

Numbers and Letters: A Synthesized Approach to Dialogue in Conversational AI in the Context of the NBA Playoffs

Andrew Chiang¹, Sunny Chung², Alex Fan², Anthony Wong¹

1. Computer Science, 2. Quantitative Theory and Method

Emory University, Atlanta, GA 30322

{andrew.lee.chiang, sunny.chung, alex.fan, anthony.t.wong}@emory.edu

Abstract

Using natural language processing, we propose a novel method to create a speculative and adaptable chatbot in the context of NBA playoffs. Our primary goal is to have the chatbot emulate an informative human conversation flow built around opinionated forecasting. There are two primary ways we will do this. The first method is by providing the bot with an opinion. This means that when given information by the user, the bot should be able to have an opinion on the player or team, and how each will do in the playoffs. The opinion should change depending on which teams and players are involved in the conversation, which would allow the user to converse with the bot as if they are talking to another knowledgeable NBA fan. The second methodology is to add speculation to the bot. This forecasting feature makes our bot unique because it is able to engage the user about events which have not yet happened. More specifically, when gathering qualitative information, the bot should present the information and explain it such that it appears to have an independent process of understanding players and teams. The bot will use statistical information to formulate its opinion based on metrics we create. This prevents the bot from creating wild opinions which could confuse the user and keeps its comments within the realm of possibility. The metrics will involve comparison of team and player statistics to league averages to spot strengths and weaknesses. To accomplish these goals, we propose a simple and generalizable flow of conversation that moves between both user opinion and system opinion on NBA topics.

The chatbot employs a myriad of techniques to achieve these objectives. First, we will construct a broad ontology tree that will allow us to capture a breadth of cases and situations. Statistical analysis will be incorporated to generate the bots opinion. Other NLP methods such as named entity recognition and sentiment analyzers will be used to account for cases not covered within the ontology. This combination of a diverse expression tree with APIs will inform the bot's feedback to opinions. These APIs will also help us gather the most recent NBA trades which will allow our bot to remain up to date with week to week developments in the league.

Our work in developing our chatbot, will give us insight into making a more human-centric program and improve our ability to convincingly engage audiences over qualitative information.

1. Introduction

Our project is a chatbot which engages hardcore NBA fans about the 2019-2020 NBA season. The bot chats with users about the hypothetical playoff picture, makes comments on various players, responds to users opinions and comments on different trades.

This project is important because current work done with sports chatbots is limited to regurgitation of statistics. Our project moves away from that paradigm and lets users talk qualitatively about the sport the way they normally would with other fans. During the season, fans talk to each other through opinionated forecasting: making predictions about who will win what game and what the

eventual playoff picture will look like. By focusing on the hypothetical, our bot is able to better emulate how a fan might talk to another. This is done in a way that cannot be done online because the response is immediate whereas typically a fan would have to hope that another fan is available to chat with them or would have to wait for analysts to write articles. Besides discussing the playoffs our bot is also able to discuss trades in context of how it has and will impact the playoffs. Integrating trades into our bot is particularly important because sports analysts often only focus on big name trades. Because our bot is able to hold a wealth of statistics, it is able to discuss trades which don't receive as much attention, but are still important to fans of particular teams. Our bot is thus able to bridge the gap between fans talking with one another and the content that is produced by the NBA network.

Our project is different from other projects because every single opinion our bot has is not hard-coded, but dynamically generated. The bot's opinion is qualitative but derived from statistics about players. These statistics range from generic stats which rate a player's overall efficiency or their age, to specific stats on their performance on key aspects of the game such as points, assists, and rebounds per game. Every single time the bot is run, the statistics are queried and analyzed on the backend to form the bot's opinion. By having a wide variety of statistics we're able to generate opinions which are specific to each player and comment on a specific aspect of their playstyle. It focuses on making comments about current season performances while it also addresses the player's experience level in the NBA and their performances throughout their career. Thus from conversation to conversation the type of opinion the bot forwards will change. But most importantly, by dynamically retrieving the statistics, the bot's opinion of a player will change each time the API is updated. This allows us to best mimic a natural conversation because fans' opinions about players change each time a game occurs. While many bots have information they only use it to comment on the past or present. Their opinions are based on products or events which have already happened. However, natural conversation is also speculative, especially in the realm of sports. By engaging in future based conversation, our bot is able to help develop more human-centric conversation in a way that other bots do not.

2. Related Work

Before this project, we took inspiration from a multitude of sources, including Emora, a social chatbot created by Professor Jinho Choi and his team. The social chatbot Emora uses a custom dialogue management framework called emora-stdm, which is available on the Python Package Index (Finch 2020). We used emora-stdm to build our chatbot as well. While our chatbot is designed to be specifically tailored towards NBA fans, the fundamental architecture of both bots remains the same.

In addition to more practical inspiration relevant to the scope of the project, we took a look at the overall field and specific approaches by different groups. These include commercial virtual assistants by Google, Apple, and Amazon. Overall, these chatbots use turn based dialogue management and have specific ontologies to catch different inputs. These chatbots are designed to provide simple information and assistance to the user, but they are not able to hold an extended conversation with the user (Google Corporation). Other than commercial virtual assistants, there have been many open source virtual assistants, one example being Almond. Almond differs from commercial virtual assistants because it leverages a "crowdsourced public knowledge base of natural language interfaces and open APIs," giving it more flexibility (Campagna et al. 2017). However, it fundamentally remains a virtual assistant, meaning that the bot is not able to have an extended conversation. Chatbots have also seen use on many websites, from social media to e-commerce, for automation purposes.

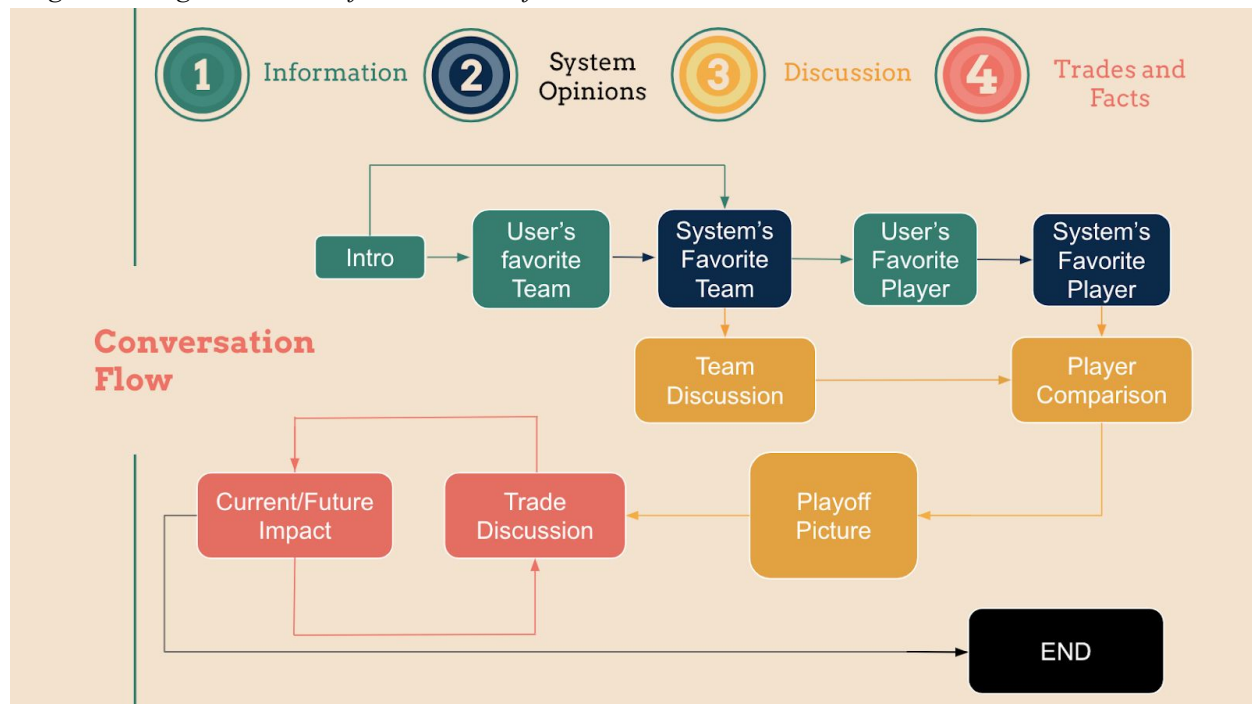
Aside from more traditional chatbots with dialogue states, there has been a movement towards using neural networks to create a true conversational AI. Gao et al. cover neural approaches to "(1) question answering agents, (2) task-oriented dialogue agents, and (3) chatbots." Instead of using a task-oriented dialog with rigid states, Gao et al. proposes a statistical (neural) model to replace the whole

interaction (Gao et al. 2019). Traditionally, the dialog would compose of natural language understanding, state tracking, external knowledge retrieval, dialog policy, and then natural language generation. We took inspiration from the usage of neural networks in the field and attempted to make our bot “smarter” and more responsive by using statistics from the NBA, but we only scratched the surface of its potential due to the scope of the project.

3. Conversation Design

To ensure a coherent interaction with the user, we had a subject matter expert architect the conversation. Since our target audience is meant to be serious NBA fans, we wanted our conversation to be a conversation that they would enjoy. As a result, we made sure to include nicknames and characteristics about players so the users would feel comfortable around our chatbot. In addition, we specifically built dialogue for the system’s two favorite players, and had specific dialogue about each. Below is a diagram of the overall conversation flow, and two example dialogues.

Diagram 1. High level view of conversation flow



The conversation flow diagram shown above is a high-level overview of how our conversation is carried out. Starting from diagram 2, we specified more in detail. However, due to the extensive number of states that cover a variety of responses and appropriate reactions, our diagram shows the main idea of each state, without delving into the variation of the syntax.

Diagram 2: Conversation Flow Diagram

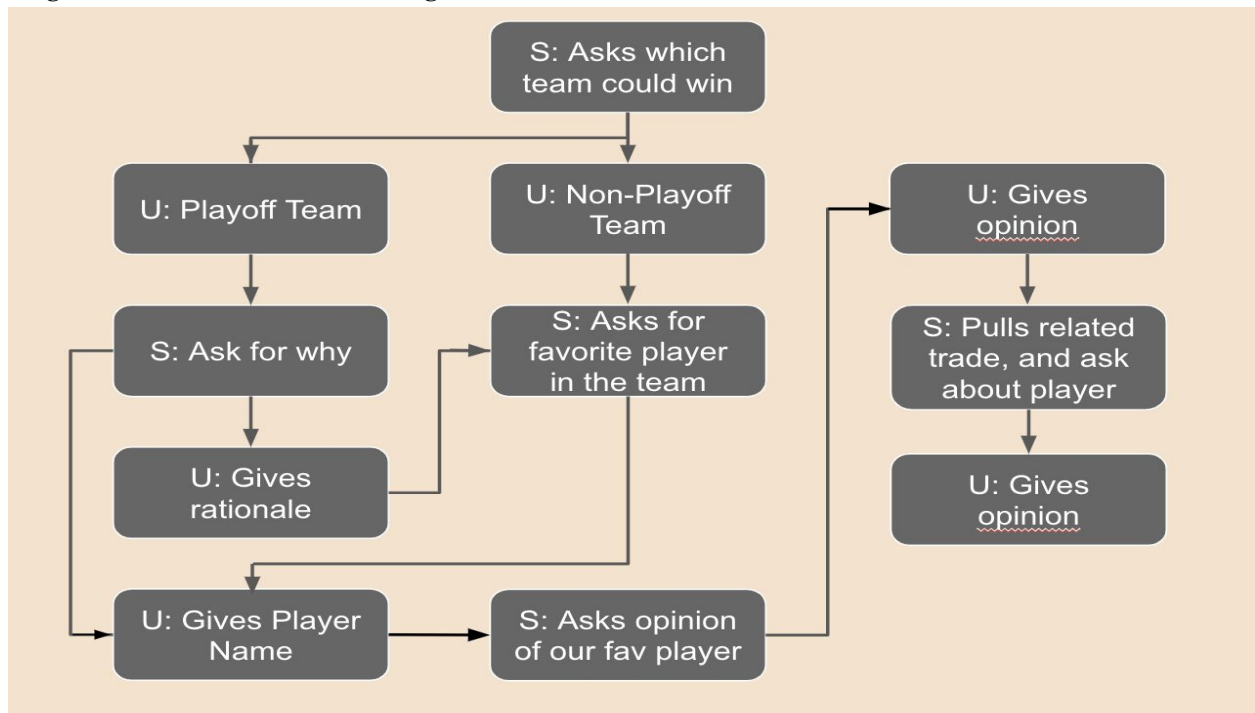
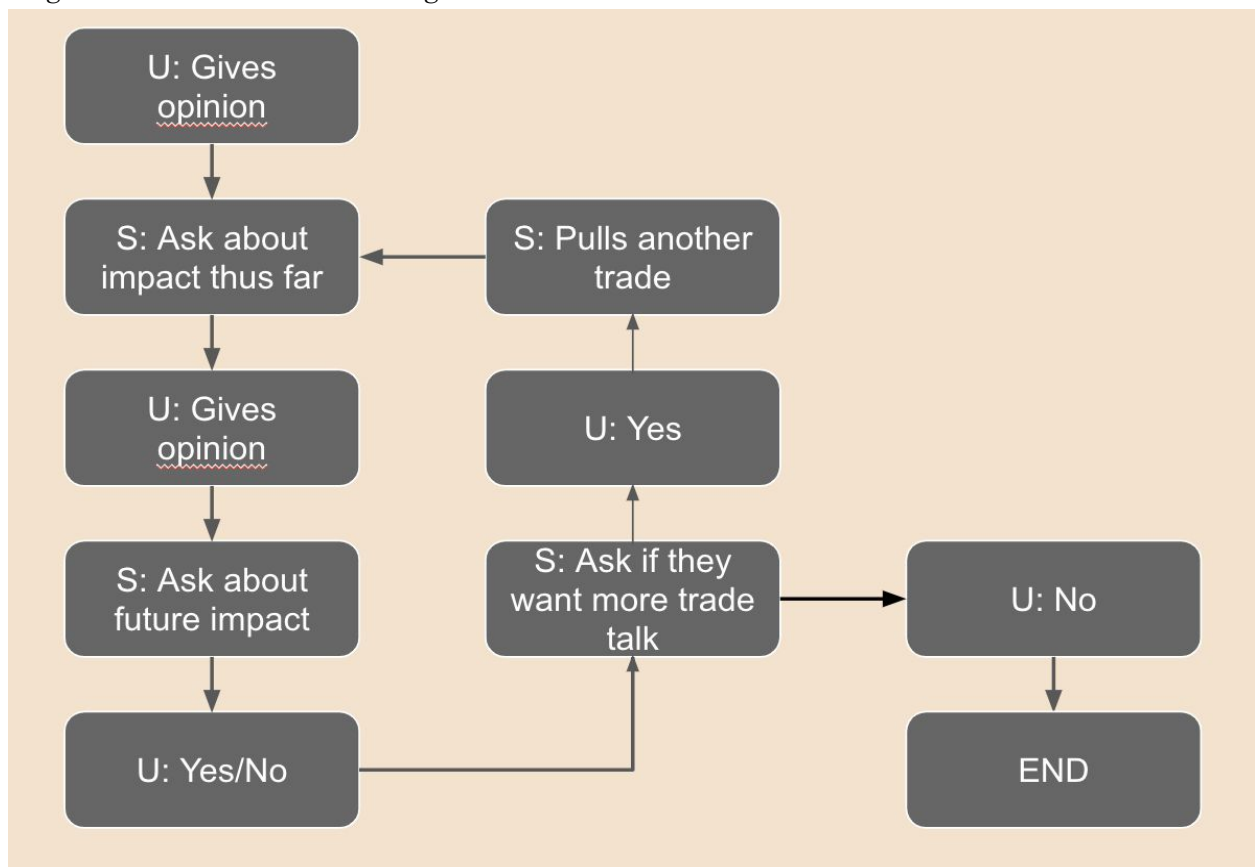


Diagram 3: Conversation Flow Diagram



Sample conversation 1:

S: Hi I am NBA chatbot. The NBA season has been suspended because of COVID-19. If we had playoffs based off the current standings , which team do you think would win?

U: i think denver nuggets will win

S: Why do you think the nuggets will win?

U: because they have good team chemistry

S: Okay, Ive also heard that the nuggets are really good . Do you think there is a player that is super important to the nuggets ?

U: joker is always my first pick

S: I see nikola jokic is having an exceptional season. Personally, I think Clippers will win because Kawhi Leonard is more clutch . Whats your opinion of Kawhi Leonard ?

U: i guess so but i think kawhi is too fragile

S: Even though Kawhi is fragile he always overperforms in playoffs. Anyways, I saw that Gerald Green from Houston Rockets went to Denver Nuggets . What are your thoughts about Gerald Green ?

U: he is always a great addition to the team

S: So you think Gerald Green is aight. I dont think he is too terrible but he doesnt seem to have a huge impact to the game. Do you think that trade influenced how the Denver Nuggets performed this season?

U: not really because they were already doing well

S: he was traded quite recently before covid shutdown so maybe the stats wont be the same but who knows? . Do you think he will play a big role in getting the Denver Nuggets to playoffs next year?

U: no not really

S: So you do not have high hopes I do not think he will play a big role. I think he is past his prime. Do you want to chat about another trade ?

U: no

S: Okay, well in that case we can talk again later. Bye!

Sample conversation 2:

S: Hi I am NBA chatbot. The NBA season has been shutdown because of COVID-19. If we had playoffs based off the current standings , which team do you think would win?

U: i think clippers have the best chance this year

S: Why do you think the clippers will win?

U: because they have really nice shooters

S: Okay, Ive also heard that the clippers have really solid shooters . Do you think there is a player that is super integral to the clippers ?

U: i would say kawhi is most influential

S: I see kawhi leonard is having an exceptional season. Personally, I think Miami Heat will win because Jimmy Butler is more clutch . What do you think of him ?

U: he is good midrange shooter

S: Right ! I also agree that Jimmy Butler is really good . Lets go back to the teams now. Earlier in the season I heard that Isaiah Thomas from Washington Wizards went to LA Clippers . What do you think about Isaiah Thomas ?

U: he is too old and lost his touch

S: Okay, I can tell that you dont like Isaiah Thomas . I dont think he is too terrible but he doesnt seem to have a huge impact to the game. Do you think that trade influenced how the LA Clippers performed this season?

U: not much

S: I agree, the trade was early in the season and he has not shown he was worth it . Do you think he will play a big role in getting the LA Clippers to playoffs next year?

U: maybe

S: I do not think he will play a big role. I think he is past his prime. I am a little uncertain too. You can never account for randomness like injuries sidelining players. would you like to chat about another trade ?

U: no thank you

S: Okay, well in that case we can talk again later. Bye!

4. Approach

The chatbot attempts to create a conversation about the NBA playoffs from both a quantitative and qualitative approach. Utilizing a combination of player statistics acquired from open-source APIs and hard-coded ontologies, the chatbot is able to generate its own engaging opinions about players while responding to the user's own opinions.

In terms of the quantitative side of our conversation, stats are pulled from the Sports Reference API in real-time for each player. These stats are used in our macros that compare and rate players. In particular, to compare players, we utilize a player's points, assists, and rebound over the games played this season to determine whether they are as good or worse than the system's favorite player (who is chosen based on the best player efficiency rating prior to the system's execution). The player rating macro takes this idea further, essentially building a comprehensive evaluation of a player being traded/talked about within the conversation. In addition to leveraging similar statistics like the previous macro, the player rating macro also takes into account the career-long statistics of the player and the position of the player (this is especially important because different players focus on different aspects of the game from shooting to defending, which can skew the statistics). By utilizing these data analysis-driven macros, the chatbot is able to generate a system opinion that is unique to an individual player that the user might care about. Moreover, these opinions are not hard-coded to the players themselves; therefore, over time the bot's opinion is able to change depending on how much the player plays and improves through the season and their career. While there may not be a specific challenge tied to this aspect of our approach, the minutiae of the macros can be refined further; for example, perhaps putting weights on specific stats for different positions may be a better methodology in the long-run than what our current approach does.

Bringing focus back to a more general outlook though, throughout the creation of this chatbot, our goal was to emulate human speech as much as possible. With that framework in mind, for instance, the chatbot specifically does not utter any numbers back to the user, even in the quantitative-heavy segments such as comparing player statistics or generating an evaluation about the player. This is because even the most dedicated basketball fan would likely never recall exact statistics in conversations, so while the chatbot does leverage exact statistics, it summarizes how those statistics should be interpreted. Similarly, the chatbot has an ability to recognize nicknames of famous players (for example, "lbj" for LeBron James), once again modeling realistic conversations where the user does not refer to their favorite player using the full name. Alongside this framework of human-like interactions, the chatbot tries to employ some techniques from active listening such as asking open-ended questions to the user and repeating key phrases from the user utterances. In doing so, the user is able to feel as though they are conversing with someone who is actually listening and engaging with them, rather than a poorly-designed questionnaire-like bot.

Now, this methodology of asking open-ended questions and understanding user utterances brings with it its own set of challenges: specifically, the breadth of possible user inputs becomes infinitely large in both content and structure/syntax. In order to address this challenge, what we now term as the qualitative aspect of this project, we utilize a litany of resources to try and capture user opinions and

sentiment. Using an ontology of hard-coded adjectives (divided by the sentiment valence) and rationales (ex. Shooting, dribbling, efficiency, stamina etc.), the chatbot attempts to recognize and match the user input. Once matched, the chatbot is able to respond accordingly by either parroting back the rationale and agreeing or disagreeing with the user about that rationale. However, without more advanced techniques such as a legitimate implementation of phrase structures or word2vec (both, at this point, beyond the scope of the project, the available resources corpus-wise, and the technical knowledge of our group) to proactively predict the words based on the context of the user utterances, this challenge becomes nearly impossible to solve. To mitigate the likelihood of errors as much as possible, the chatbot employs a pre-built sentiment analyser from the Natural Language Toolkit (NLTK) to act as a fail-safe in case the ontologies do not match to a user utterance. The sentiment analyser, called the Valence Aware Dictionary and sEntiment Reasoner (VADER), is a pre-trained model that parses a user utterance and returns sentiment valence scores on a positive-to-negative scale based on a rule-based approach of evaluating each token in the sentence (Hutto and Gilbert 2014). This means that having more words that are considered “negative” by the model (the model’s lexicon labelled prior) appear in the user utterance leads to a higher the negative valence score returned (and vice versa with positive words). This approach to error handling is not ideal, for example, VADER does a poor job with taking into account context-specific terminology and considers phrases such as “he is good at stealing,” which in the context of basketball is considered a positive phrase, and returns a large negative score because “steal” is considered negative. Even so, this is better than returning a generic, catch-all comment back to the user about their opinion in the vast majority of cases.

The chatbot incorporates a couple periphery technologies as well. One instance is where it uses the GATE macro to determine whether the player indicated as the user’s favorite is actually on the team they indicated as their favorite. Because macros by themselves cannot determine which state the system transitions to, we leverage the GATE macro to forcibly close the path that the system can go through if the player is not on the team (this matching was done in the background through a macro). Another periphery technique used was retrieving and accessing trade news through the Sports Reference API. These news were stored into a json, which can later be updated as trade news are updated, and the trades are pulled randomly to have a good variety of conversation with the user.

5. Analysis

Based on the sample conversation above, we can see every turn transitions smoothly with an appropriate response. Our bot is capable of understanding whether a user made a positive or negative response with ontologies of good or bad words, and is able to capture sentiment even if certain words don’t appear in the user’s response. In addition, these ontologies are classified further into parts of speech and the system response becomes very granular to the input. The bot is capable of taking in any team that the user likes and converse about trade related to either the user’s favorite team or the bot’s favorite team and their impact on current standings as well as future outcome.

However, our chatbot does have some few situations where it can’t handle the user’s response appropriately. For example, conversation sample 2 shows that it doesn’t account for what the user said before each turn and responds to the comment about Isaiah Thomas’ age as if it was given in isolation. This is due to the fact that the bot is unable to understand the entire context of what the user said and cannot adjust its responses to take into account all comments made by the user over the course of the conversation. Another error could arise with dependency on sentiment analyzer. Because VADER sentiment analyzer is not built off of a basketball-related corpus, it does make few errors in capturing user’s sentiment unless the user clearly shows whether they like or dislike the player. It would sometimes think the user is neutral about a player when they are slightly negative. Fundamentally, our bot also has challenges with the user asking questions, since questions have such a wide range and breadth to them.

We specifically added a state to catch if the user asked a question in one specific turn, but we did not account for it in any other turn. In other words, the user could ask a question during the conversation, but the bot would be unable to catch it.

Compared to other chatbots, we are confident that we cover a wider range of responses with more accurate responses. Our data analysis on player statistics gives more domain specific knowledge of how good or bad a player is and which are their strengths and weaknesses. Our chatbot's domain knowledge in nicknames and different ways of calling specific players enable the users to speak as if real nba fans are having a conversation. Especially with our bot's favorite players who are Jimmy Butler and Kawhi Leonard, the conversation becomes more robust. It also becomes clear that our code is not hardcoded for comment about trade, because the comments are adjusted based on their strengths, weaknesses, it will show a wide range of variation.

6. Conclusion and Future Work

Our project will allow fans to have an outlet through which to discuss their passions, even if they are not with another fan. This is important because one of the ways that users engage with their interests is by talking about them with other people. By making the bot's feedback qualitative as opposed to stats-intensive we allow users to have a more natural conversation.

Our chat bot needs to be able to respond to player specific comments, not just generic statements which could be applicable to any player. What we mean by this is that our bot currently has two favorite players. It is able to respond to comments which are specific to those two players' in game play styles and out of game attitudes. However, each player in the NBA is different and has distinct positive and negative traits. The bot should be able to engage users over their specific characteristics of their favorite players as opposed to just a general commentary. A standard to which this should be measured is, if the bot made a comment about a player, would you be able to make a reasonable guess about which player that is, or would that comment be applicable to any player. By creating a more specific opinion, the bot will feel more like a fan with a deep knowledge base as opposed to a casual viewer. This is important because the folks who are most likely to use the bot multiple times are hardcore fans.

In addition to being able to engage with specific opinions, another feature to alter would be the sentiment analyzer. VADER plays a significant role in our bot because it is in the majority of the error handlers. However, because VADER was trained to analyze the sentiment of a generic conversation, it may miscategorize basketball specific terms and respond in a way that is unnatural for a basketball fan. Developing a different version of VADER or some other sentiment analyzer, which is trained to recognize basketball slang, would allow our bot to more accurately respond to users.

Another key feature that could be developed in Emora_stdm is the relationship between states and macros. Our bot is heavily reliant on macros in everything from error handling to generating opinions. This means that any limitations in how macros interact with states directly impacts the quality of the bot. If there was a way for macros to dictate which state was transitioned into next, the code would then be more efficient and readable. This would allow us to better utilize the API and be able to query more directly between a user's input and an API endpoint. Such an integration would allow our bot to engage with the user's opinions more specifically and perhaps even discuss player performance pre-2019.

NBAchatbot is a bot meant to provide NBA fans with an outlet to engage their passion. Our use of statistics with emora_stdm has allowed the bot to engage in distinct conversations that emulate natural conversations which occur in the domain space of the NBA. By using a sentiment analyzer in error handling the bot is able to make users feel like their opinions are acknowledged even if the response is not specific. Its changing opinions appeal to the users interest while preventing the conversation from becoming stale. Overall, this bots dynamic generation of opinions as a player's performance changes distinguishes it from all others.

References

Finch, J. (2020). emora-stdm 1.64. <https://pypi.org/project/emora-stdm/>

Giovanni Campagna, Rakesh Ramesh, Silei Xu, Michael H Fischer, Monica Lam (2017). Almond: The Architecture of an Open, Crowdsourced, Privacy-Preserving, Programmable Virtual Assistant. <https://dl.acm.org/doi/abs/10.1145/3038912.3052562>

Google Corporation. Conversation design process.
<https://designguidelines.withgoogle.com/conversation/conversation-design-process/write-sample-dialogs.html#>

Hutto, C.J. & Gilbert, E.E. (2014). VADER: A Parsimonious Rule-based Model for Sentiment Analysis of Social Media Text. Eighth International Conference on Weblogs and Social Media (ICWSM-14). Ann Arbor, MI, June 2014.

Jianfeng Gao, Michel Galley and Lihong Li (2019), Neural Approaches to Conversational AI, *Foundations and Trends® in Information Retrieval*: Vol. 13: No. 2-3, pp 127-298.
<http://dx.doi.org/10.1561/15000000074>