Assignment 2: Emotional ChatBot

CS886– Winter 2017

Out: March 1, 2017
Due: March 30, 2017

Submit your assignment *in PDF only* by email to the insructor at `jhoey@cs.uwaterloo.ca`
Use the email subject line: `CS886 Assignment 2 FIRSTNAME LASTNAME`
and the PDF file name: `CS886-A1-LASTNAME-FIRSTNAME-STUDENTID.pdf`

Description

Conversational Agents (or "chatbots" as they are commonly called[1]) are programs that engage a human in a repeated dialogue. Chatbots are increasingly being used online by businesses who seek to provide information or assistance to their clients. For example, the furniture store IKEA has a chatbot on its webpage called "Anna" who can answer questions about the company products, etc[2]. Apple's SIRI is another example, a general purpose chat engine that is meant to provide a voice interface to a smartphone.

Chatbots are notoriously difficult to build in a way that makes the indistinguishable from humans, thus passing the Turing test. Each year, the Loebner Prize [3] (of $1,000,000) is awarded to the chatbot that can "fool" a set of human judges into thinking it is human. Thus far, no one has won the Loebner Prize. A consolation prize (of around $10,000) is awarded each year to the bot that fools *the most* judges.

The winning chatbots from last year's Loebner prize are described on this page:
`http://www.aisb.org.uk/events/loebner-prize`,
and the top three chatbots from 2016 are linked here:

1. Mitsuku `http://www.mitsuku.com/`
2. Tutor `http://www.rong-chang.com/tutor_mike.htm`
3. Rose `http://brilligunderstanding.com/rosedemo.html`

In this assignment, you will analyze a chatbot's responses according to their emotional content, and propose an idea for improving the emotional responses of a chatbot.

PART A:

First, pick a chatbot. You can use one of the Loebner prize winners, or pick another one online[1]. Then, pick an emotions theory, e.g., ACT, OCC, etc. Finally, chat with your chosen chatbot and find two different instances of a chat sequence that is *consistent* with your chosen emotions theory, and two that are not consistent. For the inconsistent examples, give a possible chat that would be consistent. You don't have to use the same chatbot or emotions theory for all four examples, they can all be different combinations if you want. Use your own definition of consistency: simply explain why you think these chats are consistent or inconsistent with the emotions theory, and whether the theory is wrong, or the chatbot is wrong.

[1] see `chatbots.org`
[2] although I checked recently and it seems it has been removed.
[3] See `http://www.aisb.org.uk/events/loebner-prize`

Here is an example of a consistent chat with Mitsuku. I used ACT and modeled it as an interaction between two *friends*.

- Human: `I'm sad and Lonely`
  (*whine to* EPA:[-1.4,-0.8,-0.5])
- Mitsuku: `You've come to the right place.  I will keep you company`
  (*comfort* EPA:[1.5,1.7,-0.6]. Optimal is *make up with* [2.2,1.3,0.2], distance 1.13).
- Human: `I'm laughing`
  (*laugh with* EPA:[2.5,1.8,2.1])
- Mitsuku: `I'm glad you are having a good time`
  (*congratulate* EPA:[2.5, 1.7, 0.9], optimal is applaud=[2.15, 1.63, 1.62] distance 0.8)

Both of these seem fairly close in EPA space (distance close to 1.0).

Here is an example of an inconsistent chat with Tutor Mike. I used ACT and modeled it as an interaction between a *student* and a *tutor*.

- YOU: `you look so silly`
  (*scoff at* EPA: [-1.6,-0.9,-1.0])
- Mike: `Really?`
  (*question* EPA: [0.4,0.8,0.5] but optimal ACT is [1.7,2.1,-0.8] , much more powerful and good, but less active (distance 2.25), closer to *counsel* e.g. "I recommend you ask me a question instead", or *comfort* e.g. "Its ok, don't worry about it".)
- YOU: `yes`
  (*agree with*) EPA: [1.7,1.3,0.9])
- Mike: `All right!`
  (*agree with*) EPA: [1.7,1.3,0.9] but optimal ACT is [2.0,2.1,-1.0], more powerful and much less active (distance 2.03), closer to *comfort* again.)

Now, the EPA distances are further (great than 2.0), so seems inconsistent. In this case, I believe both the chatbot and ACT are off. I would expect a tutor to say something more angry or reproachful (e.g. "don't talk to me like that" reprimand EPA: [-0.6,0.6, 0.7], which Mike is closer to than ACT). In the second case (if Mike had said "Really?"), then I see the ACT solution to be more in line with what I would expect, although I'd expect the tutor to express something more sad (more negative).

PART B:

Although chatbots have come a long way since Weizenbaums' ELIZA (1966), they are still not able to fool humans, and are largely easy to distinguish from humans. One thing that is missing from modern chatbots is *emotion*. In one paragraph, propose a method for endowing a chatbot with emotions. Your method must be grounded in some theory of emotions, must make a convincing case for why it will enhance or make better a chatbot, and must be reasonably creative. Make sure to justify your idea: why will it work? What about it will make your chatbot more human-like?

What to hand in:

6 marks A detailed description of your four example chats. There is no right or wrong answer here. The point is to get you to think about chatbots and how they express emotions (and how hard it is to do this well!)

4 marks A paragraph with your description of how a chatbot could be emotionally aware. I am looking for a good justification and something a bit creative.
