北京師範大學

学士学位论文

论文题目： **An Analysis of Writing Skills in Machine Written Poetry by Case Study**

作 者：孙载勋

导 师：**Andrew Cardoso**

系别、年级：外国语言文学学院 2014 级

学科、专业：英语语言文学

完成日期：2018 年 5 月

北京师范大学教务处

Abstract

The focus of this research stays between science and literature. Combination work of these two areas has not been done much in the past, and this fact makes this study more valuable. There are numbers of poems written by Artificial Intelligence (AI) that already passed the Turing test, but unlike other tasks done by machines, literature writing is more complicated than only getting a high score for being humanness. This paper aims to discover this complicated area by using literature elements. The emotional respond and overall well-shaped poem are not only about getting a high score for humanness, but other features also make a good poem. This paper was able to discover these elements. Started from the small details of looking through each line of the two lyrics, one written by a human and the other written by a machine. For the result, incomplete and confusing sentences helped to get positive reactions by making the readers able to fill in the gap with their thoughts, but it turned negative when the term became too vague. When it comes to the whole poem, the low score of humanness did not impact much on the overall poem, but the poem as a whole has to make the readers feel like it is written from a human to bring out an emotional response. This study also found that machine written poems already reached the level of human written versions. Although, AI written poems failed to get highly positive or profoundly negative values from the readers. Most of the readers spoke out the level of the poem was mediocre.

Keywords: Artificial Intelligence (AI), tradition, poetry, machine learning

Contents

Part I Introduction

1.1 Background

Today, one of the hottest scientific topics is AI, and they are quickly reaching over all the industries. AI started to take over the works that were considered as ‘human only duty,' and this is same for the literature world. According to Schaub, an AI co-written short novel was able to pass through the first stage of one literary contest (2017). A considerable number of companies are showing interest in this area, and they are making outstanding successes. AI's today is capable of finishing the most artistic works of human including poetry. Google has completed building AI which can write a poem if the first and last line is given (Gibbs, 2016). There are other companies trying to build an AI capable of writing poetry besides Google. Creating an AI that is able for writing a perfect poem can directly prove the possibility of making a human minded machine. For this achievement numbers of companies including Microsoft are driving into this area. According to Chen (2015), AI named Xiao ice wrote over one million poems in less than three thousand hours. Among those poems, 139 were selected and published to the public. Substantial research laboratories and companies are digging inside this field, but the information about the current AI situation is entirely lacking. Lack of knowledge brings unrealistic hope and frightening to the unscientific academic fields. Insofar, papers that are not heavily technological and more targeted to the humanities field is a must at this moment.

Today's science field is publishing an enormous amount of AI-related papers. Based on these papers and researches AI's are getting more intelligent day by day. Even though AI is getting smarter at writing creative works, there is not enough research looking through AI writers in the literature viewpoint. Most of the current research is wholly based on the scientific or literature side. Since today studies are lacking collaboration between the two parties, this paper would draw a connection to the two academic fields. For the scientific team of documents, huge numbers are speaking out that AI's can write as good as humans and sometimes even better. Some scholars also speak out that it is not far from the day that Pulitzer prize winner will become a computer program (Schaub, 2017). Becoming better at writing does not mean the machines are going through the same process with humans of writing. The answer to this question is not solved yet. The reason it

is hard to prove this factor is because today science cannot "emulate real nervous systems in a detailed way" (Reeke, & Edelman, 1988, p. 169). It is mostly clear that machines do not share lots of commonness with human neuro-system (Reeke, & Edelman, 1988). Insofar, today's papers are pointing out that it is possible to make an AI that is same or even better than humans shortly with using a similar or different process of learning. On the literature side, there are not many papers that speak out about AI written poetry. There are only papers that can be used the machines are getting better at creative writings today. Huge numbers of them including "Tradition and the Individual Talent" by T. S. Eliot (1982) speaks out that reading the traditional works from the past is one of the most important parts when it comes to poetry. Learning through books is a part that AI's are much better at than human beings with speed. T. S. Eliot also said that there is no such thing as pure creativity it is just creating a new version by using the elements of the past. His words are an important statement because it proves humans might also use similar path with computers in writing a poem. This paper will combine these two ideas and find out if the level of today's machine written poetry is nearby human written poetry and will also search for the elements that make poem feel like humanistic.

1.2 Purpose and Content of the Research

This paper would gather up two viewpoints writing and, AI algorithm, to find out the humanistic aspects of verse. Collaboration between the two perspectives will carry the first half of the article to generate ideas from previous researchers. This paper would avoid using direct studies in this field and use studies from these two fields. There were two reasons for not using any direct content related to this study. One was because the growth of this field, AI writing is one of the most fast-moving industries. Therefore even one or two-year-old papers of this field was strongly outdated with such a different process with today's machines. The second reason was that this is a new field. Same with other new growing academic branch there was no in-depth research in this area. Not because it is a useless realm to study, but because it was just born with lack of knowledge to itself. Insofar, ideas from the previous existing province were needed. The second half is the significant part that leads this paper with the survey and the interview. It would be able to analyze a poem written by a Google AI by using these two methods of research. Before the examination, there would be a Turing test ran by Asian students to collect more data for this research. The actual

questionnaire would be able to bring out data about how well the AI passed the Turing test, which lines make this poem rank high score of humanness, and which lines feel misplaced to the readers. This study aims to have a better understanding of the current AI algorithms and poetry as a whole. With the better understanding of machine-written poetry and human written poetry, the factors and lines that make the readers give a high score of humanness would be able to appear. Also, this paper will research the data of awkward lines that provide readers with a feeling of displacement, and the elements that give the readers emotional response. With all this data it would be able to draw a picture about the relationship between all the critical reaction shown by the readers. By this, it would be possible to improve the ability of AI writing skills and introduce new algorithms of the writing process to the human writers. Therefore, this paper would be able to give more writing fundamental knowledge on both sides of writers, Human and AI.

Part II Literature Review

2.1 Information of Turing Test

More than 60 years ago scholars started to wonder the possibility of the machines to think as human beings. Turing is the first scholar who brought out this idea publicly in his paper "Computing machinery and intelligence" in 1950. Turing brought the theory of imitation game to find out if the machine can or cannot think like humans. Imitation game is processed efficiently with two chatrooms 'x,' and 'y.' A machine leads one of the chatrooms, and interrogators have a conversation in both of them and have to guess which one is human. The interrogator can ask questions to both of them and have to figure out which one is human, based on the questions tossed by interrogators own selection. One thing important in this process is that the questions and answers have to be passed through computer typing to avoid the impact of the voice and tone in the game. Making more than 30% of the interrogator choosing the wrong answer is the golden goal for the programmers who made the machine. Which means, less than 70% of the interrogators selecting the right answer leads to passing the Turing test for the particular device. Passing the Turing test implies that machine can think like a human being (Turing, 1959). Todays, AI's are deeply in the daily life of people, and they are more intelligent than human beings in lots of sectors but being able to think, and process like human is a more difficult factor. These difficulties were holding back the process of building a humanistic program. These difficulties delayed the first machine that passed this test until 2015.

According to Warwick (& Shah, 2015), an AI chatbot named Eugene Goostman had passed the Turing test. The imitation game was done with the same style that Truing written in his paper in 1950. 5 minutes of questions towards the machine and human about any random subjects was made, and it was not allowed to ask questions that show specific information in the chatroom. One session carried five games with five judges and held six separate meetings in 2days, which gave 30games in total for the result. The Royal societies attention took this test. The effect achieved a 33% rate of the judges speaking out the machine as a human being, which just surpassed the 30% range. However, this does not mean there are no downsides to this achievement. The judges were told that Eugene is a 13-year old Ukrainian boy to explain the poor English grammar and wrong understandings that were made by the machine (Warwick, Shah, & Comput 2015). Because of

these lacks and other limits of this AI, a vast number of scholars are showing their opinion that this result does not indicate that Eugene passed the Turing test. Hern (2014) did a great job these different arguments from scientists. Harnad a professor of cognitive sciences even said "This is complete nonsense. We have not passed the Turing test. We are not even close (para. 11)." Another scientist who works with Eugene's team said, they passed one kind of Turing test, but not a general one that is perfect on every side. These negative opinions came out because the limitation of Eugene above, but there were also scientists that thought the entire idea of Turing test should be shifting and changed. They said the concept of the Turing test should not be just passing this specific numbers but more about making actual humans feel like the machine is a human being for a lifetime, not only for a short conversation of five minutes. These opinions show that to honestly pass the Turing test the critical feature is not 5 minutes or the pre-given information of the test set but making a machine that makes people think its humanized.

This means until today a device that can speak like humans in every subject is not yet possible to build, but when the topics are narrowed down, it is quite simple to create one model. It has already been a while that machines started to write breaking news. Moses (& Peterson, 2017) said, Washington Post begun to use writing bots for their report beginning from 300 short journals about the Rio Olympics, and from then around 850 articles are made from the writing bot Heliograf every year. Heliograf was never running through some specific test but has been passed the Turing test the natural way by writing numbers of an article without making the readers find out that it is a bot.

2.2 The Function of Google Writing Machine

The critical point of AI is the fact that machine can learn by itself. Every computer program has to be written in code to operate. By using these codes, programmers can open a gateway to speak with the computers and make them do the things needed (Kreider, 2002). AI is also one part of computer programing, so the underlying components do not shift dramatically, but the main difference is the fact that AI can build up the algorithms by itself. When it comes to writing poetry, regular programs have to code in what to do each line individually. This means exact words inside the poem has to be written by the programmer. Building an AI that can write a poem is completely different. The programmer has to prepare enough data of lyrics to make the machine read through

and write the right algorism. The programmer only has to tell the computer to go through all of the data and proved a basic version of a math algorism. After going through all the data, the program outputs a math algorithm that can write a poem. This procedure is an entirely different process with a natural way of writing lyrics. Blake (1992), even speaks out it does not follow any morality that every poetry should follow. The output algorithm is too complicated for a human to catch the meanings of each function, but even without the knowledge of that, the readers can know it is a well-made machine by reading the results that have come out from the computer. Domingos (2017) speaks out; this process is so powerful and sophisticated that it soon becomes an algorithm that cannot be understood by a human brain. It is the humans that give AI lives, but after it comes out to the world we cannot truly understand the reason for the result it made, but only guess. An entirely different system of AI with actual humans are causing this problem.

The Google-built poetry machine follows similar steps with the algorithm used in machine learning in general. According to Burgess (2017), more than 11,000 books were fed into the computer before the writing process started and 3,000 books were romance books. Gibbs (2016) said it used a technology called recurrent neural network language model(RNNLM). Translation and image captions commonly use this technology. After finishing the learning data process, the programmer inputs the very first and last line of the poem. The machine goes through these two lines and fills up its sentences between by using the algorism it made by itself. According to Bowman (et al., 2016), this is one type of neural network that uses algorithms similar with our neural systems.

2.3 Tradition in Poetry

There are numbers of reasons why AI algorithm are showing its power in creative writing, and one way of looking this is by looking through "Tradition and the Individual Talent" by T. S. Eliot (1982). It is true that machines are showing their power inside literature, but this does not mean it is a simple step to make, and for poetry, it is even harder. Because of this cause, some scholars speak out that the last step of creative writing for AI's will be poetry. Because, not like other kinds of writing poetry has to move readers heart and carry sympathy (Johnston, 1978). Moving one's mind is a lest easiest thing to approach for a heartless machine. One of the major ideas of Eliot's theory is the term "Traditional." The reason this is one of the most concerned areas on AI writing

is that the key element of AI is self-learning by reading through a huge number of books in a massive speed. The specialty of AI comes from the ability to learn from the huge amount of data and creating a new idea by using the data it has swallowed. When it comes to poetic writing "tradition" is the key element it has to swallow. This term cares more meaning than just the "past work," and it is a needed element to write a poem. Miles (1981) speak out modern poetry is "becoming less derivative and more original." (P. 552) According to Eliot (1982), having the tradition is knowledge about the works from not only the past but also today. Historical sense is not a dead body in the past but breathing with us at the moment. White (2007), argues his thoughts about the term "tradition" as "At various points in his lecture Eliot seems positive to cast about for alternatives, preferring at one moment the somewhat less economical phrase our heritage of literature, at another, the similarly wordy expression 'the whole history of Poetry'" (p. 373). It is evident that AI writers have its benefit on quickly reading through a considerable number of books, but it is not sure the deep learning process contain thoughts of itself towards the materials it has swallowed. AI's are fundamentally different from other computer programs since artificial intelligence is based primarily on symbolic ideas instead of using numeric information like the regular programs (Rich, 1985). Therefore, there is a high possibility to make AI's think and process similarly as humans.

Part III Research Methods

3.1 Research Question

The purpose of this research is to find out if the poetry written by a machine can pass the Turing test done by Asian university students who are majoring in English. This particular type of Turing test will not take the limit on figuring out the level of AI poetry and the similarity between human writers, because there are already vast numbers of machine-written poetry that scored higher than human written ones for being humanness. General Turing test for machine written poetry is already in a level that surpassed most of the human work, and plenty studies prove this fact. For a more significant leap, this paper would go more in-depth on what are the specific elements that make the readers think as a human written version. This paper will use literature viewpoint over linguistics and scientific resources. For this achievement, each line for two poetries will be analyzed to find out which sentence and elements of poetry bring out the specific result. This study will also go more in-depth of human mind side and find out the relationship between high scored humanness and emotional response. This part of the research would not only analyze the two poems as a whole but go further to break down each line and elements that bring out the result. In the full process, contents would be viewed weightier than the context of the poetry. The subject of experiment's level of English would also be chosen broadly to guarantee a broader view from readers with varying degrees of English and understanding of poetry.

3.2 Subjects

The subjects of this research would be limited to Asian students who are majoring in English. More likely, Asian students had less understanding of the structure and elements that vast numbers of past written poetries carry. It is true that this lack of understanding would make them hard to catch the small details inside the poetry, but this study believes the absence of this knowledge will give more focus on the feelings they get by general content taken from the poem. This status matches significantly for this paper since this research is trying to seek what are the things that bring up human-like content rather than the structure. 146 students from China BNU, China Qinghua University, Korea Seoul National University, Korea Korea University, Korea Hanyang University, and Yonsei University have been chosen to take this Turing test. For the gender 78

males, and 64 females have attended in this Turing test and questionnaire. Four students marked their gender section as undifferentiated. This study has chosen eight students who have participated in the inquiry for the interview for more profound questions and overall ideas towards the poetry. This paper selected four students from the ones who got the right answer for the Turing test, and the other four students were the ones who spread out the wrong answer. These students were separated into two groups that carry four students each. Inside each group, two students got the answer for the Turing test in advance of the interview, and the other two had their meeting before telling them the answer.

The interview was made by one to one meeting for each person, to avoid thoughts of other people affecting their own opinion. Three of the interviews were made online interviews were done for three students, but most of them were made on face to face meeting to catch more information among their thoughts. The questionnaire and Turing test was made both online and offline. This study used Google Forms and WenJuanXing(问卷星) for the online survey from first of February, 2018 to 20th of March, 2018. This paper did not use any indirect and secondary opinions for this research. Insofar, the judges and students that have attended the survey, Turing test, and interview were the only elements used for this research.

3.3 Data Collection Methods

3.3.1 Questionnaire and Turing test

The first section of the questionnaire has collected the general information of the Asian English major students. For the general information part questions about gender, English level, closeness to literature, and understanding level of poetry were asked. Gender section questions were multiple choice question of male, female, and undifferentiated. English level and understanding of poetry were based on their own opinion of themselves. A range of one to five was given, with one being poor and five being high. The amount of reading on literature work per year in their mother language or English was also taken to see the relationship and exposure towards literature work. This question did not fit in the range of one to five, therefore made the mean value of the data, six books per year equal to three of the multiple questions. After that, the information was split to fit in the same range of one to five. This process was for making all the data in the same range for

better comparison. Their thoughts towards the understanding of poetry, in general, was also collected for the same reason. By using these four questions, the public data of the students who attended the questionnaire and Turing test were able to be discovered. The second section of the poll was the Turing test. For the Turing test, students read two poetries, one written by a Google AI and the other by Allen Ginsberg. Following is the two-poetry imparted to the students.

	Google AI written poetry		Those Two - Allen Ginsberg
A 1	there is no one else in the world.	B 1	That tree said
A 2	there is no one else in sight.	B 2	I don't like that white car under me,
A 3	they were the only ones who mattered.	B 3	it smells gasoline
A 4	they were the only ones left.	B 4	That other tree next to it said
A 5	he had to be with me. she had to be with him.	B 5	O you're always complaining
A 6	i had to do this. i wanted to kill him.	B 6	you're a neurotic
A 7	i started to cry.	B 7	you can see by the way you're bent over.
A 8	i turned to him.		

After reading these two poems, the students selected one that felt more humanness.

Next four questions were used to get a more profound impression about the two poems. This paper collected questions about the feelings and overall grade among the two lyrics in a range of one to five. This collection was to find out the relationship between the answer for the Turing test and thoughts towards the actual poetry. By using these questions, data about the relationship between getting a high score for being humanness and a good poem was able to be taken. A relationship between having an emotional response and getting a high score for being humanness was also a necessary data that came from these questions.

Data towards specific lines were also taken to figure out which sentences and words make the poem less human and which line makes the readers think as written by a human. The information about the best sentences for each poetry was also taken to seek the relationship between the human-like lines and well-written lines. Lines that gives a feeling of misplacement were also surveyed to collect which specific lines give a wrong impression to the readers. These three data will help to find out the effectiveness of small elements of poetry, and it will also bring out a raw data of specific components.

3.3.2 Interview

Eight students have attended the interview. Half of them were the ones who got the Turing test right, and the other half was the ones who got it wrong. This separation was for getting a variety of answers from students who have different opinions towards these two poems. These groups of four students have separated into two groups again. The first group noticed the answer for the Turing beforehand, and the other group took the interview without getting the answer for the Turing test. The purpose of separating into these groups was to check if the fact of knowing a computer wrote poetry A affects their opinion towards the two poems. All of the interviews were made one by one right after the Turing test to avoid other views mixing into their opinions.

There was a total of eight students who have attended the interview, but each sector only carried two students each. These numbers were not enough to use the interview section to draw new ideas for the paper. Insofar, the data from the interview was only used to help out the data from the questionnaire, instead of being used individually. The most active point of the discussion was to collect more deep information and reasons for choosing one poetry over the other one for the Turing test. Another use was to get the reasons for selecting the lines for all the questions above. The data set of eight was not enough number to carry out the whole paper, but it was a reasonable number for appending some color into the raw numeric data set.

3.4 Data Analysis Methods

There were several methods used for drawing out meanings from the data collected from the questionnaire, but for the general poll did not use lots of techniques. The only thing needed for the gender data was bringing out the ratio of the numbers to know the percentage of each gender that attended the survey. The three questions about the level of English and literature knowledge was differently used. Two graphs were used to see the details. One was to compare the relationship between the three answers. By comparison, this paper was able to figure out the relationship between the actual English level, understanding towards literature, and understanding of poetry. The other graph gave the result of summation for all the data in descending sort order. This data was able to tell if the students were selected from a large variety of English level to bring out exciting results for this questionnaire.

The Turing test data also only used the ratio data to find out if the students who spoke out the right answer stays under 70%. The interview took a significant part of this section for more detailed

opinion towards the two poems. The reason for selecting each poetry and more profound thoughts towards the two poems were able to notice during the interview section.

Four questions were asked to see the reaction of the students towards each line of the poem. The first question was to find out which line made the readers think it is not written by a human, because of unexpected placement or the content in that section. To get an answer to this a question "Which line feels the weirdest?" was asked, and the students were told to select one line. The second question of this section was to find out which line gave a huge impact on choosing the poem for the Turing test and also which line got the highest score of humanness. The question "Which line do you think is the most human-like?" was asked to get the answer. Students had to choose one line among the two poems. The last two questions were to find out which of the lines were most enjoyable for each poem. The question "What is the best line for each poetry?" was asked, and students chose one line for each of the poems A and B.

Four graphs were used to analyze the data from the four questions above. First was the comparison graph to see the relationship between all the four questions. All the data were drawn individually to look at the results of each subject apparently. The purpose of this graph was for getting the general idea about the whole data result. The figure was in descending sort order over weirdness to check which lines feel the most unnatural to most of the students. Following two charts used accumulated style of graph to figure out which lines got the most positive reactions and which line gained the most interest from the readers. To find out the positive responses of each line summation of two positive questions (best line, most humanistic) was made. For the overall response, all the four questions were summed. The last graph was used to give a total score to each line. To get this result a different equation was used, which is $b + m - (w \times 3)$. (b = result of best, h = result of human-like w = result of weirdness) The reason this equation was chosen to calculate the overall score is that there were three times more data sets of positive than negative ones. Therefore, this equation was selected over others to match the equality and power of each question.

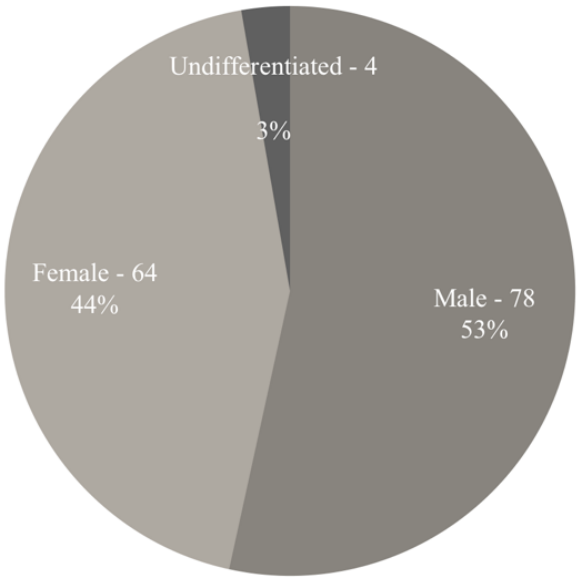
The last four questions were designed to get the opinions of the students towards the two poetries as a whole. The design of the items was to select from a range one to five, as one being low and five being high. First, two questions were to find out how much emotional response the students got from each poem as a reader. To achieve the answer for this the question "Did the piece move your heart?" was asked for both poems A and B. The other two questions were to find out

the overall score the students give to the two poems. The question "What grade would you give to the piece?" was asked for both of the poems. For the graph, poetry A and B were combined for the last four results. In conclusion, two diagrams were used in total, one about the overall score of the poem and the other about how much moved were the readers. The mean and variation value was not in shown in the graph but was separately measured for grabbing more useful outputs from the data set. The actual form of the questionnaire is attached in Appendix A.

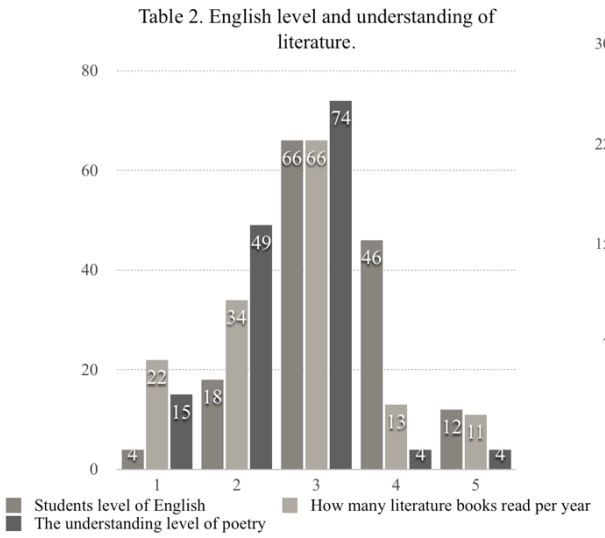
Adding more meaning to the data collected above was the purpose of the interview section. The return value from the interview was not changed to numeric type for raw data; instead, it was used for analyzing the results. It helped to figure out the reasons for the result of the questionnaire and was used broadly over the whole paper.

3.5 Results of the Questionnaire

Table 1. Gender of the students



For the gender of the students who have attended the questionnaire, males were slightly more than females with 53% and 44%, but the difference was not huge, which perfectly matched the purpose of this paper. The empty 3% came from the students who did not want to show their gender for the survey.



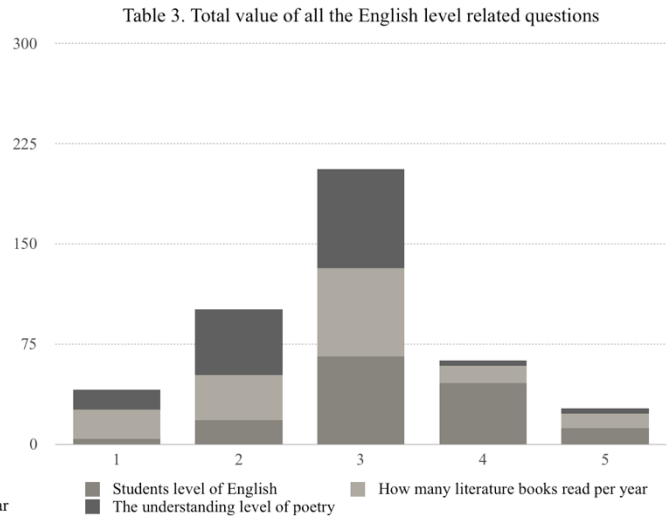
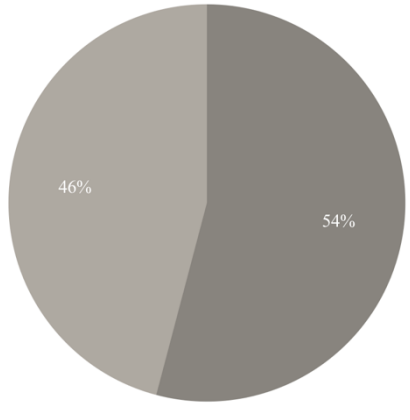


Table 2 shows the individual results of the three questions related to the level of the students. For the three English level related items the majority was staying in the middle, which means this research carries strong opinions of the English major students that remain in the middle level. This result does not carry a hundred present accuracy since the scores were based on the own thoughts of the students instead of test with raw number results. Although, it still shows the rough idea that the chosen students matched the purpose of this research with mostly middle level, with both high and low-level students. The question about literature books was taken with specific book numbers instead of getting it from a range of five. Insofar, a modification was made before drawing the graph to make that dataset fit in this graph. First, the mean value of the data set was calculated which was around 6.02. After that, the highest amount 37 and lowest value one was calculated. The middle-value 4-8 was selected for '3,' and others were spread out evenly with the highest and lowest value. Another interesting point that comes out from this graph is the fact that even though three is the value most chosen from the students, the mean value is all different with a not small gap. The mean value for the level of English stayed the highest with 3.3 and understanding of poem ranked the lowest with 2.54. This result shows that genuinely most of the students have a lower level understanding poetry than the actual English level. Table 3 shows the summation of all the three values. This graph shows that the students who choose '3' were the highest with the overall value of 206, but still not got over half. This fact proves there were lots students on the middle side, but more than half of the total students have stayed in a low or high level of English.

Table 4. Result of Turing test



The students who chosen poem A as a human writer was a bit higher with 12 more students, and this gave a ratio of 54% for poem A and 46% for poem B. Poem A was the one written by Google AI, and it over scored the limit of Turing test which is 30%. It even over scored the poem written by a human, but the difference is minimal with only 8% difference, which stays inside the lapse rate.

Table 5. Students thoughts towards each lines

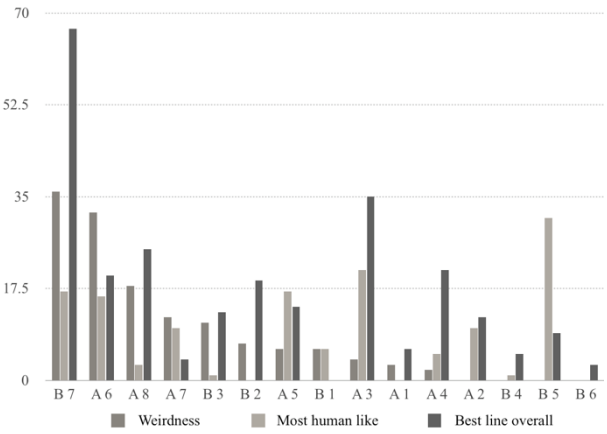
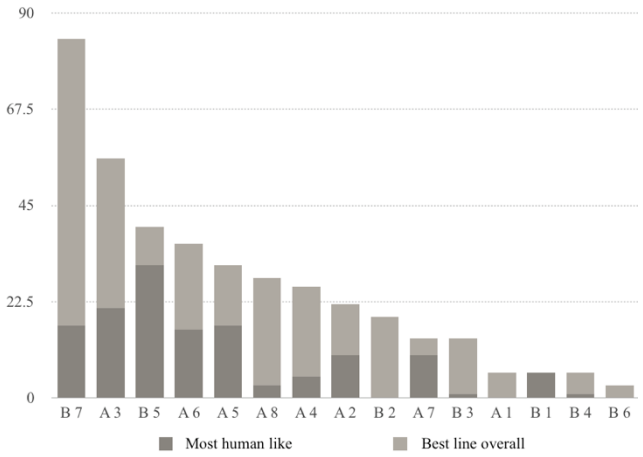


Table 6. Addition of positive lines



A considerable number of graphs were used to analyze the data taken from the questions asking about the opinion towards each line. Lots of impressive results inside these questions had to be drawn at different angles to see the outcomes. Table 5 carries all the consequences of the questions related to the lines individually. This graph showed the relationship between each issue and to see the weirdness it used descending power by the result of weirdness. It also gives a basic visually around all the output results of all the questions. The last line of poem B got the first place of being weird with a score of 36, but ironically it got the highest score in table 6 as well which shows the total positive ratings. The summation of the positive ratings for B7 was 84. By looking through these two tables, this paper can notice that B7 got the most attention both negatively and positively. There were also lines including A3 and B5 that got high scores of positive reactions without many negative ones. Line A8 also showed an impressive result with the students appearing strangeness and getting a low score for humanness but getting the third place for getting the most votes for a good overall line. This line shows that humanistic and good written poem lines might have less connection.

Table 7. Addition of all the results

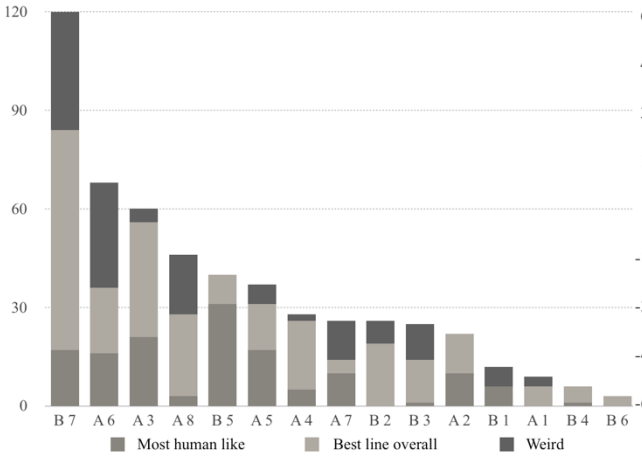


Table 8. Overall score for each lines

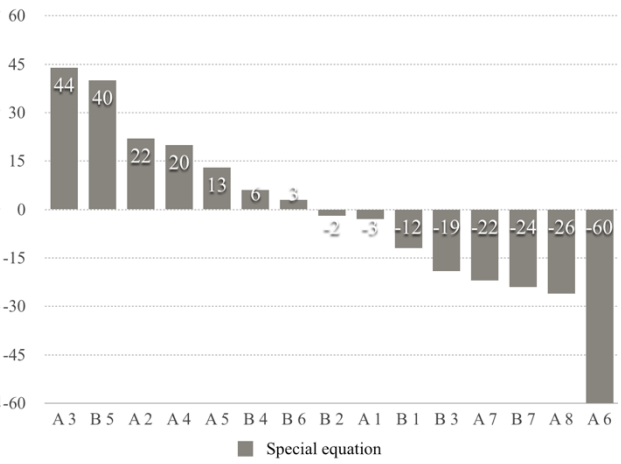


Table 7 and 8 were also for the questions for each line of the poem. Table 7 added all the results for the four questions to see the overall interest. In this table, it can tell that it is true that B7 got the most interest overall with 120 votes. This outcome proves the idea from table five, and six once again. This graph also shows the lines that got almost no attention including B4 and B6. Going through these lines and finding out the elements that prevent readers from profoundly thinking about the lines would also be an exciting factor to discuss in the next section. Table 8 uses the unique equation b + m - (w × 3), (b = result of best, h = result of human-like, w = result of weirdness) to find out the lines that got a high score for all the positive outputs. This graph shows that A3 and B5 reached the highest value with scoring 44 and 40. This result means there were relatively low numbers of students who choose these two lines for being negative and vast numbers of students have spoken out that these lines feel positive. B7 get a low ranking in this graph of getting a third place from the bottom, because of too many students showed their opinion that this line was not expected to be written in this part.

Table 9. Effectiveness of the poems

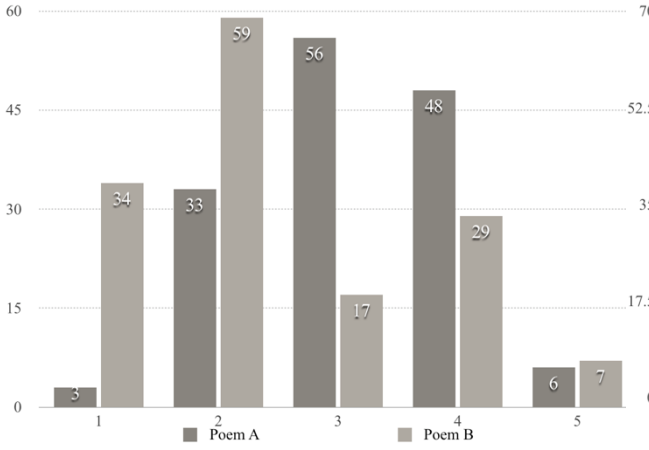


Table 10. General grades for the poems

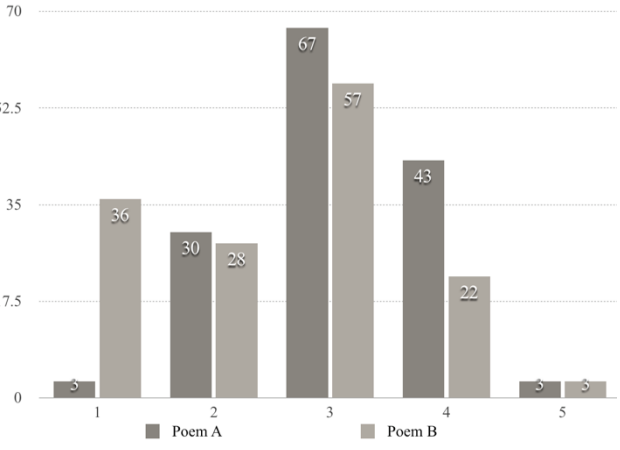


Table 9 and 10 shows the results from the last part of the questionnaire that carries the grades for poetry A and B as a whole. Table 9 responds to the question about how much emotional response did the reader's show, after reading these two poetries. By direct comparison, it is shown that poetry A got higher scores for this question. It is shown in the graph because A got much more students choosing four for the grade, and B got too many students choosing one. The mean value for two of them shows the difference between the two poems with raw numbers. Poem A was near 3.143 and B got near 2.424, which proves again that A got slightly higher scores. The exciting part of this data is the variation value of the two. Variation for poetry A was only about 0.78, which is low, but poetry B got a variation around 1.395. This result is a considerable difference and tells students had similar opinions among A but had remarkably different ideas for poetry B. Table 10 shows the results about the overall scores of the two poems given by the students, and a lot of similarly was shown. The variation value was similar too, poetry A got even less value of 0.656, and B was about 0.5 points higher which gave a result around 1.167. Even the overall scores had similar results with the first question as it is shown in table 10. For the mean value, A got near 3.089 and B got around 2.506. These two graphs show that getting a good impression as a poem and giving the readers emotional response has a close association. More detailed numbers for all the results will be in appendix B.

Part IV Analysis of the Results

4.1 Analysis of the Subjects

As shown in previous data, all the students who have attended for the questionnaire had an almost equal number of gender, and the level of English was also genuinely spread. For the level of English, there were more elements to analyze. One is the difference between the overall level of English, exposure towards literature, and understanding of poetry. The mean values were 3.3, 2.71, and 2.54, which gave the whole English level the highest and knowledge of poetry the lowest. This result shows not vast numbers of student's study English by using any literature books, and their interest towards poetry and literature as general are staying in a low status. The Ministry of Culture, Sports and Tourism announced the mean value of the books read by the adults in Korea, and the mean value was 8.3 books per a year for 2017. Since these numbers of books include non-literature works, so it is not possible to say the Korean and Chinese university students majoring in English reads less than the average adult in Korea, but it is true that it is still staying in a low number. Since this study needed students who are not so familiar with the components of poems, they perfectly matched the purpose of this study, but increasing the books read by major English students in North-East Asia students are a need. Variation value for the raw data of the books read per year was also worth analyzing since the number was high with the amount 37, this means there is a massive gap between the students who read a lot and less.

4.2 Analysis of Turing Test Results

Numbers of poetry have surpassed the ratio of 30% for tricking the judges who have attended the Turing test, but the problem of this is that most of the judges were people with high-level English and great interest towards poetry. In this kind of condition, it is almost impossible for the judges to make their decision without any stereotypes among the traditional poems. This fact makes it harder to pass the Turing test for the machines, but the problem is that the focus point turns to find the verse that seems like human written and have elements of past lyrics. In this condition, a great human written poem with a brand-new style will have a hard time defeating a machine written poem that follows all the elements of lyrics. Things including rhymes, rhythmical language, and verse become too focused. The problem is that these are not the factors that are

highly important in poetry. According to Masters (1915), none of the following style matters, since the critical thing in a poem is to inspire an ordinary man. Anderson (1999) goes even further and speaks out poetry has excellent power for healing that it can also be used in Psychotherapy. These are the crucial factors in a poem, and this Turing test was specially designed to measure the level of current AI in these parts.

For the actual Turing test Google AI poetry has surpassed the 30% range and got even higher score than the one written by Ginsberg with making 54% of the students choosing poetry A over B for being more human-like. Even though poem A got 12 votes more than poem B, it is not possible to say it is better since the difference is too small, but the data still shows us not many students were able to tell the difference between the two. The reason for choosing one poem over the other was different from the individuals according to the interviews with the students. Student one said it was tough to pick one over another since both of the poetry felt awkward. The most exciting discussion came out from two students who have chosen poem B, students said poem B felt more human-like than A because it was more coherent. Student 3 also selected poem B for similar reasons she said that the lower case 'i' throughout the whole poem was annoying and that was the critical point of choosing B over A for being more human-like. This was an unexpected result. The Google writing AI did not go through poetry data, it only learned by reading novels; therefore it is understandable for not entirely following elements of poetry, but not being coherent in general is a different story. The machine swallowed writings that were contents of the human writing. Therefore, this paper has predicted the result would be in human writing from; the interviews have shown different results. The more interesting fact is that the reason for the students choosing AI written poetry was because it was moving. The reason student 5 and six chosen AI poem as more human-like was because they felt it has touched their emotions. Touching emotions is the hardest goal to achieve for an AI, and Google writing bot showed the possibility of this in the questionnaire and interviews.

4.3 Analysis on the Poetic Elements in Each Lines

4.3.1 Lines with Negative and Positive Results

Lots of interesting facts came out from analyzing each line of the poem based on the results of the questionnaire. One interesting factor of the survey are the lines that got lots of votes for both

positive and negative reactions. The last line of poetry B, sixth line of poem A, and last line of poem A are the ones that got the most votes for not associating with the previous lines. Around 62.77% of the students have chosen these three lines as the weirdest, but this did not lead these lines of getting less positive reactions. These three lines reached 37.85% for the summation of positive value score. Among these three lines, the last line of poem B got the most exciting score by getting the first place in both sectors. It gained 26.28% for negative votes and 21.48% for the affirmative votes. Student one said the reason the last line of poem B felt so weird is that it is out of grammar and the meaning of the sentence did not make sense. It is true that this line feels like out of grammar and carries complex ideas that make the readers hard to understand, but why did it also get lots of affirmative votes? It might be because it is the crucial line of the poem. Over the three lines that were chosen for grabbing the most interest of the students two of them were the last line of each verse, and both lines carried meanings that are not truly clear. Poem A had a climax of the story right before the last line, and this last line wrapped up the story, but just by showing the action made from the girl, without expressing any emotions of her. The last line of poem A is the only part that expresses action instead of feelings throughout the whole poem A. This leaves a space for the readers to fill in their own emotions and imaginations inside the sentence. This imagination would brighten up the line more than what it is. Unlike formal writings, most of the literature works prefer to be a little bit abstract. Abstractness brings in the mind of the readers inside the work so they can feel like it is their own story. The last line of poem A follows this simple rule of literature, and this might be the reason it was able to gain lots of attention. The last line of poem B also follows a similar style. It seems to carry all the language the poet wants to spread out, but it is written callously. The sentence "you can see by the way you're bent over." only shows the result without any given clue about the reason. These are contents that are hard to swallow since it is hard to understand the deep meaning inside of it. This situation is same for the sixth line of poem A. This line does not match perfectly with the previous line. The mood of the poem suddenly changes to horror from romantic. This shift makes the line a confusing one. All the students who have attended the Turing test were humans, and they made their definition towards what are the elements that make the piece feels like it is written by a human inside their minds. It seems one of the definitions caused by lots of students is being understandable. Not being able to understand would be the reason why so many students chosen these lines for being awkward, but

this involved content also makes an area for the imagination to hop in. Since there are a lot of empty spaces inside of the material the readers can fill the information with their own experiences. This process makes the line feel fitter right into the readers. For these three lines they have succeeded to create equivocalness to make the readers fill in their thoughts, but it went a bit too far which made some of the students lost entirely inside the sentence.

The last line of poem A is unique in the broader data. This line got lots of attention on the negative and positive side, but only three people spoke out this line feels the most human written. The overall votes were high with making 25 students think it is the best-written line in poem A, and it even got lots of attention for feeling like it is misplaced with 18 votes. Being humanness was the only score it scored low. The reason for getting lots of votes for the two question is apparent. It is because some number of students did not understand the meaning of the sentence and thought it is misplaced, but some other students pushed their thoughts into the sentence and made their version of ending. The reason it got such a low score for being humanness might be more complicated, but one statement can be made. It is because the readers do not think they would have ended the story in this way. Understanding the sentence and feeling it is a different two factor. Humans can recognize situations that they never experienced before, but they cannot empathy a case that they have never been. This paper considers this same situation has happened to this last line of poem A. Some numbers of readers were able to draw their picture for this sentence, but few numbers of students were able to empathize the meaning inside the phrase. Lack of understands towards humans would be one of the reasons this line failed to score a high score for being humanness.

4.3.2 Lines with Positive Results

There was also some line that had lots of positive reactions without many negative ones. From line two to four in poem A and fifth line of poem B are the ones. They got the first to fourth place for the individual equation score which was for finding out the best lines overall. B5 and A2 got non-votes for being weird and other two lines also only got four and two votes with being one of the high rankers of the addition of positive lines graph. All of these lines carry meaningful contents, without any hard languages that are hard to understand. This result proves that each line must move significant meanings inside to grab the attention of the readers. By demonstrating the effects

the other side, it becomes more evident. Some lines got shallow attention for all the questions. Lines including "That tree said," "there is no one else in the world," and "That other tree next to it said" are those lines. Only three to six positive and six to non-for negative votes. Therefore, they stay on the lowest part of the addition of all the results graph. The reason they did not get any negative response is that there are no grammar issues, and there was nothing hard to understand with straightforward sentences. They also got less attention for the positive side, because there was nothing much to think about in these lines. These are lines that are merely explaining the situation without any rich vocabularies. It also did not carry any essential meanings. Sentences that carry these elements did not grab the attention of the readers.

4.3.3 Relationship Between Individual Lines and Poem

According to the individual results and interviews, one good line or weird line did not influence the thought of the whole poem that much. This fact is proved by looking up the individual data sets of students. Almost 30% of the students have selected the most human-like line from the poem they thought an AI wrote it. This fact proves the influence of a single line is not that active throughout the whole poem.

4.4 Analysis of Two Poems Overall Scores

4.4.1 Connection Between humanness and Overall Score

Two questionnaires asked about the overall poetry, one was about the emotional response, and the other asked the general grade for both poetry. The data for these two questions showed that these two questions and the Turing test were closely related. The mean values for A for were 3.14 and 3.09. For B it was 2.42 and 2.51. When these two results transpose the ratio, poetry A gets 56% and 55% which were staying at the similar value with the Turing test that outputted 54%. With the error bound this can be shown that all the grades came out the same. This result indicates that scoring high for humanness has extraordinary influence with emotional response and overall well-written poem for the students. It is okay to have some lines that make the readers feel it is misplaced, but the whole poetry has to feel like it is written by a human and also understandable. Without these two elements, most of the readers would feel like it is a poorly written poem.

4.4.2 Different Variation of the Two Poems

One other exciting output came from the variation of the two results. The mean value of the two poems did not show a huge difference, but the variation value has proved an enormous difference. The variation for poetry A stayed low with 0.78 and 0.65 which meant lots of students did not give high or low scores. Most of the students remained in the middle. Not much students thought this poem is well written, but not many students thought it is deplorable either. Poem B showed the result of 1.39 and 1.17. This value indicates many students have chosen original scores for these poems. Some students loved this poem, but also lots of them thought it is bizarre. It is an exciting factor by only the result without knowing the reason, because this shows machine written poems stay in the middle giving similar emotions for most of the people, but human written ones carry lots of likes and dislikes. This paper shows it limits analyzing the reason for this because more data would be needed to explain this cause, but it is clear that human writers are still in a higher level than machines in touching the hearts of humans vigorously.

4.4.3 Different Opinions Towards the Same Poem

Interviews around these questions were able to show lots of opinions of students, numbers of them spoke out for the reason of giving high scores for the poems. The idea stayed in similar factors with the Turing test, which provided a result of the poem A being more moving and B as more coherent. One of the elements that made the student think that A is out of consistent was the repeatable lines. One particular opinion towards poem A was that it is scary. Student two said poem A made him frightening because he thought A was written by a machine and the lines written by it was malicious about the world. Students who have chosen A over B had entirely different answers student five, and six said they were able to see loneliness and feelings of human in poem A. Especially the gloomy mood inside of it made them feel more moved. The other common factor these two students shared was the negative feelings towards poem B. Student six said it was robotic, and student five said poem B misses the critical element of a poem which is saying some words to the reader. It is interesting that students have entirely different opinions towards the same words and lines. Having different views is one of the critical factors of the poem, and the thing that makes poem an exciting style of writing that lasted for ages.

Part V Conclusion

This paper found out that a high scored poetry for humanness does not mean it is a well-written poem, but it takes a massive roll of moving the heart of humans. With this big idea, it was able to find some elements that make Asian students think it is humanness poetry. When the lines became too confusing, the readers lost the whole meaning and a massive number of students though it is unnatural and misplaced, little abstractness inside the sentence gave positive reaction towards the lines. There was no colossal harm in writing confusing lines if it carries big ideas inside because those confusion and abstract sentences leave some space for the readers to take their imagination inside the poem. It was an essential factor for the lines to toss words with deeper meaning instead of giving out the straight information. These straightforward sentences block the readers genuinely engaging with the poem because it made the readers hard to put themselves in the poem and be the character. This study discovered the difference between understanding and empathizing. If the action of the line is not fully understood, lots of students failed to empathize on that line and did not think a human writer sufficiently reproduced it. When a line had lots of emptiness for the thoughts of the readers to jump in without the right content it made the readers felt less humanistic. Interestingly different students had entirely different opinion on the same lines in the interview, but lots of students agreed poem A had more elements that touched their feelings and poem B was more coherent. When it comes to poetry as a whole, it was okay for some of the sentences to feel less humanness, but the whole poem had to have a high score for humanness to make the readers moved and impressed at the poetry. When it came to the three overall scores of both poems they had entirely different scores of variations, which showed human written lyrics brings more various opinions from the Asian students.

This paper also found some negative impact on the Turing test style of the questionnaire. Student four said it was hard to concentrate on the poetry itself because too much attention went on finding out which one is written by a human. It shows that giving out the questionnaire without the information about the machine might be able to collect more accurate data for what kind of feelings do Asian students get from poems. Not being able to find out the exact reason for getting such a different value of variation for the two verses is another limitation for this paper and would be a right area for the future studies to search. Numbers of humanities researchers have different

opinions appearing from the same data. The problem is there are no wrong answers among them; they are all right in some way. Insofar, it is not the answer that humanities researchers are seeking. It is not the answer that is needed; the actual value comes out from a massive amount of opinions and the right tuning of those thoughts. This idea comes same for this paper, the data sets are all raw, but the viewpoint of looking through these data would be different for every scholar. These would be all valuable viewpoints and since humanity side researchers have a massive lack of machine-written poetry at the moment, to genuinely bright up this branch lots of related studies will be needed.

Appendix A

Google AI poetry Turing test

Poetry A
there is no one else in the world.
there is no one else in sight.
they were the only ones who mattered.
they were the only ones left.
he had to be with me. she had to be with him.
i had to do this. i wanted to kill him.
i started to cry.
i turned to him.

Poetry B
That tree said
I don't like that white car under me,
it smells gasoline
That other tree next to it said
O you're always complaining
you're a neurotic
you can see by the way you're bent over.

What is your level of English?(Add the grade and name of the English level test if you took one)

- ○ 1
- ○ 2
- ○ 3
- ○ 4
- ○ 5
- ○ Other: _____

How do you define your gender?

- ○ Male
- ○ Female
- ○ Prefer not to say

How many literature books do you read per a year?

Your answer

Do you think you have a good understanding of poetry?

1	2	3	4	5
○	○	○	○	○

Which poetry do you think is written by a human writer?

- ○ A
- ○ B

Which line feels the most weird?

	Line 1	Line 2	Line 3	Line 4	Line 5	Line 6	Line 7	Line 8
Poetry A	☐	☐	☐	☐	☐	☐	☐	☐
Poetry B	☐	☐	☐	☐	☐	☐	☐	☐

Which line do you think is the most human like?

	Line 1	Line 2	Line 3	Line 4	Line 5	Line 6	Line 7	Line 8
Poetry A	☐	☐	☐	☐	☐	☐	☐	☐
Poetry B	☐	☐	☐	☐	☐	☐	☐	☐

What is the best line for each poetry?

	Line 1	Line 2	Line 3	Line 4	Line 5	Line 6	Line 7	Line 8
Poetry A	☐	☐	☐	☐	☐	☐	☐	☐
Poetry B	☐	☐	☐	☐	☐	☐	☐	☐

Did piece A move your heart?

1	2	3	4	5
○	○	○	○	○

Did piece B move your heart?

1	2	3	4	5
○	○	○	○	○

What grade would you give to piece A?

1	2	3	4	5
○	○	○	○	○

What grade would you give to piece B?

1	2	3	4	5
○	○	○	○	○

SUBMIT

Never submit passwords through Google Forms.

Appendix B

Gender of the students

Gender	Number of Students	Ratio
Male	78	53%
Female	64	44%
undifferentiated	4	3%

The English level of understanding of literature of the students

Score	Level of English of the students	How many literature books read per year	The understanding level of poetry	Summation value
1	4	22	15	41
2	18	34	49	101
3	66	66	74	206
4	46	13	4	63
5	12	11	4	27
Mean value	3.30136986301	2.70547945205	2.54109589041	2.84931506849
Variation value	0.785888534434	1.1392850441	0.672968662038	0.972727841371

Exact number of books read per year

Number of books read	Number of students	Mean value	Variation value
1	22	6.02054794521	36.9105366861
2	14		
3	20		
4	16		
5	18		
6	20		
7	12		
8	1		
10	4		
11	1		
12	1		
13	1		
14	2		
15	3		
16	3		
17	2		
20	1		
26	1		

Number of books read	Number of students	Mean value	Variation value
27	1		
30	1		
32	1		
37	1		

Result of Turing test

	Machine written poetry	Human written poetry
Number of students	79	67
Ratio	54%	46%

Result of scores for each line in the poetry

	Weirdness	Most human like	Best line overall	Summation of positive values	Summation of all the values	Overall score
A 1	3	0	6	6	9	-3
A 2	0	10	12	22	22	22
A 3	4	21	35	56	60	44
A 4	2	5	21	26	28	20
A 5	6	17	14	31	37	13
A 6	32	16	20	36	68	-60
A 7	12	10	4	14	26	-22
A 8	18	3	25	28	46	-26
B 1	6	6	0	6	12	-12
B 2	7	0	19	19	26	-2
B 3	11	1	13	14	25	-19
B 4	0	1	5	6	6	6
B 5	0	31	9	40	40	40
B 6	0	0	3	3	3	3
B 7	36	17	67	84	120	-24

Result of scores for each line in the poetry by ratio

	Weirdness	Most human like	Best line overall	Summation of positive values	Summation of all the values
A 1	2.190%	0.000%	2.372%	1.535%	1.705%
A 2	0.000%	7.246%	4.743%	5.627%	4.167%
A 3	2.920%	15.217%	13.834%	14.322%	11.364%
A 4	1.460%	3.623%	8.300%	6.650%	5.303%
A 5	4.380%	12.319%	5.534%	7.928%	7.008%

	Weirdness	Most human like	Best line overall	Summation of positive values	Summation of all the values
A 6	23.358%	11.594%	7.905%	9.207%	12.879%
A 7	8.759%	7.246%	1.581%	3.581%	4.924%
A 8	13.139%	2.174%	9.881%	7.161%	8.712%
B 1	4.380%	4.348%	0.000%	1.535%	2.273%
B 2	5.109%	0.000%	7.510%	4.859%	4.924%
B 3	8.029%	0.725%	5.138%	3.581%	4.735%
B 4	0.000%	0.725%	1.976%	1.535%	1.136%
B 5	0.000%	22.464%	3.557%	10.230%	7.576%
B 6	0.000%	0.000%	1.186%	0.767%	0.568%
B 7	26.277%	12.319%	26.482%	21.483%	22.727%

Effectiveness of the poetries

	A poetry	B poetry
1	3	34
2	33	59
3	56	17
4	48	29
5	6	7
Mean value	3.1438	2.4246
Ratio of the mean value	56.458%	43.542%
Variation value	0.780681178457	1.39500844436

General grades of the poetries

	A poetry	B poetry
1	3	36
2	30	28
3	67	57
4	43	22
5	3	3
Mean value	3.0890	2.5068
Ratio of the mean value	55.202%	44.798%
Variation value	0.656455244886	1.16776130606

Appendix C

Transcripts

Transcript 1 - Right answer for the Turing test, with the information of the answer

Interviewer: Did you enjoy the Truing test and other questions?

Student: Kind of, but it was hard for me because all of the poems felt slightly weird.

Interviewer: What was the reason for choosing B over A for the Turing test?

Student: It was almost just entirely a guess, that is why I was surprised after getting the right answer. Not because the computer poetry was human-like, but because the human written one was so computer-like.

Interviewer: You have chosen the last line of poetry B as the weirdest line, even though you have selected poetry B for the Turing test. What is the reason for that?

Student: I think the last line of poetry B was out of grammar and did not get the meaning of the sentence. That is why I have chosen it over the other phrases. As I said the reason for selecting poetry B was a guess, but it was too repeatable. I think that is the most significant reason I dropped out poetry A as a human version.

Transcript 2 - Right answer for the Turing test, with the information of the answer

Interviewer: What was the reason for choosing B for being more human-like?

Student: It seems more coherent than A. It feels to me that A is more like an emotional catharsis, but b has more cohesion.

Interviewer: Did any of the poems moved you?

Student: None of them moved me, A scared me.

Interviewer: Why did it scare you?

Student: For I regard this one as the one written by a machine. I feel that the computer is malicious about the world. So this brought about bad associations.

Interviewer: Among all which line do you like the most?

Student: I turned to him. The one scares me the most.

Interviewer: Why was it scarier than the line that killed the guy?

Student: No like the line implicated that it is going to do it. The lines before means it has considerations for it. Giving the notion some thought or something like that, but the last line left room for imagination. This is the scariest part

Transcript 3 - Right answer for the Turing test, without the information of the answer

Interviewer: Did you enjoy the Truing test and other questions?

Student: Yes, I did.

Interviewer: Good! Did you enjoy the Truing test and other questions?

Student: I enjoyed poetry A better than B, but the lower case 'i' was kind of annoying to me. I know that in poetry you can allow grammar errors for poetic license, but I could not find any purpose for that 'i.' That was the biggest reason for choosing poem A as an AI written one.

Interviewer: You said that poem A was more enjoyable than B. What is the reason for that?

Student: I think I have to say I liked the mode of the poetry. I like gloomy, and love stories and poem A followed that line. Another reason was that it was more like a story than a poem. Maybe it is because it is too poetic, but I did not get the meaning for poem B. I did understand what the poet wants to speak out through the poetry. However, for poetry A even with some wired structure I was able to grab out the image of the story. I think that is the most significant reason why I graded the overall score and more moved score higher for poem A, even though it I thought it was a computer version.

Transcript 4 - Right answer for the Turing test, without the information of the answer

Interviewer: What was the biggest reason for choosing B as the more human-like a poem?

Student: First is because I felt B is more coherent. I did not like poem B that much but still thought it is much more consistent than poem A. Another reason is that I thought A is way too repeatable. It seemed like it is trying too hard to make the rhythm by repeating the same words again and again.

Interviewer: Were there any difficulties with the Turing test?

Student: It was hard to concentrate on the poem itself, because of the Turing test. I had to focus on finding out which is written by a human, rather than truly enjoying the poetry.

Interviewer: Did any of the poem move your heart?

Student: No, as I said I was thinking about which is the human version, so was not able to enjoy and feel the poem itself.

Interviewer: What was the reason for choosing the 6th line of the poem A for both weird and human-like?

Student: I do not have a particular reason there is violent expression maybe?

Transcript 5 - Wrong answer for the Turing test, with the information of the answer

Interviewer: what was the reason for choosing A for being more human-like?

Student: There seems to be more emotion in the first poem. Could I see the loneliness? I think the third phrase from the beginning of A expressed feelings that would be felt if it were a human being.

Interviewer: Ok, where your own emotions moved by reading poem A?

Student: Yes!

Interviewer: What about poem B?

Student: Not really, it was interesting to express the trees in a personified way. But feel like doesn't make sense

Interviewer: Ok, so you liked the idea of thinking the stuff in a new way, but did not like the content of the storyline that much right?

Student: Yes. In particular, I think poetry should have something to say to me, but I do not know what to say in the second poem. Just tree complaining.

Interviewer: Were there any lines that broke your mood while reading A?

Student: I had to do this.

Interviewer: Is it because it did not make sense that much of the previous line?

Student: I think that sentence is not needed. In the next sentence, I say I'll kill him, but that sentence interrupts the flow...? Mood?

Interviewer: Ok, did it change your mind towards poem an after knowing that a machine wrote it?

Student: It does not matter who wrote it. I am not going to change my mind Because the feelings this poem gives me are unchanged.

Transcript 6 - Wrong answer for the Turing test, with the information of the answer

Interviewer: what was the most significant reason for choosing that one over B?

Student: A was sentimental, and B was unnatural, I just felt B was robotic.

Interviewer: Ok, did the sentimental stuff in poetry A moved you?

Student: I was deeply moved.

Interviewer: Which line made you feel moved the most?

Student: The first line, there is no one else in the world. I think that line is the sick one that makes me fall into loneliness. Don't know why I just longed for solitude when I read A for the first time. I felt an untold agony, and I explored space at that time.

Interviewer: Ok, did poem B also moved you?

Student: As I told you B is unnatural, apparently, it is rectangular, but I am a compassionate person, so absolutely, definitely B moved me as well.

Interviewer: Which was the weirdest line of both poems?

Student: The poem 'A' produces an atmosphere of terror, but which is hilarious, because the line "I wanted to kill him" indicates in detail the narrator wanted to kill him? I think the line is not suitable for the characteristic of poems. I feel like B's personification makes me feel unrealistic. I know the author wants to give us a lesson like "Broaden your horizon so that as you become more able to take care of yourself."

Interviewer: Very last question, did it change your mind towards poem an after knowing that a machine wrote it?

Student: Hmm well I do not know...

Transcript 7 - Wrong answer for the Turing test, without the information of the answer

Interviewer: What was the biggest reason for choosing A for being more human-like?

Student: The sentence (I had to do this I wanted to kill him) makes me feel like the first poem is more human-like. It is rather emotional. Like I thought AI is not possible to feel any emotion

Interviewer: Ok, did the poem make your personal feelings moved?

Student: Actually, I cannot assure that the poem moved me but the poem A is more readable than the last one.

Interviewer: Ok, which line do you think is the weirdest among all?

Student: You can see by the way you are bent over.

Interviewer: What is the reason for it?

Student: Why did the author suddenly come up with this sentence? It does not match with former sentences, I mean the content

Transcript 8 - Wrong answer for the Turing test, without the information of the answer

Interviewer: What was the reason for choosing poem A as more human-like?

Student: Because I could not get any meanings from poem B, I did not get what the writer wants to say. That made me feel like a robot defiantly wrote poem B.

Interviewer: Then for sure poem B did not move your feelings right?

Student: Yes, it did not move me at all.

Interviewer: Than what about poem A did that move you?

Student: It did, but not that much. I think it is because I did not like the style of it though. Not because it is a poorly written poem.

Interviewer: Which line do you think is best among all the lines?

Student: O you are always complaining was the best line for me.

Interviewer: Why was it?

Student: I just felt like, that particular line was speaking towards me.

Interviewer: You thought that an AI wrote that poem, didn't that fact bother your moods of liking that line?

Student: No, I do not think that matters that much.

Reference

Anderson, R. N. (1999). The healing powers of creativity: Using poetry in psychotherapy. Individual Psychology, 55(2), 256.

Blake, R. W. (1992). Poets on Poetry: The Morality of Poetry. The English Journal, 81(1), 16.doi:10.2307/818330

Bowman, S. R., Vilnis, L., Vinyals, O., Dai, A., Jozefowicz, R., & Bengio, S. (2016, May 12). Generating Sentences from a Continuous Space. Proceedings of The 20th SIGNLL Conference on Computational Natural Language Learning. doi:10.18653/v1/k16-1002

Burgess, M. (2017, October 4). Google's AI Has Written Some Amazingly Mournful Poetry. WIRED. Retrieved from www.wired.co.uk/article/google-artificial-intelligence-poetry

Chen, Y. (2015, August 20). Microsoft relaunches chatbot Xiaoice. Global Times. Retrieved from.http://www.globaltimes.cn/content/938219.shtml

Domingos, P. (2018). The master algorithm: How the quest for the ultimate learning machine will remake our world. New York: Basic Books, a member of the Perseus Books Group.

Eliot, T. (1982). Tradition and the Individual Talent. Perspecta, 19, 36-42. doi:10.2307/ 1567048

Gibbs, S. (2016, May 17). Google AI Project Writes Poetry Which Could Make a Vogon Proud. The Guardian, Guardian News and Media. Retrieved from www.theguard ian.com/technology2016/may/17/googles-ai-write-poetry-stark-dramatic-vogons.

Hern, A. (2014, June 9). Scientists Dispute Whether Computer 'Eugene Goostman' Passed Turing Test. The Guardian. Retrieved from www.theguardian.com/technol ogy/2014/jun/09 /scientists-disagree-overwheth er-turing-test-has-been-passed

Johnston, E. T. (1978). Forward to the Basics with Poetry. English Education, 9(4), 228-231.

Kreider, T. (2002). A.I.: Artificial Intelligence. Film Quarterly, 56(2), 32-39.

Marshall, R. (1999). Caribbean Healers and Healing: Awakening Spiritual and Cultural Healing. Powers. Integrating Traditional Healing Practices into Counseling and Psychotherapy, 73-84.doi:10.4135/9781452231648.n7

Masters, E. (1915). What Is Poetry? Poetry, 6(6), 306-308. Retrieved from http://www.jstor.org/stabl e/20570534

Miles, J. (1981). Poetry and Tradition. Pacific Coast Philology, 16(2), 1-8.

Moses, L., & Peterson, T. (2017, September 17). The Washington Post's Robot Reporter Has. Published 850 Articles in the Past Year. Digiday. Retrieved from https://www.digiday.com/media/washington-posts-robot-reporter-published-500-articles-last-year

Reeke, G. N., & Edelman, G. M. (1988). Real Brains and Artificial Intelligence. Daedalus, 117(1), 143-173.

Rich, E. (1985). Artificial Intelligence and the Humanities. Computers and the Humanities, 19(2), 117-122. Retrieved from http://www.jstor.org/stable/30204398

Schaub, M. (2017, April 04). Is the future award-winning novelist a writing robot? Los Angeles. Times.

Turing, A. M. (1950). I.—Computing Machinery And Intelligence. Mind, LIX(236), 433-460.doi:10.1093/mind/lix.236.433

Warwick, K., & Shah, H. (2015, June 29). Can Machines Think? A Report on Turing Test Experiments at the Royal Society. Journal of Experimental & Theoretical Artificial Intelligence, 28(6), 989-1007. doi:10.1080/0952813X.2015.1055826

Warwick, K., Shah, H., & Comput, H. C. (2015, December 28). Passing the Turing Test Does Not Mean the End of Humanity. Cognitive Computation, 8(3), 409-419. doi:10.1007/s1255 9-015-9372-6

White, P. (2007). 'Tradition and the Individual Talent' Revisited. The Review of English Studies, 58(235), new series, 364-392. Retrieved from http://www.jstor.org/stable/ 4501601

文本复制检测报告单(简洁)

№:ADBD2018R_20180425123324424298661155 检测时间：2018-04-25 12:33:24

检测文献： Human Writing Skills in A.I Written Poetry by Analyzing a Poem Written by Google AI.

作者： SON,JAE HOON

检测范围：
中国学术期刊网络出版总库
中国博士学位论文全文数据库/中国优秀硕士学位论文全文数据库
中国重要会议论文全文数据库
中国重要报纸全文数据库
中国专利全文数据库
图书资源
优先出版文献库
大学生论文联合比对库
互联网资源(包含贴吧等论坛资源)
英文数据库(涵盖期刊、博硕、会议的英文数据以及德国Springer、英国Taylor&Francis 期刊数据库等)
港澳台学术文献库
互联网文档资源
CNKI大成编客-原创作品库
个人比对库

时间范围： 1900-01-01至2018-04-25

指导教师： Andrew Cardoso

检测结果

总文字复制比： 0.3% **跨语言检测结果：** 0%

去除引用文献复制比：**0.3%** 去除本人已发表文献复制比：**0.3%**

单篇最大文字复制比：**0.3%**(|--)

去除本人已发表文献复制比：

重复字数：	[224]	总字数：	[74128]	单篇最大重复字数：	[224]
总段落数：	[5]	前部重合字数：[0]		疑似段落最大重合字数：	[224]
疑似段落数：[1]		后部重合字数：[224]		疑似段落最小重合字数：	[224]

指　标： ☐疑似剽窃观点 ☐疑似剽窃文字表述 ☐疑自我剽窃 ☐一稿多投 ☐过度引用 ☐疑似整体剽窃 ☐重复发表

表格：**0** 公式：**没有公式** 疑似文字的图片：0 脚注与尾注：**0**

0%	(0)	Human Writing Skills in A.I Written Poetry by Analyzing a Poem Written by Google AI._第1部分（总16054字）
0%	(0)	Human Writing Skills in A.I Written Poetry by Analyzing a Poem Written by Google AI._第2部分（总16399字）
0%	(0)	Human Writing Skills in A.I Written Poetry by Analyzing a Poem Written by Google AI._第3部分（总16474字）
0%	(0)	Human Writing Skills in A.I Written Poetry by Analyzing a Poem Written by Google AI._第4部分（总16535字）

2.6% (224) Human Writing Skills in A.I Written Poetry by Analyzing a Poem Written by Google AI._第5部分（总8666字）

（注释： 无问题部分 文字复制比部分 引用部分）

1. Human Writing Skills in A.I Written Poetry by Analyzing a Poem Written by Google AI._第1部分 总字数：**16054**

相似文献列表 **文字复制比**：0%(0) **疑似剽窃观点（0）**

2. Human Writing Skills in A.I Written Poetry by Analyzing a Poem Written by Google AI._第2部分 总字数：**16399**

相似文献列表 **文字复制比**：0%(0) **疑似剽窃观点（0）**

3. Human Writing Skills in A.I Written Poetry by Analyzing a Poem Written by Google AI._第3部分 总字数：**16474**

相似文献列表 **文字复制比**：0%(0) **疑似剽窃观点（0）**

4. Human Writing Skills in A.I Written Poetry by Analyzing a Poem Written by Google AI._第4部分 总字数：**16535**

相似文献列表 **文字复制比**：0%(0) **疑似剽窃观点（0）**

5. Human Writing Skills in A.I Written Poetry by Analyzing a Poem Written by Google AI._第5部分 总字数：**8666**

相似文献列表 **文字复制比**：2.6%(224) **疑似剽窃观点（0）**

1	| - - -《网络（http://www.anyv.net/）》- 2017	2.6%（224） 是否引证：否

说明： 1.指标是由系统根据《学术论文不端行为的界定标准》自动生成的

2.本报告单仅对您所选择比对资源范围内检测结果负责

3.Email：amlc@cnki.net http://e.weibo.com/u/3194559873 http://t.qq.com/CNKI_kycx

http://check.cnki.net/