CS Capstone Progress Report

February 15, 2018

How to Make an Effective Robot Comedian

Prepared for

Oregon State University

Heather Knight

Prepared by

Group 13

AKA Robotics

Arthur Shing

Kevin Talik

Anish Asrani

Abstract

Human-Robot interaction can learn a lot from stand-up comedy. A stand-up set has scripted jokes, statements that are predetermined, as well as improvisational statements, that give the performance a sense of liveliness and character. A comedian can observe an audience and improvise a delivery of a joke to connect the audience to the content. This makes the experience more authentic and genuine for the observer. The purpose of this project is to to discover what makes an entertaining interaction by studying a robot that performs comedy. We propose that a performance is enhanced when (1) the comedian interacts spontaneously with the audience, (2) the comedian has and conveys a coherent, well-developed character, and (3) the comedian adapts its act to cater to an audience based on their reaction. These propositions will be tested locally, remotely, and in a real stand-up setting.

CONTENTS

1 RECAP

A lot of the machines that surround us aren't very engaging to interact with. They serve their purpose, people get what they need, and the interaction is over. People do not consider robots as entities. That is the gap we are trying to close by performing stand-up comedy with a robot. Stand-up comedy is a casual and entertaining way for people to get more exposure to robots and see that robots are not just objects, but they are much more than that.

An effective robot comedian should be able to entertain the audience and generate laughs. We hypothesize that the effectiveness is dependent on three major aspects - crowd-work or the ability to integrate the audience in the performance, portraying a coherent and convincing character, and the ability to adapt the performance based on audience feedback. We will base our performances and studies around these three areas.

2 PROGRESS SO FAR

We have started to execute a number of jokes on the robot; from one-liners to story-telling with a punchline. We have also started to make use of various sensors on the robot. There is an audio tracker which makes the robot look in the direction where any sound is coming from. This can be used to make the robot feel more natural during the performance and make relevant comments when there is a sudden spike in volume levels.

3 LEFT TO DO

While we have managed to execute many jokes on the robot, we have to determine a solid script that connects all these jokes together. Once we do that, we can go ahead and perform shows with the robot. This will enable to us get relevant data from the audience and analyze it to answer our research questions.

4 PROBLEMS

5 EXPERIMENTAL DESIGN

The comedian system that is implemented will test three critical areas of a comedic performance corresponding to our research. The first question is about **adaptation** of a performance – how the robot and interpret an audience response. The second question studies **crowd work** during a show, and the choices a Comedian can make to engage the audience. Lastly, the third considers the implications of a perceivable **character** that the robot can portray, in particular robotic versus humanlike.

5.1 Adaptation

5.1.1 Goal

We hypothesize that audiences will prefer a robot that acknowledges them, and integrates their data and responses into its set. To test this hypothesis we propose two tests: (1) to transition to topics dependent on the audience response, and (2) to present a crowd report upon completion of the set. The goal of this portion of the project is to determine if incorporating the audience into the set will enhance the overall performance of the comedian. To evaluate these hypotheses, we will conduct live studies in which people experience different versions of the software described below.

5.1.2 Methods

Generally, the performance will have two to three parts: the Seed Jokes, the Middle Content, and the Close. The Seed will influence the Middle Content (which will be chosen themed jokes and basis of the show). The Middle Content will transition to a intelligent or generalized Closing Joke when it is time to end the show.

Figure 2 depicts how a joke will be represented by the robot. It will perform the joke, collect audience feedback information, and branch to the joke that will best fit the response. At the end of the set, the robot will present a summary of what it thought that audience liked.

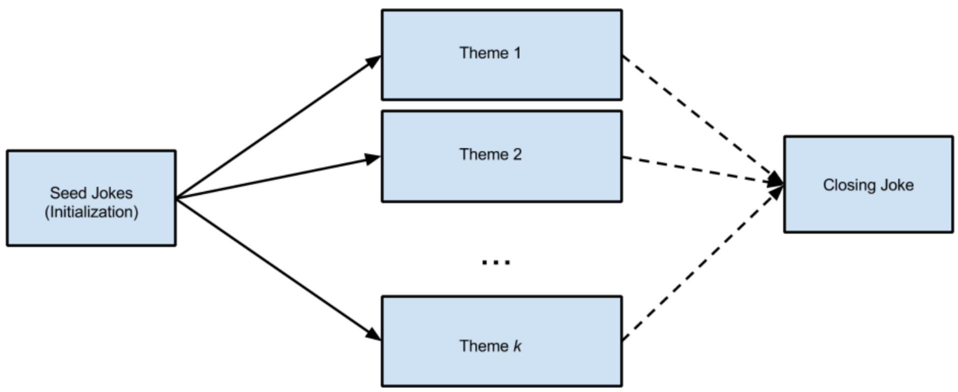

Fig. 1. This shows how the algorithm will have up to k Themes to choose from, determined from the seed joke. The closing joke is a subset of the set of all jokes, and may be outside of a specific theme. There could be different spanning trees of jokes that end at the same closing joke

TOPIC TRANSITIONS: In the beginning of the set, the comedian will present an ”initialization” procedure, known as the ”Seed Jokes” to test the response of the audience to different jokes. Depending on their response, the comedian will transition to a theme that is evaluated to be the best fit. The robot comedian will have many jokes to choose from that contain different material, but not all audiences will like all of the jokes. Figure 1 shows how the theme will be chosen from a set of up to k themes. From the Seed Jokes, one of the themes will be chosen. If there is time, we may also explore the choice of strategic closing jokes. These jokes might be stronger jokes than some of the others, and is helpful in ending the show on a stronger note.

For example, if two of the seed jokes are about ”food” and ”Mindfulness”, the performance will branch to the respective theme that matches the audience response (Branch 1 ”Food” or Branch 2 ”Mindfulness”). If jokes with a theme of ”food” are not landing with the audience, the algorithm will need to know when to transition to a new theme, or when to end the set. When the robot tells a joke, it needs to be able to analyze the feedback and choose the next joke to perform. This needs to be done quickly, so that the robot is not spending noticeable time (for the audience) choosing a joke. There may not be a lot of jokes to choose from, but the choice needs to be made fast.

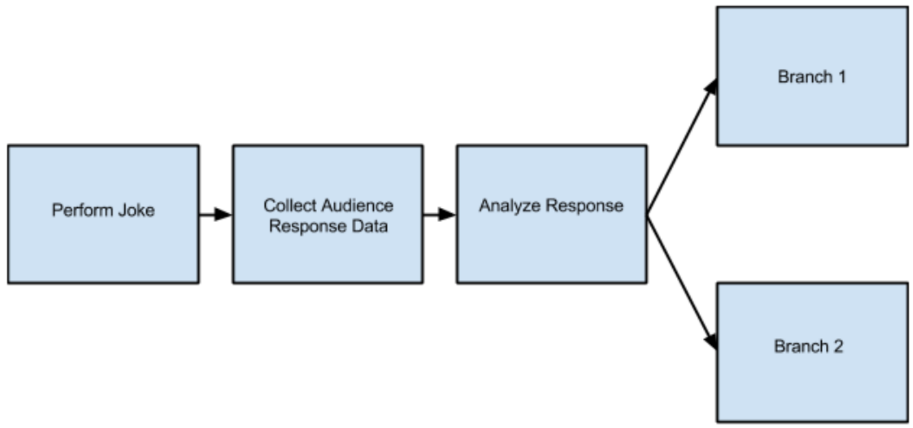

Fig. 2. This flow-chart depicts how, once the robot delivers a joke, will wait for feedback, interpret data, and then make a joke decision (Branch 1 and Branch 2)

AUDIENCE REPORT: The close of the robot's act will include the robot's report of what attributes the audience responded to most. The hypothesis here is that getting insight into the robot's algorithms will increase the audience's perception of the robot's intelligence, and, second, that it will make them laugh. People enjoy hearing about themselves.

All of the above behaviors, from adaptation to the end of performance audience report need to be evaluated with real people. Initial tests will be done on campus with small groups of people, the final test will be done in conjuction with the crowd-work and character manipulations, testing the entire algorithm together with a larger crowd, e.g., 10-25 people.

5.2 Crowd-work

5.2.1 Goals

Similar to the crowd report described above, we hypothesize that generally interacting with the audience (a.k.a crowd-work) throughout the performance, will improve the audience'soverall enjoyment of the show. We want to analyze the importance of this crowd-work relating to the central design of the project. Crowd-work should make the audience feel like they are a part of the show. This can be done in different ways - calling out and talking to the audience, watching the audience and incorporating them into the jokes, and asking them questions to keep them engaged, or to build off to make new jokes.

5.2.2 Methods

There are various kinds of crowd-work we want to test: one research question is whether crowdwork matters at all, and the second is does the crowdwork needs to be real or robot can just pretend it is paying attention?

To answer these questions we suggest three research conditions: (1) no crowdwork, (2) fake crowdwork, (3) real crowdwork. The first one would be no crowd-work whatsoever. The robot goes about performing its set and does not directly address the audience at all. The second could spanover-the-top and inaccurate crowd-work, or best-guess crowdwork, with the possibility of being real (e.g., predicting that most people in the audience were from Oregon, even if it didn't really hear what they said). The third case would integrate actual robot sensing. It would be important for conditions #2 and #3 to be parallel to assess whether crowdwork really matters.

As condition one is fairly obvious, let us discuss deeper possibilities for condition #2. In the obviously fake research condition, the robot will talk to the audience directly but it will be completely wrong in its observation. The absurdity of a robot trying to understand the audience and being completely off could be entertaining for the audience, or it may not connect with the audience at all. The exact reception of this sort of crowd-work is something we are trying to study.

The other version of condition #2 is realistic but premeditated. For example, pre-known facts about the audience could be built into the robot or guessed. These pre-known facts could include the location of the performance, age demographics of the audience. For example, if the audience is known to be college-aged, the robot could be fed input to make comments about things relevant to college students.

Using actual robot sensing data is condition #3, and is certainly the ideal model, but requires sensing capabilities, processing power and hardware, so it would be good to know if it is really necessary. In this condition, the robot would be actually looking for cues from the audience during certain situations. For example, one example is asking questions and capturing words from the audience, then using that same word later. For example, the robot could ask a simple question about the weather, or the audience member's hometown. In this case, the robot can listen for specific words and ask another question about that specific town or city.

Another real sensing capability the robot could use is audience volume levels after the delivery of jokes. The robot will keep track of the audience input. The robot could then acknowledge if the audience enjoyed the joke or did not enjoy the joke using these inputs. Additional sensors and processing abilities on the Nao robot include face-detection and bumper detection, so the exploration of audience sensing could potentially include speech, volume, vision, and touch.

All of these condiditons will be assessed with live audiences (even if its just a few people in a classroom) to check to what degree is crowd-work important for a robot comedian. The audience's response will be used to see if they enjoy a humanized robot or if they prefer a more robotic one, or maybe even a combination of both. As crowd work is just a form of human interaction, we expect it will improve the audience's perception of the robot's intelligence, add surprise to the show, and increase audience enjoyment levels. On the other hand, perhaps faking it can get 80% of the effect of the real version. That will be part of the evaluation.

5.3 Character

5.3.1 Goal

The goal of this section of the project is to examine whether or not robot comedy can benefit from having jokes delivered from a robot's perspective. Our hypotheses are that a robot presenting jokes about technology or being a robot will be funnier than a human telling the same jokes, and that robots will be less funny than humans at telling jokes from a human perspective.

In Jerry Palmer's *Taking Humor Seriously*, comic meaning is argued to depend on the interrelated factors of a joke's context and setting, its delivery, the identity of the deliverer, and the audience [**?**]. Of specific interest to us are the factors of a joke's delivery and the identity of the deliverer. In previous studies, robot comedy has been used to analyze effective aspects of joke delivery. However, little has been done in discovering effective aspects of a joke's content as it relates to the identity of the deliverer. For example, Sjöbergh and Araki [**?**] found that jokes were perceived as funnier when delivered by a robot, rather than being delivered in text form. However, Sjöbergh and Araki used word-play jokes that were gathered from the internet, and delivered them through a robot by using a flat, machine-like sounding text-to-speech tool called AquesTalk. This form of delivery does not take into account the importance of effective joke delivery. While Sjöbergh and Araki did not implement measures for analyzing non-verbal delivery, other work has examined the importance of non-verbal signals in delivering jokes [**?**] [**?**]. Despite this, there is little to no existing literature on the effectiveness of jokes related to the identity of the deliverer. In our context, this means examining the effectiveness of robot-specific jokes in robot comedy.

5.3.2 Methods

To address this goal, jokes will be written from a human or robot perspective. The jokes written from a human perspective will have a corresponding robot version, ideally with as much one-to-one correspondence as possible in regards to cadence, length of joke, parrallel content, similar motions, and so forth. These jokes will be subject to intense scrutiny by members of the project and by the client, such that revisions and edits can be made to create funny jokes with a definite correspondence between the two versions. For example, a human version of a joke might look like the following (lines with a definite correspondence with the robot version are highlighted):

Hey, hey, I got news. This is big. Ok, quiet down. Get this. That's RIGHT folks. I'm no longer single. *throws hands up* (*@ I met a man on tinder. @*) (*@ His name's Sebastian. He's a math nerd. @*) (*@ Swiped right as fast as my fingers could move. @*)

Whereas, the robot version of the above joke is shown below:

Hey, hey, I got news. This is big. Ok, quiet down. Get this. That's RIGHT folks. I'm no longer single. *throws hands up* (*@ I met a robot on tinder. @*) (*@ His name's Data. He's a really geeky robot. @*) (*@ Swiped right as fast as my motors could turn. @*)

5.3.3 Development process of joke writing

These jokes will be scripted in Choregraphe, where adjustments to vocal tones and pausing will be made. Then, animating the robot for non-verbal gestures will be done to enhance the delivery. The overall process may look similar to Figure **??**.

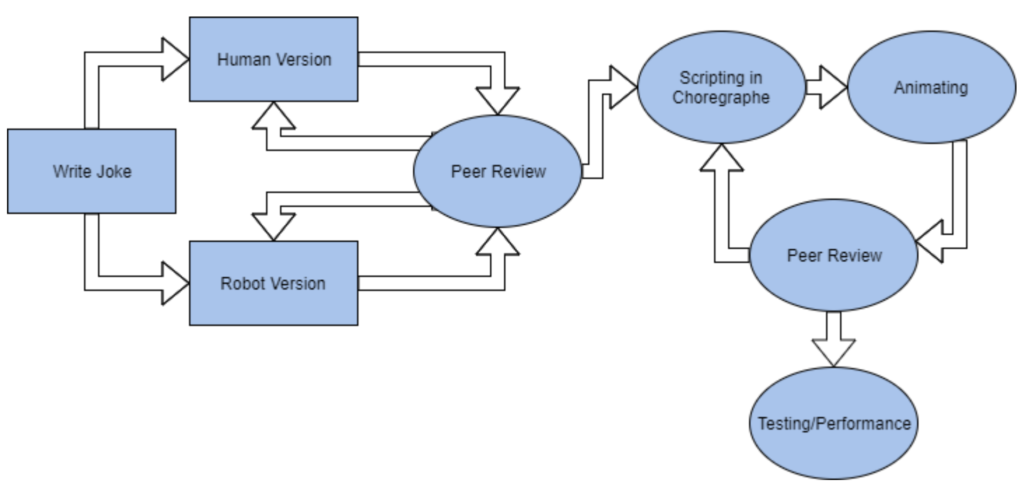

Fig. 3. The work flow from joke writing to testing.

5.3.4 Experimentation

The stand-up routine of the robot will comprise of text of the joke themselves, the motions that the robot uses to accompany them, and the way the robot surveys the audience after each punchline, e.g., in a human or robotic fashion. To determine the differences in audience response to the routines, studies will be done first on Amazon Mechanical Turk, and later with co-located audiences. Participants will be shown a video of the robot's stand-up routine, and then presented with a short survey. The routines will be between 5 and 10 minutes long, and the survey will include questions pertaining to each joke or routine. Participants will be compensated with standard rates for watching brief videos and answering survey questions.

6 Conclusion