KTH Computer Science
and Communication

A data-driven approach for a chatbot using transcripts from a TV-series

NIKLAS BLOMQVIST
HENRIK HYGERTH

Degree Project in Computer Science, First Level, DD143X
Supervisor: Gabriel Skantze
Examiner: Karl Meinke, Örjan Ekeberg

April 29, 2014

Abstract

This paper explores a data-driven approach for a chatbot where it utilises a database of transcripts from a TV-series in an attempt to hide its lack of linguistic knowledge and familiarity of the human language. A built-upon version of the naive implementation with rules prioritizing responses from the same scene to increase coherency is implemented and is compared to the original. Both versions of the chatbot shows some coherency for the individual instance but this diminishes in longer conversations. No significant di fference can be found between the two versions.

Referat

En datadriven metod för en chatbot som använder manusskript från en TV-serie

Denna rapport undersöker en datadriven metod för en chatbot där den använder en databas av repliker från en TV-serie i ett försök att dölja sina bristande språkkunskaper. En utbyggd version av den naiva implementationen som har regler för att prioritera repliker från samma scen som föregående svar implementeras och jämförs med den ursprungliga. Båda versionerna av chatbotten uppvisar koherens för enskilda fall men detta förminskas i längre konversationer. Ingen signifikant skillnad mellan de två versionerna påvisas.

Contents

Chapter 1

Introduction

In 1950 Alan Turing published the article "Computing Machinery and Intelligence" in which Turing describes a game for testing if a machine has intelligence to make itself indistinguishable from a human. [1] The game describes an interrogator that is supposed to question a human and a machine through a conversation, often through a chat program, to determine who is the human. This is now known as the Turing test and if the interrogator cannot distinguish the human from the machine, the machine has passed the test. [2] The first chatbot implementation able to play this game was published by Joseph Weizenbaum in 1966 and was called ELIZA. [3]

Today, there exists different types of chatbots and the two most common ones are either rule-based or data-driven. Rule-based chatbots processes the user input through a large collection of rules. An example of a rule-based chatbot is A.L.I.C.E. which uses AIML, an XML language created specifically for chatbots to match the input pattern to predefined response templates. It then replaces the keywords in the responses with keywords extracted from the input to make the context of the response seem relevant. [4] There are many different types of data-driven chatbots, some scrapes the internet for data and some searches a database for data.

This paper will explore a data-driven approach where the chatbot uses a database of transcripts from a TV-series. Transcripts from TV-series contain human interaction and behaviour and with the use of this the chatbot will try to hide its lack of knowledge in the human language. A built-upon version of the chatbot will be implemented using a few rules in an attempt to increase coherency.

1.1 Problem statement

Implementing two versions of chatbots using transcripts. Both versions will be data-driven but the second version will have rules for prioritizing lines from the same scene as the previous response in an attempt to increase coherency. Version 2 will also have a minimum matching score, and if this is not satisfied, a default answer will be replied to avoid answering based on low matchings.

Can the chatbot produce relevant answers and will there be any noticeable difference between the two versions? If there is a difference, will version 2 be more coherent than the first version?

Chapter 2

Background

Chatbots or chatterbots, as they were originally called, have come a great way since the first implementation was published in 1966 by Joseph Weizenbaum, the so called ELIZA chatterbot [3]. This was 16 years after the publication of "Computing Machinery and Intelligence" by Alan Turing [1]. In the article Turing introduces "The Imitation Game", a game consisting of three people, a man, a woman, and an interrogator. The interrogator has no idea which of, lets call them A and B, is the man and which of them is the woman. A:s object in the game is to confuse the interrogator as to what he is, a man or a woman, while B:s object is to try to help the interrogator. The conversation between A,B and C is carried out by text or a courier as to not give any obvious clues(e.g. voice, appearance etc). The question is then asked "What will happen if a machine takes the part of A in this game?". Will the interrogator decide wrongly as often when the game is played like this as he does when the game is played between a man and a woman? The first machine able to play this game was ELIZA by using it's tricks in an attempt to fool the interrogator. The most important trick was to answer a question with a question.

The primary goal with a chatbot is to fool the user into thinking the responses made are from a human and not from a machine, and thereby creating a false sense of intelligence, an artificial intelligence (AI). After ELIZA there have been many chatterbots to follow and in 1991 Dr. Hugh Loebner pledged a grand prize of $100000 and a gold medal for the first computer whose responses were indistinguishable from a human's in an instantiation of the The Imitation Game, also known as the Turing Test. [5] No implementation of a chatterbot has been able to take home this prize yet but each year an annual cash prize and a bronze medal is awarded to the most human-like computer.

Many chatbots today are rule-driven, that is, they interpret the user input and with different rules decide how to build the response. Data-driven chatbots on the other hand connects to some database and the user input string is then interpreted

and matched to a relevant response found in the database. This could either be in the form of a web search where the data returned is something found online or in the form of a text from an existing database. A data-driven chatbot matches the input with an existing response while a rule-driven chatbot tries to generate a response from the input.

A common problem previously with using data as a source was that the storage capacity was low and it was hard to maintain data divided over multiple volumes. Since the beginning of the 21th century storage capacity have been rapidly increasing and the prices has decreased which means that collecting big amount of data can be done on a smaller budget. [6] Using high volumes of data as a source for different project is now maintainable. Data-driven chatbots have benefited from this development and more data-driven chatbots has started to emerge.

An example of a data-driven chatbot is AI_AGW, a chatbot which analyses tweets(messages) made on twitter; if it comes across a tweet which contains an anti-science text about global warming it makes a response to the user with a counter argument contained in a database. To select the counter argument it matches it to the argument made in the original tweet. [7]

2.1 Natural Language Processing

Natural Language Processing or NLP is the computerized approach for analyzing text based on theories and technologies and its goal is to accomplish human-like language processing. The field of NLP was in the early days of AI referred to as Natural Language Understanding, NLU and while it is well-agreed upon that the true goal of NLP is NLU, the goal has not been accomplished. Natural Language Processing is in fact divided into two distinct focuses, language processing and language generation where the first of these is equivalent to the role of reader/listener and the second is equivalent to the role of writerspeaker. They share much of the underlying theory and technology but language generation also require a planning capability. This paper will focus on the language processing aspect of NLP for handling the user input of the chatbot. [8]

2.2 Approximate string matching

There are many different areas where matching the similarity between two strings is useful. Areas such as text searching, computational biology, pattern recognition, signal processing etcetera. Many times errors of input needs to be handled, an example of this is when spell checking user-generated words to a dictionary. Approximate string matching is the technique of matching strings approximately rather than exactly. [9] There exists a variety of ways for matching strings approximately and a particularly widely-used method is the edit distance, also known as

the Levenshtein distance. It calculates the minimum amount of character operations (insertions, deletions and replacements) to make the two strings equal. [10] Another related approach is finding the length of the longest common subsequence of the two strings. Other similar approaches are based on calculating the number of shared substrings of length n, so called n-grams. [11]

2.3 N-gram

N-gram, also known as q-gram, is the method of dividing strings into an array of substrings of length n. This method can be used to approximately match two strings with each other. [11] The standard n to use is 3, creating so called trigrams. The string “how are you” is then divided into substrings of 3. See Table 2.1.

Trigrams
how
ow_
w_a
_ar
are
re_
e_y
_yo
you

Table 2.1. Example of trigrams for the string "how are you"

The following formula describes the percentage of similarity between the trigrams of two strings, X and Y, which can be used for comparing matchings. See Figure 2.1 [11]

$$\frac{2 \times |trigrams(X) \cap trigrams(Y)|}{|trigrams(X)| + |trigrams(Y)|}$$

Figure 2.1. Formula for calculating similarity between two strings trigrams

Chapter 3

Implementation

The implementation consists mainly of three parts. The first part is the creation of the database. The transcripts from the TV-series is converted to, by the chatbot, a readable format and the trigrams for each line is generated. The second part of the implementation is where the algorithm of the chatbot tries to match the trigrams of the user input to the trigrams of the lines in the database. The third part is an extension of the first implementation of the chatbot where ways to increase coherency is explored.

3.1 Tools

The chatbot was implemented in Node.js using JavaScript and it was chosen on the grounds that it is good for building fast, scalable network applications. The chatbot uses a MongoDB database because it integrates well with Node.js. [12] [13]

The TV-series transcripts is parsed to a JSON-file using a small program written in Java. The JSON-file is processed by the Node.js application, where trigrams are generated from the lines in the TV-show transcripts and saved into a MongoDB database. Using trigrams(n=3) as n-grams is a tradeoff between performance and accuracy. Using bigrams(n=2) will increase the size of the database whilst using quadgrams(n=4) will decrease the matching accuracy.

3.2 Generating trigrams

Transcripts from The Big Bang Theory [14] was written into files and was afterwards parsed into a JSON-file with a small Java program. See Figure 3.1 for JSON-file structure.

Line
season
episode
scene
line

Table 3.1. JSON-file data structure

The JSON-file was then parsed into a database adding the trigrams of each line to every object of the database. The trigrams generated using the following algorithm.

---

line ← remove special characters from line
line ← transform line to lowercase character
if *length of line* < 3 **then**
  line ← add underline as padding so the line is at the minimum length of 3 characters
end
line ← replace spaces with underlines
line ← add underline characters at the beginning and at the end of the line
trigram ← [*String*]
$i \leftarrow 0$
while $(i + 2) <$ *length of line* **do**
  $t \leftarrow line[i] + line[i + 1] + line[i + 2]$
  trigram ← push t to trigram
  $i++$
end
return trigram

Figure 3.1. Algorithm for creating trigrams for the string object line

3.3 Database structure

A MongoDB [13] database was used to store the lines and trigrams. The database contains one table (see Figure 3.2) which includes the information about the lines. The size of the database is three seasons of transcripts from the TV-series The Big Bang Theory which corresponds to over 14000 lines.

Line
_nextline
season
episode
scene
line
_id
trigrams

Table 3.2. The database structure

3.4 Matching algorithm

The algorithm (for flowchart see Figure 3.3) for matching the input with the lines in the database works as following. Firstly, all of the lines in the database are collected and trigrams for the input is generated. Secondly, the trigrams generated from the user input is compared with the trigrams from every line in the database to see how many trigrams they have in common. Thirdly, a score is calculated with the formula:

$$\frac{2 \times |trigrams(X) \cap trigrams(Y)|}{|trigrams(X)| + |trigrams(Y)|}$$

Figure 3.2. Formula for calculating similarity between two strings trigrams

The scores are saved and sorted ascending the highest to lowest in a list. The best match, the one with the highest score, is chosen to see if it has a reply e.g. that it is not the last line in a scene or in an episode. If it has a reply it will be used as the response of the chatbot, if not, the list is iterated until a match with a reply is found.

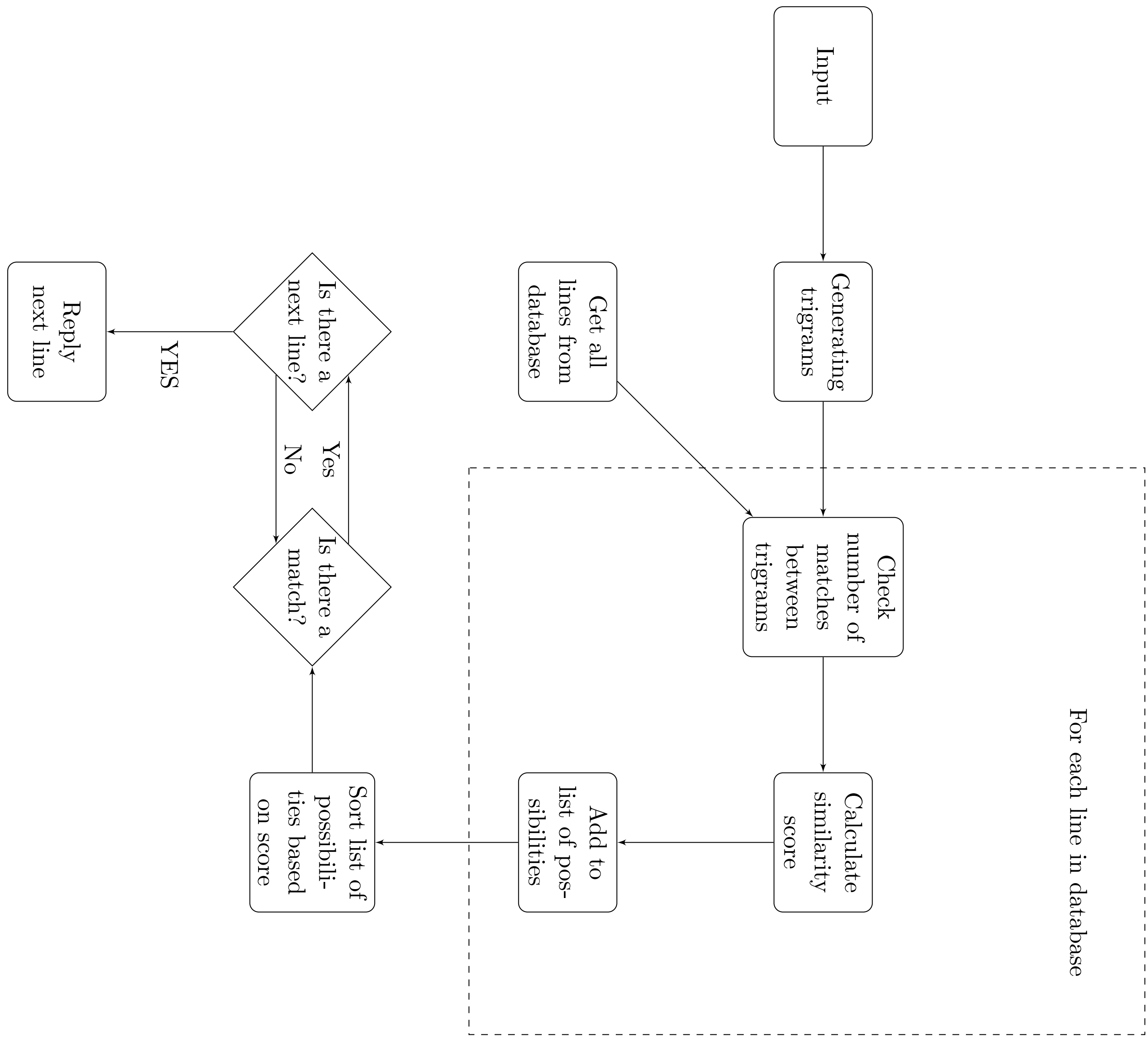

Figure 3.3. Flowchart over version 1 algorithm

3.5 Version 2

A built upon version of the chatbot is implemented to explore ways to increase coherency (see Figure 3.4). To do this, a threshold for valid matchings is added and matchings from the same scene as the previous response is prioritized. If the highest match from the same scene has a probability score of 0.6 or higher, it will be the reply, otherwise the database is iterated until a match is found. The highest match needs to have a probability score of 0.4 or higher to be valid or a default answer will be replied.

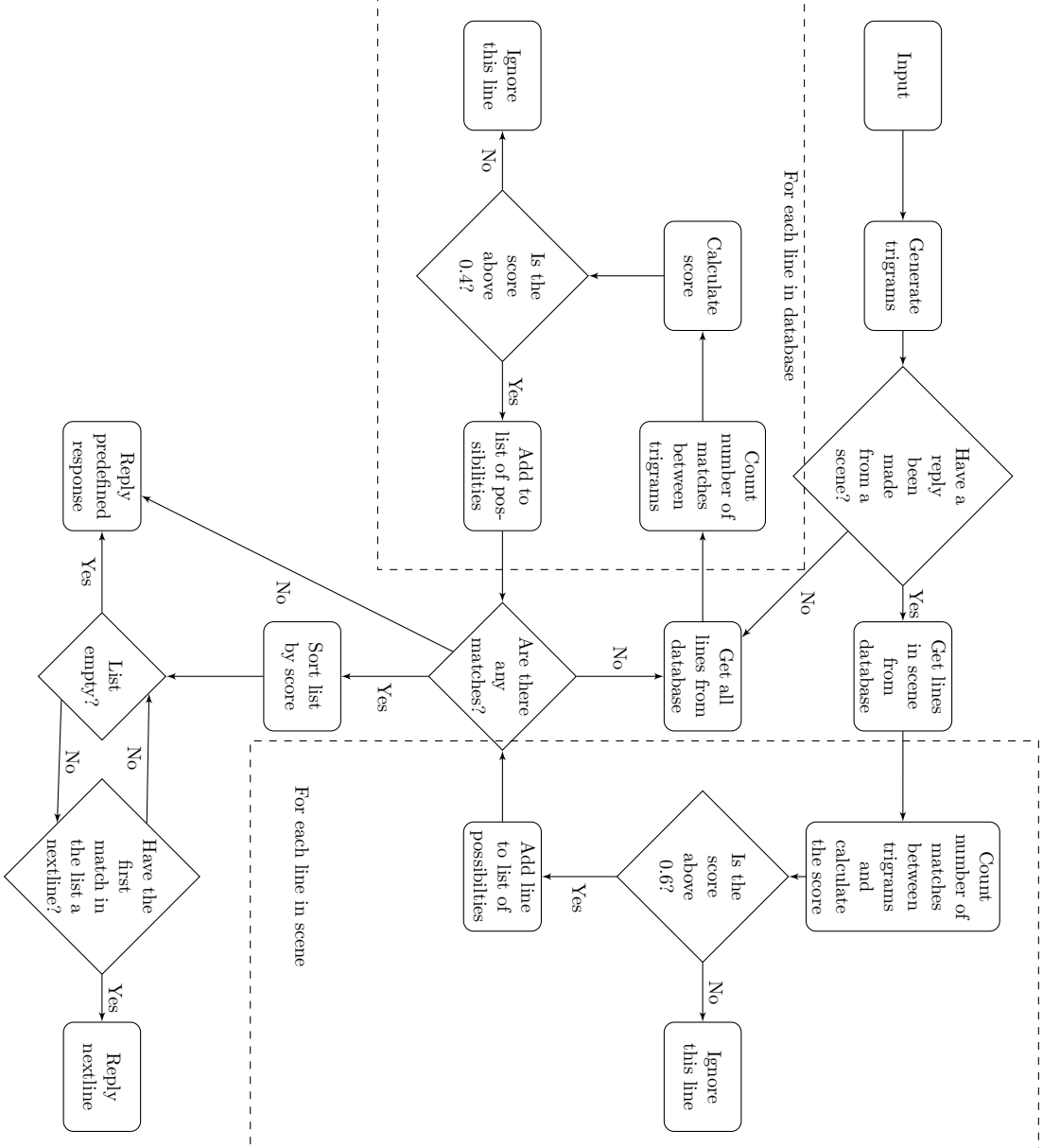

Figure 3.4. Flowchart over version 2 algorithm

Chapter 4

Method

A problem that can occur when matching strings is that the user input could contain errors. It is common to mistype words and use wrong sentence structure. This paper uses trigrams on a character level and not on a word level in an attempt to avoid this. The strings with words which are shorter than three characters are padded with underline characters to bring the length of them up to three characters. For example the string “hi” becomes “hi_”. To increase accuracy in the matchings all special characters such as commas, hyphens, question marks, etc. are removed from the trigrams. To further help the matching process all characters will be matched on their lowercase representations.

Since the chatbot uses itself of dialogues from a transcript the chatbot will generate human-like sentences. The behaviour of the chatbot and relevance of the replies can be questioned though. A challenge with this implementation is to keep the chatbot consistent and coherent and not straying away from the subject. This is a problem that will be especially noteworthy when user input consists of general questions such as “what is that?”. The chatbot will lack the recognition of context and will not know what “that” is. Some sort of strategy to prioritize the current subject is needed to establish a desired level of coherency. A solution tried in this paper is to force the algorithm to prioritize the lines in the same scene as the previous reply to raise the probability that the chatbot will respond with a relevant reply.

In this implementation of a chatbot the scope of the database will be the limiting factor of the chatbots ability to reply to input. This paper will not cover the optimization of the string matching algorithm for speed and efficiency since there already exists papers on this subject. A strategy for replying is therefore needed when the algorithm has not found a good match to the input. The chatbot will use a default answer when this happens and reply with “I don’t know”.

An examination of the coherency of the two versions is made. 24 transcripts with 10 lines each, 5 from the user and 5 from the chatbot, is created and evaluated by giving them scores(0-5) on their coherency levels. An average of the versions is then calculated and compared. The two authors of this paper created 12 transcripts each and evaluated each others transcripts to avoid biases.

Chapter 5

Result

5.1 Algorithm test

Following are four basic examples of individual instances from version 1 of the chatbot. The scores are calculated with the formula in Figure 3.2, which describes the similarity between two strings. See Figure 3.3 for a flowchart of the procedure.

Type	Line	Score
User input	What is your name?	
Matched line	What´s your name?	0.813
Chatbot reply	Rajesh Ramayan Koothrappali	

Table 5.1. Instance 1

Type	Line	Score
User input	Where are you from?	
Matched line	Hi, Raj. Where are you from?	0.837
Chatbot reply	The mysterious subcontinent of India.	

Table 5.2. Instance 2

Type	Line	Score
User input	wahts your manes?	
Matched line	What´s your names?	0.438
Chatbot reply	I´m Howard.	

Table 5.3. Instance 3

Type	Line	Score
User input	What color is a banana?	
Matched line	This is banana bread.?	0.429
Chatbot reply	This is a door knob.	

Table 5.4. Instance 4

5.2 Coherency test

These are the results from the coherency test (for the used test material, see Appendix A):

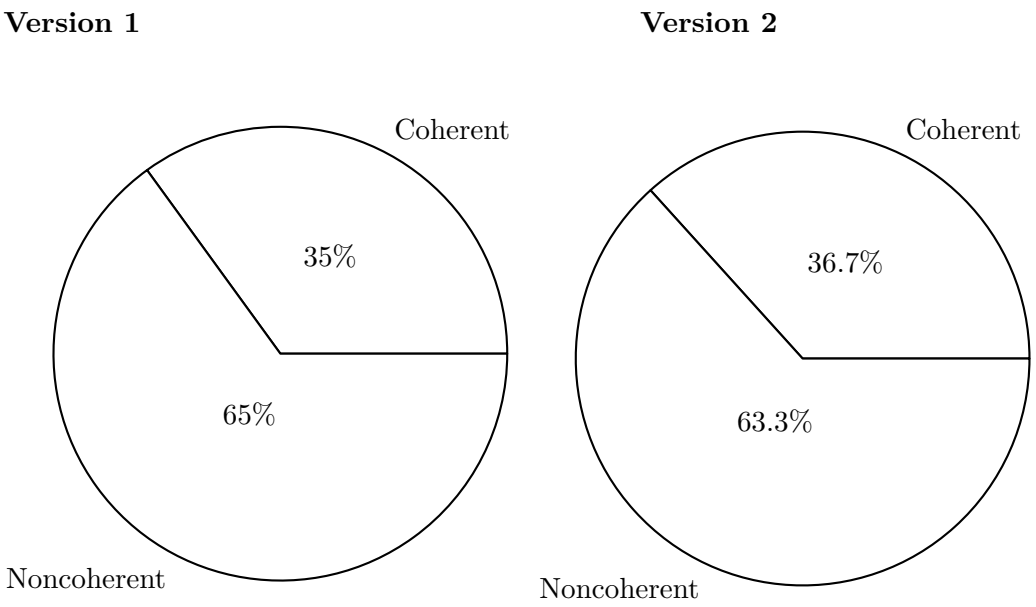

Figure 5.1. The percentage of experienced coherency in the replies, left is version 1 and right is version 2

Chapter 6

Discussion

The matching algorithm used to compare similarity between two strings is accurate, as seen in Instance 1 (see Table 5.1) and Instance 2 (see Table 5.2), and can handle smaller mistypings as seen in Instance 3 (see Table 5.3). Without any type of filtering though, the chatbot will not know which words are important and which words are "noise" (words that doesn't affect the meaning of the string). In Instance 4 (see Table 5.4) for example, the user asks for a color but the algorithm can only handle string similarity and finds the highest match with the words "banana" and "is". The resulting reply becomes irrelevant and inconsistent with the user input.

One might argue that a bigger database will result in better responses but in this case, it might not be the truth. Without some sort of linguistic capability or contextual rules the chatbot will have little to no knowledge of the context of the conversation. With a large enough database, a relevant human-like response to the individual instance of user input will most likely occur but in a longer conversation the behaviour of the chatbot will start to seem odd and irrelevant.

A problem with increasing the size of the database is the timeframe for the algorithm to find all matchings. In this implementation of the chatbot, with a database consisting of over 14000 lines, it takes the algorithm 2 seconds to find all the matchings. This is acceptable for a chatbot. An optimization of the chatbots algorithm and database structure is needed if the size of the database is to be increased.

As shown in the examination of coherency with the two versions (see Figure 5.1), version 2 displays a very small increase in coherency from the first version. This increase is too small to be judged significant for this evaluation and both versions can be seen as equally coherent. The level of coherency of the two versions is about 35% which is low. A conversation between two grown human beings will have a coherency level closer to 100%. If the coherency level is at around 80%, 1 of 5

statements or questions will fall out of context and be completely irrelevant which is a huge amount. A good chatbot will need to surpass the 90% limit for it to resemble a conversation between humans.

One factor for the low coherency level of the two implementations is because of the chatbots lack of understanding of context. Input referring to previously mentioned context will have no effect on the chatbot because it cannot remember what have been said in the conversation. The input will only be matched to a similar line in the database and the reply to that specific line is then returned. To increase coherency the user input needs to be as specific and standalone as possible.Another factor of the low coherence level is that the dialogues in the transcript can be between more than two people and this can make the conversation with the user seem incoherent. This is a problem related to the source of dialogues rather than to the chatbot itself.

There are many ways to improve the chatbots coherency. The improvement that would probably give the best increase of coherency would to give the chatbot some knowledge of what words in an sentence is important. This would mean that the chatbot could filter words that acts as "noise" in the sentences and prioritize matching the important words. Another improvement would be to categorize the lines by the subject of the dialogues in the transcript before adding them to the database. This way the chatbot could analyze the subject of the input and only match the input against lines in the database which cover the same subject. This would improve the response time of the chatbot and possibly give more relevant answers.

Chapter 7

Conclusion

This paper have explored the possibility of using transcripts from TV-series as a source for a data-driven chatbot. Two versions have been implemented to test if the chatbot can make coherent responses in a conversation with a user. One version that matches the input with lines in the database and a built-upon version that prioritizes lines in the same scene as the previous response.

From tests on coherency it was found that there was no noticeable difference between the two versions and although the chatbot can produce relevant answers for the individual instance it lacks consistency and coherency in longer conversations.

Bibliography

[1] Alan M. Turing. *Computers & Thought*. chapter Computing Machinery and Intelligence, pages 11–35. MIT Press, Cambridge, MA, USA, 1995.

[2] Graham Oppy and David Dowe. *The Turing Test*. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Spring 2011 edition, 2011.

[3] Joseph Weizenbaum. *ELIZA—a Computer Program for the Study of Natural Language Communication Between Man and Machine. Commun. ACM*, 9(1):36–45, January 1966.

[4] R. Wallace. *The Anatomy of A.L.I.C.E.* In R. Epstein, G. Roberts, and G. Beber, editors, *Parsing the Turing Test*. Springer, 2009.

[5] *What is the Loebner Prize? - Home Page.* [Online]. Available from: `http://www.loebner.net/Prizef/loebner-prize.html`, 2013. Accessed: 2014-03-04.

[6] Phillp Russom. *Big Data Analytics. TDWI Best Practices Report*, (4), 2011.

[7] Christopher Mims. *Chatbot Wears Down Proponents of Anti-Science Nonsense.* [Online]. Available from: `http://www.technologyreview.com/view/421519/chatbot-wears-down-proponents-of-anti-science-nonsense/`, November 2010. (2014-03-05).

[8] Elizabeth D Liddy. *Natural language processing. Encyclopedia of Library and Information Science*, 2001.

[9] Ricardo A. Baeza-Yates and Gonzalo Navarro. *Fast Approximate String Matching in a Dictionary.* In *SPIRE*, pages 14–22, 1998.

[10] *The Levenshtein-Algorithm - Home Page.* [Online]. Available from: `http://www.levenshtein.net/`. (2014-04-08).

[11] Grzegorz Kondrak. *N-Gram Similarity and Distance.* In Mariano P. Consens and Gonzalo Navarro, editors, *SPIRE*, volume 3772 of *Lecture Notes in Computer Science*, pages 115–126. Springer, 2005.

[12] *Node.js - Home Page*. [Online]. Available from: `http://nodejs.org`. (2014-04-09).

[13] *mongoDB - Home Page*. [Online]. Available from: `https://www.mongodb.org`. (2014-04-09).

[14] *Big Bang Theory Transcripts - Website*. [Online]. Available from: `http://bigbangtrans.wordpress.com/`. (2014-04-09).

Appendix A

Coherency test material

A.1 Transcripts over dialogues from version 1

- User:
  – Hi
- Chatbot:
  – Hi.
- User:
  – What are you doing today?
- Chatbot:
  – I'm packing up all my collectibles and taking them down to the comic book store to sell.
- User:
  – Good for you.
- Chatbot:
  – Hah, yeah! Because he didn't want Sheldon to be embarrassed, and there is nothing to be embarrassed okay, every family in America has a relative holed up in a garage somewhere huffing paint thinner.
- User:
  – What has that to do with the collectibles?
- Chatbot:
  – Engaging my olfactory sense helps make it real for me.
- User:
  – Um, okay...
- Chatbot:
  – Well that's disappointing. (End of flashback.)

- User:
  – Are you hungry?
- Chatbot:
  – I could eat.
- User:
  – What would you like to eat?
- Chatbot:
  – Uh, sure, why not?
- User:
  – It is your decision.
- Chatbot:
  – It is.
- User:
  – That is what I said.
- Chatbot:
  – Doesn't concern you. You'll be going to the university every day?
- User:
  – I also want to eat and yes I will be going to university every day.
- Chatbot:
  – Yes.

- User:
  – What is up with you today?
- Chatbot:
  – I'm a romantic.
- User:
  – How nice for you. Are you doing anything romantic today, then?
- Chatbot:
  – . . . Bernadette! Thank you, Cheesecake Factory!
- User:
  – What? Were you supposed to meet Bernadette at the Cheesecake Factory?
- Chatbot:
  – Oh, uh, not much. We do have a chocolate key lime that's moving pretty well.
- User:
  – That sounds nice.
- Chatbot:
  – Hold on. Bi-monthly is an ambiguous term, do you mean move it every other month, or twice a month?

- User:
  – By Bi-monthly I mean every other month.
- Chatbot:
  – No.

- User:
  – Where do you work now?
- Chatbot:
  – I don't understand.
- User:
  – Where do you spend your days doing stuff?
- Chatbot:
  – You know, psychology journals, internet research, and there's this great show on VH1 about how to pick up girls.
- User:
  – So you are currently unemployed?
- Chatbot:
  – No.
- User:
  – So what do you do then, if you are not unemployed?
- Chatbot:
  – Oh, you know, goth stuff. Goth magazines, goth music.
- User:
  – So you work as a goth? Is that really a paying job?
- Chatbot:
  – You're tricking me. You tell me the truth, what do we get?

- User:
  – Where do you live these days?
- Chatbot:
  – I don't understand.
- User:
  – Where do you go to sleep during the night?
- Chatbot:
  – I don't understand.

- User:
  – Do you not sleep? Are you a vampire?
- Chatbot:
  – Oh, yeah, sorry, I'll have it cleaned.
- User:
  – What is it that is dirty? Was that a yes or a no on the whole being a vampire thing?
- Chatbot:
  – Hey, Leonard.
- User:
  – Please answer my question and I am not Leonard!
- Chatbot:
  – You answered the multiple choice questions in a pattern. A B B A C. A B B A C.

- User:
  – Hi
- Chatbot:
  – Sheldon says you're going to the North Pole.
- User:
  – Am I going to the North Pole? How cool, I didn't know..
- Chatbot:
  – Yes.
- User:
  – So, uh.. What am I going to do there?
- Chatbot:
  – Uh hey, it's your business. Do whatever you want.
- User:
  – Okay, I can build snowmen all day
- Chatbot:
  – Me either
- User: – Will you come with me to the North Pole?
- Chatbot:
  – Yes

- User:
  – So tell me something about yourself
- Chatbot:
  – You were a colossal ass-hat.
- User:
  – That's about me, I want to know something about you.
- Chatbot:
  – Uh, what do you mean?
- User:
  – You know, just tell me something about you.
- Chatbot:
  – What is it?
- Chatbot:
  – Okay I give up, obviously you don't want to tell me anything about yourself.
- Chatbot:
  – Dennis, we discussed this, we're in the process of updating our equipment, and we welcome your input on our research goals, and we've agreed to look the other way if you want to use up to 20% of the grant money you attract to smuggle your grandfather out of Pyongyang. We want him here boys, make it happen.
- User:
  – Who's Dennis?
- Chatbot:
  – Mr Kim was not only the valedictorian at Stamford University, he is also the youngest recipient of the prestigious Stephenson Award.

- User:
  – Hello
- Chatbot:
  – Hi, Howard.
- User:
  – My name is not Howard!
- Chatbot:
  – Doctor and Mrs. Koothrappali, namaste. I understand your concern, but if it'll make you feel any better, my mother is fine with me going, and this is a woman who kept a safety rail on my bed until I was 17.
- User:
  – What are you talking about?

- Chatbot:
  – Oh.
- User:
  – Are you drunk or confused or something?
- Chatbot:
  – Zack was a perfectly nice guy, and then you ruined him!
- User:
  – Clearly you are very confused.
- Chatbot:
  – Sorry, Mommy, Papa, but with all due respect I really can't go through...

- User:
  – Where are you?
- Chatbot:
  – I'm... uh... at work.
- User:
  – Oh, where do you work?
- Chatbot:
  – I don't understand.
- User:
  – What do you work with?
- Chatbot:
  – I need access to the Cheescake Factory's walk-in freezer.
- User:
  – Why? Is that where you work?
- Chatbot:
  – I don't know. I mean, you have to admit things seemed simpler when we were just friends.
- User:
  – We aren't even friends!
- Chatbot:
  – Then I hereby invoke what I'm given to understand is an integral part of the implied covenant of friendship. The favour.

- User:
  – Hello, I'm your creator. Hello, Sheldon.

- Chatbot:
  – But that's not my name.
- User:
  – And we agreed to never speak of it again.
- Chatbot:
  – I haven't agreed to anything like that. Good, 'cos they don't work.
- User:
  – What don't work?
- Chatbot:
  – No, if we fill up on dumplings we'll need to eliminate another entree.
- User:
  – Food? IS THERE FOOD?! Um, yeah, Halloween.

- User:
  – Hello
- Chatbot:
  – Hi, Howard.
- User:
  – Not my name, but fine. I heard you are senile. I am in every way.
- Chatbot:
  – How is that working out for you? Leonard, ready to go?
- User:
  – That's not my name either, but fine. Where are we going? We're gonna have sex.
- Chatbot:
  – No.. freaking.. way!
- User:
  – This is incredible! I'm so happy, I'm not even going to question their judgment in picking you. I'm just going to run home and start packing.

- User:
  – Hello. How are you today?
- Chatbot:
  – Mmm, menopausal.

- User:
  – Eww. Too much information.
- Chatbot:
  – Leonard, you've got one on your tail
- User:
  – I've got one menopausal on my tail? This makes real sense.
- Chatbot:
  – That's alright, my tails prehensile. I'll swat him off.
- User:
  – Are you not healthy in your head?
- Chatbot:
  – My head, what about yours?
- User:
  – I am healthy.
- Chatbot:
  – Yeah, me, too.

A.2 Transcripts over dialogues from version 2

- User:
  – Hi
- Chatbot:
  – Hello.
- User:
  – What is your name?
- Chatbot:
  – When I lie through my teeth to a woman, you nod and agree.
- User:
  – Okay, where are you from?
- Chatbot:
  – The mysterious subcontinent of India.
- User:
  – That sounds nice.
- Chatbot:
  – Hold on. Bi-monthly is an ambiguous term, do you mean move it every other month, or twice a month?
- User:
  – I mean twice a month.

- Chatbot:
  – Then no.

- User:
  – Are you hungry?
- Chatbot:
  – I could eat.
- User:
  – How about a pizza?
- Chatbot:
  – Yeah, we can always divide soup.
- User:
  – Split fifty-fifty?
- Chatbot:
  – Really?
- User:
  – If you don't mind?
- Chatbot:
  – Feel is a kind of a... it's a strong word. Um, I just think it would be a nice gesture on your part.
- User:
  – Let's do it then.
- Chatbot:
  – Alwight then. Weady. Set. Go.

- User:
  – Nice weather today.
- Chatbot:
  – Yeah, Stephanie got it for me. It's kind of fun.
- User:
  – How nice of her.
- Chatbot:
  – I don't know, but apparently I'm in some kind of relationship, and you seem to be an expert at ending them.

- User:
  – I don't see myself as an expert.
- Chatbot:
  – Shh!
- User:
  – Why do I need to be quiet?
- Chatbot:
  – No you didn't. Look, you are a great guy, and it is things you love that make you who you are.
- User:
  – Okay, I guess.
- Chatbot:
  – It would take the pressure off.

- User:
  – Where do you work?
- Chatbot:
  – I don't understand.
- User:
  – Why is that?
- Chatbot:
  – Doesn't concern you. You'll be going to the university every day?
- User:
  – Alright. Yes I will be going to the university every day.
- Chatbot:
  – Yes.
- User:
  – So I will see you tomorrow morning?
- Chatbot:
  – See ya. 'Bye, Penny. Have fun. (Bangs head on door. Sheldon opens it.)
- User:
  – Thanks.
- Chatbot:
  – I'm telling you, Leonard, video slot machines, free drinks brought to us by a bosomy barmaid and all you can eat shrimp for $3.95 Disneyland can suck it. This is the real happiest place on earth.

- User:
  – Hi there!
- Chatbot:
  – Hello. Sorry I'm late. But I was in the hallway, chatting up Penny.
- User:
  – No problem.
- Chatbot:
  – I had no idea it was so heavy. The thing just fell right over on me, didn't it?
- User:
  – What thing did fall on you?
- Chatbot:
  – Nothing. Diet coke. And yoghurt. A little nail polish.
- User:
  – Are those things really that heavy?
- Chatbot:
  – You're not jealous of Leonard, are you?
- User:
  – Why would I be jealous of Leonard?
- Chatbot:
  – Me? No. I may have mentioned that it's a little inappropriate to be asking another man's girlfriend to his experiment without first discussing it with said man.

- User:
  – What is up with you?
- Chatbot:
  – I'm a romantic.
- User:
  – Can you prove that you are a romantic?
- Chatbot:
  – Cathedra mea, regulae meae. That's Latin for my chair, my rules. Now, you said on the phone that your area of study is physics.
- User:
  – What chair? Yes, I study physics on my spare time.
- Chatbot:
  – I don't know.
- User:
  – I am simply asking you about that chair you just mentioned.

- Chatbot:
  – I don't know.
- User:
  – Why is it a secret chair?
- Chatbot:
  – Well, I just, I thought...

- User:
  – How is work at the cheesecake factory?
- Chatbot:
  – Come on, I'll show you the trick with the shower.
- User:
  – What? Wait, are you a magician?
- Chatbot:
  – My friend, Justin.
- User:
  – Is Justin an magician?
- Chatbot:
  – Your turn, Leonard.
- User:
  – Are we playing a game now?
- Chatbot:
  – Peppermint schnapps.
- User:
  – Oh, we are playing a drinking game.
- Chatbot:
  – Peppermint schnapps.

- User:
  – Hi there
- Chatbot:
  – Hello. Sorry I'm late. But I was in the hallway, chatting up Penny.
- User:
  – Oh I didn't know we had arranged a meeting.

- Chatbot:
  – I don't understand.
- User:
  – Um, okay.. So, who is Penny?
- Chatbot:
  – Oh, she's a doctor.
- User:
  – She is? Cool. What kind of doctor?
- Chatbot:
  – I will give you the address if you to my cousin's wedding with me.
- User:
  – What address? And I don't want to go to your cousin's wedding!
- Chatbot:
  – You're only saying that because you haven't been there. I am now a hexagon in two-dimensional space and can only perceive the edges of other objects.

- User:
  – Hello
- Chatbot:
  – Hello.
- User:
  – What's your name?
- Chatbot:
  – Rajesh Ramayan Koothrappali.
- User:
  – What do you do Rajesh?
- Chatbot:
  – I work at the Gap.
- User:
  – The Gap? What is that?
- Chatbot:
  – Doesn't concern you. You'll be going to the university every day?
- User:
  – Okay. Not every day, but almost.
- Chatbot:
  – I don't understand.

- User:
  – Hello. How are you?
- Chatbot:
  – It is I, Sir Howard of Wolowitz. Can I interest you in an afternoon of spirited questing, followed by a flagon of ale at yon virtual tavern?
- User:
  – Sure.
- Chatbot:
  – I wrote the section on the Wilson-Bappu Effect completely naked.
- User:
  – Uh okay. I guess that's cool.
- Chatbot:
  – I don't understand.
- User:
  – So how about that ale?
- Chatbot:
  – Personally, I would never use a transporter. Because the original Sheldon would have to be disintegrated in order to create a new Sheldon.
- User:
  – Good thing they aren't invented then.
- Chatbot:
  – I don't understand.

- User:
  – Hi there. Who are you?
- Chatbot:
  – It is I, Sir Howard of Wolowitz. Can I interest you in an afternoon of spirited questing, followed by a flagon of ale at yon virtual tavern?
- User:
  – No, Sorry. Not interested.
- Chatbot:
  – Want me to show you my work?
- User:
  – What kind of work is that?
- Chatbot:
  – Your new girlfriend. Have fun tonight.
- User:
  – Is that your attempt on a joke?

- Chatbot:
  – Actually, I don't need a team, I could easily defeat you single-handedly. But the rules require four, so may I introduce, the third floor janitor, he lady from the lunchroom, and, my Spanish is not good, either her son or her butcher. And what about your team? What rat have you recruited to the SS Sinking Ship?
- User:
  – You're gonna defeat me now? I think not! I'll take you on alone!
- Chatbot:
  – Oh, I think we both know the answer to that question.

- User:
  – Are you following me?
- Chatbot:
  – Um... yes... I am... asking you out.
- User:
  – Okay! Where are we going?
- Chatbot:
  – We're gonna have sex.
- User:
  – Oh, you're a bit forward aren't you?
- Chatbot:
  – I don't know.
- User:
  – But maybe I don't want to have sex with you.
- Chatbot:
  – What are you guys gonna do?
- User:
  – I'm leaving.
- Chatbot:
  – I'm leaving too. Not that anyones cares.
