot. Dialogue management is important to gather the user's intent, select a task to satisfy the user's request, and make a QA session to perform a specific task, all of which should be performed by natural conversation. We proposed a multi-scenario task-oriented dialogue system based on finite state machine(FSM). Our state machine has common state transitions and they enable users to write scenarios easily and flexibly. FSM based dialogue scenarios are suitable for common state transitions. Multimodal dialogue also requires to care the elderly for dementia diagnosis and daily conversation. We applied the proposed system to diagnoses dementia.

Keywords

multimodal dialogue, dialogue system, state machine, HDS-R, health care

1. Introduction

Dialogue systems are a flexible and easy way for humans and robots to communicate. Task oriented dialogue can be used to solve specific domain tasks (e.g. hotel reservation, guides, etc.)[1]. Spoken dialogue systems such as Siri, Alexais widely used. However, human conversation occurs not only through speech, but also through gestures and facial expressions[2]. In addition, in recent years, the field in dialogue system is mainstream by end-to-end approach, which enables flexible interaction with people[3]. However, this disadvantages are that it requires a large amount of data for training and that it is difficult to change input/output modules according to the user's environment. Therefore, we propose a rule-based and multimodal dialog system. Our system allows for natural conversations while it requires people to write scenarios. Although our system is a general dialogue system, the scenarios focus on the elderly care, specifically dementia diagnosis in this paper.

2. Social Responsible AI for Well-being

Recently, the number of patients is increasing due to an aging society. Early detection of dementia is one of important factors in terms of elderly care. Although specialists diagnose dementia, this is a problem that is increasing the burden on specialists in an aging society. To solve this problem, dialogue system requires to diagnose them instead of specialists.

HDS-R[4] and MMSE[5] are methods to diagnose dementia and have clearly defined questions, order, and conditions. The diagnostic scores can also diagnose the degree of dementia. Thus, rule-based dialogue system is suitable, but simple implementation of QA is inflexible and unfriendly. It cannot be deal with exceptional behavior.

```
<scenario name="hdsr">
    <sequential name="sequential_hdsr">
    <state name="q1">
      <request timeout="15">
        <slot name="hdsr.age"/>
      </request>
      <actions>
        <action event="text-to-speech">お年はいくつですか？</action>
      </actions>
    </state>

      ...

    </sequential>
</scenario>
```

Figure 1: Scenario example of HDS-R

3. Definition of Dialogue Scenarios

Domain-specific tasks rely on pre-defined rules based on finite state machines. A task is defined by a XML file defined uniquely. In the state, it is defined that slots requested from users and system's actions, and cooperation with KB and external services. Figure 1 and 2 show an example of scenario and dialogue in HDS-R.

AAAI 2023 Spring Symposia, Socially Responsible AI for Well-being, March 27–29, 2023, USA

*Corresponding author.

†These authors contributed equally.

s2131085@uec.ac.jp (T. Suzaki)

https://taka-suzaki.github.io/ (T. Suzaki)

Figure 2: Dialogue example of HDS-R

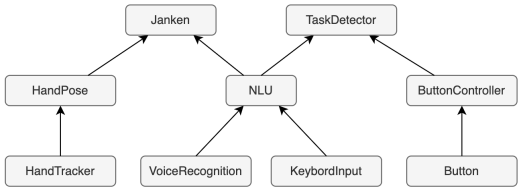

Figure 3: Graph of input modules

Table 1
Contents and results of the survey

No	Contens	Mean	STD
Q1	Did you achieve the objectives of the dialogue?	4.06	0.97
Q2	Were you able to interact with them in a natural, human way?	3.59	1.28
Q3	Did the dialogue go smoothly?	4.06	1.14
Q4	Would you like to interact with the robot again?	3.82	1.01
Q5	Did the robot perform as expected?	3.59	1.23
Q6	Comments	-	-

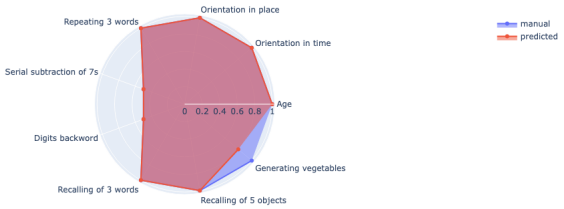

Figure 4: Results of HDS-R scoring by manual and system

4. Common State Transition

The system normally chooses the next state when the required slot is satisfied in that state, although we propose special state transition to have flexibility in state transitions.For example, "repeat" transition repeats the same state, "skip" transitions to the next state regardless of whether slots are filled, "cancel" terminates task's state machine.

5. Multimodal Interface

The role of input events is to perform slot filling and the role of output events is to generate actions. Figure 3 shows the graph of input modules. Multimodal process is possible by linking each module in a graph.

6. Experiments

Two experiments were carried out to evaluate success rate and fluency of dialogue and validity of HDS-R diagnose. First, a survey were conducted to evaluate success rate and fluency of dialogue on 18 people. The survey was rated on 5 scale. Table 1 shows contents and results of survey. The results showed high success rate was achieved while the fluency needs to be improved.

Second, our system diagnosed HDS-R to 15 subjects pretending to be elderly and compared the results of manual and system scoring. Figure 4 shows a radar chart of the results for a subject. The RMSE of the total score was 2 points, confirming that it can be scored as well as a human.

7. Conclusion

We proposed a multimodal dialogue system based on state machine. Our system enables to write scenarios easily and flexibly because of common state transitions. However, our system is inferior to the end-to-end approach in terms of dialogue fluency. In the future, we also plan to conduct a demonstration test to confirm the validity of the HDS-R scoring.

Acknowledgments

This work was supported by JSPS KAKENHI Grant Number JP20H04289 "Functional Independence Measurement System based on ADL Ontology for Aged Person"

References

[1] H. Chen, A Survey on Dialogue Systems: Recent Advances and New Frontiers, volume 19, 2018.

[2] Wanner, L. André, Kristina: A knowledge-based virtual conversation agent, in: Advances in Practical Applications of Cyber-Physical Multi-Agent Systems: The PAAMS Collection, Springer International Publishing, Cham, 2017, pp. 284–295.

[3] M.-H. YANG, J.-H. TAO, Data fusion methods in multimodal human computer dialog, Virtual Reality Intelligent Hardware 1 (2019) 21–38.

URL:https://www.sciencedirect.com/science/article/pii/S209657961930004X. doi:https://doi.org/10.3724/SP.J.2096-5796.2018.0010.

[4] S. Kato, H. Shimogaki, A. Onodera, [development of the revised hasegawa simplified intelligence assessment scale (hds-r)] kaitei hasegawashiki kannichino hyoka suke-ru (hds-r) no sakusei(in japanese), Ronen seishin igaku zasshi(Japanese Psychogeriatric Society) (1991).

[5] M. F. Folstein, "mini-mental state". a practical method for grading the cognitive state of patients for the clinician, J Psychiatr Res (1975).