

# The Literary Quality of ChatGPT's Productions

Donald Gillies

Department of Science and Technology Studies, University College London, London WC1E 6BT, UK.

Email: donald.gillies@ucl.ac.uk

Manuscript received November 16, 2024; accepted January 17, 2025; published April 16, 2025.

doi: 10.18178/JAAI.2025.3.2.122-131

---

**Abstract:** ChatGPT is capable of writing extended passages in natural language in response to prompts. This raises the important problem of trying to evaluate the literary quality of these productions. Can this quality be improved or are there inherent limitations? This paper attempts a preliminary investigation of this problem by focussing on just one kind of literary production, namely poetry. The question is related to Turing's classic 1950 paper in which he introduced the Artificial Intelligence (AI) project. Turing considers the example of whether computers can fall in love or enjoy strawberries and cream. Accordingly, the present paper gives a poem about love and strawberries and cream written by a human (an Elizabethan poet) and compares it to a poem on the same subject by ChatGPT. It is then argued that ChatGPT produces some poems which, because they are lacking in authenticity, are skilful imitations of good poems rather than good poems. This is contrasted with AI generated proofs in mathematics which establish the truth of theorems as well, if not better, than the proofs of human mathematicians.

**Keywords:** Turing, human poems, computer poems, human mathematics, computer mathematics

---

## 1. Introduction

ChatGPT released on 30 November 2022 is capable of writing extended passages in natural language in response to prompts. It can reply to questions, compose essays, answer emails, draft reports and job applications, etc. This is a really remarkable achievement for the AI community, but it is important that while admiring this success, we give careful consideration to the limitations of ChatGPT and similar systems. It is known that ChatGPT can make mistakes. Sometimes it will produce a pure invention when asked for some factual statement. Then, as will be shown later in the paper, ChatGPT can make logical and mathematical mistakes. When designing applications for ChatGPT and similar systems, these limitations have to be kept in mind or else unfortunate consequences will follow.

In this paper, I want to raise a general question about the limitations of ChatGPT. ChatGPT produces literary documents but what is the literary quality of these documents? Is ChatGPT's output of as good a quality as that of any human writer? Might it even be of better quality than that of all human writers? Or are there some defects in the literary quality of documents produced by ChatGPT which can only be overcome by a skilled human writer?

As with many general problems, it is better to begin to tackle these questions by considering a particular example of ChatGPT's literary production, and a convenient one to choose is ChatGPT's production of poetry. ChatGPT can indeed write poems, but how good are these poems? Are they inevitably of poor quality? Or might ChatGPT be able to generate poems of a very high quality, perhaps as good as, if not better than, those

of human poets? These are the questions I will now consider.

Questions of this sort are not new but have arisen already in other areas of AI. About twenty-five years before ChatGPT, AI systems were devised which were capable of producing proofs of theorems which human mathematicians had been unable to prove. This raised the question of the quality of AI mathematics in comparison to human mathematics. It is interesting to compare the situation with mathematics to that with poetry, and this will be done later in the paper.

The evaluation of the quality of ChatGPT's poetry turns out to be closely connected with a much earlier theme in the history of AI, namely the imitation game introduced by Alan Turing in his classic paper of 1950 on Computing Machinery and Intelligence, a paper which marks the beginning of the AI project. We will show the connection between this classic work and ChatGPT's production of poetry in the next section.

## **2. Some Relevant Passages from Turing's 1950 Paper**

Turing starts his paper by introducing [1] the famous imitation game. I will give a description of this by a paraphrase which stays close to Turing's original words. The game is played by a computer (A), a human (B) and an interrogator (C). The interrogator cannot see A or B and can communicate with them only by teleprinter. The interrogator knows A and B only by the labels X and Y, and his or her problem is to decide whether X is A and Y is B or X is B and Y is A, or, in other words, to decide which one is the computer and which is the human. To make this decision, the interrogator can ask any question and X and Y have to answer. B, the human, tries to help the interrogator to make the right decision, but the computer (A) has been programmed to imitate a human as closely as possible and hence to deceive the interrogator into making the wrong decision. This game was an invention of Turing's, but it has now entered everyday reality. If you ring up a large corporation with some query, a friendly voice may answer: "how can I help you today?", but it may not be clear whether behind this voice is a human being or a computer.

The connection between Turing's imitation game and ChatGPT's production of poems is easily established. After introducing the game, Turing gives a list of specimen questions which might be asked in the course of the game. The first of these is [1].

Q: Please write me a sonnet on the subject of the Forth Bridge.

Turing imagines that the answer, whether from a human or a computer might be:

A: Count me out on this one. I never could write poetry.

However, ChatGPT could respond by producing the required sonnet. Of course, ChatGPT does not play the original imitation game, for it would write its sonnet far faster than a human, and so be easily detectable as the computer. However, it is imitating human poets and so playing an imitation game in a more general sense.

Turing thought [1] that in fifty years' time it would be possible to programme computers to play the imitation game very convincingly. The release of ChatGPT seventy-two years later, though as pointed out it does not play Turing's imitation game in the strict sense, is still a remarkable vindication of his ideas. However, in 1950, Turing's views were considered by many to be rather far-fetched, and in his paper, he considers and tries to answer objections. One such objection was that computers will never be able [1] to "fall in love, enjoy strawberries and cream". As regards strawberries and cream, Turing comments [1]:

The inability to enjoy strawberries and cream may have struck the reader as frivolous. Possibly a machine might be made to enjoy this delicious dish, but any attempt to make one do so would be idiotic.

If a computer is playing the imitation game, however, there is no reason why it should not say that it enjoys strawberries and cream. The machine does not have to tell the truth when playing the game. Indeed, Turing gives an example where the machine has to lie in order to pretend to be a human. Computers can of course

do complicated arithmetical calculations much faster and more accurately than human beings. So, Turing says [1]:

The machine (programmed for playing the game) would not attempt to give the right answers to the arithmetic problems. It would deliberately introduce mistakes in a manner calculated to confuse the interrogator.

This is the point at which we can return to the questions about poetry posed by ChatGPT. My idea is to take Turing's example of falling in love and enjoying strawberries and cream, and to compare two poems on this subject, one produced by a human being and the other produced by ChatGPT. This will be done in the next section.

### **3. Poems about Love and Strawberries and Cream**

My first task was to find a poem about love and strawberries and cream produced by a human being. A search through *The Oxford Book of English Verse* [2] brought one to light quite easily. It was written by George Peele, an Elizabethan poet. Here it is [2]:

Poem 1 by a human in late 16<sup>th</sup> century

#### *A Summer Song*

When as the rye reach to the chin,  
And chopcherry, chopcherry ripe within,  
Strawberries swimming in the cream,  
And school-boys playing in the stream;  
Then O, then O, then O my true love said,  
Till that time come again,  
She could not live a maid.

The next task was to get ChatGPT to write a poem on similar subject. After some experimentation, I instructed ChatGPT to write a sonnet in Elizabethan English about summer which expressed enjoyment of strawberries and cream, and which used the expression 'my true love' once but only once. The result is given below.

Poem 2 by ChatGPT in December 2023

#### *A Sonnet for Summer*

In summer's realm, where blooms the fairest rose,  
The meadows dance, a vibrant tapestry grows.  
Amidst the green, a feast for senses rare,  
Sweet strawberries entwined with cream declare.

As golden sun in heaven's vast array,  
My true love's gaze, a radiant display.  
Her eyes, the stars that grace the midnight sky,  
In summer's glow, a love that ne'er shall die.

Each berry's taste, a kiss upon the tongue,  
With cream, a melody by bards unsung.

Yet sweeter still, the words my love imparts,  
In whispers soft, she captures longing hearts.

So in this sonnet, summer's tale is spun,  
With strawberries, cream, and love beneath the sun.

Before discussing the relative merits of these two poems, it will be helpful to return to Turing's 1950 paper. Turing claims that his imitation game, or at least something quite like it, is often used by examiners to test whether students have understood the subject. When so used it goes by the name of a *viva voce*. Turing writes [1]:

The game ... is frequently used in practice under the name of *viva voce* to discover whether someone really understands something or has "learned it parrot fashion." Let us listen in to a part of such a *viva voce*:

Interrogator: In the first line of your sonnet which reads "Shall I compare thee to a summer's day," would not "a spring day" do as well or better?

Witness: It wouldn't scan.

Interrogator: How about "a winter's day." That would scan all right.

Witness: Yes, but nobody wants to be compared to a winter's day.

The *viva voce* continues in an amusing fashion with mentions of Christmas and Mr. Pickwick, and then Turing asks what we would say "if the sonnet-writing machine was able to answer like this in the *viva voce*." By speaking of a "sonnet-writing machine", Turing is being remarkably prophetic. Now we have such a machine. So, following his suggestion, I conducted a *viva voce* with ChatGPT about its sonnet, styling myself as 'Critic'. Here is the result.

Critic: Is 'strawberries entwined with cream' a good expression in a sonnet?

ChatGPT: Yes, 'strawberries entwined with cream' could be a beautiful and evocative expression in a sonnet, especially if you are aiming to create a sensory and romantic atmosphere. Sonnets often rely on vivid and poetic language to convey emotions, and the phrase captures a sense of indulgence and richness that can enhance the overall imagery.

Critic: Isn't it rather banal and commonplace in a sonnet to compare 'my true love's gaze' to the 'golden sun'?

ChatGPT: The effectiveness of comparing 'my true love's gaze' to the 'golden sun' in a sonnet depends on how it's executed and the overall context of the poem. The comparison itself might be considered a bit conventional or commonplace, as the metaphor of comparing one's beloved to the sun is a classic and often-used poetic device. However, the success of the line will ultimately depend on the freshness of your language, the depth of your imagery, and how well it contributes to the emotional impact of the poem.

Despite ChatGPT's able defence of its sonnet, George Peele's poem seems to me, and would, I suspect, seem to most people, much superior. 'Strawberries swimming in the cream' is a much better expression than the somewhat dubious 'strawberries entwined with cream'. The comparisons of 'my true love's gaze' to the 'golden sun' and 'her eyes' to 'the stars' are indeed banal and commonplace, whereas the line 'And school-

boys playing in the stream' is both realistic and evocative.

Although George Peele's poem is undoubtedly better than ChatGPT's, some words could be said in praise of ChatGPT's effort. In the late 16<sup>th</sup> century in England, it was quite standard for members of the gentry to write sonnets to their beloveds. Naturally only a few of these, such as the sonnets by Sir Philip Sydney, are still read today. ChatGPT might, however, have been trained by being fed with all the surviving Elizabethan sonnets, and its own production is probably as good as many of the sonnets produced by gentry of average to low poetic abilities. In that sense, it would seem to be a reasonably commendable performance. Moreover, an enthusiast for AI might say that, if AI programs are already so good, it perhaps requires only a few more decades for them to be able to produce sonnets which not only reach the standard of the top human performers such as Sir Philip Sydney, but perhaps even go beyond it. Such a claim would, in my view, be mistaken because there is a deeper reason why at least some of the poems of ChatGPT cannot be good. I will explain this reason in the next section.

#### 4. The Criterion of Authenticity

There are some poems in which the poet expresses his or her feelings and/or describes some of his or her experiences. For such a poem to be good, I would claim that it is essential that the feelings and experience conveyed be genuine and not fake or pretended. This is what justifies introducing the criterion of authenticity. 'Authenticity' has been much discussed by philosophers, but in nearly all these discussions it is applied to human beings. However, authenticity has also been used by literary critics as applied to poems. There is even an online course on how to write authentically offered by [thepoetrysalon.com](http://thepoetrysalon.com). In the literary context, the concept of authenticity has not been analysed very precisely. I propose the following definition:

If a poem expresses the feelings and/or experience of the poet, then the feelings and/or experience conveyed must be authentic for the poem to be good.

There are a number of points to note about this definition. First of all, the criterion of authenticity does not apply to all poems, but only to those in which the feelings and experiences expressed are attributed to the poet himself or herself. Secondly, for this class of poems, authenticity is a *necessary* condition for the poem to be good. Authenticity is never a *sufficient* condition for a poem to be good, since it is perfectly possible for a poem to express genuine feelings but in an appalling manner.

The criterion of authenticity does apply to Poem 2 by ChatGPT. Perhaps the best line of the poem is: "Each berry's taste, a kiss upon the tongue". This may be adapted (or perhaps even lifted, since ChatGPT is not very concerned with questions of plagiarism) from some English poet of the Elizabethan age. That poet really would have had experience of the taste of strawberries and the feeling of kisses. In his or her mouth, the line could have been authentic. But ChatGPT has never experienced the taste of strawberries or the delights of a kiss. Thus, as produced by ChatGPT, the line cannot be authentic and ChatGPT's poem cannot be a good one for this reason.

This is an argument based on a single poem, but it does extend to the majority of the poems to be found in a standard collection, such as *The Oxford Book of English Verse* [2], or similar collections in other languages. Love poems may well be the largest category of poems, but no love poem by ChatGPT can be authentic since ChatGPT never fell in love. Poems often express grief and other emotions. For example, Milton in his *Lycidas* laments the loss of his friend who was drowned in a sea voyage. Milton writes [2]:

But O the heavy change, now thou art gone,  
Now thou art gone, and never must return!

However, ChatGPT could not write an authentic poem lamenting the death of a friend since it never had a friend. Another common poetic genre consists in poems describing the beauties of the natural world, such as Wordsworth's [2]:

When all at once I saw a crowd,  
A host of golden daffodils;  
Beside the lake, beneath the trees,  
Fluttering and dancing in the breeze.

However, ChatGPT could not write an authentic poem about daffodils, since it has no sense organs and has never seen a daffodil.

Some poetic genres might seem more suitable for ChatGPT, namely those which are based on travellers' tales and fantasy. However, even here there are difficulties. Consider Shelley's *Ozymandias*, which begins [3]: "I met a traveller from an antique land, Who said: ...". Of course this is literally false. We are not supposed to think that Shelley met a traveller who recited the rest of the poem to him. Yet the first line of the poem would hardly be authentic unless Shelley had met some travellers from an antique land. Almost certainly Shelley did meet travellers from Egypt, but ChatGPT has never met any travellers from Egypt or elsewhere. Coleridge's *Kubla Khan* is possibly, from among the great English poems, the one most based on fantasy. Yet this poem contains the lines [2]:

A damsel with a dulcimer  
In a vision once I saw:

ChatGPT could never authentically write such lines, since it has never experienced opium induced visions.

Against this analysis of Coleridge's lines, it might be objected that perhaps the poet never literally had a vision of a damsel with a dulcimer but liked the expression because of its alliteration.<sup>1</sup> This is true, but Coleridge certainly did have opium induced visions and probably one of them was sufficiently similar to a damsel with a dulcimer to make this line authentic in a way it could not be if produced by ChatGPT.

It seems then that the 'authenticity argument', as it might be called, applies to quite a large number of poems, but it would be wrong to argue that it applies to the whole of poetry. An exception to the authenticity argument can be found by a consideration of the contrast between Shakespeare's sonnets and his plays. In the sonnets, Shakespeare does claim to be expressing his own feelings. The first two lines of sonnet 97 are [4]:

How like a winter hath my absence been  
From thee, the pleasure of the fleeting year!

We do expect these lines to be an authentic expression of Shakespeare's feelings during his absence from his beloved.

On the other hand, when Hamlet says [5]: "To be, or not to be: that is the question:", we do not have to assume that Shakespeare himself ever entertained suicidal thoughts. He may or may not have done, but here he is attributing suicidal thoughts to an imaginary character, not to himself. So, the criterion of authenticity does not apply. This obviously can be generalised to all poetic drama.

The same point applies to many narrative poems. Chaucer's *Canterbury Tales*, leaving aside the *Prologue*, consist of a series of stories in verse told by imaginary characters invented by Chaucer. So, for example, *The*

---

<sup>1</sup> I owe this objection to Terry Eagleton.



*Knight's Tale* [6] is supposed to be told by an imaginary knight, whom Chaucer has created to be typical of the knights who existed at that time. If feelings and emotions are expressed in the tale, which Chaucer never had himself, we cannot accuse the poem of being inauthentic.

We are suggesting then using authenticity to evaluate poems of a particular type, namely those in which the poem claims to be expressing the emotions and experiences of the author rather than of some imaginary character. It is worth therefore saying a little about the nature of this evaluative criterion. Here it is important to distinguish between the intrinsic and relational properties of a text. Intrinsic properties are those such as the rhythm of the lines, the presence or absence of alliteration, rhyme, stylistic properties and so on. However, a relational property such as authenticity differs because it concerns not just the text but the author as well. We have to consider whether the text is a genuine expression of the author's feelings and/or experience, as it claims to be. Does this mean that we have to know details of the author's life in order to judge the merit of a poem? Not necessarily. Consider again Milton's lines from *Lycidas* [2]:

But O the heavy change, now thou art gone,  
Now thou art gone, and never must return!

We know from Milton's biography that Milton was indeed suffering grief at the loss of a friend drowned in a sea voyage, but, even if this fact had been lost historically, it would still be possible to judge that these lines, and indeed the rest of the poem, expressed an authentic emotion of the poet.

The historical information about Milton is in fact missing for Shakespeare. In the first two lines of Shakespeare's sonnet 97 [4], we do not know who Shakespeare's beloved was, or what exactly Shakespeare's relation with his beloved was. Scholars have proposed various theories, but the historical record does not decide the matter. Still, most people would judge that these lines are a genuine expression of the poet's feelings. It is hard to imagine anyone writing such lines if they had been glad that their beloved was away and had enjoyed their absence. Of course, the ascription of authenticity based entirely on the text is always somewhat conjectural, but the ability of recognising authenticity in the utterances of our fellow human beings is one which is acquired in everyday life. It is often important to know whether a story told by someone about themselves is the truth or a lie, and whether the expression of emotion is genuine or hypocrisy.

As long as poems were written only by human beings it was not too difficult, on the basis of the text, to make judgements of authenticity. However, the development of ChatGPT has perhaps changed this situation. Since ChatGPT can produce such convincing imitations of any type of poem, maybe we now will have to know at least that a poem was produced by a human rather than an AI before we can judge its quality.

One way of better understanding the nature of the poems produced by AI is to compare AI's production of poems with its production in other areas. This I will do in the next section by comparing AI's production of poems with its production of proofs in mathematics.

## **5. A Comparison with AI Generated Proofs in Mathematics**

To explain the nature of AI generated proofs in mathematics, we cannot rely on ChatGPT. In his very interesting 2023 article on the strengths and limitations of ChatGPT, Akoudas shows that ChatGPT is very weak when it comes to logic and mathematics. For example, Akoudas shows [7] that ChatGPT accepts the following syllogism

All Italians are happy  
Alberto is happy  
Therefore, Alberto is Italian

However, it is obvious even to students who are just beginning in logic that this is a fallacy. Even more remarkably Akoudas gives a case where [7] ChatGPT argues that “1 is greater than 1”. This mistake arises when ChatGPT is dealing with numbers written in decimal notation. If we take 11 for example, then the 1 on the left is greater than the 1 on the right, since the 1 on the left stands for 10. ChatGPT seems to have concluded from this that the 1 on the right of a particular number might be greater than the 1 on the right of another number.

This failing of ChatGPT makes it seem more human than before. After all, everyone is familiar with humans who can display great skill, insight and intelligence in literary matters while being hopeless at formal logic and mathematics. As we remarked above, Turing thought that a computer playing the imitation game would have deliberately to make mistakes in mathematics in order to pretend to be a human being. However, ChatGPT, though a computer programme, quite spontaneously makes mistakes in logic and mathematics. Akoudas [7] considers the possibility that the problem might be overcome by “outsourcing subtasks to “plugins” that can perform calculation and reasoning much more effectively than LLMs”, but he issues the following warning:

However, this is exceedingly challenging in the general case and there is good reason to doubt that such an approach can be effective.

Other AI systems exist, invented long before ChatGPT, which are very good at handling logic and mathematics. We will now consider one of these.

In 1996 a group of AI researchers (Larry Wos and his colleagues at Argonne, Illinois) achieved a spectacular success in automated theorem proving. They managed to produce a computer proof of the Robbins conjecture. This had been proposed in 1933 but, human mathematicians had failed to prove it for 63 years, despite the best efforts of the famous logician Tarski and his students. It is worth noting that AI’s success in theorem proving was more than twenty-five years before its success in writing poems. This is perhaps not surprising. Mathematical theorems and proofs can be translated into the kind of formal language used in mathematical logic. Computers were able to handle such formal languages long before they could deal with the natural languages of poetry and everyday life. It took considerable increases in computing power, as well as new approaches and techniques, before computers became able to handle natural language satisfactorily.

In the previous section, I argued that some poems generated by AI cannot be good poems because they lack authenticity. They are skilful imitations of good poems rather than good poems. So, what about AI generated proofs in mathematics? Are they satisfactory? Or is there something dubious about them which might make us prefer proofs produced by humans? Here a striking difference shows itself between the case of poetry and the case of proofs in mathematics, for a computer-generated proof definitely establishes the correctness of a mathematical theorem. Indeed, it establishes it with greater certainty than a human generated proof. Even the most brilliant human mathematicians make mistakes in their proofs, while the computer checking each step rapidly and mechanically is much less likely to go wrong.

This means that AI generated proofs can be used in a variety of circumstances. Suppose for example that the proof is of a theorem which states that from a set of axioms (SA say), a result (R say) follows according to standard (or classical) logic. SA may be the axioms of a theory which has been proposed by a scientist (Mr C), while R may be something whose truth can be ascertained by observation and experiment. Using the theorem established by the AI generated proof, Mr C can test his theory by finding out whether R is true or false. In the light of the theorem, if R is false, his theory is refuted, while if R is true, it is confirmed.

From examples such as these, it looks as if AI generated proofs in mathematics are just as good, if not better, than proofs produced by human mathematicians. However, the situation is more nuanced than this, as can



be seen from an interesting discussion of the computer proof of the Robbins conjecture by the philosopher of mathematics, David Corfield. This discussion is in Section 2 of Corfield's 2003 book, a chapter significantly entitled: "Communicating with automated theorem provers".

Corfield makes the following insightful observation about automated theorem proving [8]:

I believe it is no accident that the most successful approach to date has been one that has deliberately avoided closely imitating human problem solving techniques. Computers have their own inhuman strengths which need to be harnessed.

Corfield has in mind here Larry Wos's approach which led to the proof of the Robbins conjecture. This involved translating the mathematics into the clausal form of logic – a form which is particularly easy for computers to handle, and using methods of inference and search which are suited to computers rather than humans.

Corfield points out that there is a problem with computer proofs generated in this manner. As he says [8]:

What mathematicians are largely looking for from each other's proofs are new concepts, techniques and interpretations. Computer proofs certainly give information concerning the truth of a result, but very little beyond this.

The problem is that, although each step in a computer proof can easily be checked and shown to be valid, the proof can still remain incomprehensible to a human. The proofs of a human mathematician are always based on an idea which is first grasped before the details are filled in. There may be no obvious idea behind a computer proof.

EQP is the name of a program which proved the Robbins conjecture, and from EQP we certainly learn that the Robbins conjecture is correct, but can human mathematicians learn anything else from EQP? From proofs produced by human mathematicians, other human mathematicians can, as Corfield says, learn new concepts, techniques and interpretations. Now EQP produced a new proof. Can human mathematicians learn anything extra from it, in the way they learn new things from the proofs of human mathematicians?

The unintelligible nature of the computer proof seems to block the possibility of learning new concepts, techniques and interpretations from it. Corfield does however indicate a possible way out of this difficulty which consists in translating the computer proof into a language which is more accessible to humans. This approach was taken for the computer proof of the Robbins conjecture by Louis Kauffman. He translated the proof into a special symbolic notation involving geometrical elements, which had been introduced by Peirce and developed by Spencer-Brown. In this notation the proof made more human sense than it had done before.

Here then we have an entirely new pattern of development in mathematics. Automated theorem provers produce a humanly incomprehensible computer proof of a theorem. Human mathematicians then try to make sense of the computer proof by translating it into a humanly more accessible language.

This naturally raises the question of whether human poets can make a similar use of poems generated by ChatGPT. ChatGPT can produce poems with quite striking lines such as 'Each berry's taste, a kiss upon the tongue'. Such lines cannot be authentic when generated by an AI programme, but might they become authentic when adapted by a human poet? This is a tricky question which I will leave for further discussion.

## **6. Conclusion**

The conclusion is that, even with many technological improvements, ChatGPT will not be able to produce a good poem, which claims to express the feelings and/or experience of the poet. It will certainly be able to produce skilful imitations of such poems, but that is a different matter. This contrasts with AI generated

proofs in mathematics which give results which are useful for most scientific applications. The arguments of this paper have been based on English poems and on ChatGPT. It would be interesting to see how well they apply to poetry in other languages such as Chinese, and also to other LLMs such as DeepSeek.

### Conflicts of Interest

The author declares no conflict of interest.

### Acknowledgment

I would like particularly to thank Terry Eagleton who was kind enough to read an earlier version of this paper and made comments, one of which is described in a footnote, and which were very useful in developing and improving the paper. I also received helpful comments from some anonymous referees.

### References

- [1] Turing, A. (1950). Computing machinery and intelligence. *Mind*, 236(October), 433–460.
- [2] Quiller-Couch, A. (1939). *The Oxford Book of English Verse 1250–1918* (New ed.). Clarendon Press.
- [3] Palgrave, F. T. (1954). *The Golden Treasury of the Best Songs and Lyrical Poems in the English Language* (C. Day Lewis, Rev.). Collins.
- [4] Corfield, D. (2003). *Towards a Philosophy of real Mathematics*. Cambridge University Press.
- [5] Shakespeare, W. (1966). *Histories and Poems* (Everyman's Library). Dent. (Original work published c. 1591–1613)
- [6] Chaucer, G. (1949). *The Canterbury Tales* (The World's Classics). Oxford University Press. (Original work published c. 1386–1389)
- [7] Arkoudas, K. (2023). ChatGPT is no stochastic parrot. But it also claims that 1 is greater than 1. *Philosophy & Technology*, 36(3), Article 54.
- [8] Corfield, D. (2003). *Towards a Philosophy of Real Mathematics*. Cambridge University Press.

Copyright © 2025 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).