Designing and implementing an emotion-aware chatbot

Supervisor: Kuldar Taveter, Associate Professor of Information Systems (kuldar dot taveter at ut dot ee)

The research problem to be tackled by this proposal is revising some foundations of designing autonomous software agents and particularly chatbots with the emergence of the theory of predictive probabilistic processing in the brain (Clark, 2013), which is increasingly considered to be as important for neuroscience as the theory of evolution is for biology. According to the theory of predictive probabilistic processing, a human brain invisibly constructs everything one experiences by matching the given situation with the most similar situation from the past, which is found by very fast predictive simulations in the brain, and by "storing" each new situation. Among everything else, also emotions are constructed in the brain, in concordance with the goals aimed to be achieved (Barrett, 2017a). For example, if the goal is romantic love, the emotions *Passionate*, *Longing* and *Lustful* might be constructed which make this goal more attainable. Differently, if the goal to be attained is tough love or brotherly love, respective instances of the emotions *Disciplined* and *Bonded* might be constructed (Barrett, 2017b).

The theory of predictive probabilistic processing can fundamentally change the way software agents and chatbots are designed and implemented from traditional symbolic architectures towards probabilistic cognitive architectures for predictive processing (Pfeffer & Lynn, 2018). However, computational implementation of predictive probabilistic processing is still faced with serious problems of intractability (Kwisthout & van Rooij, 2019). Different solutions have been proposed to overcome this challenge, such as sampling the environment by the agent (Friston, et al., 2012) and heuristic solutions in the predictive architectures of software agents (Alt, Baez & Darken, 2011).

This MSc project will continue the work started in the MSc project by Kirikal (2020), where a Minimal Viable Product of a chatbot capable of reflecting the emotions by its users was designed and implemented. The resulting chatbot will serve as a prototypic solution for a social robot that can emotionally relate to its owner and for a software agent to be embedded in an Interactive Digital Narrative (Koenitz, 2015).

References

Alt, J. K., Baez, F., & Darken, C. J. (2011). A practical situation-based agent architecture for social simulations. In *2011 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support* (CogSIMA) (pp. 305-312). IEEE.

Barrett, L. F. (2017a). The theory of constructed emotion: an active inference account of interoception and categorization. *Social Cognitive and Affective Neuroscience, 12*(1), 1-23.

Barrett, L. F. (2017b). *How emotions are made: The secret life of the brain*. Houghton Mifflin Harcourt.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences, 36*(3), 181-204.

Friston, K., Adams, R., Perrinet, L., & Breakspear, M. (2012). Perceptions as hypotheses: Saccades as experiments. *Frontiers in Psychology, 3*, e151.

Kirikal, A. (2020). *Computational simulation of how emotions are processed in our brain according to the theory of constructed emotion*. M.Sc. Thesis, Institute of Computer Science, University of Tartu.

Koenitz, H. (2015). Towards a Specific Theory of Interactive Digital Narrative. In: *Interactive Digital Narrative: History, Theory and Practice*, p. 91.

Kwisthout, J., & van Rooij, I. (2019). Computational Resource Demands of a Predictive Bayesian Brain. *Computational Brain & Behavior*, 1-15.

Pfeffer, A., & Lynn, S. K. (2018). Scruff: A Deep Probabilistic Cognitive Architecture for Predictive Processing. In *Biologically Inspired Cognitive Architectures Meeting* (pp. 245-259). Springer, Cham.