

Design of AI Systems (DAT410)

Assignment 7: Dialogue systems and question answering

Luca Modica
Hugo Manuel Alves Henriques e Silva

March 5, 2024

1 Reading and reflection

The paper describes motivations, issues, technical details and demonstrations of one of the first computer systems in the research field of language understanding: GUS [1]. Developed by the research team at Xerox Palo Alto Research Center, the frame-driven dialog system is designed to manage a more-or-less realistic conversations in English, restricted to a specific domain; in this case, for the sake of the demonstration, GUS is acting as a travel agent to help planning return trips to a single city in California. The main goal, as the researchers stated, is to show the potential of a system attempting to model a natural dialog with a dedicated implementation.

The author first discusses GUS's motivations and design issues, trying to highlight the challenges that can be encountered by developing such systems. Most of the problems encountered are related to the natural dialog itself and represent situations difficult to understand or to be tackled by the system. Some of them are listed below.

- In a dialog we can encounter a **mixed initiative** from both the participants, where each of them has different expectations about the answers from the other.
- A human could write **indirect answers** to questions asked by the system, leading to a situation where making the inference necessary to answer becomes complex.
- Some conversations can even conform to **specific patterns**, special ones in specific contexts (in this case, those in a conversation with a travel agency). Those are only poorly understood or require very sophisticated strategies to be processed.

To handle some of these challenges GUS was developed such that the system's structure is as modular as possible. All the components, from the morphological to the syntactic analyzer, are decoupled from each other: this led to several advantages, from greater control over the conversation to better system monitoring and ease of use of external databases.

In the second part of the work, the paper goes deeper into the technical details of GUS, especially from the knowledge bases and processes point of view. Several key components are mentioned, based on their usage in the dialogue system: from a morphological analyzer to a frame reasoner, with a language generator at the end of the pipeline which is guided by a query map to generate appropriate answers. Another important core block is the concept of **frame**: It represents collections of information at several levels of the system, used to guide a conversation through **slots** that will be filled up with information provided by the users. Lastly, an overview of the procedures of GUS and their main types are presented: in general, they are associated with slots to indicate which kind of operation has to be performed.

After comprehensive examples that show how frames can be used to direct the course of a conversation, the author points out the limitations of the system related to a real conversation. While GUS can handle realistic conversations in general, It cannot determine for example when answers are referring to previous parts of the conversation, leading to unexpected behaviors and problems. To wrap up, the paper concludes by considering the potential improvements of such conversational systems: increasing the power of the already mentioned components, rather than improving the system for the language analysis and the knowledge representation, might lead to a better language understanding also in multiple domains.

2 Implementation

The proposed task is to build a chatbot that is able to have a natural conversation with the user. Despite the complexity of the task and the relatively short time schedule, we built a chatbot that tries its best not only in answering domain-specific questions, but also to guide the flow of the conversation. Later on, we will refer what we could have done to further improve the dialogue system.

The chatbot will be able to identify 3 different contexts: **Weather and Temperature**, **Food and Restaurants**, and **Transportation and Travelling** (flights).

2.1 Data

First things first, there are multiple approaches as to what data to use. It is referenced that for the purpose of this task we can simply use fictitious or random data. To make it as realistic as possible, we could rely on APIs, such as weather or transportation ones to provide us with real-time information. However, for our work, due to lack of time, we either generate a small synthetic dataset to exemplify use cases or used a pre-existing dataset.

- The Weather Dataset - Dataset which contains daily temperatures for different seasons of the year for multiple cities.
- Transportation Dataset - Generated dataset (with a python script) with flight numbers, origins, destinations and departure times.
- Restaurant Dataset - Generated dataset (with a python script) with the city, the cuisine the restaurant serves, the street name and the rating.

These datasets serve as examples of how we could approach the data problem in the given task.

2.2 Method

The chatbot heavily relies on keywords matching. Not only was the task to implement a dialog system, but also to implement a non-trivial task.

The keywords used for each context are the following:

- Weather: weather, temperature, forecast, season, rain, sun, clouds, wind, humidity, precipitation, thunderstorm, rainbow, snow;
- Food: restaurant, food, cuisine, meal, dinner, lunch, breakfast, dining, eat, drink, menu, dish, snack, taste, flavour, delicious, yummy, tasty, hungry, thirsty;
- Traveling: tram, bus, transportation, flight, train, car, bicycle, walk, drive, ride, commute, journey, trip, travel, commute, commuting, commuter, commuters, transit

Furthermore, when questions where only affirmative or negative answers can be given, we use the following words to know which of them the answer is: yes, yep, yeah, sure, ok, okay, fine, of course, absolutely, definitely, indeed, aye, yea, yah, yahs, yap, yup, ye, yessir, yes ma'am, yessiree, yessum, yea, yessuh.

As the non-trivial task, we have the flow of the conversation being controlled by the bot. It will proceed to ask the user questions, just like in a dialog, until it has enough context about the user to create a travel plan for him. Despite not being the most realistic plan, this serves as a way to demonstrate the logic behind the chatbot. Essentially, the user may start on any of the contexts: Weather, food, or traveling. Depending on the first context given by the user, the conversation flow will be different. The bot will try to continuously ask questions until it has gone through these main 3 topics unless the user decides not to continue the conversation.

As backup plans for user answers that the bot does not recognize, it will simply re-ask the question or assume that the user is not comfortable enough to provide the information about him.

2.3 Examples

As a demonstration of the chatbot capabilities and traits described above, down below we will show examples of conversations. This for each topic of expertise of the system.

```

Chatbot: Hi! I am your assistant for the day! I may not be the best, but never underestimate me! How can I help you, friend?
Me: Hey! I would love to know the temperature in Berlin
Chatbot: Uhhh, let's talk about weather, I love thunderstorms and rainbows!
Chatbot: Would you like to know the weather for a specific season? If so, which one?
Me: Sure, now that we are in the Summer might as well know for it, right?
Chatbot: The average temperature in Berlin during Summer is 19.957608055217 in 2023.
Chatbot: Talking about seasons, do you have a favorite one?
Me: I Love Spring to be honest
Chatbot: Just like me! Spring is the best season, isn't it? The flowers, the birds, the sun... it's all so magical!
Chatbot: Is that the city you live in by any chance?
Me: It is actually, yes
Chatbot: Great! I'll remember that for next time!
Chatbot: Question! Question! Question! I have soooo many questions for you! Let's see...
Chatbot: Have we talked about cuisine? I know soooo many restaurants. What's your favourite cuisine? I love Italian personally.
Me: Unpopular opinion but Belgian is my favourite!
Chatbot: Great choice! I know plenty of Belgian restaurants around the world!
Chatbot: I found 4 restaurants serving Belgian cuisine.
        Belgian cuisine is served in Paris and it has a rating of 4.2.
        Belgian cuisine is served in Warsaw and it has a rating of 4.9.
        Belgian cuisine is served in Edinburgh and it has a rating of 3.9.
        Belgian cuisine is served in Venice and it has a rating of 3.6.
Chatbot: Eventually, if I get to know you better I could recommend a restaurant for you...
Chatbot: Question! Question! Question! I have soooo many questions for you! Let's see...
Chatbot: Where would you like to go? Like a... dream place? I can find you a flight!
Me: Honestly, Paris is just the best, I really want to go there
Chatbot: Here are the flights departure times from Berlin to Paris for today:
Flight Number: EUN499, Departure at: 06:00
Flight Number: EUN497, Departure at: 10:00
Flight Number: EUN498, Departure at: 14:00
Flight Number: EUN499, Departure at: 18:00
Flight Number: EUN499, Departure at: 22:00
Chatbot: I have an idea! Now that I know you well enough, I think I can make a plan for you! Ideally, I would make you travel to your favourite city, eat your favourite cuisine and during your favour
like season, but I'm not sure that is possible, let's see...
Chatbot: I found a flight from your city to your favourite city! It leaves at 06:00! GO GET READY!
Chatbot: And I found a flight back from Paris to Berlin for the next day! It leaves at 06:00! Don't miss it!
Chatbot: Uff... That was a heavy task... I'm not sure I can keep up with you for much longer...
Chatbot: Do you have any other questions for me?

```

Figure 1: Examples of conversation the chatbot, by starting the conversation on the weather topic.

```

Chatbot: Hi! I am your assistant for the day! I may not be the best, but never underestimate me! How can I help you, friend?
Me: Hello! I have a passion for food and restaurants, do you know where I can get some Finnish food?

Chatbot: Someone's hungry! Now that you mention it, I could go for a bite too... Anyway...
Chatbot: I found 6 restaurants serving Finnish cuisine.
    Finnish cuisine is served in London and it has a rating of 4.7.
    Finnish cuisine is served in Paris and it has a rating of 4.3.
    Finnish cuisine is served in Berlin and it has a rating of 4.2.
    Finnish cuisine is served in Helsinki and it has a rating of 4.5.
    Finnish cuisine is served in Frankfurt and it has a rating of 4.9.
    Finnish cuisine is served in Helsinki and it has a rating of 4.6.
Chatbot: Eventually, if I get to know you better I could recommend a restaurant for you...
Chatbot: Question! Question! Question! I have soooo many questions for you! Let's see...
Chatbot: I wonder what the weather is like where you live... Where exactly do you live?
Me: As of now, I live in Paris

Chatbot: Would you like to know the weather for a specific season? If so, which one?
Me: Paris in the Spring is great, tell me the temperatures for that season!

Chatbot: The average temperature in Paris during Spring is 11.68434762088956 in 2023.
Chatbot: Talking about seasons, do you have a favorite one?
Me: Winter has to be it, I just love snow

Chatbot: Ohh, interesting choice! My favourite one is by far Spring! You get to see the flowers bloom and the birds chirp! It's magical! Smells nice too!
Chatbot: My favourite city in the world is Lisbon! Have you ever been there?
Me: I sure have, such a great city! And the people are also amazing, everyone is so helpful there

Could not agree more! Lisbon is a beautiful city!
Chatbot: Do you have a favourite city? Tell me about it!
Me: Despite the weather there, London is my favourite place to always travel to

Chatbot: So deep... I can feel the love you have for this city... I bet the weather is not as good as it is in Lisbon though... Let's check it out!
Chatbot: Would you like to know the weather for a specific season? If so, which one?
Me: Seems like you are joking with me! Sure, tell me the temperatures during Winter

Chatbot: The average temperature in London during Winter is 6.2932283898385 in 2023.
Chatbot: Question! Question! Question! I have soooo many questions for you! Let's see...
Chatbot: Do you want to know the transportation schedule from your city to your favourite city?
Me: Sure

Chatbot: I have an idea! Now that I know you well enough, I think I can make a plan for you! Ideally, I would make you travel to your favourite city, eat your favourite cuisine and during your favourite season, but I'm not sure that
is possible, let's see...
Chatbot: I found a flight from your city to your favourite city! It leaves at 06:00! GO GET READY!
Chatbot: And I found a flight back from London to Paris for the next day! It leaves at 06:00! Don't miss it!
Chatbot: Uff... That was a heavy task... I'm not sure I can keep up with you for much longer...
Chatbot: Do you have any other questions for me?

```

Figure 2: Examples of conversation the chatbot, by starting the conversation on the food topic.

```

Chatbot: Hi! I am your assistant for the day! I may not be the best, but never underestimate me! How can I help you, friend?
Me: I feel like going on a trip right now! Can you find any flights from Paris to London today?

Chatbot: Transportation, huh? I'm not a big fan of buses. Trains though... something about trains, I swear...
Chatbot: So you want to travel from Paris to London?
Me: Yes, absolutely

Chatbot: Alright, let me check the transportation schedule for you.
I found 5 flights from Paris to London.
Flight Number: E08246, Departure at: 06:00
Flight Number: E08247, Departure at: 18:00
Flight Number: E08248, Departure at: 22:00
Flight Number: E08249, Departure at: 18:00
Flight Number: E08250, Departure at: 22:00
Chatbot: Just out of curiosity, what are you planning to do in London? What's the occasion? A wedding hehe? I know! Must be your favourite city, right?
Me: London is in fact my favourite city, yes

Chatbot: I know it! I'm so good at this... I'm like a mind reader or something...
Chatbot: Question! Question! Question! I have soooo many questions for you! Let's see...
Chatbot: Do you want to know the weather in your favourite city?
Me: Weather is never great in London, but sure go ahead

Chatbot: Would you like to know the weather for a specific season? If so, which one?
Me: It's Summer right now, so tell me for the Summer!

Chatbot: The average temperature in London during Summer is 18.18195823799332 in 2023.
Chatbot: Talking about seasons, do you have a favorite one?
Me: Not many people enjoy Autumn as much as me, but I love seeing the leaves fall!

Chatbot: Ohh, interesting choice! My favourite one is by far Spring! You get to see the flowers bloom and the birds chirp! It's magical! Smells nice too!
Chatbot: Question! Question! Question! I have soooo many questions for you! Let's see...
Chatbot: Have we talked about cuisine? I know soooo many restaurants, what's your favourite cuisine? I love Italian personally.
Me: French food is one of the greatest

Chatbot: Great choice! I know plenty of French restaurants around the world!
Chatbot: I found a restaurant serving French cuisine.
    French cuisine is served in Helsinki and it has a rating of 3.4.
    French cuisine is served in Glasgow and it has a rating of 3.9.
    French cuisine is served in Paris and it has a rating of 4.2.
Chatbot: Eventually, if I get to know you better I could recommend a restaurant for you...
Chatbot: I have an idea! Now that I know you well enough, I think I can make a plan for you! Ideally, I would make you travel to your favourite city, eat your favourite cuisine and during your favourite season, but I'm not sure that
is possible, let's see...
Chatbot: I'm sorry, I couldn't find a restaurant that serves your favourite cuisine in your favourite city... But you can still go there and enjoy a meal!
Chatbot: I found a restaurant in your favourite city! It serves Finnish food and is located at Lily Place 763!
Chatbot: Uff... That was a heavy task... I'm not sure I can keep up with you for much longer...
Chatbot: Do you have any other questions for me?

```

Figure 3: Examples of conversation the chatbot, by starting the conversation on the traveling topic.

2.4 Further development

Building a chatbot can represent a very complex task and, as already mentioned, having a longer time for the development of such a system can bring interesting further developments. We will now describe several possible improvements to

make our dialogue system more advanced: they will be especially focused on *Semantic Role Labeling* and *Discourse Coherence*.

- **Natural Language Understanding (NLU)**: one of the first main improvements we can apply is related to how the bot can understand a user prompt. This can be done in several ways: it can be possible to use NLP libraries such as `SpaCy` to tokenize the input text and associate tags related to both the type of words (name, article, pronouns, ...) and the type of entity they represent (places, people, object, ...). This, alongside more advanced input utterances mapping, from using regular expression patterns to dedicated machine learning models. The goal of these techniques is to create better discourse relations, instead of determining if the user is staying on a topic.
- **Enhanced knowledge Base**: especially in complex contexts, a more advanced and structured knowledge base can improve a dialogue system's performance in understanding the user intent, extracting meaningful information and generating a proper answer. We could use for example a suitable database, from a relational one (like `SQLite`) to a graph-base (such as `Neo4j`).
- **Sentiment Analysis**: detecting emotional cues in the user's word choice and phrasing can be essential to improve not only the quality of the bot answers, but also to adjust its tone accordingly; moreover, taking into account and classifying different emotions can help enhance the emotional intelligence of the bot, allowing more targeted responses. Pre-trained models such as `BERT` or `Llama` can be used to achieve such a goal: we can compute a document (prompt) sentiment and aggregate by sentence-level scores, followed by classifying sentiment on aspect using attention mechanisms.

References

- [1] A. Author and B. SecondAuthor. Gus, a frame-driven dialog system. In *Proceedings of the Conference on Dialog Systems (DIALOG)*, pages 123–134, 1977.