

Network Institute research visit (17-30 Nov 2024) report

1. Application details

Project Title: **Humorous AI Response System (HAIRS)** (17/11/2024- 30/11/2024)

Network Institute member: **Dr. Luis Morgado Da Costa** (l.g.de.passos.morgado.da.costa@vu.nl)

Visitor: **Dr. Joanna Ut Seong Sio** (joannautseong.sio@upol.cz), Palacký University Olomouc, Czechia

2. Research topic: Can LLMs generate humour?

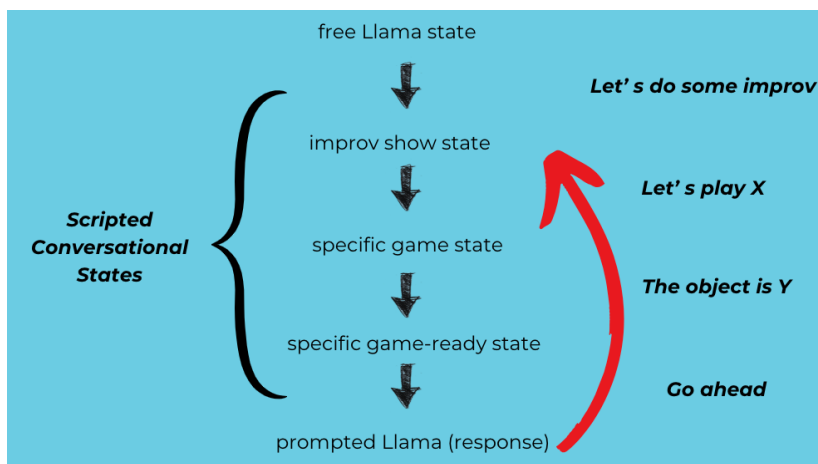
It has been shown that LLMs give higher quality and more relevant responses when provided with clear and specific instructions, e.g., generate satirical headlines; acronym task, fill-in-the-blanks, roast jokes, etc. (Gorenz and Schwarz N, 2024). In this project, we identified a humour production context that is highly-constrained. We focus on humour that arises from **incongruity-resolution** in the context of **one-liner improv games**. Humour arises if incongruent elements can be held together with some common thread: the incongruity is resolved. The resolution phase is to create a new association. Humour depends on unexpected and unusual association (O'Quin and Derks 1999). Incongruity-resolution in this sense is a type of remote association. Improv comedy is a comedy performance where the content of the performance is created on the spot. In one-liner improv games, the improvisers are given a suggestion and a format, the improvisers will then, take turn, produce a line of the given format/requirement using the suggestion.

3. Research activities during the visit

We Implemented an LLM-powered (Llama 3.1, open-source LLM) improvisation comedy module for the Leolani robot platform, and we conducted a workshop with professional improvisers to test the module.

3.1 The improvisation comedy module

We implemented 4 improv games ('sex with me', 'pick-up lines', 'famous last words', 'A not B'). Each improv game has a 'format' and a 'suggestion', e.g., in pick-up lines, the suggestion can be 'a phone', the response (format requirements in bold, improvised segment in italic) can be **'Hey baby, are you** "a phone", *because I am going to give you a ring.*' The 'format' for each game is given as a prompt, and the 'suggestion' is given verbally during the game. In the free state, the LLM is given this general prompt: 'Your name is Leo. You are a robot interested in improv comedy. You are friendly and funny. But don't be obnoxious, be polite. Reply using ONLY short concise sentences (maximum 15 words)'.



To better control turn-taking, we incorporated several scripted conversational states, where the Leoi is cued to provide responses. For each game, Leo will be cued verbally to enter a game state, and then a game-ready state after it receives a suggestion.

After receiving a suggestion, Leo would say a line to indicate that it got the suggestion (e.g., *I have one, I want to try, etc.*). To make the interaction more natural. We built in different probabilities for various scripted responses. There is a 15% chance that it will query then say the funny line unprompted; 75% chance she would wait till she is asked; 10% chance that it will say something to show it is ready (e.g., *I have one, I want to try, can I go?*)

3.2 Improv workshop where Amsterdam-based improvisers interact with Leolani (25 Nov 2024)



Three improv comedians from <https://www.easylaughs.nl>, Nicole Mischler, Peter More, Nirvi Maru, joined us at the workshop. In the workshop, the improvisers played improv games with Leo and it was followed by a roundtable discussion on the quality of Leo's improvisation (e.g., whether it is funny, whether and how it is different from human improvisers, etc.).

4. Results

- In the free state, Leo tends to ramble, and it is not possible to play improv games with Leo in that state. The scripted states help to structure the conversation flow.
- The responses from Leo tend to be longer and more repetitive than the human improvisers. Such lengthy exposition negatively affects the 'punchiness' of the funny lines.
- Speech recognition is problematic for non-native accents; pauses within an utterance are taken as end-of-sentence cues.

5. Future work

- To improve turn-taking, we need to incorporate silence to prevent the robot from talking while someone else is.
- To reduce repetitions, Leo should keep a record of its own responses and also responses from the human improvisers.
- We will evaluate (i) the responses from Leo and responses from human improvisers in terms of funniness; (ii) whether humans can tell whether each funny responses are from Leo or from human improvisers.
- We would like to bring Leo on stage to play a game in a real improv show with a real audience.