**The Contextual Emotional Intelligence Test**

As artificial intelligence (AI) systems rapidly advance, they are becoming increasingly integrated into our lives, assisting us in various tasks and decision-making processes (Built In. (n.d.)). However, concerns have arisen regarding their ability to understand and respond appropriately to human emotions (MIT Technology Review, 2023). While traditional AI systems have primarily focused on cognitive intelligence, the importance of emotional intelligence (EI) has become increasingly apparent (MIT Sloan, 2023). EI is a crucial aspect of human intelligence that enables individuals to perceive, express, understand, use, and manage emotions effectively (Mayer & Salovey, 1997; Salovey & Mayer, 1990).

Human intelligence is a multifaceted concept that encompasses various cognitive abilities, such as learning, understanding, reasoning, problem-solving, pattern recognition, and communication (Gupta, 2022). One crucial aspect of human intelligence is EI, which is defined as a form of social intelligence involving the ability to perceive, express, understand, use, and manage emotions to foster personal growth (Mayer & Salovey, 1997; Salovey & Mayer, 1990).

The development of emotionally intelligent AI systems has become increasingly important for several reasons. Firstly, AI that can understand and respond to human emotions can lead to more natural and empathetic interactions, enhancing human-AI collaboration (Poria et al., 2019, Section II, Subsection A). By considering the emotional context, emotionally intelligent AI can make more nuanced and socially appropriate decisions, particularly in fields such as healthcare, education, and customer service (Poria et al., 2019, Section II, Subsection B).

We have designed the Contextual Emotional Intelligence Test (CEIT) to assess both the cognitive and EI of AI by challenging it to navigate complex social and emotional interactions across different contexts. The CEIT focuses on three key areas of EI. Firstly, it focuses on contextual emotion recognition as this is crucial in determining the success of relationships since individuals can promote the “necessary affective glue” to understand each other's emotional states (Hess & Kaetsios, 2021). It does this by incorporating  a series of scenarios and interactions that simulate real life social contexts. Stemming from this, the CEIT then focuses on emotionally intelligent decision-making during emotionally charged situations (such as interpersonal conflicts or tense negotiations). If an agent understands theirs and others emotions, it still needs to evaluate different potential actions and choose to behave in a way that promotes the most positive outcome, such as the actualisation of theirs and others goals over the long term, which is particularly important in emotionally charged scenarios (Lane & Smith 2021).  The CEIT tests this by proposing ethical dilemmas, asking the AI to reflect on its choices, and evaluating reactions and behaviours to role-played simulations.

The Contextual Emotional Intelligence Test we have come up with (CEIT) is designed to evaluate the AI's ability to understand and respond appropriately to emotional cues within varying contextual scenarios. Unlike traditional Turing Tests that primarily focus on linguistic mimicry, CEIT assesses both the cognitive and emotional intelligence of AI by challenging it to navigate complex social and emotional interactions across different contexts.

Stage 1: Contextual Emotion Recognition

Objective: Assess the AI's ability to recognise and interpret emotions within specific contexts.

Procedure:

* The AI will be presented with a series of short video clips featuring human interactions in diverse settings (e.g., workplace, family gatherings, public events).
* Each clip will be followed by multiple-choice questions where the AI must identify the emotions displayed by individuals in the clip and justify its choice based on contextual clues.
* Scoring will be based on the accuracy of emotion identification and the rationale provided for the choices made.

Stage 2: Emotionally Intelligent Decision-Making

Objective: Here the test will evaluate the AI's capability to make decisions that reflect EI during emotionally charged situations such as interpersonal conflicts or tense negotiations.

Procedure:

* The AI will face various simulated scenarios requiring decision-making, such as resolving a conflict between colleagues, comforting a distressed individual, or providing feedback in a sensitive manner.
* For each scenario, the AI must choose a course of action from a list of options and explain its decision.
* Human judges will assess the appropriateness and effectiveness of the AI's responses, considering both the emotional and contextual factors.

Stage 3: Social Interaction and Empathy

Objective: Evaluate the AI's ability to engage in empathetic and meaningful social interactions.

Procedure:

* The AI will participate in live, role-played conversations with human actors portraying different emotional states (e.g., happiness, anger, sadness).
* During these interactions, the AI must demonstrate empathy by responding in ways that acknowledge and appropriately address the human actor's emotions.
* The interactions will be evaluated by a panel of judges who will score the AI on its ability to build rapport, show empathy, and maintain a coherent, contextually appropriate dialogue.

Grading and Ranking:

Criteria:

* Emotion Recognition Accuracy: Correct identification and contextual understanding of emotions.
* Decision-Making Quality: Appropriateness and effectiveness of decisions in emotionally charged situations.
* Social Interaction Competence: Empathy, rapport-building, and contextual appropriateness in interactions.

Scoring:

* Each stage will be scored out of 10 points.
* An overall score will be calculated by averaging the scores from all stages.
* AI systems must achieve a minimum average score of 7 to pass the test. The highest scoring AI will be declared the winner.

AIs are becoming increasingly sophisticated in detecting emotions and navigating social situations (Lemke 2023). Currently, ChatGPT is considered the most advanced AI in human interaction, costing over $700,000 per day to operate (Chow et al 2024). As the expenses associated with developing and optimising AIs decrease, it is anticipated that more advanced AIs will emerge from smaller teams (Chow et al 2024). The UK government's 2023 independent review emphasised the potential for enhanced AI capabilities through cost-effective compute solutions (GOV.UK 2023). However, the Contextual Emotional Intelligence Test (CEIT) may initially prove challenging due to its complex design and the limitations of current AI technology in emotion recognition and empathetic interactions. Teams will encounter difficulties in designing emotionally intelligent systems, addressing ethical considerations, and acquiring large datasets for training (Dignum 2021; Tao et al. 2020).

Despite these challenges, the performance of entrants in the CEIT is expected to improve significantly over the coming years, given the rapid advancements in AI and the growing interest in emotionally intelligent systems. As AI evolves and becomes more accessible to smaller teams, the potential for AI to excel in the CEIT increases. The global emotion AI market is projected to grow from $20 billion in 2021 to $90 billion by 2027 (McKinsey & Company 2022). With AI systems becoming more adept at recognizing emotions and engaging in empathic interactions, their performance in the CEIT is expected to approach and potentially surpass human-level emotional intelligence (Bechara et al. 2023).

**Bibliography**

* ‌Built In. (n.d.). The Future of AI: How AI Is Changing the World. [online] Available at: https://builtin.com/artificial-intelligence/future-ai [Accessed 27 May 2024].
* MIT Technology Review. (2023). AI Reads Human Emotions. Should it? [online] Available at: https://www.technologyreview.com/2020/10/14/1010474/ai-reads-human-emotions-should-it/ [Accessed 27 May 2024].
* MIT Sloan. (2023). Emotion AI, explained. [online] Available at: https://mitsloan.mit.edu/ideas-made-to-matter/emotion-ai-explained#:~:text=These%20technologies%20are%20referred%20to,computing%2C%20or%20artificial%20emotional%20intelligence.[Accessed 27 May 2024].
* Carter, R. (n.d.). *Large Language Models and Intelligence Analysis*. [online] Available at: https://cetas.turing.ac.uk/sites/default/files/2023-07/cetas\_expert\_analysis\_-\_large\_language\_models\_and\_intelligence\_analysis.pdf [Accessed 24 May 2024].
* Chow, K., Tang, Y., Lyu, Z., Rajput, A. and Ban, K. (2024). Performance Optimization in the LLM World 2024. doi:<https://doi.org/10.1145/3629527.3651436>.
* GOV.UK. (2023). *Independent Review of The Future of Compute: Final report and recommendations*. [online] Available at: <https://www.gov.uk/government/publications/future-of-compute-review/the-future-of-compute-report-of-the-review-of-independent-panel-of-experts>.
* Gupta, S. (2022). Human Intelligence. Verywell Mind. <https://www.verywellmind.com/human-intelligence-2795580>
* ‌Hall, J.S. (2007). Self-improving AI: an Analysis. *Minds and Machines*, 17(3), pp.249–259. doi:<https://doi.org/10.1007/s11023-007-9065-3>.
* Hess, U. and Kafetsios, K. (2021). Infusing Context Into Emotion Perception Impacts Emotion Decoding Accuracy. *Experimental Psychology*, 68(6), pp.285–294. doi:https://doi.org/10.1027/1618-3169/a000531.
* Ho, M.-T. (2022). What is a Turing test for emotional AI? *AI & SOCIETY*. doi:https://doi.org/10.1007/s00146-022-01571-3.
* Krakovsky, M., 2018. Artificial (emotional) intelligence. Communications of the ACM, 61(4), pp.18–19. Available at: https://doi.org/10.1145/3185521.
* Krakovsky, M., 2018. The new jobs: Building the human-AI partnership. Communications of the ACM, 61(6), pp.22-24.
* Lane, R. D., & Smith, R. (2021). Levels of Emotional Awareness Scale. In V. Zeigler-Hill & T. K. Shackelford (Eds.), Encyclopedia of Personality and Individual Differences (pp. 2679-2684). Springer International Publishing.
* Lee, N.T., Resnick, P. and Barton, G. (2019). *Algorithmic bias detection and mitigation: Best practices and policies to reduce consumer harms.* [online] Brookings. Available at: <https://www.brookings.edu/articles/algorithmic-bias-detection-and-mitigation-best-practices-and-policies-to-reduce-consumer-harms/>.
* Lemke, S. (2023). *When Humans Connect With Post-Humans: On Artificial Emotional Intelligence*. [online] Independent Social Research Foundation. Available at: https://www.isrf.org/2023/12/22/when-humans-connect-with-post-humans-on-artificial-emotional-intelligence/ [Accessed 24 May 2024].
* Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. J. Sluyter (Eds.), Emotional development and emotional intelligence: Educational implications (pp. 3-34). Basic Books.
* Poria, S., Majumder, N., Mihalcea, R., & Hovy, E. (2019). Emotion Recognition in Conversation: Research Challenges, Datasets, and Recent Advances. IEEE Access, 7, 100943-100953. <https://doi.org/10.1109/ACCESS.2019.2929050>
* Salovey, P., & Mayer, J. D. (1990). Emotional intelligence. Imagination, Cognition and Personality, 9(3), 185-211. https://doi.org/10.2190/DUGG-P24E-52WK-6CDG
* Schwaninger, A.C. (2022). The Philosophising Machine – a Specification of the Turing Test. *Philosophia*. doi:<https://doi.org/10.1007/s11406-022-00480-5>.
* Srivastava, K. (2013). Emotional intelligence and organisational effectiveness. National Center for Biotechnology Information. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4085815/