Entering Appointments: Flexibility and the Need for Structure?

Karola Pitsch
University of Duisburg-Essen
Communication Studies
karola.pitsch@uni-due.de

Ramin Yaghoubzadeh
CITEC, Bielefeld University
Social Cognitive Systems
ryaghoubzadeh@uni-bielefeld.de

Stefan Kopp
CITEC, Bielefeld University
Social Cognitive Systems
skopp@techfak.uni-bielefeld.de

Abstract

Initial evaluation of human-agent-interaction reveals how the system's dialog strategies shape the user's input.

1 Introduction

Human-machine interfaces based on natural communication become increasingly important e.g. for assisting elderly or other people with specials needs in maintaining their daily routines. We have begun to develop a calendar and remainder application, which allows users to enter their appointments into a digital calendar by talking to an embodied conversational agent (ECA), which is presented on a large TV-screen alongside a weekly calendar. Users can interact freely with the system by using means of verbal (and in future: multimodal) communication. The system is set up to work autonomously and to extract information about date (D), time (T) and activity (A) of an appointment from the user's spontaneous speech. It uses dedicated modules for speech recognition and a multi-layered approach of securing understanding as shown in Yaghoubzadeh, Pitsch, and Kopp (2015). While the first trials are highly promising, we will be interested here in those cases in which wrong information is entered into the system as this raises questions about (i) understanding the linguistic and multimodal structure of the users' input and (ii) the ways in which the user utterances might be shaped through the conversational strategies deployed by the agent. We will present initial observations from a user study, which allows to observe the implications of a linear approach to extracting information from a user's input which was designed for the important benefit of allowing for flexibility in managing understanding and organizing repair activities. To which extent would this need to be completed with deeper information about structuring information?

2 Quantification: Success & Failure

Six senior citizens (age: 77 to 86 years) were asked to enter appointments in the digital calendar by talking freely to the ECA and using an A4 sheet with pictograms of potential events as inspiration. Quantitative analysis reveals …

- a high amount of correct entries with only one user (04) experiencing difficulties.
- if problems occur, they constitute a failure in the slot 'activity' or an abortion.
- the mean time of entering an appointment to be around 30'' for 2 seniors (which corresponds to young control users) and less than 60'' for 3 seniors.

User	Σ	Correct E	False Entry			Aban doned	Ø-Time
			D	T	A		
01	**9**	**5**		2*	3*		00:58
02	**8**	**7**			1		00:37
03a	**3**					3	
03b	**8**	**7**				1	00:36
04	**6**	**1**			3	2	02:31
05	**9**	**8**			1		00:57
06	**8**	**7**			1		00:51

Figure 1: Success/Failure of entering appointments. (*) = two failures in one entry. User 3a/b is the same person the trial of whom was interrupted.

Despite the encouraging results the question arises: How are the faulty 'activity'-entries produced in the interaction between user and system?

3 Exploration: Discursive strategies

Two cases from user 01 (`[A]`, `[B]`) are examined.

3.1 Extracting information dynamically

After the system (S) initiates the sequence, users (U) formulate an appointment using natural speech (often with specific gaze coordination):

```
01 S: do you have another appointment,    [A]
02 U: yes- I have another appointment
03 U: =on WEDnesday, (.) i want to go to
04 U: the restaurant, at 15 o'clock
```


Proceedings of the Int. Conference of the German Society for Computational Linguistics and Language Technology,
pages 140–141, University of Duisburg-Essen, Germany, Sep 30–Oct 2 2015.

The system shows its understanding of the user's input step by step both through talk and by highlighting the corresponding slot in the calendar (Fig. 2a, b). The user takes up this stepwise procedure either ratifying each step individually ([A] 07, 09, 12) or adding missing pieces of information dynamically ([B] 06, 08).

```
05 (2.0)                                          [A]
06 S: [#2a] then on wednesday, (.) right,
07 U: yes;
08 S: [#2b] then at 15 o'clock, (.) right,
09 U: yes;
10 S: good; there then is restaurant, (.)
11 S: right,
12 U: yes;
```

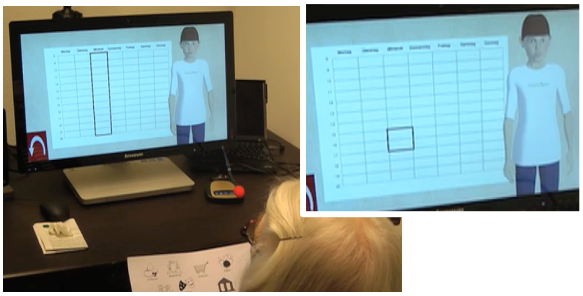

Figure 2a and 2b: Setup and highlighting slots

```
03 S: saturday, right;                            [B]
04 U: yes;
05 S: good;
06 U: at 13 o'clock
07 S: then at 13 o'clock, right,
08 U: yes; for lunch;
09 S: okay;
```

This flexible design of the dialog allows, in these cases, to extract all relevant information.

3.2 Expansions: Co-construction and the emergent nature of the user's input

Despite the availability of the correct information, the final calendar entries contain a faulty 'activity'. Analysis reveals these problems to result from users expanding their utterance incrementally adding more information about the activity. In [A], the agent ratifies the last entry (activity) (13: "good"). The user treats the ensuing 1.0-second pause and absence of information uptake on the screen as a moment to add an increment to his utterance (15:"in <city>"), which is repeated by the system (16) and ratified (17).

```
13 S: good; (1.0),                                [A]
14 (1.0)
15 U: in <city>
16 S: then there is <city>; right,
17 U: ye=hes;
```

In [B], the system has problems understanding "lunch" as the desired activity and seeks for clarification (10). The user again offers this piece of information (12) and the system picks it up (13). The system does not show immediate uptake in the calendar, which the user (in combination with the repeated questions) treats as a moment to expand her utterance in two steps (14-15). The system picks up the last piece of information "his vacations" (16) which is ratified (17).

```
10 S: okay; what did you plan there,              [B]
11 S: what did you plan there;
12 U: lunch;
13 S: then there is lunch; right,
14 U: lunch, yes:; (-) and a film screening;
15 U: (.) of his vacations;
16 S: then there is his vacations; right,
17 U: yes:;
```

This way, the agent's verbal conduct and the calendar display contribute to inviting the user to expand her utterance and to co-construct an input, which is difficult to handle for the system.

3.3 Overwriting entries & its acceptance

In both cases, the initially correctly extracted information (slot 'activity') is overwritten in the final entry by the information extracted from the users' expanded utterance ([A],19, [B],19-20).

```
18 S: okay; then I add the following              [A]
19 S: wednesday at 15 o'clock is <city>;
20 S: right,
20 U: .hh ye=hes;
```

```
18 S: okay; then I add the following              [B]
19 S: saturday at 13 o'clock there is his
20 S: vacations; right,
21 U: ye=hes;
```

While it has not been transparent to the user – when locally ratifying the 'activities' "<city>" ([A],15) and "his vacations" ([B],17) – that these would overwrite the 'activity'-information provided before ([A],10, [B],8) – it is noteworthy here that the user accepts both calendar entries.

4 Implications & Discussion

These observations point to investigating further (i) strategies for technical system to shape user conduct, and (ii) procedures for detecting completeness of information and managing the end of sequences. This will require also a discussion about linking such type of structural information with generic flexible computational approaches.

References

Ramin Yaghoubzadeh, Karola Pitsch and Stefan Kopp. 2015. Adaptive grounding and dialogue management for autonomous conversational assistants for elderly users. *Proceedings IVA 2015*.