ACCURACY OF TEXTUAL INTERFACES USING COMPARATIVE QUESTIONS TO ELICIT PERSONAL VALUE-RELATED INFORMATION

CSE3000 Research Project

Responsible Professor
Catholijn Jonker

Author: **Beatrice Vizuroiu**
B.Vizuroiu@student.tudelft.nl
EEMCS, Delft University of Technology, The Netherlands

Supervisor
Pei-Yu Chen

1. Background & Motivation

Behavior support applications (BSA):

- introduced to our daily lives extensively
- assist individuals in modifying their behaviors to achieve specific goals [1]

Values = intangible drivers that influence the way we form opinions and carry out actions [2]

Why do we need values from the users?

- gain mutual understanding between the system and the user throughout the conversation and personalize it

Why is personalization important?

- maximize agent's effectiveness for the goal of the user [3]

2. Research Question

How accurate is comparative questioning in eliciting personal value-related information through textual interfaces?

Related sub-question: *How usable is the system that employs comparative questioning through a textual interface?*

3. Methodology

Step 1: Create Scenarios & Questions

Step 2: Create Textual Interface via Chatbot

Step 3: User Study with 15 participants

4. Textual Interface

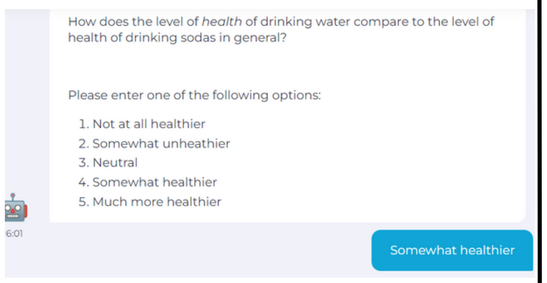

Figure 2: Textual interface

5. Behavior Tree as User Model

Context (Party) influences the preference profile (edges from the activities)

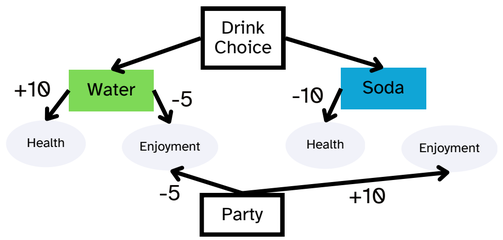

Figure 1: Part of Behavior Tree including preference profile of user for the firstscenario

6. Results

Common Particpants

Hamming Distance (HD) shows that interfaces utilizing comparative questioning require more manual changes of the behavior trees than the interfaces using in isolation questioning.

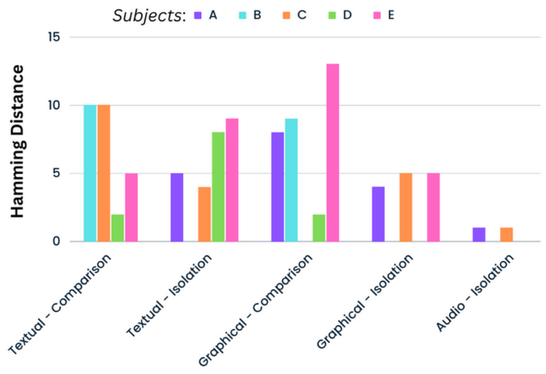

Figure 3: Hamming Distance

The absolute weight (AWC) between the computed and the manually entered value is greater for interfaces employing the comparative questioning context accentuating the result obtained from HD.

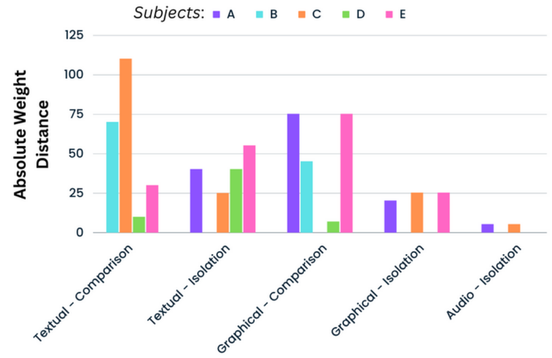

Figure 4: Absolute Weight

Individual Particpants

7 out of 15 participants indicated mistakes in the generated behavior tree resulting in 53% overall accuracy in constructing the user model.

The more values involved in a scenario, the less accurate the user model becomes.

Scenario 1	Scenario 2	Scenario 3	Scenario 4
93%	73%	80%	93%

Table 1: Accuracy of User Model per Scenario

78% average usability score indicates good usability [4].

Lowest Score	Highest Score	Average Score
52.5%	100%	78%

Table 2: Key Usability Scores Observed

Alternative Interfaces

A. Textual Interface and Comparative Questioning Context (this study)
B. Textual Interface and in Isolation Questioning Context
C. Graphical Interface and Comparative Questioning Context
D. Graphical Interface and in Isolation Questioning Context
E. Audio Interface and in Isolation Questioning Context

Tables 3 and 4 show that the condition of this study (A) required the least amount of manual changes (lowest HD). The standard deviation (SD) of AWC, suggests that the condition B would outperform this condition (A).

A			
HD		AWC	
Mean	SD	Mean	SD
0.8	0.98	9.66	14.2

Table 3: Statistics Summary of Own Condition

B				C			
HD		AWC		HD		AWC	
Mean	SD	Mean	SD	Mean	SD	Mean	SD
5.07	1.81	30.87	13.43	5.33	3.53	36.87	21.92
D				E			
HD		AWC		HD		AWC	
Mean	SD	Mean	SD	Mean	SD	Mean	SD
1.33	2.19	8	13.1	3.6	6.2	13.5	15.6

Table 4: Summary of Statistics for Alternative Conditions presented in section 4.3

7. Insights & Future Work

1. Usable system if deployed in current state (78%)
2. System not too accurate for multiple goals at once (53%) -> focus on one goal at a time
3. Second most accurate interface to build the user model from, but could benefit from variation with in isolation questioning
4. Size of sample (15) could lead to Type I errors results -> increase it

Related literature

[1] I. Kola, "Enabling social situation awareness in support agents," 2022.
[2] H. S. Shalom, "An overview of the Schwartz theory of basic values." Online readings in Psychology and Culture2, 1 (2012), pp. 1–26, 2012
[3] M. B. Van Riemsdijk, C. M. Jonker, and V. Lesser, "Creating socially adaptive electronic partners: Interaction, reasoning and ethical challenges." Proceedings of the 2015 international conference on autonomous agents and multiagent systems, pp. 1201–1206, 2015
[4] "System usability scale," june 13, 2023. [Online]. Available: https://blog.hubspot.com/service/systemusability-scale-sus