Author: Melvin Schop (M.T.Schop@student.tudelft.nl)
Supervisor: Agathe Balayn
Responsible Professors: Ujwal Gadiraju and Jie Yang

CSE3000: Research Project
02/07/2021

TU Delft Delft University of Technology

Collecting Tacit Knowledge through Competitive Text-Based Games

Background 1

- **Problem:** Blind spots in Machine Learning models due to lack of tacit knowledge. How to be more human?
- **Tacit knowledge:** Generality of Prompts and Answers.
- Somehow gathering this tacit knowledge for usage in Machine Learning models through **gamification methods**.
- Games With a Purpose [1] (GWAPs) are tools to perform a 'computation' through **human interactions with a game to solve a problem** or gather information.

[1] Luis von Ahn and Laura Dabbish. "Communications of the ACM". In: 51.8 (2008),pp. 57–67.doi:10.1145/1378704.1378719

Method 2

- **Previous work:** (Tacit) data collection through crowd sourcing
- **Game Design:** Multiplayer, Competitive, Text-based, Prompts and Answers, Party of 3-10 players

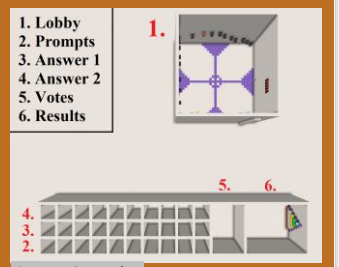

Figure 1. Game Flow

- Implemented in Minecraft to be engaging and form an accessible space of play.
- **Collected Data:** Given a topic, what P-A pairs can we expect?
- **Focus:** Engagement, Quality of the data and Expandability.

Research Questions 3

- How effective is a text-based multiplayer competitive game at acquiring tacit knowledge about humor from crowds of people for use by machine learning models?
- Is the game sufficiently engaging for the involved players?
- What tacit knowledge can we extract from free-text answers?
- How reliable is this method?
- How does this compare to other methods?

Results 4

- 15 different players contributed to the raw data
- 121 distinct prompts with 371 total answers to these prompts
- Collected over the course of 40 games
- The total of votes that have been cast/received is 582
- 8% of votes were skipped, due to bad prompts/answers
- Common biases within bubbles of players

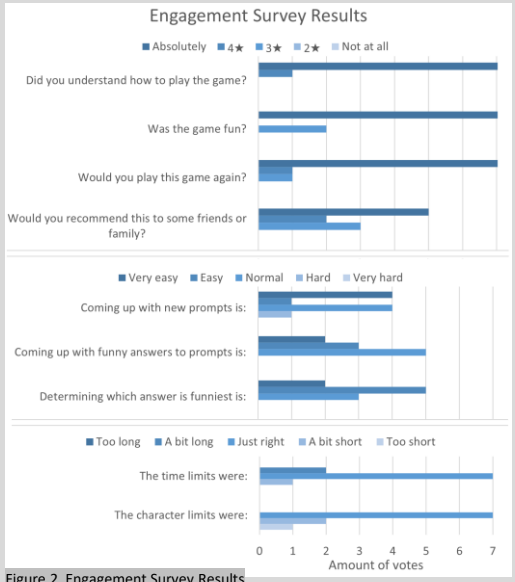

Figure 2. Engagement Survey Results

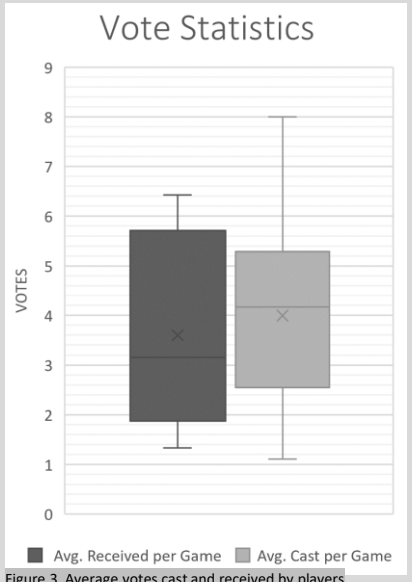

Figure 3. Average votes cast and received by players

Conclusions and Future Work 5

- The game is sufficiently engaging compared to other similar works.
- We collect: Relations between topics, prompts and answers, and knowledge about which answers are funny within certain bubbles.
- Reliability depends on size and diversity of the players involved. Cheating or manipulation of data is punished.
- Compared to other methods: high levels of engagement, data collected remains relatively unrefined.
- Improvements are possible by using third party tools and upscaling.