

Fall 2020

Improv Theater and Artificial Intelligence

Izzy Michels

Kenyon College, michels1@kenyon.edu

Follow this and additional works at: https://digital.kenyon.edu/dh_iphs_prog



Part of the Digital Humanities Commons

Recommended Citation

Michels, Izzy, "Improv Theater and Artificial Intelligence" (2020). *IPHS 200: Programming Humanity*. Paper 25.

https://digital.kenyon.edu/dh_iphs_prog/25

This Poster is brought to you for free and open access by the Digital Humanities at Digital Kenyon: Research, Scholarship, and Creative Exchange. It has been accepted for inclusion in IPHS 200: Programming Humanity by an authorized administrator of Digital Kenyon: Research, Scholarship, and Creative Exchange. For more information, please contact noltj@kenyon.edu.

Improv Theater and Artificial Intelligence

Izzy Michels
Programming Humanity
Final Project

Introduction

Improvisational theater is an art form where unscripted theater is performed. Dialogue, characters, and actions are created on the spot. Errors made within an improvisational theater scene are encouraged, and can form an input to how the scene evolves. Ultimately this project focuses on the evolution and creation of artificial intelligence bots interacting with the world of improv theater.

Chatbots Versus Improv Bots

A chatbot is a software application used to conduct an online chat conversation via text or text-to-speech, in lieu of providing direct contact with a live human agent. There are many different types of chatbots ranging from a regular expression chatbot like Eliza, who was designed to imitate a therapist, a slot-response chatbot such as Amazon's Alexa, who responds and acts on commands, or even neural nets like GPT-2, BERT, or XLNet all of which are used for various elements of natural language processing and text classification tasks. The Artificial Improvisor is a form of artificial conversational agent, or chatbot, focused on open domain dialogue and collaborative narrative generation. Using state-of-the-art machine learning techniques, spanning from natural language processing and speech recognition, to reinforcement and deep learning, these improv bots provide a completely new and exciting asset to this technology that is different from these other types of chatbots. Below is an example of each type of chatbot listed in order from left to right.

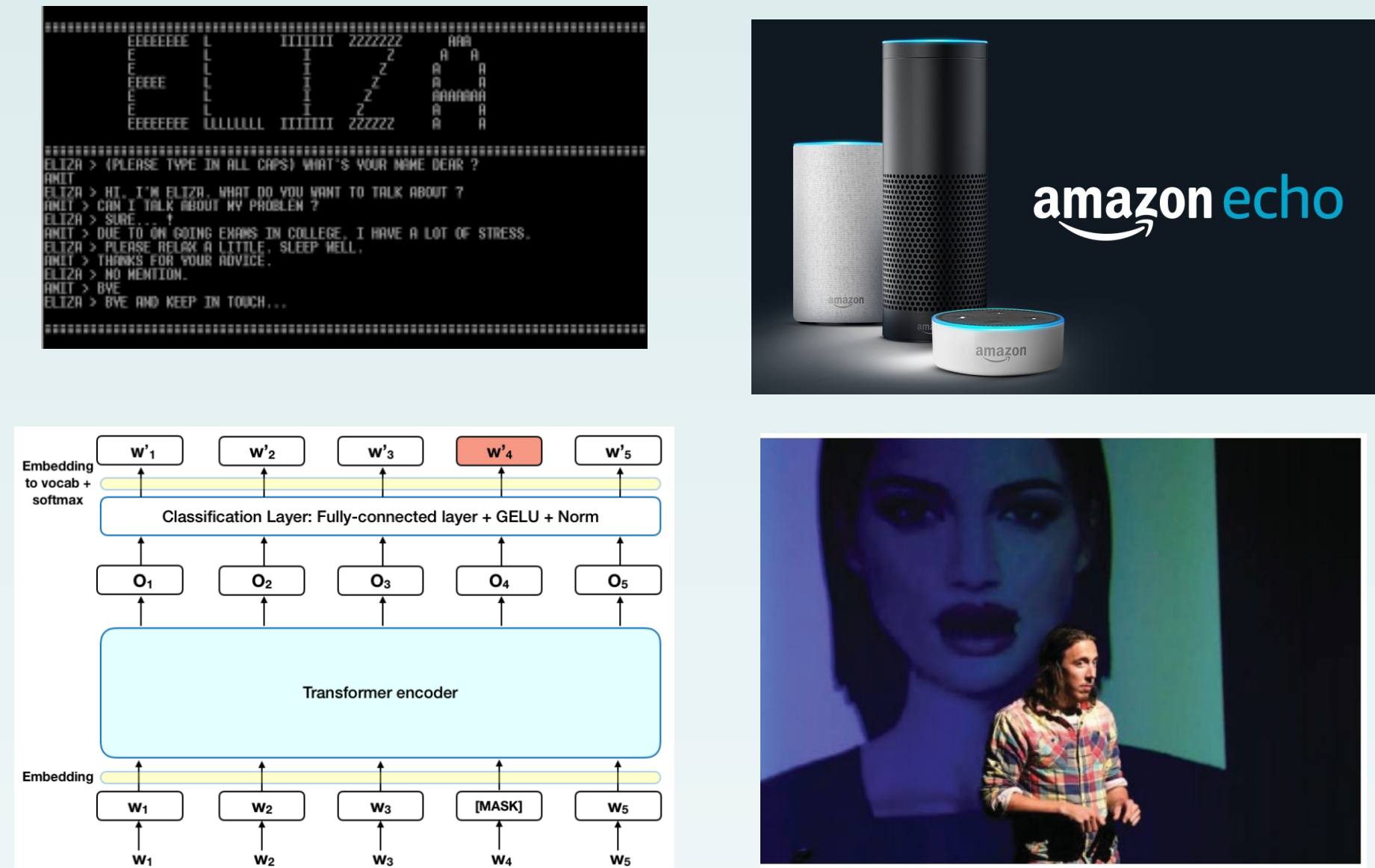


Figure 1: Human performing on-stage with Pyggy. The avatar of the AI is projected behind the human performer.

Analysis

Positives

Generally, improvisational theater is a domain where experimentation is encouraged, where interaction is paramount, and where failure flourishes. It allows artificial intelligence agents to be effectively tested, and therefore, audience reaction can provide a subjective measure of improvement and cognizance. Improv welcomes chaos and mess ups. AI chatbots and their randomness in response generation, can sometimes be considered quite comical.

Negatives

With that said, there are still some issues. While the system is trained to perform dialog, it is not trained to tell a cohesive story with a narrative arc. The addition of memory network advancements may improve callback. In addition, new engineering and training will be necessary to collaboratively build a narrative arc. Additionally, while chatbots can be improved using natural language feedback, converting feedback to natural responses that fit in the conversation is a whole other ballgame. This work could be further advanced with better deep matching models for retrieval based systems, as well as end-to-end generative models for improv chat with sequence-to-sequence learning. Finally, when it comes to "yes and," although Cho and May have made great strides in that department, it is clear our models are still inferior at producing good yes-and's when compared to professional improvisers. Finally, it is quite difficult to re-create a sense of physical humor in robots. Improv comedy is very much encompassed by the physical choices, voices, and actions of the actor. This is something not talked about very much in the research and an integral part to comedy in general, which is near impossible to replicate in chatbots.

Methodology/Technology

SOTA: If we're talking about "the state of the art" of Artificial Intelligence, then it's deep neural networks in deep learning. Deep learning models deal with non-textual data such as voice and image recognition. NLP is an applied artificial intelligence program that helps the chatbot analyze and understand the natural human language communicated with its users. Chatbots are able to understand the intent of the conversation rather than just using the information to communicate and respond to queries. These features give the machine an intelligent feature that can recognize object. Creating artificially generated human dialogue is a classical task within the research field of artificial intelligence. The use of artificial intelligence as a part of improvisational theater has recently been explored by Mathewson and Mirowski along with many others. Their work included the creation of a dialogue system that allowed a human actor to communicate with a robot that produced lines in response to lines uttered by the human. An artificial improviser dialog system is composed of three major building blocks: 1) speech recognition, 2) speech generation, and 3) a dialogue management system. In another study done by Furu Wei, he explains that existing research on response generation for chatbots focuses mostly on first response generation, which aims to teach the chatbot to say the first response appropriate to the conversation context. His paper introduces a new task, second response generation, termed as improv chat, which aims to teach the chatbot to say the second response after saying the first response, with respect to the conversation context. In another study, two versions of the robot dialogue were constructed by Maria Skeppsted and Magnus Ahltop. One version that selected existing lines in a training corpus, and one version that relied on text generation techniques. Dialogue could either be created in the form of a goal-driven dialogue system that was meant to be used to perform a specific task, or in the form a non-goal-driven system, for which no such task was given. One implementation method for the task of generating dialogue is to use actual lines, which is not completely in the nature of improv, but gives the robot a basis for vocabulary and context. Another solution is to generate new sentences that do not necessarily have to have been present in the corpus used for training. For this task, neural network techniques are typically applied. They used speech recognition and a text-to-speech system, which functioned in real-time in front of an audience. Finally, according to Hyundong Cho and Jonathan May in their study, modern dialogue systems are not explicitly trained to build common ground, and therefore overlook this important aspect of communication. Improvisational theater intrinsically contains a high proportion of dialogue focused on building common ground, and makes use of the yes-and principle, a strong grounding speech act to establish coherence and an actionable objective reality. Open-domain neural dialogue systems, by contrast, specifically lack coherence. Many works have identified the same issues of repetitive or non-committal responses generated by neural conversational systems that are at least partially related to the lack of sufficiently high quality yes-and's. Inspired by yes-and's in improv, Cho and May constructed SPOLIN, a collection of dialogue pairs with responses that are not only coherent with dialogue context, but also initiate the next relevant contribution. Below is a graph explaining the use of SPOLIN and "yes and" within this improv technology.

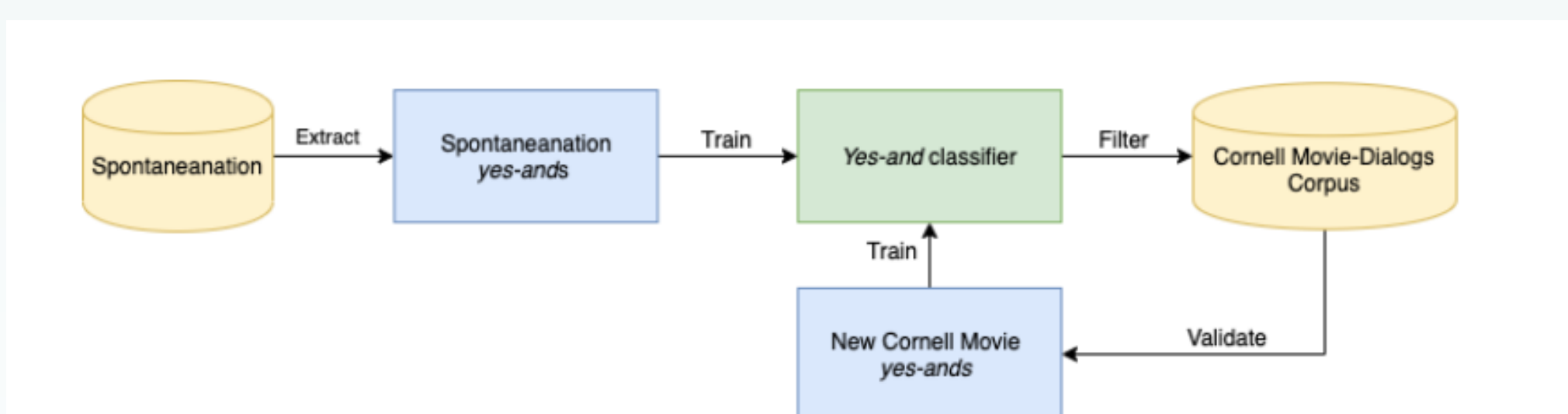


Figure 2: An illustration of the yes-and collection workflow. The core SPOLIN corpus comprises Spontaneatation yes-and's and Cornell yes-and's (in blue boxes). However, SPOLIN can be augmented by including other general-purpose dialogue corpora in place of Cornell in this workflow, as described in Section 5.

Will Chatbots Surpass Humans? For now, chatbots won't replace humans, but that could easily change quite quickly. While certain functions could be replaced to provide a faster, more efficient interaction with customers, AI chatbots will continue to be limited for a number of reasons. Until chatbots develop empathy and human problem-solving skills, they won't be able to replace us, especially in the world of improv. Additionally, as consumers get more used to interacting with chatbots, there will be new opportunities for phishing, hacking, and general mischief. With all that said, this technology is constantly improving upon itself. For example, Eugene Goostman is a chatbot that some regard as having passed the Turing test, a test of a computer's ability to communicate indistinguishably from a human. Mitsuku, is a record-breaking, five-time winner of the Loebner Prize Turing Test and the world's best conversational chatbot. A great chatbot has conversational maturity, is omni-capable, integrates with CRM, is emotionally intelligent, is free to explore, has autonomous reasoning and is pre-trained. The continuous development of these aspects of chatbots could make it easier and easier for them to pass the Turing test and eventually surpass human intelligence.

Future Directions

The future work in this area should focus on strategies and mechanisms to generate humor in different human-computer interaction settings (social networks, virtual agents, robots and smart physical spaces) with the objective of improving the overall user experience. Future work should incorporate recent advances in deep reinforcement learning for dialogue generation. Three useful conversational properties recently shown to improve long-term success of dialogue training are: informativity, coherence, and ease of answering. In terms of "yes and" and improv specifically, expanding the dataset through an iterative data collection process, with other larger text-based open-domain dialogue, in order to gather longer conversations exhibiting more complex improv-backed turns, is the overall goal. In addition, this work could be further advanced with better deep matching models for retrieval based systems, as well as end-to-end generative models for improv chat with sequence-to-sequence learning. The addition of memory network advancements may improve callback as well. Finally, additional engineering and training will be necessary to collaboratively build a narrative arc when it comes to storytelling.



Figure 2: Human performing onstage with A.L.Ex., personified as a robot. The user interface, showing speech recognition results and machine-generated response, is projected behind the performers. Image credits anonymised.

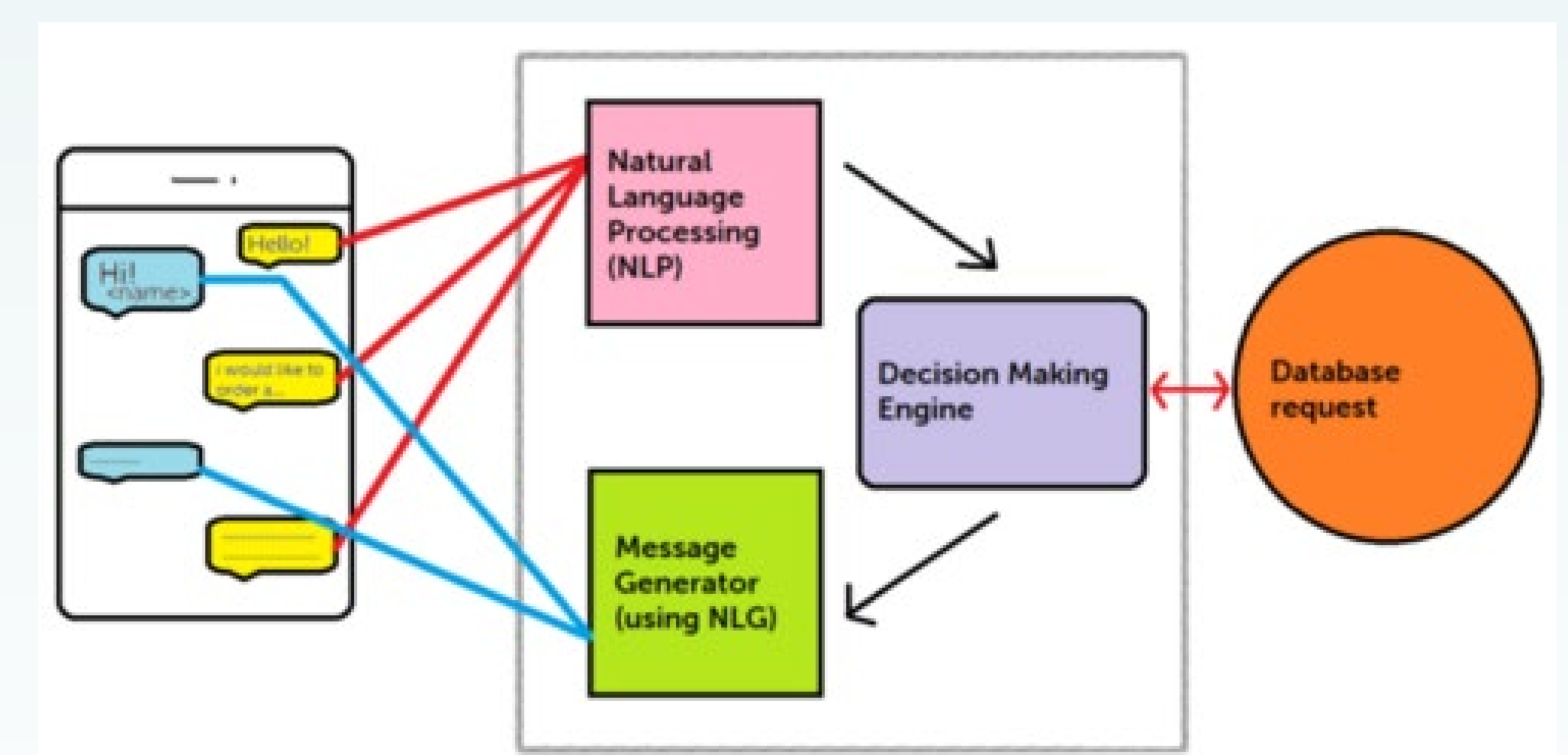
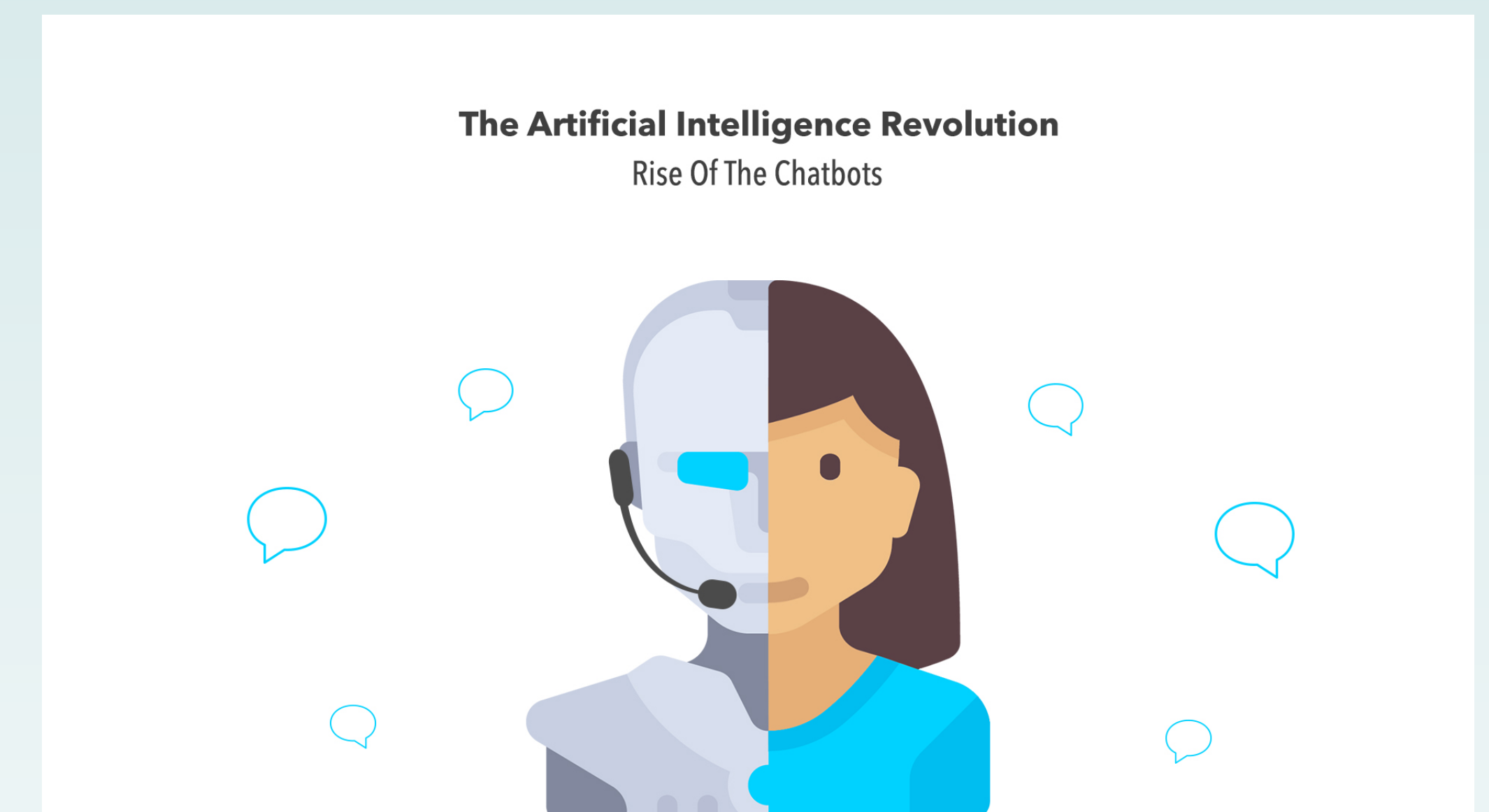


Figure 2: An illustration of the yes-and collection workflow. The core SPOLIN corpus comprises Spontaneatation yes-and's and Cornell yes-and's (in blue boxes). However, SPOLIN can be augmented by including other general-purpose dialogue corpora in place of Cornell in this workflow, as described in Section 5.



Conclusions

Humor is a paramount indicator of socially desirable and positive interrelationships. The increasing use of human-computer interfaces seems to be projecting us into a dark era of human isolation. Providing them with humor, instead, they will likely enhance our humanity. When it comes to improv, an amazingly unique form of storytelling, the introduction of artificial intelligence to an already complicated art form creates new and exciting possibilities. Although it seems that right now, according to the majority of this research, there are obviously complications that will never amount to exact human ability, the future is bright. Physically, emotionally, and narratively, improv theater is difficult to recreate through robots. With that said, all these special dialogue systems, yes and generators, and intricate study of improv has greatly changed the way computer scientists are looking at performance and AI in general. Although I am wedded to the original human forms of improv, since it is something I do and care about, I do believe it is important to open our eyes to the new and exciting possibilities artificial intelligence can and does hold. Chatbots already exist in our everyday lives. Performance AI is not a typical norm in the present moment, but once it becomes one human beings will have to adjust accordingly.



Sources

- Cho, Hyundong, and Jonathan May. Grounding Conversations with Improvised Dialogues, Information Sciences Institute University of Southern California.
- Mathewson, Kory W, and Piotr Mirowski. "Improvised Theatre Alongside Artificial Intelligences." Proceedings, The Thirteenth AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment (AIIDE-17), 1University Of Alberta Edmonton, Alberta, Canada, korymathewson.com/assets/2017-aiide-mathewson-mirowski.pdf.
- Posted August 12, 2019 | Atlanta. "AI 'Performers' Take Center Stage and Get Creative with People in Public Spaces." Ivan Allen College of Liberal Arts at Georgia Tech. 12 Aug. 2019, www.iac.gatech.edu/news-events/stories/2019/8/ai-performers-center-stage-creative-people-public-spaces/624365.
- Skeppstedt1, Maria, and Magnus Ahltop. "Towards a Structured Evaluation of Improv-Bots: Improvisational Theatre as a Non-Goal-Driven Dialogue System." Computer Science Department, Linnaeus University, Växjö, Sweden.
- Wei, Furu. "Improv Chat: Second Response Generation for Chatbot." Microsoft Research Asia, Beijing, China.
- Braun, Ken. 2018. "Council Post: Does The Evolution Of Chatbots Include Replacing Humans?" Forbes. Retrieved December 15, 2020 (https://www.forbes.com/sites/forbesagencycouncil/2018/10/16/does-the-evolution-of-chatbots-include-replacing-humans/?sh=27aee487459b).