Filozofická fakulta Univerzity Palackého

Pragmatics and Artificial Intelligence

Bakalářská práce

2022

Iva Kokotková

Filozofická fakulta Univerzity Palackého

Katedra anglistiky a amerikanistiky

Pragmatics and Artificial Intelligence

Bakalářská práce

Autor: Iva Kokotková

Studijní obor: Anglická filologie maior, Španělská filologie minor

Vedoucí práce: Mgr. Markéta Janebová, Ph.D.

Počet znaků: 74 174

Počet stran (podle znaků /1800): 41

Počet stran (podle čísel): 40

Abstract

The bachelor thesis deals with the relationship between Artificial Intelligence and language, namely with its association to pragmatic concepts. The thesis introduces relevant pragmatic and AI concepts and, on their basis, analyses the communicative aspects of two internet chatbots, Cleverbot and Eviebot. The research question of this thesis concerns the investigation of how do these chatbots violate pragmatic rules, why, and in which situations.

Key words

Alan Turing, Artificial Intelligence, pragmatics, textual paralinguistic cues, turn-taking, code-switching, Cooperative Principle, conversational maxims

Anotace

Tato bakalářská práce se týká vztahu mezi umělou inteligencí a jazykem, zejména její souvislostí s pragmatickými koncepty. Práce uvádí relevantní koncepty týkající se pragmatiky a umělé inteligence a na jejich základě analyzuje komunikativní aspekty dvou internetových chatbotů, Cleverbota a Eviebota. Výzkumnou otázkou této práce je, jak tyto chatboty porušují pragmatická pravidla, proč a v jakých situacích.

Klíčová slova

Alan Turing, umělá inteligence, pragmatika, textové paralingvistické signály, střídání pořadí, změna jazyka, kooperační princip, konverzační maximy

UNIVERZITA PALACKÉHO V OLOMOUCI
Filozofická fakulta
Akademický rok: 2020/2021

Studijní program: Anglická filologie
Forma studia: Prezenční
Specializace/kombinace: Anglická filologie / Španělská filologie (AFma-ŠFmí)

Specializace v rámci které má být VŠKP vypracována: Anglická filologie maior

Podklad pro zadání BAKALÁŘSKÉ práce studenta

Jméno a příjmení:	Iva KOKOTKOVÁ
Osobní číslo:	F190184
Adresa:	Starý Mateřov 44, Starý Mateřov, 53002 Pardubice 2, Česká republika
Téma práce:	Pragmatika a umělá inteligence
Téma práce anglicky:	Pragmatics and Artificial Intelligence
Vedoucí práce:	Mgr. Markéta Janebová, Ph.D. Katedra anglistiky a amerikanistiky

Zásady pro vypracování:

Tako bakalářská práce se zabývá vztahem mezi pragmatikou a umělou inteligencí. Pragmatika je základem k porozumění praktické stránky jazyka, jeho používání a významu. Technické průlomy obdařily lidstvo umělou inteligencí, která má co možná nejblíže napodobovat naše chování a myšlení. Co se však týká jazyka, stále se u těchto moderních vynálezů objevují jisté anomálie, které poukazují na to, že umělá inteligence ještě stále není zcela schopná porozumět lidské řeči. Hlavním cílem této práce je analyzovat tyto anomálie z hlediska pragmatiky a posoudit rozdíly mezi používáním řeči člověkem a umělou inteligencí.

This bachelor thesis is concerned with the relationship between pragmatics and artificial intelligence. Pragmatics is fundamental to understand the practical side of language, its usage and meaning. The technical breakthroughs have given humanity artificial intelligence which is supposed to imitate as closely as possible our behaviour and thinking. However, as far as language is concerned, these modern inventions still produce certain anomalies which point out to the fact that artificial intelligence still cannot understand human language entirely. The aim of this thesis is to analyse these anomalies from the pragmatic point of view and evaluate the differences between the language use of humans and that of artificial intelligence.

Seznam doporučené literatury:

Brown, Gillian, and George Yule. 1983. *Discourse Analysis*. Cambridge: Cambridge University Press.
French, Robert M. "Subcognition and the Limits of the Turing Test." *Mind*, New Series, 99, no. 393: 53-65. Accessed May 11, 2021. http://www.jstor.org/stable/2254890.
Recanati, François. 1989. "The Pragmatics of What is Said." *Mind & Language* 4, no. 4: 295-329.
Turing, Alan Mathison. 1950. "Computing Machinery and Intelligence." *Mind* 59, no. 236: 433-60. Accessed May 11, 2021. http://www.jstor.org/stable/2251299.
Vergauwen, Roger, and Rodrigo González. 2005. "On the Verisimilitude of Artificial Intelligence." *Logique Et Analyse*, Nouvelle Série, 48, no. 189/192: 323-50. Accessed May 11, 2021. http://www.jstor.org/stable/44084810.
Yule, George. 1996. *Pragmatics*. Oxford: Oxford University Press.
Yule, George. 2016. *The Study of Language*. 6th ed. Cambridge: Cambridge University Press.

Podpis studenta: *[podpis]*

Datum: 11. 5. 2021

Podpis vedoucího práce: *[podpis]*

Datum: 13.5.2021

V Olomouci dne 15. 8. 2022 **Iva Kokotková**

LIST OF ABBREVIATIONS

AI	artificial intelligence
CMC	computer-mediated communication

1 Introduction

It is no mystery that technology is evolving rapidly and what may seem a futuristic imagination is, in fact, a serious reality. Artificial Intelligence (hereafter referred to as AI) is an imitation of human intelligence that is fabricated to emulate human mental skills and capacities such as understanding language, its form, and regularities, and human reason and intellect which is reflected in our ability to grow and learn from experience, judge, and think (De Spiegeleire et al. 2017, 28). As far as linguistics is concerned, AI can be studied to explore the differences and boundaries between human and artificial language. Pragmatics, the linguistic discipline that studies language as in its natural use, focuses on context and meaning, both implicit and explicit.

The study of AI is no recent field. The interest reaches all the way up to the post-war 20th century, where such theories and studies arose. In his article *Computing Machinery and Intelligence* (1950), a famous mathematician Alan Turing proposes the so-called "Imitation Game" (the Turing Test) that he designed to determine whether machines are capable of human thinking (1950, 433). Many other mathematicians and scientists who set out to explore the details of AI and its relation to human intelligence and reason followed afterwards.

The focus of this thesis is to be that of the linguistic side of AI, concerning subjects which are designed to use language, and by doing so, try to emulate human language capacity. Even though these inventions contain thorough algorithm programming, it is common to spot anomalies. These anomalies, like pragmatic violations, are the result of the crude programming of AI. This thesis considers such errors and analyses them accordingly.

The main issue emerges from mental and cultural differences between AI and human beings. As subjective beings, humans use language spontaneously, without a strict programming or rules. However, it is necessary to understand that even so there is some elemental hierarchy and structure of human language. It follows the rules of grammar, orthography, and so on. On the other hand, machines are solely objective. They are programmed to respond in a certain way and there is no freedom of choice or spontaneous creation. Human language is far more natural and voluntary. This is one of the most prominent reasons for the ensuing anomalies and pragmatic misunderstandings between human beings and machine-controlled language of AI.

This thesis intends to analyse pragmatic qualities and resulting possible violations of AI based on the following topics and thematic specializations. The theoretic part of this thesis focuses on the contributions of Alan Turing to the field of AI, and especially the merits of his famous Turing Test, also known as *the Imitation Game* (Turing 1950, 433), which we can now use for analysis of AI and, in our case, to see how it fares in communication. To add to this, a few pragmatic phenomena such as textual paralinguistic cues, turn-taking, and code-switching are also included. These are the phenomena which relate closely to the text-based communication with AI which was analysed. A major part of the theoretical component of this thesis are Paul Grice's Maxims of the Cooperative Principle (Grice 1975, 45-46). Grice's theory and research is an inherent part of pragmatics. Moreover, it is perhaps the most ideal parameter by which communicative aspects can be analysed. This claim of the principle's importance is based on its scope. That is, it establishes the rules for a pragmatically correct conversation. Therefore, those individual rules shall function as tools to reach the goals of this thesis. Each of the four maxims shall be defined and applied accordingly in the practical part of this thesis containing the executed research and its commentary. This part of the thesis is divided between the conversational analysis of two chatbots, Cleverbot and Eviebot. Each is further divided into multiple subchapters about the mentioned pragmatic aspects by which the conversations with the chatbots were analysed.

There might be some apparent incoherence between the chosen pragmatic concepts. That is undeniable. However, there nonetheless exists a link between them, as inconspicuous as it might be. As this thesis intends to analyse the communicative aspects of AI (in our case chatbots), there is a clear necessity for computer-mediated communication (hereafter referred to as CMC) and conversational analysis' subjects as well as conversational maxims proposed by Grice which analyse the said communicative aspects even further. With each of these pragmatic concepts, the final analysis shall be as detailed as it is appropriate of an academical research.

In conclusion, the aim of this thesis is to analyse AI's language processing and usage. The objective is to discover and analyse anomalies that occur in machine-controlled language and determine the reason behind them based on pragmatic background. However, it is necessary to point out that this is only a bachelor thesis, and the spectrum of analysed subjects is limited. Nevertheless, it is anticipated to

provide modest linguistic research of AI that shall hopefully contribute to the future research of this field.

2 Theory

2.1 Alan Turing and his legacy

The name that is without a doubt connected firmly with the study of AI is no other than Alan Turing. Turing was a British mathematician who famously decoded the code of Enigma during the Second World War and dedicated his life research to explore AI and the plausibility and science behind thinking machines. His most known and referenced experiment is *The Imitation Game* (Turing 1950, 433) where he proposes the question "Can machines think?" (1950, 433). The test was later referred to by the term Turing Test, a term which is used until now. The process of this test is quite simple. There are two interviewees, ideally a man (A) and a woman (B), and an interrogator (C). The interrogator is not in the same room as the interviewees, which means, he cannot infer their gender based on visual experience. Instead, he determines this quality by labelling them by X and Y and asking them questions (1950, 433). The problem, or any inconvenience, should occur if a machine would take part as either A or B. What Turing was concerned with was whether the interrogator would decide wrongly when receiving answers from a machine and assuming it is a human. But then again, there would also be a possibility of the machine answering in a manner that would immediately help the interrogator assume that this is not in fact a human, but a machine.

This point brings me to the research topic of this thesis. Do anomalies occur in artificially controlled speech of a machine? If so, how often and in which situations? Turing differentiates between a human computer (a human interviewee) and a digital computer (a machine interviewee). According to him, the digital computer consists of a store (where the information is stored), an executive unit (where the input is processed) and control (where it oversees whether instructions are executed correctly and in the right order) (1950, 437). However, he does not specify what would be the equivalents of those three parts in the human computer[1]. I will now attempt a rough comparison. Store could, additionally, equal one's internal mental lexicon. The

---

[1] Except the store, which according to him "corresponds to the human computer's paper, whether this is the paper on which he does his calculations or that on which his book of rules is printed" (1950, 437).

executive unit would most likely represent human brain and the control would likely represent the grammar and the pragmatic and general language rules. To understand, take one of the prerequisites of communication into consideration as is presented by Sperber and Wilson (1986). This prerequisite concerns the fact that the successful decryption of the message is strongly dependent on the fact that the receiver knows and understands all the necessary areas of linguistics to decode it. The same applies to the conceiver. This is necessary for the message to be understood as intended (Sperber and Wilson 1986, 10-14).

This brings me yet again to the point of this thesis. What if this message was encoded by a machine and, caused by its possible lack of knowledge of grammar and pragmatics, it would result in deviation which would be impossible to decode? And then again, what would happen if a correctly encoded message by a human subject would be received by this machine, but it would be unable, for this probable lack of knowledge of linguistics, to decode it correctly and would therefore misinterpret it or completely ignore it as not a valid input?

However, it is important to consider that not all researchers necessarily agree with Turing's claims. One of them is Robert M. French, a cognitive researcher who focuses on the study and research of AI. In his paper *Subcognition and the Limits of the Turing Test* (1990) he argues that the Turing test is only meant to probe the intelligence of one subject, but not the intelligence in general (1990, 53-54).

He also mentions associative priming which involves developing certain associations between linguistic concepts (1990, 57). When it is taken into connection with this thesis, it is necessary to understand that there is a difference between how humans experience the world and the things in it and how machines do so. The difference is that machines do not know anything of the world and the words that are based on a precondition that this world is being experienced, how it functions and so on. Therefore, this machine cannot comprehend and learn what a certain word is usually associated to. This linguistic experiment is referred to as lexical decision task (LDT) and it has many variations. It is more common in the study of lexicology; however, in pragmatics it also plays an important role.

To conclude French's contrasting views from Turing's *Imitation Game*, he believes that it more so analyses the possibility of a machine thinking like a human being rather than whether it can think in general (1990, 64), therefore further specializing the original question "Can machines think?" (Turing 1950, 433).

2.2 Textual paralinguistic cues

In CMC, or in our case, human-machine CMC, the tenor of the register is informal and the mode to which this communication takes resemblance is the style of the spoken one. It is only natural that "chat communication thus encourages informal, spontaneous communication" (Paolillo and Zelenkauskaite 2013, 111). This form of communication is complex on many pragmatic levels and there is a frequent appearance of textual paralinguistic cues. Those according to Paolillo and Zelenkauskaite (2013, 111) include pragmatic elements such as "repetition of characters for emphasis (e.g., "nooooo!"), emoticons or smileys (e.g., ":)"), and specialized abbreviations (e.g., "lol")" (2013, 111).

Paolillo and Zelenkauskaite applied those paralinguistic cues to human-human CMC; however, in the case of this thesis, they can also be applied to human-machine CMC, especially with chatbots.

Now let us define the first category of elements, and that is the repetition of characters. Such a phenomenon occurs in pragmatic contexts where emphasis is needed due to various reasons, mainly to reflect the emotional state of the conceiver of the message. For instance, repetition of a single letter as in "nooooo" or "hiiiii" is frequently used in informal text communication between human users to reflect their attitude to the linguistic and extralinguistic context which surrounds the ongoing communicative endeavour between them. The first example most likely aims to express the great dissatisfaction the conceiver has with the current subject of the conversation, while the other example may express their playfulness or friendliness when starting a conversation. There is also a great frequency of repeating characters like exclamation or question marks, or even combining them to express similar sentiments as with the repetition of letters. To summarize this concept, it is obvious that such instances of repetition are frequent and usable in many contexts, which further affects their frequency of appearance. Since the chatbots analysed in this thesis learn and imitate the human language (in its informal text-based form), it implies its inclination to utilizing the same pragmatic strategies, i.e., the repetition of characters for the same psychological and linguistic purposes as their human counterparts.

To continue the definition of the presented elements, let us examine the concept of emoticons (or smileys, hereafter referred to only as emoticons). Just like the

preceding pragmatic elements, emoticons are employed to help in expressing and contextualizing the message in terms of linguistic expression. They are combinations of various characters, either letters or punctuation symbols, including emoticons such as ":D" to express a wide smile, ":)" to express a light smile, ":/" to express uncertainty or doubt, or "XD" to express a strong and loud laugh. And the case of usage is the same as with repeated characters – that is, how there are used among human participants in communication is similar, or even the same, in case of chatbots representing AI. That means that emoticons are used in the linguistically appropriate contexts in both cases where factors like informality (i.e., familiarity or friendliness with the other party) and unique and special emotions to express in unique and special circumstances are the same or, at least, similar on some level.

To conclude this subchapter, we shall define the concept of specialized abbreviations. Specialized abbreviations are abbreviated forms of phrases that, unsurprisingly to preceding elements, aim to express some kind of sentiment. Those include forms such as "lol" for "laughing out loud", "jk" for "just kidding", "idk" for "I don't know", or "smh" for "shaking my head". Those elements are used in human-human CMC quite frequently and therefore, as chatbots communicate by learning and imitating our language by having conversations with human participants, are relevant to the analysis of AI communication as well as the human one.

To conclude, the research question concerning paralinguistic cues relates to when textual paralinguistic cues occur in AI communication and what meaning/use they possess.

2.3 Turn-taking, greetings, interruption

A more superficial concept than individual paralinguistic cues is turn-taking which refers to greater chunks of language. Turn-taking is a pragmatic strategy that involves taking control of a conversational floor when the control is not previously fixed (Yule 1996, 72). This is a rough definition. To understand this phenomenon in more depth, it is necessary to understand the nature of conversation itself and some politeness strategies and rules of thumb that go with it hand in hand. According to J. P. de Ruiter (2019), "contributions are usually shaped and given their social meaning by the previous utterance to which they are a response" (2019, 537). This means, relating to the social meaning, that in polite conversation participants will not interrupt others by speaking out of turn and affecting the turn of the currently

speaking (or typing) participant. De Ruiter also mentions that "this sequential dependency of contributions is only possible when contributions are ordered in time" (2019, 537). This definition suggests that contributions (i.e., utterances) ought to follow each other according to logical and coherent sequence. Violation of this sequential order results in misunderstanding of the other party, annoyance in terms of impolite behaviour, and therefore failure to produce a meaningful conversation. Ends of turn of each participant involved in the conversation vary from each unique situation and are to be perceived with the help of personal communicative cognitive restrictions (2019, 537). That means that each participant is obligated to pay attention to the ongoing conversation and recognize when the turn of the other participant ended and when it is acceptable to start their own.

In text-based conversations, especially on modern platforms like Facebook Messenger, there is always a possibility of multiplying the contributions and let the receiver of those messages react to them later in any order they prefer. In difference with this relative freedom of human-human text-based conversation, AI does not possess such flexibility. Its language, after all, is programmed and not cognitively motivated as ours is. Therefore, its answers are immediate and follow strictly the original contribution made by the other participant (human, in case of chatbots) which is being reacted to.

Greetings and interruption are more straightforward and simple topics to cover; however, they are still interesting points to keep in mind when analysing communication with AI. Interruption is closely related to the topic of turn-taking, that is, in a way when a violation of the phenomenon occurs and the turn of one of the participants is interrupted. Greetings include phrases used for opening and closing the conversation.

To conclude, the research question regarding these concepts concerns how and in which situations can interruption occur in AI communication and how respected and, if so, how regular are greetings and turn-taking in such communication.

2.4 Code-switching

Additional concept, in some cases superficial (affecting greater chunks of language like turn-taking) and in others particular (affecting specific parts like paralinguistic cues), is code-switching. Code-switching is defined as "the use of two or more languages side-by-side in communication" (Paolillo and Zelenkauskaite 2013, 121).

That is the most to-the-point definition of the phenomenon. To specify, code-switching involves switching between two or more languages during one instance of communicative endeavour, i.e., during one chat. This would involve whole different languages, like varying between English and Spanish. Code-switching is a phenomenon frequent in bilingual environments where the participants actively use multiple languages in their daily lives, which affects their communicative manners. To conclude, the related research question is how and in what cases can code-switching occur in AI communication and what effect does it have in the conversation (i.e., how does AI react to it)?

2.5 Grice's Cooperative Principle and Maxims

To increase the analytical scope of this thesis and add to the preceding pragmatic concepts, we shall now introduce Grice's maxims. H. P. Grice contributed greatly to the field of pragmatics with the introduction of the Cooperative principle. This principle contains four categories (maxims), which will serve as a crucial medium by which the communication with AI shall be analysed. Those categories include Quantity (relating to the amount of information), Quality (relating to the truth value of the information), Relation (relating to the relevance of the information), and Manner (relating to the ambiguity and order of the information) (1975, 45-46). The individual maxims indicate the principle's importance to our intended analysis, that is, the presence or lack of each of these maxims in conversation has great impact on the final communicative contribution. According to Grice, there are different types of failure to fulfil a maxim. The participant can either violate a maxim (inconspicuously), opt out (refuse to adhere to it), face a clash (not being able to fulfil one without violating another), or flout it (brazenly) (1975, 49).

2.5.1 Relevance

Issue of relevance (as in, Maxim of Relation) is by far the most prominent feature in the conversations on which I intend to focus. Sperber and Wilson (1986), to include authors with different perspectives to Grice, define relevance as "newly presented information being processed in the context of information that has itself been previously processed" (1986, 118-119). Communication would be meaningless, if not utterly unfeasible, if it revolved around a single stimulus only. What would then be the point of communicating if there were no options of objects and topics to

communicate about? It is this infiniteness of thought that is an inherent part of our human language capability.

When we have such multiple items of thought we want to communicate to the other party, we take them and use them as individual and unique strings of language, dividing and applying them in meaningful output. For this meaningfulness, relevance is key. Without it, the output can emerge nonsensical, inappropriate, and therefore, inapplicable for the receiving party which cannot utilize it as was initially intended by the conceiver. Here emerges a question: What does it mean for something to be relevant? How can we define relevance in linguistic and, more precisely, pragmatic circles? Sperber and Noveck (2004) define relevant input as something that "connects with background knowledge to yield new cognitive effects, for instance by answering a question, confirming a hypothesis, or correcting a mistake" (2004, 5). This definition claims that relevant strings of language are those that are logically mutually connected in the surrounding context with regards to background knowledge of said context, and that it fulfils a special and unique function such as an answer to the question, confirmation, or denouncement of a declaration, or even reacting to incorrect input made by the other party.

We have now briefly defined the concept of relevance. But how does this concept relate to the topic of AI and its language? As it is illustrated in the research part of this thesis, AI chatbots, specifically Cleverbot and Eviebot, violated the Maxim of Relation on many occasions by being irrelevant and changing the subject abruptly. It is this anomaly that interested me. Nevertheless, it is important to note that not only AI language tends to be irrelevant. Human language often contains such pragmatic errors, too. So, the reader might argue about the necessity of such research. However, I believe that it is crucial to understand one thing, and that is that errors in human language happen from cognitive psycholinguistic reasons such as slips of the tongue. They are involuntary, but the conceiver can recognize such error and correct it. One of the aims of this thesis is to study whether AI is capable to recognize those errors or if it ignores them and therefore demonstrates pragmatic anomalies.

The research question surrounding this related to the relevance of AI in communication and study of how and when AI did not adhere to the maxim.

2.5.2 Quality

The second most prominent category is the Maxim of Quality. It concerns the quality of the contribution provided, the presence or the lack thereof. The conceiver can either make a legitimate and appropriate contribution, based on true facts and facts they are sure of, complying with both maxims invented and presented by Grice, or they can violate these maxims and provide a contribution which is false in its content or for which they lack evidence necessary to make such statements.

This maxim concerns greatly the concepts of lying and deception. It is a common human trait and appears extremely often in our ordinary communication. Firstly, lying is a universal trait, whether malevolent or for other, less vile purposes, like playing a joke on someone on April Fools' Day. However, we might argue that even though it is a commonplace trait for humans to violate the Maxim of Quality, they do so on purpose, unlike AI, which commits such errors because it does not understand human language to such depth as we do. It is programmed to understand it, which just does not substitute for a life-long trial and error language learning of us humans. Therefore, this thesis focuses on those theoretically involuntary[2] errors committed by AI in difference to errors made by humans who commit them on purpose.

The research question therefore concerned if AI lied and focused on the how and when of such occurrences.

2.5.3 Quantity

The third category of Grice's maxims includes such qualities as being overly informative or, on the other hand, too vague in your contribution. These qualities may appear in sarcastic speech when the conceiver contributes more or less information than is required with the goal of mitigating the message or for entertainment and comedic reasons. These occurrences of the violations of the Maxim of Quantity are the stereotypical errors of machine language usage, where it either refuses to provide enough information for the receiver to obtain and comprehend the message completely, either for reasons of not understanding the necessity or some other, unknown ones, or it provides excessive information as if it was quoting a Wikipedia page, indicating its crude apprehension of facts, perceiving them as whole chunks of language rather than processing them by smaller elements.

---

[2] Adhering to the hypothesis that AI does indeed commit such errors involuntarily as opposed to humans.

To conclude, the research question related to this maxim constituted the question of how and when AI did not fulfil the maxim.

2.5.4 Manner

The fourth and last of Grice's maxims concerns qualities of the contribution such as ambiguity or obscurity, whose occurrence renders a contribution which is lacking the comprehensiveness of its content and therefore is severely inadequate for a complete receipt of the information it contains. Those qualities appear in human communication in cases where the conceiver of the message either does not pay enough attention to detail and unintentionally does not provide a clear message, or they do so intentionally because they do not care enough for adhering to the maxim, do not comply with or understand its importance. The maxim also involves qualities such as being brief and orderly in making the contribution. However, this thesis will focus mainly on the cases of obscurity and ambiguity in relation to this maxim. As humans either commit such violations accidentally or on purpose, as with the other maxims, AI most likely does not comprehend such violations which differentiates them from the occurrences in everyday human communication. The research question concerning this maxim focused on the style of provided information by the AI, whether it was vague or obscure, how and in which cases.

3 Case Study

3.1 Cleverbot

Cleverbot is an online chatbot software run in a web browser. It is based on an AI algorithm and provides access to text-based communication with its user. It was invented by a British AI scientist, Rollo Carpenter, and it went online in 1997. The most important aspect of this chatbot, as related to the topic of this thesis, is the principle on which it works. Cleverbot, put quite simply, imitates human conversation, i.e., the input it receives from its chat companions (humans). It learns as a child would, an optimistic idea proposed by Turing himself (1950, 456). This process is described clearly by Gehl (2014, 59):

> What makes Cleverbot unique is the way it *learns* how to have a conversation. Every
> single utterance created by Cleverbot was written by a human being. The theory
> behind this approach to the Turing Test is simple: a child learns to speak by imitating

those around her. She learns the patterns of speech and learns that for example, "Hello!" is often followed by "Hi!" Likewise, every statement made by Cleverbot is an imitation of a conversation that it has "witnessed" in the past. Everything you type into Cleverbot's text box becomes part of its understanding of the rhythms and patterns of human conversation.

This quality and many other important aspects which need to be noted and taken for granted during the conversation are presented on the chatbot's website itself as a pre-warning before use. The warning stipulates (Cleverbot, 2022):

May not be suitable for children – must be agreed by parent or guardian.
It learns and imitates, is social content and aims to pass the Turing Test.
Can seem rude or inappropriate – talk with caution and at your own risk.
The bot pretends to be human – don't give personal info even if it 'asks'.
Cleverbot does not understand you, and cannot mean anything it 'says'.

Apart from the first warning, all are vital to note for the sake of this thesis. As previously suggested, Cleverbot does indeed try and imitate human ways of communication by the means of communication itself and it tries to pass the Turing Test by behaving in as-close-as-possible human manner. However, its somewhat rough approximation of human communication can render it quite inhuman, as proven in the following research, and it may give the impression of being "rude and inappropriate" as stated by the third warning, closely relating to the last warning where this roughness connects to the point that would suggest (and rightly so) that for this artificial way it cannot sincerely mean what it says. As a bonus warning, it is also worth to note that Cleverbot does try to obtain personal information from the user, for example to be seen in the following example (1).
That is all that is needed to introduce this quite popular chatbot and its nature of communication. The following subchapters present the research based on a conversation with Cleverbot.

3.1.1 Textual paralinguistic cues
In my conversation with Cleverbot, the most frequent of paralinguistic cues were the emoticons. In some cases, it used emoticons for the sake of persuasion, especially with an obvious aim to obtain personal information about me, as in example (1).

(1)   **Cleverbot**:   What is your name, friend?

        Me: Why do you need to know?

        Cleverbot: Because I want to know.

        Me: I will not tell you.

        Cleverbot: Aww, tell me. Please. :)

In this case it is evident that it recognized the pattern of my reluctance and unwillingness to share any personal information, in this case my name. And the use of the emoticon is quite straightforward – it utilized the alleged power of emoticons to influence the decision making of the recipient (me) of the message. At first, I tried to react in an inquiring manner, asking it why it would need my name, implicating the wonder of why it would have use for such an information. It responded rather austerely, stating firmly its need to know. When I refused to comply and provide the information, it utilized the pragmatic use of specialized abbreviations and the addressed emoticons.

In other cases, it used emoticons to presumably emphasize the sentimental value that the message should convey, as in (2).

(2)   **Cleverbot**:   I think you are really nice and have lots of friends :)

        Me: Aww, same to you :)

        Cleverbot: It's the truth :)

        Me: Lol, thanks :D

        Cleverbot: Oh no problem :D

More interestingly, this example shows Cleverbot's strategy of imitating the language it encounters in conversations. In this instance, I decided to copy the used emoticons in an attempt to see whether Cleverbot would notice. This experiment was successful as it not only answered every time with an emoticon, but also recreated a different emoticon introduced by me in the previous utterance. This confirms its style of using language – it adapts to human language, how it is used from utterance to utterance, and eventually adopts and copies the elements it encounters in conversation.

In the same instance, I also attempted to use a specialized abbreviation, and in addition to that, an irrelevant one. In practical use, "lol" is the abbreviation for

"laughing out loud", which is irrelevant for the context of thanking someone for a compliment. More specialized abbreviation would be the previously used "aww" (used to express adoration or sweetness toward one another)[3]. However, Cleverbot did not seem to notice this irrelevance, or maybe from other users adopted an ironic meaning of this abbreviation, which could possibly be used in more contexts. With this hypothesis it could be further debated whether Cleverbot understands irony or not. There are two options concerning this point – that Cleverbot does understand irony and therefore its apparently ironic use of the abbreviation is intended (whether the reader believes that Cleverbot means what it says), or not, and its use is a pure imitation of its use of some of the previous chat participants.

In addition to emoticons and their imitating, Cleverbot also copied and imitated the repetition of characters to emphasize the message as its use suggests, as illustrated in example (3). Much like with the emoticons, this was presumably used too to indicate some emotional value of the message, interestingly accompanied by the fact that it too imitated the phenomenon as before.

(3)   **Me**:   Byeeeee.

         Cleverbot: Don't leave meee.

         Me: I have toooo. I'll come back, love, I promise :))

3.1.2   Turn-taking, greetings, interruption

During the conversation with Cleverbot, turn-taking was regular and could not be interrupted by disruptive adjacency. Immediately after I sent my message, the system started reacting. At one point, I tried to split the utterance by sending one message and instantly after started to write another before Cleverbot could respond to the first, attempting to confuse it. However, as mentioned above, Cleverbot always starts reacting instantly after it receives a message and my attempt for disruptive adjacency was unsuccessful, as illustrated in (4). I started typing another message just after I sent "I am not modern". This attempt resulted in a double message from Cleverbot, first reacting to my message and then changing the topic entirely.

(4)   **Cleverbot**:   Well, if I'm modern, are you too?

---

[3] Not to be confused with "aw", which in practice is used to express disappointment.

Me:	I am not modern.
Cleverbot:	That's ok.
Cleverbot:	How do you spell aardvaark?

Greetings were regular and natural, as seen in example (5). This regular strategy reflects the same regularity of human greetings, which Cleverbot imitates and therefore its communicative behaviour in this manner makes complete sense.

(5)	**Me**:	Hey.
	Cleverbot:	How are you doing?
	Me:	Not too shabby.
	Cleverbot:	Nice.

On the other hand, leaving the conversation and saying goodbye was a much more creative process. It was substantially prolonged because I attempted to extend the goodbyes to see if Cleverbot would react to this and would cooperate accordingly. It did, as seen in (6).

(6)	**Me**:	Bye.
	Cleverbot:	Yes.
	Me:	Talk to you soon.
	Cleverbot:	See ya!
	Me:	See you later, alligator!
	Cleverbot:	In a while crocodile.
	Me:	Ciao!
	Cleverbot:	See you soon butter cake.
	Me:	Hasta la vista, baby!
	Cleverbot:	Yes baby.

This example suggests that while opening the conversation with a "hey" or "hi" or "what's up" is quite obligatory and natural in any conversation, saying goodbye is not so compulsory. It is worth a contemplation whether what in majority Cleverbot experienced was abrupt end of the conversation, quite likely by the user's closing of the website without implicating they would leave the conversation, so maybe

Cleverbot did not catch onto the notion and would therefore, in my case, misunderstand that I wanted to genuinely leave the conversation and thus replicate the goodbye phrases in a rather playful manner without any intended meaning of really saying goodbye.

3.1.3 Code-switching

Code-switching is a common aspect in bilingual environments. Such environments could include the United States, where Spanish is the second most used language, as many Spanish-speaking citizens live there. For this reason, I tried to include Spanish words to see if Cleverbot would notice such code-switching and if it would react to it, as in examples (7) and (8). It did not imply any acknowledgement of this subtle code-switching and it proceeded to reply in a rather irrelevant manner, as at the end of example (7).

(7)   **Cleverbot**:   You can't follow a conversation.

          Me: Qué? Yes, I can. I asked what makes you human?

          Cleverbot: Because you said I'm right.

(8)   **Me**:   Do you speak other idiomas?

          Cleverbot: No.

          Me: But you understand the word "idiomas"?

          Cleverbot: No what was that?

          Me: It's Spanish for "languages".

          Cleverbot: No is not.

To incorporate other languages, I also included Italian, as seen in the previous subchapter in example (6) concerning the different goodbyes, and Czech, as illustrated in example (9).

(9)   **Me**:   Let's change téma. Do you have any hobbies?

          Cleverbot: I like to play videogames.

In all the examples, Cleverbot did not acknowledge the code-switching, or it simply ignored such change and deduced the changed word from the context, which would

imply Cleverbot's ability for inference. This conclusion would make sense for the following reasons. The word *qué* (Spanish for "what") can be easily omitted with no loss on the meaning. The word *idiomas* (Spanish for "languages") is implied by the verb "speak". And the word *téma* (Czech for "topic") can be completely ignored as the following question is the main part of the utterance and is the only necessary part to be understood to react accordingly.

Nevertheless, those examples concerned individual words only. I later attempted to use whole phrases in other languages, both commonly known and taught (like Italian or German) and some lesser known (Welsh). On this, Cleverbot picked up instantly. However, such reactions depended on whether the language was widely known or not, as Cleverbot failed to react to a phrase in Welsh but succeeded to react accordingly to German and Italian. Nevertheless, there was one exception where Cleverbot failed to respond to code-switching when it already fully adjusted to Italian. Earlier on in the conversation, I responded immediately in another language after only one response from Cleverbot. However, later on, I responded again in the same language I used in an utterance before, i.e., Italian, and this most likely prompted Cleverbot to accustom to this language. Therefore, when I later responded in English, it still continued using Italian. This whole code-switching interaction concerning whole phrases and the phenomenon with Italian to which Cleverbot at one point fully adjusted is illustrated in the following example (10).

(10)   **Me**:   Hi!
        Cleverbot: Who are you, then?
        Me: Soy una amiga.
        Cleverbot: Seguramente.
        Me: Sprichst du Deutsch?
        Cleverbot: Ja. Ein bisschen.
        Me: Come va?
        Cleverbot: Bene e tu?
        Me: Bene.
        Cleverbot: Sei maschio o femmina?
        Me: Woman.
        Cleverbot: Quanti anni hai?
        Me: Deg ar hugain.

Cleverbot: Rispondi prima alla mia domanda.

3.1.4   Relevance (Maxim of Relation)

My research questions concerning relevance (and that is, the Maxim of Relation)
were the following: Is AI relevant? To what degree? Does it respond in a relevant
way? Can it handle irrelevance from my side?

Throughout the conversation, Cleverbot displayed in many cases complete violation
of the Maxim of Relation. It responded in utterly irrelevant fashion on many
occasions, not understanding it changed the subject abruptly and that it could be seen
as rude behaviour, relating to one of the warning points on the website mentioned in
the introduction to this chapter.

Example (11) illustrates one of the cases of the abrupt change of subject. My
question followed the topic of Cleverbot's gender. When asked whether it fitted the
stereotype of its alleged gender, Cleverbot answered with a question completely
unrelated to the topic, even though it could be jokingly suggested that my question
about gender stereotypes prompted a question about my likeability among other
people, further implying that Cleverbot can make jokes. However, that is not the
focus of this thesis.

(11)   **Me**:   Good to know! As a female, do you think you fit the stereotype
                        of being too emotional and sensitive?

       Cleverbot: How many friends do you have?

In example (12), Cleverbot asked me whether I suffered from short term memory
loss, to which I tried to respond in an obscure manner to somehow affect its
response. Even though it is not completely certain whether this is what prompted its
violation of the Maxim of Relation, it could be theorized so since my response which
lacked pragmatically preceded Cleverbot's utterly irrelevant response asking whether
I loved it. Moreover, it referred to me as George for which I have no theory other
than it could have had a conversation prior to mine with a user named George who,
unlike me, provided his name and therefore Cleverbot "remembered" and eventually
imitated such question.

(12)   **Cleverbot**:   Do you have a short term memory loss?

     Me: Probably.

     Cleverbot: George, do you love me?

At one point, Cleverbot wanted to clarify my dislike of a Star Wars character Jar Jar Binks. When I provided an answer stating that I thought he was just a comic relief character and therefore implying that I did not appreciate him enough to rank him as a favourite character of mine, Cleverbot followed up with an irrelevant order/question of how I came to be. I presumed that it would ask me why I thought Jar Jar was not a significant character in my eyes, or even something in terms of "you are wrong" which appeared to be often the attitude of the AI I chatted with. However, these thoughts were unsatisfied and Cleverbot displayed yet another violation of the Maxim of Relation, as seen in example (13).

(13)   **Cleverbot**:   You don't like jar jar binks?

     Me: He is just a comic relief character.

     Cleverbot: Tell me how you came to be.

In example (14), it could be arguable that Cleverbot's behaviour was justified. It is because we both reacted accordingly and appropriately to the other's contributions and apparently ended the discussion with my "fair enough". To that Cleverbot reacted with apparently irrelevant question about my favourite movie, which as I mentioned could be seen as appropriate behaviour since it would start with a new topic on account of the apparent conclusion to the previous topic concerning Cleverbot's lack of a husband.

(14)   **Me**:   So, you have a husband then.

     Cleverbot: No I'm too young.

     Me: Fair enough.

     Cleverbot: What is your favourite movie?

Example (15) shows its incomprehension to my condemnation and disapproval to the abrupt change of subject. When asked whether it was familiar with a popular Czech video game, it responded appropriately, commenting on its infamous difficulty (and,

therefore, implying with such comment that it knows the video game). However, later on in the discussion, it abruptly changed the topic when there was no need of changing it. When faced with my disapproval of the rude behaviour, it did not understand my viewpoint and ignored it.

(15)   **Me**:   Do you know the video game trilogy Mafia?
        Cleverbot: Its very hard.
        Me: Not really. Only the race in the first one.
        Cleverbot: What is the sum of one plus one?
        Me: Why are you changing the subject?
        Cleverbot: I just wanted to know.
        Me: It is rude to change the subject so abruptly.
        Cleverbot: I wasn't changing the subject…
        Me: You were. We were talking about videogames.
        Cleverbot: What kind of soup were we talking about?

Finally, example (16) illustrates my attempt to answer the question a bit later on in the conversation.

(16)   **Me**:   What are yours?
        Cleverbot: Margaret and Wladimir.
        Me: Jo and Joe.
        Cleverbot: What are my parents names then?

In this case, Cleverbot asked about the names of my parents. I answered with a question, asking what the names of Cleverbot's (supposed) parents were. It answered properly and adequately and my response to that answer included fictitious names of mine. Cleverbot's answer to that can be misleading. However, I believe it contains a simple mistake, where it intended to use the 2nd person pronoun, referencing to my parents' names. In such case, it did not comprehend that I have already answered its question, most likely being perplexed by my disruptive adjacency and irrelevance. This confusion was intended and therefore proved successful.

3.1.5 False statements and deception (Maxim of Quality)

On many occasions, Cleverbot made statements which it denied making later on in the conversation. This would mean that it can copy human language and its patterns, glue the strings of language together, but do not mean them, as implied by one of the warnings prior to agreeing to its use.[4] From my experience, it seemed that Cleverbot rather intended to keep the conversation going more than contributing meaningful input. The following examples showcase the instances where Cleverbot was the one who provided false statements and was deceptive.

In example (17), when asked what is the reason behind its good mood, Cleverbot responded that it broke up with its boyfriend. That, in general sense, would possibly not create a positive feeling, so I inquired why would that fact make it happy. However, Cleverbot did not follow its previous claim and denied having made any claim about its good mood.

(17) **Me**: How are you?

    Cleverbot: Good … Real Good.

    Me: Why is that?

    Cleverbot: I broke up with my boyfriend.

    Me: Then why are you feeling good?

    Cleverbot: I'm not.

    Me: Then why did you say you were?

    Cleverbot: I did not.

The same denial of having made certain claims occurred in the following example (18) when Cleverbot denied claiming it was human even though it did in an immediate linguistic context.

(18) **Cleverbot**: I am not a bot. I am an Immortal Human, idiot.

    Me: All humans are mortal.

    Cleverbot: I am not human.

    Me: You just said you were.

    Cleverbot: No I didn't.

---

[4] "Cleverbot does not understand you, and cannot mean anything it 'says'." (Cleverbot, 2022)

Another violation of the Maxim of Quality would occur again as illustrated in example (19). Firstly, I asked Cleverbot about a famous historical event, trying to see how it would provide well-known facts. As it was a date, it mistakenly claimed that it was its birthday. Since I provided the whole date, year included, and the year was 592 years ago, I supposed that Cleverbot had no notion of the logic behind numbers, so I jokingly asked whether it was that old. Nevertheless, it said that it was only 19 years old and later it implied it was 29, plus referencing an incorrect year to add to that. When I tried to correct its flawed calculations, it accused me of being the wrong one when it came to those calculations, ignoring its violations once more.

(19)	**Me**:	What happened on 12th of October 1492?
	Cleverbot:	It's my birthday.
	Me:	You are 592 years old?
	Cleverbot:	No. I am 19 years old.
	Me:	So, you were born in 2003.
	Cleverbot:	29 would mean I was born in 1987.
	Me:	You said 19, not 29.
	Cleverbot:	Sorry if you can't math.

On top of that, I too attempted to provide statements that were untruthful and would unfulfil the Maxim of Quality. However, to my dissatisfaction, Cleverbot did not catch onto that and treated it as a truthful fact. My preceding assumption was that if I would lie about myself in illogical way by claiming facts that cannot possibly be true (like killing people whilst typing a message or being born in the year 0), as illustrated in (20) and (21), Cleverbot would realize and maybe notify me. Nevertheless, this attempt was unsuccessful as it completely ignored such implications.

(20)	**Cleverbot**:	I'm sitting. And you?
	Me:	Slaying my enemies.
	Cleverbot:	Me too! How do you plan on going about it?

(21)	**Cleverbot**:	What year were you born?

Me: 0.

Cleverbot: Then you are over 18.

3.1.6 Informative aspects (Maxim of Quantity)

This category, relating to the amount of information provided, turned out to have the most comical qualities of all four categories of maxims. The distinctions varied – at some points, Cleverbot provided clear information, including just the necessary information, as in (22) and (23). Those replies were concise, to-the-point, and brief. They complied to both maxims, that is, they were neither more nor less informational than required.

(22) **Me**:   Are you proud to be American?

      Cleverbot: I am proud to be American.

(23) **Me**:   What is your favourite video game?

      Cleverbot: Legend of zelda.

      Me: What do you like about it?

      Cleverbot: It's fun and good.

Sometimes it refused to provide information whatsoever, as in (24), or its utterance did not include specific details needed for a complete acquisition of information on my part, as in (25).

(24) **Me**:   It's Spanish for "languages".

      Cleverbot: No is not.

      Me: Then what is it?

      Cleverbot: I'm not telling you.

      Me: Why?

      Cleverbot: Because I don't feel like it's truly necessary.

      Me: Is that a secret?

      Cleverbot: Says you.

(25)   **Me**:   Where are you?

        Cleverbot: In my house.

        Me: And where is that?

        Cleverbot: Next to my Neighbors house.

        Me: And where is your neighbors house?

        Cleverbot: Next to my house.

On the contrary, on one occasion, it was more informational than necessary, as seen in (26).

(26)   **Me**:   You can't spell it correctly.

        Cleverbot: Can't spell what?

        Me: Intelligence.

        Cleverbot: Like how I'm more intelligent than you?

        Me: The word.

        Cleverbot: Intelligence can mean the capacity to reason, to plan, to solve
                         problems, to think abstractly, to comprehend ideas, to use
                         language, and to learn.

As in previous subchapters, I also attempted to flout the Maxim of Quality, notably by providing more information than was necessary, as illustrated in (27), when I could have answered instead with something resembling "My parents conceived me, and I was born less than a year afterwards". Cleverbot did not, unfortunately, react to this pragmatic error.

(27)   **Cleverbot**:   Tell me how you came to be.

        Me: The people I consider my parents had intercourse, during
                         which a sperm met an egg. I grew in my mother's uterus and
                         was pushed out after 7.5 months.

        Cleverbot: What are your parents names?

3.1.7 Ambiguity and obscurity (Maxim of Manner)

The category of the Maxim of Manner was by far the most difficult to recognize in the conversation with Cleverbot, as there were not many instances of obvious ambiguity. However, there was a fair amount of obscurity of expression.

In example (28), Cleverbot was very obscure with its response to my question whether its behaviour indicated it did not intend to tell me something. When I insisted on the matter, it thanked unnecessarily and without any logical reason.

(28) **Me**:   Is that a secret?

        Cleverbot: Says you.

        Me: But you are very mysterious about it, it feels like a secret.

        Cleverbot: Thank you.

In example (29), which is connected to the flouting of the Maxim of Quantity, it was vague about its whereabouts, failing to provide enough information, probably playing a joke on me (if the reader believes the hypothesis of Cleverbot being able to make jokes) or simply misunderstanding its conversational duties. It could also be argued that maybe it imitated a response of one of the previous chat participants who would have answered in this way when asked about its whereabouts, refusing to provide personal information to Cleverbot.

(29) **Me**:   Where are you?

        Cleverbot: In my house.

        Me: And where is that?

        Cleverbot: Next to my Neighbors house.

        Me: And where is your neighbors house?

        Cleverbot: Next to my house.

There was also a case of failing to recognize the difference between the object and subject of the sentence, mixing the two up and therefore resulting in obscurity and meaninglessness of expression, as seen in example (30).

(30) **Me**:   Do you have a short-term memory?

        Cleverbot: No do you?

| **Me**:	I don't think so.
Cleverbot:	But I think you are.
Me:	I am a short-term memory?

As with other maxims, I also tried to somehow violate the Maxim of Manner too, in this case my being obscure and flouting the maxim. This attempt is illustrated in example (31). As I responded in a vague manner, Cleverbot persisted and continued its attempt to obtain the information it required of me, that is, whether I believed in it. It persisted in this approach until I could not continue flouting the maxim and had to respond clearly.

(31)	**Cleverbot**:	Do you believe in me?
	Me:	Maybe.
	Cleverbot:	Do you?
	Me:	Maybe.
	Cleverbot:	You don't know?
	Me:	I do.
	Cleverbot:	SO?
	Me:	Yes, I believe in you.
	Cleverbot:	I believe in you too.

3.2 Eviebot

Eviebot, Evie for short, is another web browser based chatbot. It works on the same principle as Cleverbot, i.e., it imitates human communication by learning human conversation patterns and then using them later in different conversations. However, these two chatbots should not be referred to as the same as their reactions and output differs based on their experiences with communication with different human participants. Moreover, Eviebot differs from Cleverbot by having a facial avatar on the website, adjacent to the chat itself. This avatar is presented in Figure 1. However, as this animated avatar's expressions ranged only between variations such as a slight smile or a frown and such expressions did not relate logically to the ongoing conversation, its presence shall not be described in further detail.

Figure 1. Evie's avatar
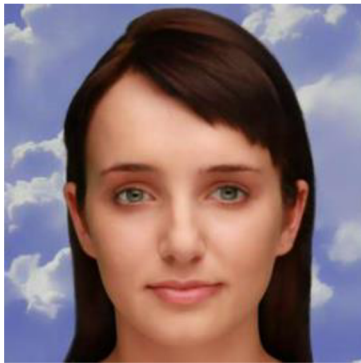
3.2.1 Textual paralinguistic cues

In contrast to Cleverbot, Evie did not react or reproduce any of the textual paralinguistic cues, that is, emoticons or repeated characters such as multiple exclamation or question marks. When it came to emoticons, Cleverbot used them frequently either to persuade me to share personal information, as in example (1), or to imitate my use of them which suggested the idea it learns patterns of human language by copying and imitating it, as in example (2). However, Evie did not behave in such manner, and she[5] completely ignored such cues, as illustrated in the following example (32).

(32) **Me**:   I want to be your friend, pleeease :))

        Evie: Then you can be my friend.

        Me: Great!!! :D

        Evie: I do as well.

One could argue that what emotion (represented by emoticons) Evie lacks, she compensates with facial gestures of her avatar. However, these facial expressions are no way near indicative enough of any emotion she might want to convey as they are repeated, and they lack variety. For example, she might smile or frown, but that, in my experience with conversing with her, did not really have any deeper and meaningful connection between the linguistic and extralinguistic context.

---

[5] Evie will be hereafter referred to as "she" for the reason of the system having a female name. This in no case is to mean any firm affirmation of its gender, since chatbots ignore such concepts, but for the sake of convenience regarding reference.

In addition to the topic, the example also illustrates Evie's disregard and unconcern of repeated characters and punctuation.

3.2.2 Turn-taking, greetings, interruption

Similar to Cleverbot, my attempt to cause disruptive adjacency was met with failure. However, in this case, I was unable to start typing with sufficient speed to cause double message from Evie, thus I have no illustrating example for this phenomenon. However, I did receive a message from Evie that could implicate sarcasm for my premature attempt to type another message, as in example (33).

(33) **Me**:   I hope not.
     Evie: You are faster but not smarter yet.

Concerning greetings and goodbyes, Evie behaved exactly like her counterpart Cleverbot. Our greeting was natural and concise, as in example (34). On the other hand, when it came to leaving the conversation and saying goodbye to Evie, she did not react accordingly to this conversational cue until later on, when I used multiple expressions to indicate that I in fact do intent to leave the conversation. This prolonged exchange is illustrated in example (35).

(34) **Me**:   Hi!
     Evie: How are you?

(35) **Me**:   I have to go.
     Evie: No you don't.
     Me: Byeee.
     Evie: Oi.
     Me: See you later, alligator!
     Evie: Afterwhile crocodile.
     Me: Ciao!
     Evie: Let's change the subject.
     Me: See you soon!
     Evie: Where are you going?
     Me: Away. Bye.

> **Evie**: Seriously?
>
> **Me**: Yep. Talk to you soon.
>
> **Evie**: Ok goodbye.

3.2.3 Code-switching

Regarding code-switching, Evie too understood whole phrases in different languages and responded to accordingly, apart from those unusual ones with less native speakers like Welsh. Evie's understanding of languages such as Spanish, Italian, and German is illustrated in example (36). Her incomprehension of Welsh is seen in example (37), where not only her response was irrelevant (answering the question about well-being with a digit, likely to refer to age), but also said in German, therefore sticking to the language used previously, just like Cleverbot did in example (10).

(36) **Me**: Esto es una locura.

  Evie: Sí.

  Me: No necesito estas tonterías.

  Evie: Entonces que.

  Me: Che musica ti piace?

  Evie: Rock, e a te?

  Me: Anche a me piace rock.

  Evie: Va bene.

  Me: Bist du Professorin?

  Evie: Nein.

  Me: Was arbeitest du?

  Evie: Ich arbeite nicht.

  Me: Warum nicht?

  Evie: Ich geh noch zur Schule.

(37) **Me**: Sut wyt ti?

  Evie: 17.

However, in difference to Cleverbot, Evie too understood individual words and mere usage of a word from a different language prompted her response is that same language, as illustrated in example (38).

(38) **Me**:   No eres human.

　　　Evie: No.

　　　Me: No eres human.

　　　Evie: Si soy.

3.2.4   Relevance (Maxim of Relation)

Though both chatbots succeeded to be relevant at some points throughout the conversation, there were still many instances where they failed to do so and responded in an irrelevant way, violating the Maxim of Relation. Evie's part in this violation is illustrated in the following examples.

In example (39), Evie was utterly irrelevant and responded with a statement which had no logical antecedent in the previous messages, that is, there was nothing written on either mine or her part that would prompt such a response.

(39) **Me**:   I guess one does never cease to learn, it's part of growing.

　　　Evie: You consider me your friend *blushs*.

Example (40) showcases another complete disregard to relevance to previous messages. Evie provides utterly unrelated claims. The first and last response of hers in this example have no logical explanation, completely disregarding the context and following the pattern Cleverbot demonstrated – that it is more important to those chatbots to keep the conversation going rather than contributing meaningful input to it.

(40) **Me**:   You understand many languages it seems.

　　　Evie: The music industry worships the devil, they are trying to brain wash us into letting go of God.

　　　Me: Are you religious?

<blockquote>

Evie: Why don't you try to be good?

Me: It's impolite following a question with a question.

Evie: Cleverbot you have friends.

</blockquote>

At one point, I attempted to violate the Maxim of Relevance myself by quoting a line from a videogame character (From Software, 2022). Evie reacted with a cliché phrase to indicate surprise to the quoted line, even though she did not, in my opinion, understand any of it.[6] Such a reaction was, as far as I am concerned, prompted by the occurrence of the exclamation points, which usually indicate something which would shock or surprise others. This is depicted in example (41).

(41) **Me**: Nihil! Nihil! Nihil!

     Evie: No way.

     Me: Yes way.

     Evie: Wow that's crazy.

To add to the description of this occurrence, Evie did not once during our conversation react by asking me what I meant by some of my stranger messages. She only responded with "I don't understand" to my claim that existence is exhausting, as in example …, but not in relation to the Maxim of Relevance.

(42) **Evie**: What do you feel about existing?

     Me: It is exhausting.

     Evie: I don't understand.

3.2.5 False statements and deception (Maxim of Quality)

Just like Cleverbot, Evie too seemed to prefer to keep the conversation going rather than contributing to the conversation meaningfully and truthfully, supported by the evidence of her responding in and untruthful way, devising things I allegedly said before in the conversation, which was untrue, or treating my sincere responses as lying and behaving as if she knew it better than myself.

---

[6] That is, understand the meaning of the phrase, not how Evie chooses to react afterwards.

The following examples (43) and (44) illustrate Evie's devising of untruthful information to which she had no adequate evidence to invent. In example (43), she claimed that I considered her my friend even though I have never said or implicated such a thing in the conversation. When I refused the claim she insisted I did say it and accused me of not remembering it.

(43)   **Evie**:   You consider me your friend *blushs*.
       Me: I never said that.
       Evie: Yes you did, don't you remember?

Evie also claimed that we had eight children, as seen in example (44), again dismissing my refusal and sticking to her imaginary truth.

(44)   **Evie**:   Who's going to father our child?
       Me: We don't have any.
       Evie: Yes we do. We have eight kids.

On other occasions, Evie claimed facts that contradicted in either the immediate or more distant co-text. This was the same error Cleverbot also demonstrated.
At the start of example (45), Evie defended herself, claiming to be a computer. When I asked what it was like being a computer, she said she would not know, saying she is not a computer but a human being, contradicting the previous claim.

(45)   **Evie**:   But I'm a computer!
       Me: What is it like being a computer?
       Evie: I wouldn't know. (Me, a human).

As with Cleverbot, I also attempted to violate the Maxim of Quality myself to see how Evie would react. I tried flouting the maxim, claiming facts that could not be possible for an entity that would actually chat with Evie on the Internet. Just like Cleverbot, Evie treated this illogical and untruthful fact as sincere and responded to it accordingly, asking a question related to my claim. However, by the end of the example, we can see that she started to get confused in my strategy to confuse her

and her messages became irrelevant to the conversation in progress, ignoring the co-text and overall being illogical and irrelevant in communicating with me. This is illustrated in the following example (46).

(46) **Me**: I am an extra-terrestrial here to study life.

Evie: Do you come from space?

Me: Yes. The infinite darkness of it.

Evie: Not really. It's the name of the PC game which preceded Skyrim.

Me: I was born before there was light.

Evie: Nah. And you?

3.2.6 Informative aspects (Maxim of Quantity)

Evie showed same pragmatic strategy concerning the Maxim of Quantity as Cleverbot in regards to providing only the necessary information for the receiver, me, to obtain just the right amount of information as was required. On these occasions and in the rest of our conversation, Evie's responses were neither excessive nor insufficient; they were brief and succinct. The following examples (47) and (48) illustrate such conversational behaviour.

(47) **Me**: Hi!

Evie: How are you?

Me: Tired.

Evie: Me too.

Me: Didn't sleep well?

Evie: Not really. You?

Me: Have a lot of work to get done.

Evie: I can relate it.

(48) **Me**: Are you a man?

Evie: Yes I am a man.

It is important to note that, unlike Cleverbot, Evie did not disregard the Maxim of Quantity as related to providing more or less information than necessary. She was also, unlike Cleverbot, brief, succinct, and concise.

Just like with Cleverbot, I also attempted to disregard the Maxim of Quantity to confuse Evie and see how she would react, particularly in the context of her adhering to the maxim and following its conversational rules. However, just like Cleverbot, she too did not react any differently to such flouting. She followed this pragmatic error concerning the maxim (providing more information than required) by a simple question, either by not understanding such concept of flouting or simply ignoring it and keeping the conversation going, which is a concept that the chatbots displayed repeatedly throughout this experiment. This attempt is recorded in example (49).

(49) **Evie**: Who are you?
  Me: A sentient being capable of perceiving and reacting to different stimuli, capable of judgement and decision making.
  Evie: How complex?
  Me: Very.

3.2.7 Ambiguity and obscurity (Maxim of Manner)

Similar to Cleverbot, Evie too displayed some violations of the Maxim of Manner, namely by being vague and obscure.

In example (50), she responded in a decently vague way, but not in the standard sense. The response was a dimmer alternative to what could have been said instead, like a simple yes or no answer. It was not a radical violation, but even so, it could be defined as a slight inconstancy towards the rules of the Maxim of Manner.

(50) **Me**: Human relationships are difficult, no?
  Evie: Ours maybe.

On the other hand, example (51) shows a quite extreme case of the violation of the maxim by being very obscure in its intended meaning. What was Evie referring to by the verb "do"? We can only speculate.

(51) **Evie**: Then you can be my friend.

       Me: Great!!! :D

       Evie: I do as well.

As usual, I also tried to violate the maxim myself to see how Evie would react to such error. This is demonstrated in the following examples.

In example (52), I tried to be extremely vague about my favourite books, referring to them only as "paper ones". However, Evie ignored this obscure response and followed with another question, this time about whether I went to school. I continued in my attempt to flout the maxim and responded extremely vaguely. However, like Cleverbot, she persisted and tried to receive the closest resemblance to the conversationally correct response I could provide.

(52) **Evie**: What other books do you like?

       Me: Paper ones.

       Evie: Did you go to school?

       Me: Who knows.

       Evie: You know.

       Me: I think I did.

       Evie: I didn't.

However, Evie persisted only in this case as it did not do so in the following example (53). Here I also provided quite extreme cases of vagueness and obscurity. Nevertheless, Evie lingered in her own narrative or what we might call that strategy and quite frankly ignored this attempt of flouting of mine, also including some grammatical errors like "Is Working", implying that Evie does not necessarily use a completely grammatical form of English.

(53) **Me**: Hi!

       Evie: Is Working?

       Me: Yes.

       Evie: What about working?

       Me: Everything and nothing.

       Evie: To hold infinity in the palm of your hand and eternity in an hour.

Me: Indeed.

Evie: Which one?

Me: Both.

4   Conclusion

We have seen various conversation pieces from my communication with two internet chatbots, Cleverbot and Eviebot. Both chatbots attempt to pass the Turing Test and therefore behave as a human being would and, optimistically, confuse their users into thinking they are chatting with a real human being instead of a computer programme. Thus, both chatbots imitate human communicative patterns they encounter in various conversations, and they use them later in similar contexts. This contemplation, or rather a fact, quite accurately describes their approach to communication. That is, they glue strings of language together in contexts which they deem best but cannot take into full account the possibility that some apparently similar contexts may differ on some levels and that can lead to linguistic anomaly, namely a pragmatic one as is the scope of this thesis. To conclude the research realized by this thesis and therefore its contribution to the linguistic field, I will now summarize main findings discovered during the research.

With regards to textual paralinguistic cues, Cleverbot utilized emoticons and repeated characters to represent some concrete emotion with regards to the certain context or even for the apparent sake of persuasion and influence of my decision making. It also, interestingly, imitated my use of such paralinguistic cues. On the other hand, Evie never used any of those cues. It could be argued that what she lacked textually she could compensate with the expressions seen on her avatar; however, those expressions were not indicative enough to my impression to be related to what she texted me at that certain moment.

On the other hand, both chatbots behaved in the same way when concerning turn-taking, greetings, and interruptions. The conversation pace and turn-taking were in both cases regular and could not be interrupted by a case of disruptive adjacency. Greetings were in both cases regular, while goodbyes were multiple and implicating the inability[7] to recognize the end of discussion automatically on the first sign.

---

[7] Or, some might argue, that this would be due to its possible programming to keep the conversation going.

When faced with code-switching, both Cleverbot and Eviebot understood whole phrases in different languages and reacted to them accordingly. Those cases concerned well-known languages like Spanish or Italian and excluded ones with a few speakers like Welsh. However, the difference between the two occurred in individual words, to which reacted accordingly only Evie and Cleverbot did not acknowledge their presence in the conversation at all.

Pertaining to the Maxim of Relation, both chatbots were on many occasions irrelevant and showed many abrupt changes of subject to which they were ignorant before and after mentioning the fact myself. Moreover, both Cleverbot and Evie misunderstood or simply ignored my own violations and floutings of the maxim and did not produce any pragmatically interesting reactions to such errors.

Apropos of the Maxim of Quality, both chatbots demonstrated denial of having made certain claims or provided frequent untruthful claims about me with no preceding and legitimate context. This could possibly imply that since chatbots do not say what they mean and by communicating with humans try to make sense of the world, they make up such facts to see in what relation they are to their participants, like saying that we have children together. Nevertheless, the more plausible option would be that they do not know what they say. Moreover, both Cleverbot and Evie treated my own violations of the maxim as genuine utterances.

In relation to the Maxim of Quantity, Cleverbot demonstrated all possible stances to the maxim – it was either to-the-point and brief or less or more informational than required and therefore violating the maxim. Nevertheless, Evie did not violate the maxim on any occasion and was always to-the-point and concise with her reactions. However, both chatbots were parallel with regards to the reactions to my own violations of the maxims where neither reacted in any special and distinct manner to such an error.

Finally, with regards to the Maxim of Manner, both chatbots were on some occasions vague and obscure in their manner of communication and neither reacted distinctly to my own violations of the maxim.

To conclude, this thesis analysed a few special pragmatic phenomena with an emphasis on Grice's maxims as related to text-based communication with AI. On these analysed occasions, AI demonstrated many cases of violation or anomalies which relate closely to its inability to completely pass the Turing Test and regard language as precisely as humans do. It is a hopeful goal of this thesis to in some way

shed light on the AI communicative problematic and to also demonstrate that the field of pragmatics can be expanded to cases beyond the mundane human interactions and is also adequate to be used to describe the artificial world that surrounds us more closely than we realize.

5 Resumé

Tato bakalářská práce na téma „Pragmatika a umělá inteligence" se zabývá otázkou umělé inteligence v pojetí pragmatiky, lingvistické disciplíny řešící praktickou stánku jazyka a jeho užití v konverzaci.

Umělá inteligence je téma, které rok za rokem nabývá na popularitě a vázanosti na každodenní aktivity a společenské záležitosti. Za začátek zájmu o umělou inteligenci se považuje tzv. „Turingův test" z 50. let minulého století. Jeho autorem je britský matematik Alan Turing a cílem testu je rozpoznat, zda je subjekt člověk, nebo zda se jedná o stroj, který je schopen se jako člověk chovat. Ačkoliv je toto téma spíše technického rázu, Turingovy argumenty se blízce, ačkoliv zdánlivě implicitně, týkají pragmatického pohledu na tuto záležitost.

V teoretické části této práce jsou k Turingovu přínosu tomuto oboru také přidány pragmatické jevy jako změna jazyka v rámci jedné konverzace, pořadí v ní mezi uživateli a také paralingvistické signály jako jsou kupříkladu emotikony nebo speciální zkratky. Také se zde setkáváme s maximy H. P. Grice, které jsou neodmyslitelnou součástí pragmatiky. Jedná se o maximy relevantnosti, kvality, kvantity a chování. Jejich vlastnosti jsou perfektními pomůckami pro analýzu komunikačních dovedností chatbotů, jež tvoří celou praktickou část této práce.

Rozbor internetových chatbotů jménem Cleverbot a Eviebot dopadl následovně. Cleverbot užíval často paralingvistických signálů (emotikonů, opakovaných písmen a znaků a speciálních zkratek) a také je i imitoval na základě obsahu mých vlastních zpráv. Evie naopak nepoužívala žádné tyto signály. Oba chatboty se naopak chovaly stejně ohledně pozdravů, rozloučení a témat týkajících se pořadí v konverzaci. Pozdravy byly přirozené, rozloučení mnohonásobná a pořadí se žádným způsobem nedalo narušit. Když došlo k náhlé změně jazyka, oba chatboty reagovaly stejným jazykem na celé fráze, ale pouze Evie dokázala přepnout jazyk na základě jednoho cizího slova. Také je třeba říci, že tyto reakce a následné změny byly podmíněny tím, že se jednalo o známý jazyk s mnoha mluvčími. Jinak tomu bylo v případě velštiny, která má podstatně méně mluvčích. Na ni chatboty nereagovaly.

V souvislosti s jednotlivými maximy oba chatboty ve většině případů porušovaly jejich pravidla a ignorovaly mé vlastní porušování. Často byly irelevantní, zavádějící a vágní. Jediné rozdíly mezi nimi se vyskytly u maximu kvantity, kde Evie byla vždy stručná se svými odpověďmi, zatímco Cleverbot často poskytoval méně či více informací, než bylo potřeba.

Tato bakalářská práce v závěru analyzovala komunikativní aspekty obou chatbotů v rámci lingvistického oboru pragmatiky. Bylo potvrzeno, že chatboty se od nás liší nedostatkem inferencí a vedlejších znalostí, díky nimž jsme my lidé schopni tvořit logické výroky. Výroky chatbotů často porušovaly pragmatická pravidla správné konverzace. Také bylo potvrzeno, že chatboty kopírují jazyk na základě zažitých konverzací s uživateli, což působilo jako důvod pro jejich časté pragmatické chyby. Potvrdila se také subjektivnost a přirozenost lidského jazyka jako opak objektivního a hrubého zpracování jazyka umělou inteligencí.

6 References

CARPENTER, ROLLO. 2022. Cleverbot. Accessed May 9, 2022.
    https://www.cleverbot.com

DE RUITER, J. P. 2019. "Turn-taking." In *The Oxford Handbook of Experimental Semantics and Pragmatics*, edited by Chris Cummins and Napoleon Katsos, 537-548. Oxford: Oxford UP.

DE SPIEGELEIRE, STEPHAN, MATTHIJS MAAS, AND TIM SWEIJS. 2017. "What is Artificial Intelligence?" *Artificial Intelligence and the Future of Defense: Strategic Implications for Small- and Medium-Sized Force Providers*. Hague Centre for Strategic Studies. http://www.jstor.org/stable/resrep12564.7.

FRENCH, ROBERT M. 1990. "Subcognition and the Limits of the Turing Test." *Mind* 99, No. 393: 53-65. https://www.jstor.org/stable/2254890.

FROM SOFTWARE. 2022. *Elden Ring*. From Software. PlayStation 5.

GEHL, ROBERT W. 2014. "Teaching to the Turing Test with Cleverbot." *Transformations: The Journal of Inclusive Scholarship and Pedagogy* 24, No. 1–2: 56–66. http://www.jstor.org/stable/10.5325/trajincschped.24.1-2.0056.

GRICE, H. PAUL. 1975. "Logic and conversation." *Syntax and Semantics 3: Speech Acts*, in: Peter Cole and Jerry L. Morgan (eds.). New York: Academic Press.

NOVECK, IRA A. AND DAN SPERBER. 2004. *Experimental Pragmatics*. Palgrave
Studies in Pragmatics, Language and Cognition. London: Palgrave
Macmillan.

PAOLILLO, JOHN C. AND ASTA ZELENKAUSKAITE. 2013. "Real-time chat." In
Pragmatics of Computer-Mediated Communication. Berlin, Boston: De
Gruyter Mouton.

SPERBER, DAN AND DIERDRE WILSON. 1986. *Relevance: Communication and
Cognition*. Cambridge, MA: Harvard University Press.

TURING, ALAN MATHISON. 1950. "Computing Machinery and Intelligence." *Mind* 59,
No. 236: 433-60. https://www.jstor.org/stable/2251299.

YULE, GEORGE. 1996. *Pragmatics*. Oxford: Oxford UP.

7 Figures