Designing Graphical User Interface To Elicit Personal Values
How does a graphical user interface that uses comparative questioning influence the accuracy of user value model?

Martynas Krupskis m.krupskis@student.tudelft.nl | Supervised by: Catholijn M. Jonker, Pei-Yu Chen

1. Background

- Behavioral support agents are systems designed to support individuals with behavioural challenges **(Kola et al., 2020)**.
- AI systems need to comprehend a user's current priorities, situation, and the impact of contextual factors on their behavior. To do that user model's have been developed, but changing nature of human values require run-time updates. **(C. Calman 2004)**
- Comparative questioning: "How much do you enjoy running compared to watching movies?"

2. Methodology

Sample size: 15 people of working age

- **Five conditions**: Graphical UI (User interface) & Comparative Questioning, Graphical UI & Isolated, Text UI & Comparative, Text UI & Isolated, Audio UI & Isolated.
- **Four behavioral challenges:** drinking more water, eating more nutritious food, exercise more, going to sleep earlier.
- Walkthrough of the prototype to generate a *value tree*, tweak each weight to build a *baseline tree, representing user's own perceived values.*
- Send SUS survey for a participant to complete.

3. Evaluation Metrics

$$Hamming\ Distance = \#\ of\ weight\ changes\ between\ G\ and\ B\ trees$$

$$AWC = \sum_{e_{i,j} \in E} |w_{i,j}^{baseline} - w_{i,j}^{generated}|$$

References

1. Ilir Kola*, Catholijn M. Jonker, M. Birna van Riemsdijk. M.B. 2020). Who's That?- Social Situation Awareness for Behaviour Support Agents. In: Dennis I... Bordini R. l espérance Y (eds) Ensineerine Multi Agent Systems. Phas 2019 ecture Notes in Computer Science vol 12058. Springer Cham. 2. C Calman. Evolutionary ethics: can values change. Journal of Medical Ethics, 30:366–370, 08 2004. 3. A. Bojanowska, Ł. D. Kaczmarek, M. Koscielniak, and B. Urbańska, "Changes in values and well-being amidst the COVID-19 pandemic in Poland," PLOS ONE, vol. 16, no. 9, p. e0255491, Sep. 2021,doi:https://doi.org/10.1371/journal.pone.0255491. 4. V. R. Krishnan, "Impact of MBA Education on Students' Values: Two Longitudinal Studies," Journal of Business Ethics, vol. 83, no. 2, pp. 233–246, Nov. 2007, doi: https://doi.org/10.1007/s10551-007-9614-y.
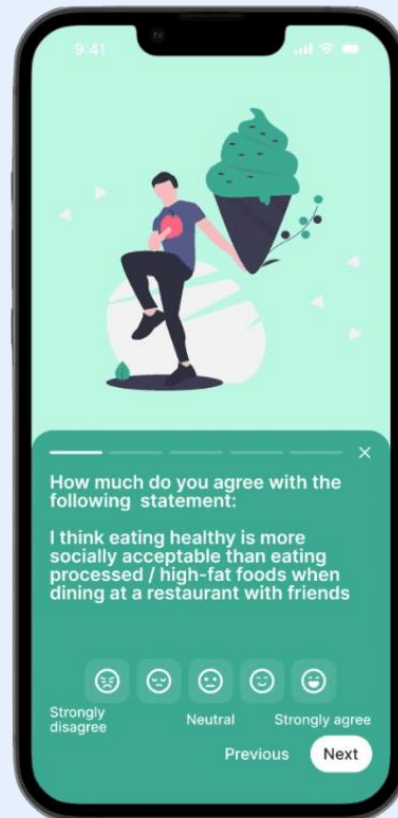
4. Prototype & Behavior Tree

Image 1: One of the questions for increase of water intake support agent
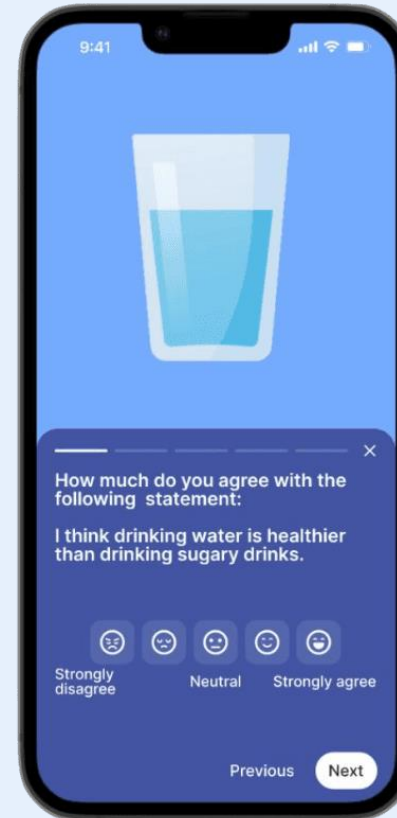

Image 2: One of the questions for an increase of water intake support agent
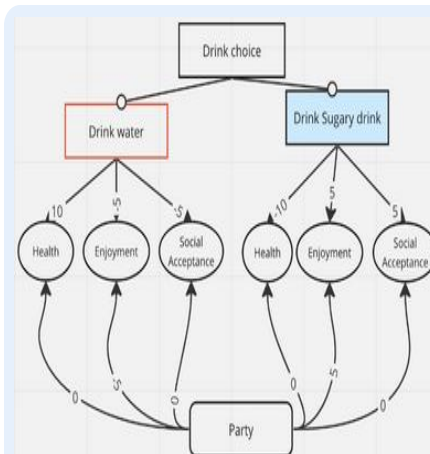
5. Behavior Tree

Image 4: An example of a behavior tree for the drinking more water behavioral challenge. Choice-value relationship shows general profile, while context(Party)-value relations show changes to the general profile.
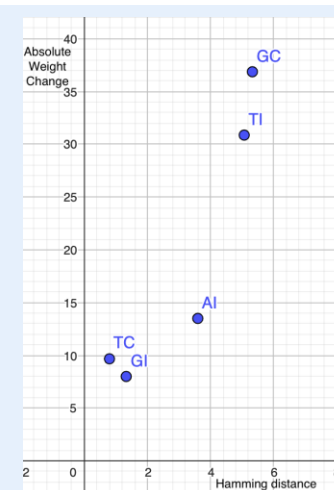
6. Results I

Image 4: Each condition's Average Absolute Weight Change and Hamming distance. Points closer to the origin are more accurate.

7. Results II

- SUS score of 78.6 puts prototype in 80-80th percentile.
- Absolute weight change of 36.87 and Hamming distance average of 5.33. Average change per changed edge is 6.92, bigger than 1 step in a scale.

8. Conclusions

- Statistically significant higher accuracy distances than other conditions. Due to incompatibility of comparative questions and behavior tree model and possibly discrepancies between experimentation procedures.
- GI more accurate than GC, but TI less accurate than TC. Contradictory conclusions about questioning technique relation to accuracy.

9. Future work

- Future research should ensure supervised and aligned between condition experimentation process, ideally carried out by a single person or by a strict protocol.
- A value model should be adapted to store and display comparative information.
- Current method of value elicitation is not naturally integrated into human's life, future research should optimize for no fatigue in the elicitation process.
- A change in human values might be triggered by many more factors: **random shifts in desires or interests, big life events such as attending a university or living through a global pandemic.** These possibly exhibit a much larger shift in one's values than misalignment scenarios, future research should explore **re-trigger mechanisms,** or monthly/quarterly **periodic check-ins.** Given a wide adoption of system a machine learning technique that learns and predicts potential goal updates based on similar user profiles could be employed.
- Perceived values were used as a baseline, however perceived values were extracted by tweaking the generated value tree. Further research could be done if this is a scientifically sound baseline.