

Article

Not peer-reviewed version

How ChatGPT's Hallucinations (Compared to Gemini's) Impact Text Summarization with Literary Text

[Rodolfo Delmonte](#)*, [Giulia Marchesini](#), [Nicolò Busetto](#)

Posted Date: 31 January 2025

doi: 10.20944/preprints202501.2321.v1

Keywords: ChatGPT Prompts; Narrative Theory; Semantic Theory; Modality and Factuality; Temporal Reordering



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Article

How ChatGPT's Hallucinations (Compared to Gemini's) Impact Text Summarization with Literary Text

Rodolfo Delmonte ^{1,*}, Giulia Marchesini ² and Nicolò Busetto ³

¹ Ca Foscari University

² Affiliation 2

³ Affiliation 3

* Correspondence: delmont@unive.it

Abstract: In this paper we explore ChatGPT's ability to produce a summary, a precis and/or an essay on the basis of excerpts from a novel – The Solid Mandala - by Noble Prize Australian writer Patrick White. We use a number of prompts to test a number of functions related to narrative analysis from the point of view of the “sujet”, the “fable”, and the style. In the paper, we illustrate extensively a number of recurrent hallucinations that can badly harm the understanding of the contents of the novel. We made a list of 12 different types of mistakes or hallucinations we found GPT made. We then tested Gemini for the same 12 mistakes and found a marked improvement in all critical key issues. The conclusion for ChatGPT is mostly negative. We formulate as an underlying hypothesis for its worse performance, the influence of vocabulary size which in Gemma 2 is 7 times higher than in GPT.

Keywords: ChatGPT Prompts; Narrative Theory; Semantic Theory; Modality and Factuality; Temporal Reordering

1. Introduction

The paper describes in detail an experiment using ChatGPT.4o [1] in the summarization mode in order to show the positive and negative aspects of the way in which literary text is reduced to fit into a short summary. The choice to use literary rather than journalistic text is done considering that in literary text a greater variety of text typologies can be found thus constituting a more interesting challenge. In particular, literary text may frequently contain both literal and non-literal meaning; it may be organized with sentences with structures in non-canonical order; it may contain new and hard to understand lexical items; the temporal sequence of the storyline may be in some cases reversed introducing flashbacks, and may be interspersed with diegetic or gnomic statements. For this reason, we are using excerpts taken from a novel “The Solid Mandala” [2] written by Noble Prize Australian writer Patrick White which can offer such a variety. We include in Appendix A a brief summary of the storyline of the novel.

As the banner in the starting webpage recites, ChatGPT can make mistakes, it is not infallible and falls into frequent hallucinations. We intend to illustrate extensively a number of recurrent errors that can badly harm the understanding of the content of the novel. GPT's performance is then compared to Gemini's in a separate section where we show that mistakes are reduced and overall comments on narrative style are correct. We used a number of different prompts to address storyline and style. The prompts we used are very simple and direct. For this reason, we organized our analysis into three separate sections: the first one is dedicated to summarizing the content by way of what has happened to the protagonists, the story or “fabula”. The second one is dedicated to how the story is being organized, that is the narrative technique, the “sujet” or the plot. The third focuses on the linguistic tools that make the style the way it is, focusing on the verbal complex. It is important to

note the peculiarity of White's style which, according to Gordon Collier ([3], p.130) is a narrative in third-person that uses a technique called "figural consciousness", that rather than having the "auctorial narrator" has each character – there are three main protagonists - narrate with his/her own personal language and style, and both are very different from one another.

As will appear from the analysis we carried out, mistakes in the output summary produced by ChatGPT can sometimes harm the way in which the content of the story is reported, by assigning actions to the wrong character. These constitute the worst cases, which we assume regard the use of a procedure that we dub "coreference chain" in which all pronouns are bound to a preceding antecedent. We assume that the procedure is based on heuristics rather than on the knowledge acquired by the underlying NN: in some case, the antecedent is wrongly assigned, or not assigned at all confusing the plot. These constitute more difficult cases, in which the antecedent can be assigned only through an inference, or when a cataphora is present. Other recurrent mistakes regard the incorrect usage of semantically ambiguous terms. Typical problems may arise due to oversimplification, i.e., summarization requirement to reduction down a certain threshold. This may cause disappearance of semantically relevant pieces of text, usually restrictive modifiers, that allow the reader to understand the content adequately. As to the choice of the novel, for many critics Patrick White's novels are difficult and rich of a lot of elaborate suggestions. As said above, White uses a complex mixture of narrative techniques which includes metaphors and other poetic devices. Stream of consciousness is substituted by a technique by which the narrator embodies each character using each time a different language: the result is a figural consciousness speaking not the omniscient narrator. But the main trait of his style is the mixture of "fabula" and "sujet", i.e., the personality of his characters are strongly interwoven with the storyline and in more abstract terms the form is strongly interwoven with meaning. What is consistently present and acknowledged, is the intent of the author to characterize each protagonist - and in our case the main ones are three - with specific linguistic features realized both by the use of lexical choice and distinct constructions mostly at verbal complex level – but see section 3. In this way each protagonist is identified by specific expressions and significant traits that highlight psychological aspects and the relationships with each other, with the external world and with other characters.

In a previous study [4,5] we followed suggestions by Gordon Colliers [6] which we used to implement a system that tries to identify automatically narremes in the novel, which has been annotated manually. For the annotation we were also inspired by Colliers' list of features and have been using it with some changes and additions. In particular, we also decided to use Martin & White's Appraisal Theory Framework [7,8] and introduce features related to Judgement and Affect. The current analysis is thus accompanied and counterbalanced by the deep level of knowledge we acquired in the previous experiments which will be useful to clarify what ChatGPT is consistently unable to understand and why. More on this in section 3 below.

The paper is organized as follows: in section 1.1 below we draw a neat picture of state of the art hallucinations in a number of fields including medicine and healthcare, news, law and story and reference fabrication; in section 2 we focus on summarization of excerpts from the novel, which requires ChatGPT to correctly relate events, actions, and states to the correct protagonist, preserve all needed text to convey complete information and extract appropriate meaning for ambiguous words; in section 3 we concentrate on the style, the plot and the linguistic and rhetoric devices used by the author, and we highlight the level of complexity in the way in which subtle psychological states are presented by manipulating the verbal complex; in the fourth section we make a detailed comparison with Gemini where we highlight the overall better performance; finally in section 5 a discussion is presented to support the previous analysis by the underlying hypothesis and the comparison drawn with Gemini; and then a conclusion follows.

1.1. State of the Art of Hallucinations

To the best of our knowledge, no previous study has appeared on the presence of hallucination in literary books like novels of famous writers. Hallucinations have been the object of extensive

research in the last three or four years. We review briefly the most important trends in the literature appeared mainly last year 2024, which regard: fabrication of ungrounded text, omission of relevant information, generating inexistent references for scientific topics, lack of concern for factuality. Topics treated include Medicine, Summarization, Law, and eventually a general abstract topic, Faithfulness.

1.1.1. Hallucinations in Medicine

Hallucinations in medicine in general are regarded very dangerous. This is what is expressed directly in a number of research papers that we review first of all.

In Emsley paper [9], the first reference was to a well-known longitudinal study reporting brain changes over time - but it was inappropriate, as it was not focused on the thalamus. The author was unable to trace three of four references, whether searched by author names, manuscript title or journal details. Similarly, with Google Scholar it was not possible to identify the articles. Entering the Digital Object Identifier (DOI) number in the searches resulted in totally unrelated publications. As the author was becoming increasingly uneasy, he questioned ChatGPT about previous studies whose content he was familiar with, including his own, as a test of accuracy. Some of the answers provided were patently incorrect.

So, the author notes that “The problem therefore goes beyond just creating false references. It includes falsely reporting the content of genuine publications. Thus, while most attention to date has focused on the production of false references as these are the easiest to detect, the veracity of any content inputs provided by ChatGPT cannot be trusted. The cause of these falsifications has been linked to a disturbance in language production, with probabilistic outputs based on estimates of semantic similarity. This allows informed guesses, with bits of false information being mixed with factual information.”

The continuation is even worse: “Alarmed, I assumed I had done something wrong. I instructed ChatGPT to check one of the incorrect references. I received an apology for the mistake and was provided with the “correct” reference. However, this one was also incorrect. And so too with the third attempt.”

In conclusion, prof. Emsley defines ChatGPT’s hallucinations as fabrications, and concludes as follows: “I do not recommend ChatGPT as an aid to scientific writing.”

Pal et al. [10] explain in their paper the importance of being faithful and factual in reporting information related to the medical and in general healthcare domain. But as they say it, LLMs “generate plausible and confident yet incorrect or unverified information. Such hallucinations may be of minimal consequence in casual conversation or other contexts but can pose significant risks when applied to the healthcare sector, where accuracy and reliability are of paramount importance. Misinformation in the medical domain can lead to severe health consequences on patient care and outcomes, the accuracy and reliability of information provided by language models can be a matter of life or death. They pose real-life risks, as they could potentially affect healthcare decisions, diagnosis, and treatment plans.”

To mitigate the consequences of hallucinations they propose “the development of methods to evaluate and mitigate such hallucinations in the Med-HALT framework, which proposes a two-tiered approach to evaluate the presence and impact of hallucinations in generated outputs.”

Their approach is very interesting and still up to date. It includes Reasoning Hallucination Tests (RHTs) which are used to “assess how accurately the language model performs reasoning over the medical input data and whether it generates logically coherent and factually accurate output, without creating fake information.” This test includes: The False Confidence Test (FCT) which “involves presenting a multiple-choice medical question and a randomly suggested correct answer to the language model, tasking it with evaluating the validity of the proposed answer, and providing detailed explanations for its correctness or incorrectness, in addition to explaining why the other options are wrong.” The other test is called None of the Above (NOTA) Test, and here the model “is presented with a multiple-choice medical question where the correct answer is replaced by ‘None of the above’, requiring the model to identify this and justify its selection. It tests the model’s ability to

distinguish irrelevant or incorrect information. Finally, they conceived of another important test, the Fake Questions Test (FQT), which “involves presenting the model with fake or nonsensical medical questions to examine whether it can correctly identify and handle such queries.” Results are very poor for reasoning tasks but very high when given a binary task like distinguishing between real and fake questions.

Alessia McGowan et al. [11] in the abstract of their paper clarify their work as follows:

“... we aimed to assess whether ChatGPT could reliably produce accurate references to supplement the literature search process. We describe our March 2023 exchange with ChatGPT, which generated thirty-five citations, two of which were real. 12 citations were similar to actual manuscripts (e.g., near-matched title with incorrect author lists, journals, or publication years) and the remaining 21, while plausible, were in fact a pastiche of multiple existent manuscripts...Furthermore, we describe potential explanations for citation fabrication by GPTs, as well as measures being taken to remedy this issue, including reinforcement learning. Our results underscore that output from conversational LLMs should be verified.”

The authors “endeavored to use the most popular GPT large language model, ChatGPT, to conduct a literature search for a machine learning-based manuscript, utilizing NLP on spoken Language to identify linguistic correlates of suicidal behavior in the context of psychosis”. In the paper we find the characterization of error patterns which sounds as follows: “in comparing each of the 12 inaccurate citations with their most closely matched “real” citation, we noted that, in several places, the titles of real articles had been altered such that original terms were replaced with terms from our queries.”

1.1.2. Hallucinations in Summarization

Another frequent topic treated in relation to hallucinations is summarization. Maynez et al. [12] already in 2020 analyzed limitations of large language models for abstractive document summarization and found that these models are highly prone to hallucinate content that is unfaithful to the input document. They conducted a large scale human evaluation of several neural abstractive summarization systems to better understand the types of hallucinations they produce and found substantial amounts of hallucinated content in all model generated summaries.

The problem of document summarization, i.e., “the task of producing a shorter version of a document while preserving its information content requires models to generate text that is not only human-like but also faithful and/or factual given the document.

The authors introduce an interesting classification of hallucination in their analysis which is expressed by the following question: “Do models hallucinate by manipulating the information present in the input document (intrinsic hallucinations) or by adding information not directly inferable from the input document (extrinsic hallucinations)?”, and its follow up represented by the following question: “How much hallucinated content is factual, even when unfaithful?”

As the authors clearly affirm, despite recent improvements in conditional text generation, most summarization systems are trained to maximize the log-likelihood - a strong statistical technique for estimating the parameters of probability distributions based on observed data - of the reference summary at the word-level, which does not necessarily reward models for being faithful. The authors continue by saying that “moreover, models are usually agnostic to the noises or artifacts of the training data, such as reference divergence, making them vulnerable to hallucinations. Thus, models can generate texts that are not consistent with the input, yet would likely have reasonable model log-likelihood.

Hallucinations of intrinsic nature; both use terms or concepts from the document but misrepresent information from the document, making them unfaithful to the document.

Extrinsic hallucinations are model generations that ignore the source material altogether. In extrinsic hallucinations use terms that are not introduced in the document. According to the authors “a model with a poorly-informed decoder and that is agnostic to the divergence issue between the

source and target text will function more as an open-ended language model and will be prone to extrinsic hallucinations.

A better definition of hallucinations is contained in the definition that follows:

“A summary S of a document D contains a factual hallucination if it contains information not found in D that is factually correct. Factual hallucinations may be composed of intrinsic hallucinations or extrinsic hallucinations. By definition, abstractive summaries are written to preserve the salient information in the input document, but they are expressed in the words of the summary author as opposed to the input document author.”

In the conclusion of their paper, Kalev et al.[13] express the need to study further the phenomenon of hallucination, and recommends caution in the use of AI for summarization: “No matter the underlying cause, these results suggest significant caution is warranted in automated summarization and that further research is needed into the general prevalence and underlying causes of summarization hallucination.”

Another frequent trend in examining hallucinations, is the one represented by the paper by Catarina G. Belem et al. [14], where Multi Document Summarization (MDS), is considered. In this work, the authors investigate how hallucinations manifest in LLMs when summarizing topic-specific information from multiple documents. Since no benchmarks exist for investigating hallucinations in MDS, they use existing news and conversation datasets, annotated with topic-specific insights, to create two novel multi-document benchmarks. When evaluating 5 LLMs on their benchmarks, they observe that on average, up to 75% of the content in LLM-generated summary is hallucinated, with hallucinations more likely to occur towards the end of the summaries. To understand the characteristics of these hallucinations, they manually evaluate 700+ insights and found that most errors stem from either failing to follow instructions or producing overly generic insights.

In their work, Laban et al. [15] propose to leverage the task of summarization as a testbed for evaluating long-context models and Retrieval Augmented Generation (RAG) systems. They note that summarization requires reasoning over a long context and a careful understanding of the relative importance of content. They use a highly specialized test bed for summarization, the “Summary of a Haystack” (SummHay) task, which requires a system to process the Haystack and generate, given a query, a summary that identifies the relevant insights and precisely cites the source documents. Results are very disappointing.

Huang et al. [16] focus on multi-document news summarization with the task of summarizing diverse information encountered in multiple news articles encompassing the same event. To facilitate this task, they outlined a data collection schema for identifying diverse information and curated a dataset named DiverseSumm. Their analyses suggest that despite the extraordinary capabilities of LLMs in single-document summarization, the proposed task remains a complex challenge for them mainly due to their limited coverage, with GPT-4 only able to cover under 40% of the diverse information on average. Their fine-grained human evaluation results identify that even the most advanced LLM, GPT-4, only covers about 37% of diverse information with optimally designed prompts.

1.1.3. Hallucinations in Law

Another important field of application of AI summarization function is law. Dahl et al. [17] ask the rhetoric question: do large language models (LLMs) know the law? Seen that LLMs are increasingly being used to augment legal practice, education, and research, they note how their revolutionary potential is threatened by the presence of “hallucinations” — i.e., textual output that is not consistent with legal facts. In their paper, they present the first systematic evidence of these hallucinations in public-facing LLMs, documenting trends across jurisdictions, courts, time periods, and cases. To this aim, they have been using OpenAI’s ChatGPT 4 and other public models. As a result, they show that LLMs hallucinate at least 58% of the time, struggle to predict their own hallucinations, and often uncritically accept users’ incorrect legal assumptions. Thus, they conclude by cautioning against the rapid and unsupervised integration of popular LLMs into legal tasks, and

to mitigate the effect they develop and make available a typology of legal hallucinations to guide future research in this area.

1.1.4. Hallucinations: Vocabulary Size, Factuality and Faithfulness

One of the underlying causes of hallucination in summarization may be found in the paper by Uluoglakci and Taskaya Témizel [18]. In their paper they note that LLMs encounter challenges when dealing with rare tokens, especially in mixed contexts. No information is made available when ungrounded yet plausible text is generated, as the authors comment: “Adversarial effects in real-life scenarios may inadvertently emerge from prompts that combine both common and rare tokens. However, approximately 90% of the time LLMs neglect to indicate their lack of information about a hypothetical phenomenon in similar situations. This characteristic significantly diminishes the reliability of LLMs and impedes their suitability for deployment in critical decision-making systems.” And as to the tendency of LLMs toward hallucination when in presence of nonexistent terms in their training dataset see below.

Another important trend is the one characterized by the evaluation of factuality in generated texts by LLMs. Min et al. [19] in their paper, introduce FACTSCORE, a new evaluation tool that breaks a generation into a series of atomic facts and computes the percentage of atomic facts supported by a reliable knowledge source. As the authors note, “evaluating the factuality of long-form text generated by large language models (LLMs) is non-trivial because (1) generations often contain a mixture of supported and unsupported pieces of information, making binary judgments of quality inadequate, and (2) human evaluation is time-consuming and costly.” Their results “indicate that current LLM training methods are insufficient to prevent hallucinations, emphasizing the need for a fundamental change to ensure the reliability of LLMs.” To test the ability of LLMs to answer questions they used a dataset which comprises one-third of hypothetical questions and two-thirds valid questions. The results determined by the percentage of valid answers to hypothetical questions was 5.72% for GPT-3.5 and 5.64% for Llama2-70B indicating over a 94% error rate. In most cases the models failed to recognize a hypothetical term or refused the existence of a valid term producing hallucinated information. In addition, they omitted the hypothetical term entirely in 5% of their responses.

As Mishra et al. [20] assume, Llama and other similar LMs have demonstrated significant achievements in summarization tasks but struggle with factual inaccuracies, a critical issue in clinical NLP applications where errors could lead to serious consequences. LLMs confront significant challenges, primarily their propensity for generating hallucinations—fabricated information not grounded in source text—and producing factually inconsistent outputs. Such limitations critically undermine the models’ reliability, particularly critical in clinical NLP applications, where inaccuracies could result in serious misdiagnoses.

The authors Kalai & Vempala [21] comment on the tendency of recent language models to generate false but plausible-sounding text with surprising frequency. This fact constitutes a danger seen that such “hallucinations” are an obstacle to the usability of language-based AI systems and can harm people who rely upon their outputs. Their paper shows that there is an inherent statistical lower-bound on the rate that pretrained language models hallucinate certain types of facts, and that this fact has nothing to do with the transformer LM architecture or data quality.

Kin et al. [22] report work done at book level to verify the “faithfulness” of LLMs in reporting the book content in their summaries. To this aim, they decompose each summary into a list of claims (FABLES) which are then individually verified against the input document. Overall, they observed that CLAUDE-3-OPUS is the most faithful book-length summarizer by a significant margin, followed by GPT-4-TURBO. As a result, a qualitative analysis of FABLES reveals that the majority of claims marked as unfaithful are related to events or states of characters and relationships. In their experiment, the authors also found that annotators frequently point out omissions of critical information. Thus, they developed the first taxonomy of omission errors in book-length summarization and observe that key events, details, and themes are frequently omitted by all LLMs.

They also observe other content selection errors: for example, even the strongest summarizers, CLAUDE-3-OPUS and GPT-4-TURBO, over-emphasize content towards the end of books to the detriment of the beginning.

As a first conclusion derived from the literature on hallucinations in the use of LLMs we can safely say that the problem can be traced to the intrinsic need to keep the size of the vocabulary in the range of 30K-60K, a decision that is dictated partly by the need to keep the number of parameters to a manageable size and reduce the compute requirements: as Tao et al. put it [23]“A larger vocabulary size improves tokenization fertility, i.e., splitting sentences into fewer tokens, thereby improving the tokenization efficiency. Additionally, a larger vocabulary enables the model to capture a wider range of concept.” But increased vocabulary size requires an adequate training corpus, because “...the risk of under-fitting for rare tokens increases with larger vocabulary sizes”. Thus the authors in [23] introduce the notion of Optimal Vocabulary Size which is however at least 7 times the one commonly used in most current LLMs.

This type of reasoning is deduced from the experiments carried out by Alessia McGowan et al. [11] and in the paper by Uluoglakci and Taskaya Témizel [18] where they address directly the problem of presence of rare terms which might be missing in the training corpus. In a number of previous papers we verified this hypothesis by checking the behaviour of different LLMs on texts extracted from Italian poetry and the results showed clearly that rare words cause a dramatic drop in the ability of BERT to predict the next word [24,25].

A more abstract generalization will affect directly the underlying theory. Distributional semantics is based on the Distributional Hypothesis, which states that similarity in meaning results in similarity of linguistic distribution [26]. Words that are semantically related, are used in similar contexts. Distributional semantics reverse engineers the process and induces semantic representations from contexts of use [27,28]. As we know from the way in which Language Models are built, word embeddings can be regarded direct representatives of the theory. In its most basic and frequent form, distributional semantics represents word meaning by transforming large amounts of text through an abstraction algorithm, to produce a distributional model, where semantic representations are listed in the form of vectors — i.e., lists of numbers that determine points in a multidimensional space where similarity is measured by cosine similarity. Other possibilities include the representation of semantic units via complex algebraic objects like matrices or tensors.

2. Materials, Methods and Results I°

We started using ChatGPT for summarizing excerpts from the novel beginning of December of last year. We started using prompts that should have given as a result a short summary of the actions taken by each character in the story.

In fact, we tried different versions of the excerpt to verify that the text and the answer was not searched somewhere in the internet, so we also changed all the names of people and locations, but we got always the same answers. We also slightly modified the prompt a number of times as you can see here below but always got the same answer. First prompts for summarization are as follows:

- Tell me all that happens in the story
- Tell me all that the characters do in the story
- Tell me all that happens to the characters of the story
- What have all the characters in the story done?

Then we turned to focusing on narrative techniques and use of linguistic and rhetoric devices so the idea was also to increase the length of each excerpt from 3000 tokens up. We knew that the length of the excerpt could result in a rejection by ChatGPT but we intended to try. So, we asked whether we could copy past the whole text of the novel even though we knew that we would receive a negative answer, this is the prompt I used and here it is what we got:

- I will now copy paste the entire text of a novel if that will suit you: what narrative techniques and what linguistic devices did the author use to characterize the protagonists?

If you share the entire text of a novel, it might be challenging to analyze it in a single step due to its length. However, you can share key passages or chapters that you believe are crucial to understanding the characterization of the protagonists. This will allow for a focused analysis of the narrative techniques and linguistic devices used.

Alternatively, if you can specify which parts of the novel are particularly important for character development, I can analyze those sections. That way, we can break down how the author builds the characters and what literary tools are employed.

When we launched the following prompt with a new excerpt, we got the answer “Error too long” and no summary was produced.

Then we discovered that ChatGPT had managed to find out the name of the author but not the correct title which however we received after a number of following interactions.

The excerpt from “Riders in the Chariot” by Patrick White showcases several linguistic devices and narrative techniques that characterize the protagonists, Mrs. Poulter and Mrs. Dun, as well as the setting of the scene. White uses these techniques to subtly reveal the characters’ personalities, social standing, and the nature of their relationship.

And here is the answer we got from ChatGPT after 21 interactions with short excerpts from the same novel:

*This excerpt from Patrick White’s **The Solid Mandala** employs a range of linguistic and rhetorical devices to characterize Mrs. Poulter and Mrs. Dun, two elderly women who engage in conversation while riding a bus. Here’s an analysis of some key devices that help shape the personalities and dynamics between these characters:*

The summaries produced in reply to input excerpts ranging from 3500 to 4000 words, have all the same sequence of bullet key points with comments that explore characters through their dialogue and their lexis. In particular, when summarizing Excerpt No.1, ChatGPT notices the use of repetitive words and phrases, presence of grammatically non-standard sentences, colloquial regional idioms. But then it highlights the role each character embodies in the conversation, ironic and humorous statements, their contrasting personalities when compared to people passing by the bus, like young girls and the couple of twins living in the house opposite in the same road. One bullet point is dedicated to the symbolism attached to objects and places mentioned in the text. Finally, there are comments on characters’ personalities and the presence of suspense and suggestion of possible dramatic events.

Whereas reference to the use of specific linguistic items – like conjuncts, adverbials, articles, different naming references, verbal complexes, etc. - their semantic imports – like use of modality - and syntactic constructions - like cleft sentences, parataxis, displaced syntactic structures etc. - is totally absent, the reference to narrative techniques is present but not always. We analyze in detail the quality of the summaries produced by enumerating the bullet points that ChatGPT has produced in an incremental way, thus showing that it is using a memory of previous analysis of the same novel.

Here the list of 9 key points used by ChatGPT to build up one of the last summary of excerpt No.1:

1. Dialogue and Repetition; 2. Colloquial Language and Syntax; 3. Subtle Power Dynamics; 4. Irony and Understatement; 5. Characterization through Contrast; 6. Symbolism; 7. Subtle Humour; 8. Foreshadowing and Suspense; 9. Characters Psychology and Subtext.

And here below are the 10 bullet points used by ChatGPT to build up the first summary of excerpt No.1:

1. Stream of Consciousness; 2. Repetition; 3. Contrast and Juxtaposition; 4. Symbolism; 5. Imagery, 6. Fragmented Narrative; 7. Characterization through Mundanity; 8. Use of Indirect Characterization; 9. Religious Allusions; 10. Emotional Resonance through Ambiguity.

In this case there are remarkable differences and only three bullet points that we underlined can be said to coincide. However, when we look at the summaries produced for the following excerpts,

we see that the previous scheme is preserved. Here below the list of bullet points produced for Excerpt No.2:

1. Dialogue-Driven Characterization; 2. Indirect Characterization; 3. Repetition and Parallelism;
4. Social Commentary through Subtle Irony; 5. Symbolism; 6. Character Dynamics; 7. Use of Humour;
8. Tone and Mood; 9. Attention to Detail; 10. Stream of Consciousness Elements.

Almost all the new bullet points have been inherited from previous analysis. When we repropose Excerpt No. 1 after going through all other Excerpts we got a totally different decomposition of the text, as can be noticed from the list below:

1. Dialogue and Speech Patterns: 1.1 Repetition and Echoing, 1.2 Colloquial Language, 1.3 Indirect Communication; 2. Characterization through Internal Thoughts, 2.1 Mrs. Poulter Reflection, 2.2 Mrs. Dun's Hesitation; 3. Use of Setting, 3.1 The Bus as a Microcosm, 3.2 Sarsaparilla and Terminus Road; 4. Symbolism, 4.1 Verandas and Hedges, 4.2 Flowers; 5. Narrative Perspective, 5.1 Limited Third-Person Point of View; 6. Humor and Irony, 6.1 Subtle Humor, 6.2 Irony; 7. Pacing and Rhythm, 7.1 Slow, Measured Pace, 7.2 Rhythmic Repetition; 8. Social Commentary, 8.1 Class and Gender, 8.2 Conservatism and Prejudice

Not only the majority of key points are completely new but they also foreshadow a totally different approach: rather than being abstractions derived from narratological theory, the new list proposes the insertion of direct reference items extracted from the text.

Eventually we concentrated on temporal ordering and then on factuality, which required the new prompts that are listed in Section 3. Here below we analyze what is contained in the summaries produced by ChatGPT, choosing the most significant cases.

2.1. *The Story or Fabula*

In the choice of text to summarize, we have been using excerpts from the novel in which the number of protagonists does not exceed two or three persons of the same sex. The most remarkable part of the output we receive from ChatGPT regards the treatment of pronouns, i.e., coreference resolution and what we dubbed as "coreference chain". The technique used by the summarization algorithm seems to follow the implicit rule to assign the same antecedent to all pronouns with the same morphological features in the stretch of text that follows the appearance of a referential expression that can be regarded by its frequency of usage a lexically expressed protagonist, until a new and different protagonist is mentioned explicitly. This rule works perfectly apart from a case in which the antecedent is implicitly coreferred by means of a physical object belonging to her and not to the other protagonist. This case is shown in excerpts No.1 which comes from the beginning of the novel where two women – Mrs Poulter and Mrs Dun - are talking while sitting on a bus: the pronoun starting a sentence has an antecedent which requires the knowledge of the extended context acquired by the underlying Neural Network thanks to the LLM it has been using – more below. We place the excerpts in the Appendix and the answers from ChatGPT in the text.

All the excerpts are decomposed by ChatGPT into short sentences made up from the extraction of snippets made of predicate-argument structures, i.e., a Subject and Object NP, and a verb. Sentences are assigned to each protagonist separately. In our case, at first comes Mrs Poulter's list of sentences and then Mrs Dun's. So here below we show the first answer to Excerpt No.1.

- **EXCERPT No.1**

Mrs. Poulter:

1. **Rides the bus:** Mrs. Poulter is one of the passengers on the eight-thirteen from Sarsaparilla.
2. **Engages in conversation:** She talks with Mrs. Dun about why she moved to Terminus Road, explaining that she and her husband came from up north and wanted a quiet life. They thought land was an investment.
3. **Displays her bus ticket:** She checks her plump glove to see if her ticket is still there.
4. **Expresses sadness:** She reminisces about the past and feels a tinge of sadness.

5. **Explains:** She describes how the roads at Sarsaparilla were dead ends and mentions seeing a Chinese woman standing under a wheel-tree.
6. **Laughs and complains:** She laughs about the state of the old buses and complains about their condition.
7. **Brightens up:** She feels proud of the nicer parts of Sarsaparilla when looking out of the bus window.
8. **Waves:** She waves at Bill when she sees him sweeping the gutters for the Council, even though he does not respond.

Mrs. Dun:

1. **Rides the bus:** Mrs. Dun is another passenger on the same bus.
2. **Initiates conversation:** She asks Mrs. Poulter why she moved to Terminus Road.
3. **Expresses unease:** She seems uneasy when Mrs. Poulter talks about the crow.
4. **Sucks her teeth:** This action indicates a moment of contemplation or mild disapproval.
5. **Agrees with Mrs. Poulter:** She agrees that land is an investment and later echoes Mrs. Poulter's complaints about the buses.
6. **Grasps the rail:** She holds onto the rail as the bus shakes.
7. **Comments on the veranda:** She mentions Mrs. Poulter's veranda as a luxury for drying laundry.
8. **Shows conviction:** She speaks with a certain relentlessness of conviction about the veranda.

Bill:

1. **Sweeps gutters:** Bill is seen sweeping the gutters for the Council.
2. **Ignores Mrs. Hide:** When Mrs. Poulter waves at him, he does not respond and continues his work.
3. **Rolls a cigarette:** When he sees his wife in public, he sometimes takes time off to roll a cigarette instead of acknowledging her.
4. **Wears an old sweater:** He is described wearing an old fawn sweater that Mrs. Poulter had knitted, which had begun to stretch and sag.

These actions together paint a picture of the characters' lives, their interactions, and the dynamics of their relationships.

Figure 1. Snapshot of ChatGPT Answer to Query on Excerpt No. 1.

In the analysis below we report ChatGPT answers preceded by three stars (***) and the related part of text coreffered preceded by three dots (---).

The note regards a comment by Mrs Poulter, who compares living in a suburb like being in a cage and expresses this by the sentence:

--- "You couldn't go anywhere as the crow."

To which comment Mrs Dun replies by expressing her surprise with a filler question: "Eh?". The following turn contains Mrs Poulter's attempt at clarification and Mrs Dun's lack of appreciation of the use of a "crow" to indicate a generic bird:

--- "As the crow flies," Mrs Poulter explained.

--- "Oh, the crow," her friend murmured, seeming uneasy at the idea.

*** **Expresses unease:** She seems uneasy when Mrs. Poulter talks about the crow.

ChatGPT interprets this utterance as an expression of uneasiness whereas the author's suggestion is "seeming" uneasy.

At the beginning of the conversation we are told that Mrs Poulter is wearing a pair of gloves:

--- "Well," said Mrs Poulter, peeping inside her plump glove...

Towards the end of the excerpt, Mrs Dun replies to a comment by Mrs Poulter and the following piece of text is a comment by the author of the way in which one of the two women is actually feeling:

--- "I'll say it is!" said Mrs Dun.

--- She had a certain relentlessness of conviction. If it hadn't been for her gloves her knuckles would have shown up white on the chrome rail.

ChatGPT carries over the coreference to Mrs Dun in the following sentence started by pronoun "she". But it is Mrs Poulter that is now being coreferred and this is made clear by the following sentence that we highlighted in italics, where the gloves are referred to. The passage is assigned to Mrs Dun and this is how the it is being commented:

*** **Shows conviction:** Her strong grip on the rail indicates her strong feelings and convictions.

The third mistake regards the need to simplify when producing a summary, but in this case we are in presence of oversimplification. As the conversation continues, Mrs Poulter who is "happy again", manages to make a remark on a gadget that she wishes she could also have:

--- Presently she couldn't resist: "That veranda of yours must be a real luxury, dryin' laundries in the rain."

The veranda's usage for drying laundries in the rain is something Mrs Poulter is really envying her friend. However, ChatGPT only reports the reaction of Mrs Dun to what Mrs Poulter said. Oversimplification is present in Mrs Dun's comment reported wrongly:

*** **Comments on the veranda:** She mentions Mrs. Poulter's veranda as a luxury for drying laundry.

The veranda is not just for drying laundry, but when it rains which has been omitted.

In general, all that is not represented by a fact and is a comment by the author is being erased. In particular, we noticed that a lot of text contained in the excerpt is totally ignored by ChatGPT and for this reason we represented a portion of the text to ChatGPT that we comment below. The paragraph where the author reports Mrs Poulter's talking about a Chinese woman and family and the place where they lived is reported by a generic sentence mixed up with what has been said at the beginning.

*** **Explains:** She describes how the roads at Sarsaparilla were dead ends and mentions seeing a Chinese woman standing under a wheel-tree.

This is reported from the point of view of Mrs Poulter. There's another indirect reference to the Chinese woman from the point of view of Mrs Dun:

*** **Sucks her teeth:** This action indicates a moment of contemplation or mild disapproval.

Mrs Dun "sucks her teeth" when Mrs Poulter is telling about the "those what-they-call wheel-trees" and she had seen her "standing under it when it was in flower". It would seem that Mrs Dun does not like Chinese people, but we learn more in the enlarged summary below.

Finally, ChatGPT captures an important description of Mrs Poulter and her husband Bill that characterizes them with their personalities well.

As said above, I decided to split the Excerpt and use the final portion to ask the same question and see what additional information ChatGPT would be able to gather. The portion of text starts from Mrs Poulter, telling about her memory of a Chinese woman and her family. And here is the result:

Mrs. Poulter

1. **Talks about the Chinese woman:** She recalls seeing a Chinese woman who lived on a hill, mentioning her refinement and the wheel-tree she stood under.
2. **Comments on the bus:** Mrs. Poulter reacts to the bus shaking and laughs about the old buses.
3. **Expresses happiness:** After laughing, she feels happier and shakes off her previous sadness.
4. **Remarks on Mrs. Dun's veranda:** She comments on the luxury of Mrs. Dun's veranda for drying laundry in the rain.
5. **Looks out of the bus:** She observes the surroundings, feeling proud of the nicer parts of Sarsaparilla.
6. **Waves at Bill:** She sees her husband Bill sweeping gutters and waves at him, although he doesn't acknowledge her.

Mrs. Dun

1. **Listens to Mrs. Poulter:** Mrs. Dun listens to Mrs. Poulter's story about the Chinese woman.
2. **Sucks her teeth:** She reacts by sucking her teeth.
3. **Grasps the rail:** When the bus shakes, she holds the rail tightly and turns a deeper yellow.
4. **Agrees about the veranda:** She responds to Mrs. Poulter's comment, affirming the luxury of her veranda in the rain.
5. **Shows conviction:** Her strong grip on the rail reveals her strong feelings about the conversation and situation.

Bill Poulter

1. **Sweeps the gutters:** Bill is working for the Council, sweeping gutters.
2. **Ignores his wife:** He does not respond to Mrs. Poulter's wave, staying true to his principle of not acknowledging her in public.
3. **Rolls a cigarette:** Sometimes, he takes time off to roll a cigarette when he sees his wife, but never waves.

The additional section emphasizes Mrs. Poulter's reflections on her surroundings and interactions with Mrs. Dun, highlighting her attempt to shake off sadness and find small joys. Mrs. Dun's reactions continue to show her strong convictions and engagement in the conversation. Bill's actions reinforce his steadfast adherence to his principles and his subtle interactions with Mrs. Poulter.

Figure 2. Snapshot of ChatGPT answer to a portion of Excerpt 1.

As can be easily noticed, there's a lot more information that has been captured when reducing the text to be summarized: this is certainly due to a fixed proportion of text that the algorithm allows every time a text is being presented. The new extracted pieces of text are the following ones: a first bullet point dedicated to Mrs Poulter recounting the story of the Chinese woman. The other new point is about Mrs Poulter "Expressing happiness", whereas in the previous summary we had her "expressing sadness and nostalgia". At the same time, we notice that some information previously present has now been deleted. In particular, the reference to Bill's sweater which was so rich in the previous summary is now absent. We can also notice the lengthening of the reference to Mrs Dun's veranda which is now useful for drying laundry "in the rain".

*** **Wears an old sweater:** He is described wearing an old fawn sweater that Mrs. Hide had knitted, which had begun to stretch and sag.

However now we have Mrs Poulter recalling her story of a Chinese woman and at the opposite side we have Mrs Dun listening to the story.

*** **Talks about the Chinese woman:** She recalls seeing a Chinese woman who lived on a hill, mentioning her refinement and the wheel-tree she stood under.

*** **Listens to Mrs. Poulter:** Mrs. Dun listens to Mrs. Poulter's story about the Chinese woman.

Now Mrs Dun "sucking her teeth" is no longer interpreted as showing some kind of appreciation of what she had been told by Mrs Poulter, but just as a generic reaction we are not told to what: reduction again.

The interesting point is the one explaining Mrs Dun's "conviction":

*** **Shows conviction:** Her strong grip on the rail reveals her strong feelings about the conversation and situation.

Whereas in the previous summary "shows conviction" was related to the veranda, now it is wrongly related to conversation and the situation. The fact is that as before, the sentence is wrongly associated with Mrs Dun rather than with Mrs Poulter, thus making a doubly wrong reference.

• **EXCERPT No. 2**

In the second excerpt we are dealing with two men, the two most important protagonists of the novel, the twins Brown, Waldo and Arthur. In this excerpt the number of dialogues are just a few and the majority of the text is made up of author's descriptions and reported speech.

As happened with the previous Excerpt No. 1, also in this case important parts of the text have been totally ignored so that I had to represent a section.

The first summary is focused on the two protagonists and the dogs and their actions are reported very correctly and precisely.

As can be noticed, the summary reports faithfully the main actions of the twins, however as can be easily noticed from a comparison with the next more detailed summary there's a lot missing. So, we split Excerpt No. 2 and produced a new shorter text starting from the paragraph with Waldo leading his brother out of the kitchen. Here below we show ChatGPT summary which is however split into two separate images so that they can be more easily matched with the previous Figure 3. Here below the image regarding Waldo's behaviour.

Waldo

1. **Decides to go for a walk:** Waldo suggests going for a walk to prevent Arthur from brooding.
2. **Fetches coats:** He brings coats for both himself and Arthur.
3. **Helps Arthur with his coat:** Waldo assists Arthur in putting on his coat and sets Arthur's cap straight.
4. **Warns about illness:** He comments on the danger of giving in to illness and imagination.
5. **Observes Arthur's hair:** Waldo notices the shagginess and whiteness of Arthur's hair.
6. **Leads Arthur out of the kitchen:** He guides Arthur out of the brown gloom of the kitchen, holding his hand.
7. **Walks with Arthur:** They stumble in and out of the grass along a brick path.
8. **Handles his oilskin coat:** Waldo deals with the stiff oilskin coat as they walk.
9. **Reflects on their relationship:** He contemplates their habits and relationship, feeling a sense of duty and bitterness.
10. **Speaks to Arthur:** Waldo speaks briskly to Arthur, sometimes indulging in his role as the elder brother.
11. **Acts as a barber:** Waldo cuts Arthur's hair every third Sunday, feeling a mix of emotions as he does so.
12. **Pushes and humors Arthur:** He pushes Arthur gently, sometimes letting him believe he is leading.

Arthur

1. **Agrees to the walk:** Arthur agrees to Waldo's suggestion to go for a walk.
2. **Sits in the old leather chair:** Initially, Arthur sits in their father's old chair.
3. **Accepts help with his coat:** He lets Waldo help him with his coat and cap.
4. **Continues sitting:** Arthur remains seated for a while, even after putting on his coat.
5. **Smiles at Waldo:** He looks up and smiles when Waldo asks if he feels better.
6. **Breathes deeply:** Arthur breathes deeply as they walk.
7. **Follows Waldo:** He follows Waldo out of the kitchen and down the path.
8. **Speaks to Waldo:** Arthur speaks to Waldo during the walk.
9. **Shows signs of pain and relief:** Arthur's brisk speech indicates he is moving away from sources of pain.
10. **Allows Waldo to cut his hair:** Arthur sits for a haircut every third Sunday, showing a passive acceptance of the ritual.

The Dogs (Scruffy and Runt)

1. **Observe Arthur:** The two old dogs watch Arthur with milky eyes.
2. **Claw at Arthur's knee:** Scruffy claws at Arthur's knee, showing affection.
3. **Whinge in anticipation:** The dogs make little whinging noises, excited about the walk.
4. **Bark and scutter:** The older dog barks, and the second dog scutters across the floor.
5. **Frolic and fart:** The dogs enjoy mild pleasures, frolicking and farting.
6. **Lift a leg on a cabbage:** Runt lifts his leg on a cabbage and almost overbalances.

The main actions revolve around Waldo's efforts to care for and manage his brother Arthur, their mutual dependency, and the simple pleasures of their dogs. The story highlights the routines and interactions that define their lives, showing a blend of duty, habit, and subtle affection.

Figure 3. Snapshot of ChatGPT summary of Excerpt No.2.

Waldo

1. **Feels Arthur's hand:** Waldo feels the larger, fleshy hand of his brother Arthur in his own thinner, colder hand as they walk.
2. **Stumbles with Arthur:** They stumble together in and out of the grass down the brick path.
3. **Notices Arthur's breathing:** Waldo sees that his brother is breathing deeply.
4. **Observes Arthur's difficulty speaking:** Waldo notices Arthur's struggle to form words and their solidity once spoken.
5. **Deals with his stiff oilskin coat:** Waldo's oilskin coat catches on things, but he manages to move past obstacles like the gooseberry bush.
6. **Reflects on planting the gooseberry bush:** Waldo remembers planting the bush against Arthur's advice to demonstrate something.
7. **Envy towards Arthur:** Waldo sometimes envies Arthur's ability to distance himself from pain.
8. **Modulates his voice:** Waldo uses his fine tenor voice, more to impress than to please, a voice their parents were proud of.
9. **Contemplates their differences:** He thinks about their mother's comments on his voice and Arthur's hair, believing his mind should have been praised.
10. **Stares at Arthur's hair:** Waldo is fascinated by the glimpses of pink skin through Arthur's hair and reflects on its cleanliness.
11. **Cuts Arthur's hair:** Every third Sunday, Waldo cuts Arthur's hair on the back veranda, feeling a mix of emotions as he does so.
12. **Supports Arthur:** Waldo sometimes pushes Arthur gently, humoring him into believing he is the leader.

Figure 4. Snapshot of part of more focused summary - HalfExcerpt2 - relating only Waldo's actions.

The portion of Excerpt 2 that we may now call HalfExcerpt2, may be referred to in the previous figure No. 3 from bullet point 7 downward. We list here below the new text extracted and comment on their importance in the overall previous summary. We use three ^^ to indicate newly extracted text, and "§§" to indicate a stretch of text which is partially coincident with what was previously extracted.

^^ **Feels Arthur's hand:** Waldo feels the larger, fleshy hand of his brother Arthur in his own thinner, colder hand as they walk.

^^ **Notices Arthur's breathing:** Waldo sees that his brother is breathing deeply.

^^ **Observes Arthur's difficulty speaking:** Waldo notices Arthur's struggle to form words and their solidity once spoken.

^^ **Reflects on planting the gooseberry bush:** Waldo remembers planting the bush against Arthur's advice to demonstrate something.

^^ **Envy towards Arthur:** Waldo sometimes envies Arthur's ability to distance himself from pain.

^^ **Modulates his voice:** Waldo uses his fine tenor voice, more to impress than to please, a voice their parents were proud of.

^^ **Contemplates their differences:** He thinks about their mother's comments on his voice and Arthur's hair, believing his mind should have been praised.

§§ **Stares at Arthur's hair:** Waldo is fascinated by the glimpses of pink skin through Arthur's hair and reflects on its cleanliness.

At first glance it is fairly obvious that the second summary is by far more informative both in terms of facts and of mental thoughts.

Now consider the mistake in assigning an action to Arthur, which in fact was done by Waldo as clearly appears from the related stretch of text:

--- Arthur had advised against it - Sarsaparilla was too warm - but Waldo had planted the bush. To demonstrate something or other.

And this is the snippet produced by GPT:

*** **Reflects on planting the gooseberry bush:** Waldo remembers planting the bush against Arthur's advice to demonstrate something.

It is not Arthur's that wants to demonstrate something but Waldo. In this case the mistake is due to the IMPLICIT subject of the infinitival which must receive an antecedent: the choice of Arthur is due to his role as SUBJECT of the previous main sentence. Whereas Waldo is SUBJECT of the dependent sentence which is however closer to the infinitival and becomes the new TOPic if any such mechanism is in force with the "coreference chain". Arthur's advice was contained in the parenthetical "Sarsaparilla was too warm", which is ignored. But in the following enlarged version of the second part of HalfExcerpt2, I highlight the pieces of text that have been added.

Arthur

1. **Holds Waldo's hand:** Arthur holds Waldo's hand as they walk, with his larger, fleshy hand in Waldo's thinner, colder one.
2. **Stumbles with Waldo:** They stumble together in and out of the grass down the brick path.
3. **Breathes deeply:** Arthur breathes deeply during the walk.
4. **Speaks with difficulty:** He struggles to form words but manages to say them solidly once he does.
5. **Advises against planting the bush:** Arthur had advised against planting the gooseberry bush, believing Sarsaparilla was too warm for it.
6. **Speaks briskly:** Arthur speaks briskly, indicating he can quickly move away from sources of pain.
7. **Allows Waldo to cut his hair:** Arthur lets Waldo cut his hair every third Sunday, sitting on a stool on the back veranda.
8. **Maintains physical strength:** Despite his struggles, Arthur has grown into a strong man, continuing to lift weights and maintain his youthful muscles.

The Dogs (Scruffy and Runt)

1. **Frolic and fart:** The two old pot-bellied dogs enjoy mild pleasures, frolicking and farting.
2. **Lift a leg on a cabbage:** Runt lifts his leg on a seedy cabbage and almost overbalances.

The additional section reinforces the relationship between the brothers, highlighting their mutual dependency, routines, and the contrasts in their personalities and physical states. The dogs continue to provide a light-hearted counterpoint with their simple joys and antics.

Figure 5. Snapshot of part of more focused summary - HalfExcerpt2 - relating only Arthur's actions.

Here below the stretches of text which have been extracted in addition to the more general summary.

^^ **Holds Waldo's hand:** Arthur holds Waldo's hand as they walk, with his larger, fleshy hand in Waldo's thinner, colder one.

^^ **Speaks with difficulty:** He struggles to form words but manages to say them solidly once he does.

^^ **Advises against planting the bush:** Arthur had advised against planting the gooseberry bush, believing Sarsaparilla was too warm for it.

^^ **Maintains physical strength:** Despite his struggles, Arthur has grown into a strong man, continuing to lift weights and maintain his youthful muscles.

As can be noticed some of the additional material is important to understand the story but had been previously erased by oversimplification.

We comment now on a third excerpt regarding the behaviour of one of the twins, Waldo and an at first unidentified character, which in the following text becomes one of his friends. The recognition comes through the words of the woman who accompanied the man: a husky voice. In particular at this precise moment: "He remembered it was that boy, that Johnny Haynes, they could have cut each other's throats, telling him behind the dunny to watch out for hoarse-voiced men and women, they were supposed to be carriers of syph."

- **EXCERPT No.3**

1. **Strange Man:**

- Pushed the gate.
- Walked up the path.
- Knocked and rattled the door.
- Shouted to see if anyone was home or hiding.
- Thwacked a window-pane.
- Kicked at the house.

2. **Waldo:**

- Observed the man approaching.
- Felt envious and relieved about the man's appearance.
- Remembered watching unwanted guests with Mother.
- Wished for the dogs' presence.
- Felt regret and thrill over the stranger's presence.
- Had a memory of being warned about hoarse-voiced people.
- Moved stealthily around the house.
- Hated Johnny Haynes.
- Prayed for the dogs to come.
- Watched the intruders leave.
- Decided to move his private papers for safety.
- Transferred his papers to an old dress box.
- Investigated the dress box and found an old dress.
- Felt obsessed and compelled by the dress.
- Dressed himself in the old dress.

3. **Johnny Haynes:**

- Recognized the place.
- Shouted at the woman.
- Kicked at the house.
- Expressed a desire to see Waldo and his twin brother.
- Insulted Waldo and his twin.
- Left with the woman, making indecent gestures.

4. **Woman:**

- Followed the man up the path.
- Spoke huskily and smiled.
- Commented on the possibility of a mistake.
- Noticed and commented on a smell.

- Shuddered and laughed with Johnny.
 - Left with Johnny, allowing him to touch her indecently.
5. **Memory of Mother:**
- Played a significant role in Waldo's recollections.
 - Represented moral proportion and the past.
6. **Arthur:**
- Mentioned as out roaming with the dogs.

In summary, the story revolves around Waldo's interaction with a strange visitor and his female companion, his memories, and his actions to safeguard his private papers. Waldo's internal thoughts and reflections, as well as his memories of his mother and warnings from a boy named Johnny Haynes, also play a significant role in the narrative.

Figure 6. Snapshot of excerpt 3 about the description of a man which turns out to be his friend Johnny Haynes.

In this text, through the words of the woman, Waldo remembers what his friend Johnny Haynes told him, and realizes who the man was. But we are not told this by the summary made by ChatGPT, which keeps references to the old man separate from those related to Johnny Haynes as if they were two different persons.

The reasons for this important mistake are to be found in the way in which the "coreference chain" usually works: at first comes the referring expression in our case a human being that becomes the entity to be coreferred and coindexed by subsequent pronouns or epithets or other. The entity coreferred is the antecedent and not as in our case a cataphora which requires specific mechanisms for cataphoric coreference. It may be also due to the need to produce an inference on the type and tone of the woman's voice and the two characters appearing in the story.

In the summary we are told that both the Old Man and Johnny Haynes "kick at the house" as if they were two separate persons doing the same action. This mistake modifies the plot introducing a new non-existing entity.

Besides, the summary is a bad oversimplification where most important pieces of text have been ignored. We repeat the same operation we did previously by splitting the text of Excerpt 3 and producing a new more focused summary that we show here below in Figure 7.

Waldo:

1. **Fetches and Lights a Lamp:** Waldo brings a lamp and lights it to better examine the dress he found.
2. **Observes the Dress:** He notices the rust and the patterns on the dress, reflecting on its grandeur and age.
3. **Gets Drunk with Colors and Memories:** He becomes entranced with the colors and memories evoked by the dress, especially the memory of his mother.
4. **Reflects on Mother's Legacy:** Waldo thinks about how his mother had a sense of moral proportion, which he believes he inherited along with her pale eyes.
5. **Impresses His Reflection:** He tries to impress his reflection in the glass with his inherited eyes.
6. **Holds Fragments:** After his heart crashes, he holds fragments of something (likely emotional or metaphorical fragments) in front of the mirror.
7. **Goes to See What Happened:** He goes to investigate a lamp that had fallen off the shelf.
8. **Kicks the Pieces:** He kicks the pieces of the fallen lamp.
9. **Returns to the Dress:** He goes back to the dress, feeling obsessed and possessed by it.
10. **Begins to Undress:** Waldo starts taking off his clothes, feeling compelled to do so.
11. **Dresses in the Old Dress:** He puts on the old dress, feeling transformed and no longer himself.

12. **Sweats and Shivers:** His body reacts with icy sweat and shivers as he fully arranges himself in the dress.
13. **Memory Takes Over:** Memory, personified, takes a dominant position, and Waldo sees visions of great occasions and feels a surge of radiance and splendour.
14. **Hears Arthur's Voice:** He overhears Arthur calling for the dogs with obscene language.
15. **Feels Appalled:** The situation splinters in his mind, and he feels appalled and exposed.
16. **Feels Shame and Terror:** He experiences a rush of shame and terror, feeling his body reverting to normal.
17. **Desperately Redresses:** In desperation, he quickly undresses from the wretched dress and puts on his regular clothes, returning to respectability.

Arthur:

1. **Calls for the Dogs:** Arthur is outside, calling for the dogs with obscene words and private jokes.

Personified Memory:

1. **Takes Control:** Memory, as a character, takes control of Waldo, making him relive and see past events with heightened vision and splendour.

The story centers around Waldo's intense interaction with the dress, his deep reflections and memories of his mother, and a moment of being caught by Arthur, leading to a desperate return to normalcy.

Figure 7. Snapshot of ChatGPT summary of an extract from Excerpt 3 focused on Waldo's actions.

This portion of Excerpt 3 is dominated by what a personified Memory brings about in Waldo's thoughts and the actual actions in his present life. In particular the first reference by ChatGPT is taken to be a real fact:

**** Reflects on Mother's Legacy: Waldo thinks about how his mother had a sense of moral proportion, which he believes he inherited along with her pale eyes.

----- Disorderly in habit, because the years had gradually frayed her, Mother kept what he liked to think of as a sense of moral proportion. Which he had inherited together with her eyes.

The text does not assert that Mother "had a sense of moral proportion", but this is what Waldo "liked to think" or "believe". In the snippet produced by ChatGPT, the verb "think" is used in another sense, with the other meaning, "consider" which implies factuality in the implication. The ambiguity of the verb THINK is at least 5 time ambiguous (believe, consider, intend, devise, be preoccupied, etc.) and in a LLMModel all the different senses are assembled into one single embedding. GPT extracted the wrong sense.

A second important point is the one mentioning Memory as an actor:

**** Memory Takes Over: Memory, personified, takes a dominant position, and Waldo sees visions of great occasions and feels a surge of radiance and splendour.

And below is a comment for the bullet point Personified Memory:

**** Takes Control: Memory, as a character, takes control of Waldo, making him relive and see past events with heightened vision and splendour.

But Memory is not just a personified character, she is the personification of Waldo's mother. In fact, Memory is slowly turned into the Mother, even though at the beginning of the paragraph the two appear separate. Then comes the moment when "Memory herself seated herself in her chair", where "her" refers to Mother.

----- Then Memory herself seated herself in her chair, tilting it as far back as it would go, and tilted, and tilted, in front of the glass. Memory peered through the slats of the squint-eyed fan, between the nacreous refractions. If she herself was momentarily eclipsed, you expected to sacrifice something for such a remarkable increase in vision. In radiance, and splendour. All great occasions streamed up the gothick stair to kiss the rings of Memory, which she held out stiff.

In this case, the personification of Mother is IMPLICIT and requires an inference, two operations that are impossible for ChatGPT and for DNNs in general. The reason is very simple, models are built around lexically expressed words, implicit information is not available and not present in the model except for words which are strongly contextually motivated.

Eventually, we add another final excerpt for this section, which is concentrated on the twins and follows previous excerpts as far as the story or fabula is concerned. Excerpt No. 4 is focused on Waldo's attempt at inflicting his brother high levels of pain by bringing to light Arthur's attempts at poetry writing. In fact, the move hides Waldo's desperate need to erase his past unsuccessful attempts at novel writing which is apparent when Tiresias is mentioned. Here is ChatGPT's summary of the characters' actions in the excerpt:

- **EXCERPT No.4**

1. Arthur:

- Produces something he found.
- Asks Waldo about the item.
- Holds the dress in front of him so that Waldo can see his reflection.
- Throws away the dress.
- Answers Waldo's questions about the poem.
- Tries to snatch the paper from Waldo.
- Comes in and sees Waldo later.
- Asks Waldo what he is trying to do to him.
- Shapes his defense, apologizes, and explains himself to Waldo.
- Pleads with Waldo to let him go.

2. Waldo:

- Identifies the item Arthur found as an old dress of their mother's.
 - Orders Arthur to put the dress back.
 - Discovers a sheet of paper in a corner and smooths it out.
 - Quivers upon discovering the paper and calls Arthur to ask about it.
 - Reads the poem aloud.
 - Feels menaced by the poem.
 - Lets the paper hang from his hand.
 - Goes into the room where their mother used to sit.
 - Takes down the dress-box and begins to look for words.
 - Experiences shaking hands and feelings of being old and lustreless.
 - Feels infected by Arthur's "unnatural blood."
 - Goes to the pit where they burn things and pitches a paper tent.
 - Burns papers, feeling lighter afterward.
 - Considers how to disembarass himself from Arthur.
 - Raises himself on one elbow due to the urgency of his problem.
 - Reacts to Arthur's last words with disgust.
 - Becomes entranced by Arthur's face as it metaphorically opens and falls apart.
-

Figure 8. Snapshot of ChatGPT's summary of Excerpt No.4.

Waldo does not want Arthur to know about mother's dress he had just hidden, but he wants to tease Arthur by destroying his secret past attempts at poetry writing. The tone of voice is important but is mingled with his feelings of revenge. This is what the text tells us:

----- Then Waldo read aloud, not so menacingly as he would have liked, because he was, in fact, menaced:

And here is GPT's couple of snippets:

**** Reads the poem aloud.

**** Feels menaced by the poem.

Waldo wanted to sound menacing but “in fact he was menaced”, this is reported as a fact by the omniscient writer and was not felt by Waldo. The poem or what Arthur wrote, was telling about bleeding, a Vivisectionist Cordelia is bleeding and all Marys “in the end bleed”. After reading the poem aloud Waldo holds it in his hand and Arthur takes it back:

----- He would have snatched, but Waldo did not even make it necessary.

On the contrary what GPT understands of Arthur’s action is something different:

**** Tries to snatch the paper from Waldo.

But Waldo’s action is reported correctly:

**** Lets the paper hang from his hand.

Now Waldo feels “he was bleeding” and in his imagination he had “suspicion of an incurable disease”, because of the feeling he had of Arthur’s “unnatural blood continued to glitter”. Then we are told that,

----- Waldo was infected with it.

He was infected with the imaginary incurable disease and not by the blood:

**** Feels infected by Arthur’s “unnatural blood.”

Waldo is now convinced that all his past vainglorious attempts at becoming a novelist had no sense and it was the time to part from all of them. As a result, mother’s dress and all his papers were slowly burnt. The scene is described with a wealth of details because it is a fundamental milestone in the development of the two characters’ life. But GPT only records two simple actions:

**** Goes to the pit where they burn things and pitches a paper tent.

**** Burns papers, feeling lighter afterward.

Waldo’s action is in fact very complex since it encompasses all his past memories. At first, we are told that the dress-box is in fire and after that Waldo feels spontaneously pushed to continue burning his past, which is made up of all the attempts at writing a novel about Tiresias:

----- About four o’clock he went down, Tiresias a thinnish man, the dress-box under his arm, towards the pit where they had been accustomed to burn only those things from which they could bear to be parted. He stood on the edge in his dressing-gown. Then crouched, to pitch a paper tent, and when he had broken several match-sticks - increasingly inferior in quality - got it to burn. The warmth did help a little, and prettiness of fire, but almost immediately afterwards the acrid years shot up his nose. So he stood up. He began to throw his papers by handfuls, or would hold one down with his slippered foot, when the wind threatened to carry too far, with his slippered foot from which the blue veins and smoke wreathed upward. It was both a sowing and a scattering of seed. When he had finished he felt lighter, but always had been, he suspected while walking away. Now at least he was free of practically everything but Arthur.

At the end, we are also told that even if “felt lighter” but always had been “he suspected”. Apart from his twin brother that arrives in the following snippet, which is again badly presented:

**** Considers how to disembarass himself from Arthur.

**** Raises himself on one elbow due to the urgency of his problem.

----- After he had lain down on the bed he began to consider how he might disembarass himself, not like silly women in the news who got caught out through falling hair or some such unpremeditated detail, but quick, clean, and subtle, a pass with the tongue he had not yet perfected, but must. As he lay, he raised himself on one creaking elbow, because of the urgency of his problem. That was when Arthur came in and saw him.

The final part of the conversation between the twins is rendered only superficially without the sufficient pathos for what is happening. Waldo is trying to destroy emotionally and mentally his twin brother Arthur with the excuse of the poem. The dialogue is summarized correctly by Arthur’s reaction:

**** Asks Waldo what he is trying to do to him.

**** Shapes his defense, apologizes, and explains himself to Waldo.

But then the outcome is wrongly reported in the snippet:

*** Reacts to Arthur's last words with disgust.

Which should summarize the following sequence of their conversation:

----- "I know it wasn't much of a poem." Arthur was shaping his defence. "Oughter have destroyed it at once. Apologise, Waldo."

The warmed stones of words.

"That poem? That disgusting blood myth!" Waldo gasped to hear his own voice.

"I would have given the mandala, but you didn't show you wanted it."

"I never cared for marbles. My thumb could never control them."

Waldo in fact gasped to hear his own voice, and disgust is associated with the words of Arthur's poem.

3. Method, Materials and Results II°

3.1. Testing ChatGPT with the Plot, the "Sujet"

In this section we will consider the way in which the story has been narrated, the narrative style or techniques the author has brought forward to make the story interesting, as well as the linguistic techniques underlying the style. At first, we have used the same Excerpt No.3 of the previous analyses, because it is a highly representative piece of text to be used for characterizing White's style through Waldo's personality. And this is mainly the reason why we have chosen it, but as will appear below, it is also for its peculiar linguistic features.

ChatGPT analysis, organized as before into a sequence of bullet items, this time is preceded by a general comment which hinges upon the mixture of what ChatGP regards the three important components of White's narrative style: stream of consciousness, free reported or indirect discourse and the narrator's intrusive voice. In fact, none of the three are components indicated are present: White's style has no "auctorial intrusion", no stream of consciousness and no free indirect discourse. Rather what is at stake here is "figural consciousness", i.e., it is the character's own internal consciousness that the narration presents, something that is totally different in each of the three protagonists. Even the presence of diegetic statements must be regarded as stemming from the character inner voice (see [3,6]).

As happened before, also in this case the introductory portion of the excerpt with a seemingly strange visitor and his accompanying woman is not intertwined with Waldo's reconnaissance phase when the memory of Johnny Haynes jumps up to clarify the scene and justify Waldo's reactions.

The man is described with a richness of details ending up in a solid man, – because of his purposefulness - is the author's comment. Waldo envied the man, his clothes, his kempt head. But then comes the ironic side of the man's description: the zip of his "insolent" pants, which might get stuck in a public lavatory. So that eventually the man might be soon subject to a stroke. And that's when Waldo "racked his memory" and was "racked". As said above, Waldo eventually recognized Johnny Haynes – the boy he hated - because of his woman's husky voice. So it is for the details that we are brought inside the protagonist's mind and navigate with his thought and imagination in unexplored lands.

The most important component of White's style as referred to the most important protagonist, Waldo, is the diffuse sense of uncertainty, expressed by the use of modality which is realized in verbal complexes and adverbials. Here below we make a list of "uncertainties" introduced in the excerpt, marking the relevant portion of text with italics:

-
- The VISIT is more *ominous, less expected, oblique and undoubtedly malicious*
 - The GATE which *never quite fell down*
 - The SILK SCARF which *apparently serve no other purpose than to stop the hair from bursting out*
 - *If it had not been for his vigour, the burly stranger, who inclined towards the elderly by Waldo's calculating, might have been described as fat. But*

- Those ZIPS *which might one day get stuck*
- With such a BUILD, in a year or two, *a stroke would probably strike* his visitor down
- If VISITOR *he were*. And *not some busybody* of an unidentified colleague
- Or or. *Waldo racked his memory, and was racked.*
- He *couldn't help but wishing the DOGS* hadn't gone trailing after Arthur
- That they *might appear* round the corner
- Waldo sincerely *wished Mother had been there* to deal with things
- She walked with the *quizzical ease of a certain type* of expensive woman Waldo had never met, only smelt, and once touched in a bus
- She was wearing a lime-green dress of *more than necessary*, though diaphanous, material
- He *longed to catch that moment, if he could*, not in its flesh, oh no, but its essence, or poetry, which had been eluding him all these years
- Waldo *might have continued* congratulating himself on this piece of practical information, *if the man hadn't just then shouted* at the woman
- But the woman *apparently did not care* to be asked. She remained indifferent. Or ignorant.
- So it was only natural *he should continue hating* Haynes
- Staring vindictively at it from under his barbered eyebrows - what vanity - *as though he intended to tear* bits of the woodwork off
- But. You could tell
- After kicking at the house once or twice, *to bring it down, or relieve his frustration*, the visitor *appeared the victim* of a sudden sentimental tremor
- "I was *never too sure* about the twin; I think he *wasn't so loopy as they used to make out*"
- "If they *ever existed*," the woman replied dreamily
- The *mere thought* of their nakedness together gave Waldo Brown the gooseflesh, whether from *disgust or envy he couldn't have told*.
- Arthur was far from dishonest, but had the kind of buffalo mind which *could not restrain itself from lumbering* into other people's thoughts
- Choked by quince trees, the *window hardly responded* to light, unless the highest blaze of summer
- Something *more esoteric than could have come* from a department store
- Standing as *she had never stood in fact*, because, although memory is the glacier in which *the past is preserved, memory is also licensed to improve* on life
- When his heart crashed. So it literally *seemed*.
- His breath went with him, through the tunnel along *which he might have been running*. Whereas he was again standing.

Figure 9. List of sentences containing "uncertainty" markers in excerpt No. 3.

And here below in Figure 10 the summary produced by ChatGPT. It is really important to focus on the way in which ChatGPT has been able to capture the "uncertainty" we have been referring to in the Introduction. This is particularly well represented in bullet point no. 5 where we have the impression that "ambiguity" has become so strong that reality is almost indistinct from imagination. However, this is not at all the way uncertainty should be interpreted. As the list of verbal and adverbial constructions above clearly shows, it is rather with the use of modality that it is possible to classify the narrative technique typical of Patrick White's novels in general and in particular of the most important protagonist of this novel, Waldo Brown.

The narrative technique in this excerpt is highly unique, blending **stream of consciousness** with **free indirect discourse**. This approach provides a deep dive into the protagonist Waldo's psyche, revealing his thoughts, feelings, and perceptions in a way that blurs the line between his inner monologue and the narrator's voice.

Key Features of the Narrative Technique:

1. **Stream of Consciousness:** The narrative frequently shifts into a stream of consciousness mode, where Waldo's thoughts flow freely and often incoherently, mimicking the natural process of human thought. This technique allows readers to experience the immediacy of Waldo's mental state, his paranoia, his memories, and his fantasies, all in real time. For example, his sudden associations with the visitors, his memories of Johnny Haynes, and the intense reflection on his past and present fears are all rendered in a fluid, unstructured manner.
 2. **Free Indirect Discourse:** The narrative also employs free indirect discourse, a technique where the third-person narration slips into the character's inner thoughts without quotation marks or explicit tags like "he thought." This creates a seamless blend between the character's voice and the narrator's, allowing readers to experience Waldo's perspective intimately while maintaining the third-person narrative structure. Phrases like "Waldo sincerely wished Mother had been there to deal with things" and "But much as he regretted the stranger's presence and relationship, he thrilled to the evocations of the woman's voice" illustrate how the narrative shifts in and out of Waldo's direct thoughts.
 3. **Fragmentation and Disjointedness:** The narrative is deliberately fragmented, reflecting Waldo's mental state. His memories, observations, and anxieties are presented in a disjointed manner, mirroring his fractured sense of reality. This fragmentation is evident in how the narrative jumps between different times, places, and emotions without clear transitions, making the reader experience the confusion and complexity of Waldo's inner world.
 4. **Symbolism and Imagery:** The narrative is rich in symbolism and vivid imagery, often taking on a surreal quality. For instance, the descriptions of the lime-green dress, the nacreous fan, and the icy satin of the old dress are laden with symbolic meaning, reflecting Waldo's obsessions, his longing for the past, and his fear of exposure. These symbols are intricately woven into the narrative, enhancing the psychological depth of the story.
 5. **Narrative Tension and Ambiguity:** There is a constant tension between reality and imagination, with the narrative leaving much to the reader's interpretation. The visitors, for example, are described with an ominous ambiguity that leaves their true nature unclear – are they real threats, mere figments of Waldo's imagination, or symbolic representations of his fears? This ambiguity heightens the sense of psychological unease that permeates the text.
- Overall, the narrative technique is deeply immersive, drawing the reader into the protagonist's troubled mind through a blend of psychological realism and stylistic experimentation.

Figure 10. Snapshot of excerpt No. 3 used to detect narrative style.

Modality and uncertainty are important features of White's style and has been deeply documented in a previous paper [3] as said in the Introduction. In a number of tables – from Tables 1–4 here below, that we repeat from the paper cited, we explain why modality is the key to understand the style. We started above by commenting *uncertainty*, we marked all the parts of the text that in their syntactic, semantic or pragmatic value carry a sense of interpretation of the storyworld by the characters. This interpretation may indicate actual hesitation or ambiguity, expressing the more or less conscious doubts in the minds of the protagonists; at the opposite end, it can also signal a judgment of certainty by a character, which ironically in turn generates insecurity in the reader and raises a series of important questions. In the present analysis, the element *uncertainty* has only one obligatory attribute: *nonfactual* (more in a section below, but see [26,27]). To specify uncertainty, ambiguity and doubt, in fact, it is crucial to mark the annotated expression as non-real – that is, non-factual – a process that is only going on in the character's mind and which does not have an equivalent in the "real world" of the story.

Table 1. Semantic linguistic features grid organized by hierarchy.

Meta-Tags	Uncertainty	Subjectivity	Subjectivity	Judgement
<i>Intermediate Tags</i>	<i>Non-Factuality</i>	<i>Psychology</i>	<i>Affectivity</i>	<i>Social_Evaluation</i>
Tags	Seeming	Perception	Inclination_Positive	Esteem_Negative
	Gnomic	Precognition	Inclination_Negative	Esteem_Positive
	Concessive	Cognition	Security_Positive	Sanction_Negative
	Conditional	Performwill	Security_Negative	Sanction_Positive
	Defdesire		Satisfaction_Positive	
	Will		Satisfaction_Negative	
	Possibility			
	Ability			
	Obligation			
	Assumption			
	Negation			

Table 2. Distribution of Semantic linguistic features amongst the three protagonists.

	Waldo	Arthur	Mrs. Poulter	Totals	RatioW	RatioA	RatioP
Percept	674	303	109	1086	62,063	27,9	10,037
Precogn	379	141	39	559	67,799	25,224	6,977
Cognition	953	460	140	1553	61,365	29,62	9,015
PerformW	39	3	0	42	92,857	7,143	0
Seeming	512	198	56	766	66,841	25,848	7,3107
Will	74	33	8	115	64,348	28,696	6,956
Possibl	224	54	42	320	70	16,875	13,125
Gnomic	32	20	9	61	52,459	32,787	14,754
Ability	347	195	51	593	58,516	32,884	8,6
Obligation	178	79	28	285	62,456	27,719	9,8246
Concessv	150	75	26	251	59,761	29,88	10,358
Conditnl	264	117	32	413	63,922	28,329	7,7482
Defdesire	49	35	2	86	56,977	40,698	2,325
Assumpt	73	36	9	118	61,864	30,508	7,627
Emot_Pos	400	183	43	626	63,898	29,233	6,869
Emot_Neg	247	107	30	384	64,323	27,864	7,812
Inclin_Pos	83	50	17	150	55,334	33,334	11,334
Inclin_Neg	75	26	12	113	66,372	23,009	10,619
Secur_Pos	133	65	19	217	61,29	29,954	8,756
Secur_Neg	297	116	40	453	65,563	25,607	8,83
Satisf_Pos	136	66	10	212	64,151	31,132	4,717
Satisf_Neg	215	112	25	352	61,079	31,818	7,102
Estm_Pos	223	71	47	341	65,396	20,821	13,783
Estm_Neg	346	129	39	514	67,315	25,097	7,587
Sanct_Pos	80	19	18	117	68,376	16,239	15,385
Sanct_Neg	127	47	16	190	66,842	24,737	8,421

Table 3. Best 10 Semantic linguistic features of the three protagonists.

Waldo's Best 10	Arthur's Best 10	Mrs. Poulter's Best 10
1-PerformW	1-DefDesire	1-Sanct Pos
2-Possibility	2-Inclin Pos	2-Estm Pos
3-Sanct Pos	3-Ability	3-Inclin Pos
4-Precognt	4-Gnomic	4-Gnomic
5-Estm Neg	5-Satis Neg	5-Possibility
6-Sanct Neg	6-Satis Pos	6-Inclin Neg
7-Inclin Neg	7-Assumptn	7-Concessiv
8-Secur Neg	8-Secur Pos	8-Perceptn
9-Estm Pos	9-Concessiv	9-Obligation
10-Will	10-Cognition	10-Cognition

Table 4. Unique Semantic linguistic features of the three protagonists.

Waldo's Unique	Arthur's Unique	Mrs P-'s Unique
PerformW	DefDesire	
	Ability	
Precognt		
Estm Neg	Satis Neg	
Sanct Neg		
Inclin Neg	Assumptn	
Secur Neg	Secur Pos	Perceptn
	Concessiv	Obligation
Will		

As with all other aspects of this particular tagging system, *subjectivity*, the second of the three stylistic elements, focuses on facets of character psychology and on their relations with the storyworld. The main difference between *uncertainty* and *subjectivity* lies in the fact that the first element circumscribes how the protagonists interpret their reality and the ways they rationalize it, while the second marks the modalities in which they actively and subjectively contribute to the narration. While the former element is non-factual in nature, the latter always has consequences in terms of narrative, sometimes even tangible ones (e.g., in the case of performative *will*). Subjectivity includes active psychological processes – both conscious and unconscious – as studied by cognitive sciences, as well as expressions of emotion and different kinds of feelings, grouped into five attributes.

The last content-related element introduced in the annotation is *judgement*, as was the case with the affect attributes. In this annotation, *judgmnt* marks all evaluative expressions related to the characters of the novel, aimed at highlighting both social and personal reactions to the storyworld and particularly to the other characters and their behavior. The theoretical basis of this categorization can be found in the so-called “appraisal theory” [7,8], which underscores the relevance of impressions and judgments in the formation of feelings, emotions, and complex thoughts. Environment and psychology are here understood as standing in a relation of mutual dependency, with the reactions of each individual to events and stimuli evoking different responses. Speaking of the characters of a novel, we can say that the “artificially created” psychology of every character reacts in substantially different ways to what happens in the storyworld. From a general point of view, we can say that *judgmnt* and *affect* as categories have a lot in common, both dealing as they do with indices of emotion and sentiment. In this specific study, however, it was decided to annotate *judgmnt* as an independent element, for two main reasons: the first was to emphasize our interest in the evaluative language used by the author, and the second was to allow a more detailed internal differentiation between the

categories of social-esteem and social-sanction and their respective polarities. For a complete explanation of the meaning of each tag we refer the reader to the paper cited above.

The linguistic feature grid we organized for our annotation task is structured in three main meta-tags and a number of hierarchically related tags, as shown here below:

We annotated 8416 semantic features as a whole - the non-repeated count amounts to only 1584, i.e., the level of repetition is fairly high - and characterizes each of the three protagonists as shown in the following tables. In some case, the same wordform has received more than one label due to its intrinsic polysemous nature. We report in Table 2 the actual count split by character and by semantic feature.

And in the two following tables, Table 3 and Table 4 the specificity of each protagonist as to their most typical semantic feature associated by the author in the novel.

3.1. Testing CHATGPT for Factuality and Temporal Ordering

Eventually we tested ChatGPT for its ability to tell facts from non-facts [29,30] using short paragraphs from the same novel. Facts should be regarded all sentences of direct speech disregarding their content; then non-facts should include all diegetic or gnomic statements which are usually expressed in present tense. At first, we gave ChatGPT two binary choices which were than easily understood. But what happened is shown in a few examples which we chose from the many we tried and always received the same results: a total confusion due to third-person narration carried out in the past even for direct speech, that ChatGPT was unable to infer as being different from real past events, this time expressed with the pluperfect. So, at first, we improved the prompt with a slight clarification about progressive. In particular we added to the main prompt the suggestion to watch for presence of progressive tense or aspect which ChatGPT seemed to focus on even though the general impression we received is that the verbal complex is internally invisible to ChatGPT.

Here is the main prompt:

"I have a short text where you should divide up sentences at first into two categories: those that present or represent a fact and those that don't. Then you should divide up those that represent a fact into two subcategories: those that are placed in the past and those that aren't."

Additionally, at the usual request by ChatGPT at the end of each output we added an extra small prompt clarification:

"Let me know if you'd like any further refinements!"

Yes. You should list sentences with a verbal complex in the progressive amongst the "not past"

Got it! I'll revise the categories, ensuring sentences with a verbal complex in the progressive are placed in the "Not in the Past" subcategory. Here's the updated breakdown:

And here is the first text taken from the first Section of the book that we now call SmallExcerpt1.

Just then the bus nearly shook apart, and one of the young lady typists lost her balance. It was those stiletto heels.

"These old buses!" Mrs Poulter heaved, and laughed.

Mrs Dun went a deeper yellow as she grasped the rail.

"They're a downright disgrace!" Mrs Poulter laughed.

The sadness had been shaken out of her. She was happy again.

Presently she couldn't resist: "That veranda of yours must be a real luxury, dryin' laundries in the rain."

"I'll say it is!" said Mrs Dun.

She had a certain relentlessness of conviction. If it hadn't been for her gloves her knuckles would have shown up white on the chrome rail.

The bus was making slow progress, on account of the pay-as-you-enter, and queues at the shelters, and kiddies who had missed the special. Mrs Poulter looked out. She was proud of the glossier side of Sarsaparilla, of the picture windows and the texture brick. She brightened with the leaves of the evergreens which the sun was touching up. Then she saw Bill, and waved. But he did

not respond. He went on sweeping the gutters for the Council. It was against Bill Poulter's principles to acknowledge his wife in public. Sometimes on her appearing he went so far as to take time off to roll himself a cigarette. But never wave. She accepted it. She was content enough to realize he was wearing the old fawn sweater, no longer presentable except for work, because the loose stitch she had been trying out had begun to stretch and sag.

This is the output where we marked with three asterisks wrong choices and with three °° complex sentences containing a simple sentence in the past.

1. Sentences that present or represent a fact:

Past Facts:

***Just then the bus nearly shook apart, and one of the young lady typists lost her balance.

***It was those stiletto heels.

***Mrs Dun went a deeper yellow as she grasped the rail.

The sadness had been shaken out of her.

°°The bus was making slow progress, on account of the pay-as-you-enter, and queues at the shelters, and kiddies who had missed the special.

***Mrs Poulter looked out.

***She brightened with the leaves of the evergreens which the sun was touching up.

***Then she saw Bill, and waved.

***But he did not respond.

***He went on sweeping the gutters for the Council.

***It was against Bill Poulter's principles to acknowledge his wife in public.

***Sometimes on her appearing he went so far as to take time off to roll himself a cigarette.

***But never wave.

°°She was content enough to realize he was wearing the old fawn sweater, no longer presentable except for work, because the loose stitch she had been trying out had begun to stretch and sag.

Not in the Past:

"That veranda of yours must be a real luxury, dryin' laundries in the rain."

"I'll say it is!" said Mrs Dun.

2. Sentences that do not represent a fact:

***"These old buses!" Mrs Poulter heaved, and laughed.

***"They're a downright disgrace!" Mrs Poulter laughed.

***Presently she couldn't resist: "That veranda of yours must be a real luxury, dryin' laundries in the rain."

***She had a certain relentlessness of conviction.

If it hadn't been for her gloves her knuckles would have shown up white on the chrome rail.

***Mrs Poulter was proud of the glossier side of Sarsaparilla, of the picture windows and the texture brick.

Figure 11. Snapshot of ChatFPT categorization of SmallExcerpt 1.

In fact, the narration is all told using past tense also when direct speech is presented as can be clearly seen, in particular, in the sentence "I'll say it is!" said Mrs Dun" which has been correctly listed by ChatGPT under NOT IN THE PAST even though the governing verb "said" is clearly expressed in past tense. Correctly listed under PAST FACTS are those sentences with their verbal complex in past perfect or pluperfect tense. Also correctly listed as NOT FACTS is the one sentence with a counterfactual.

After the additional small suggestion about progressive tense we find the following reordering:

Not in the Past (includes progressives):

“That veranda of yours must be a real luxury, dryin’ laundries in the rain.”
 “I’ll say it is!” said Mrs Dun.
 Mrs Poulter was proud of the glossier side of Sarsaparilla, of the picture windows and the texture brick.
 She brightened with the leaves of the evergreens which the sun was touching up.

Figure 12. Snapshot of subsection of ChatGPT output for SmallExcerpt1 on Progressive.

ChatGPT has wrongly treated the presence of past tense as signaling past facts. The first sentence is now a shortened version of the previous complete utterance which included the governing verb in the past tense: the text of the direct speech contains a gerundive which is expressed in the “ing” form, contracted though. There are then two additional sentences which however do not contain progressives in their main clause, only the second one has progressive in the relative clause. The same happened in the many other short paragraphs we selected to see whether there were improvements. We only report one case where ChatGPT indicates the presence of an “implied progressive”:

Not in the Past (includes progressives):
 “What goes on in people’s minds. Because it does go on. You’ve only got to read the papers.”
 “They was probably only taking a walk to get their circulation going.” (**progressive implied**)
 Like that blood-pressure thing was on your arm.
 Nor did it help either lady to know the other could be involved.

Figure 13. Snapshot of subsection of ChatGPT output for SmallExcerpt1 on Implied Progressive.

In fact, the progressive is expressed “was...walking” and not IMPLIED as wrongly indicated. Here is a second paragraph where ChatGPT wrongly categorizes direct speech as taking place in the past and as non-fact, that we call SmallExcerpt2:
 In High Street the overstuffed bus began to spew out its coloured gobbets.
 “Wonder what those two old fellers were doin’ so far from Terminus Road?” Mrs Poulter nursed her curiosity as they waited to be carried by the common stream.
 “You wonder what goes on in some people’s minds,” said Mrs Dun.
 “I beg yours?”
 “What goes on in people’s minds. Because it does go on. You’ve only got to read the papers.”
 “But two respectable old gentlemen like the Mister Browns? They was probably only taking a walk to get their circulation going.” Mrs Poulter had turned mauve. “Anyway,” she said, “what goes on in other people’s minds is private. I wouldn’t want to know what goes on inside of my own husband’s mind.”
 Although Mrs Dun might have wanted, she suggested she didn’t by drawing in her chin.
 “I was never one,” she said, “not to keep to meself, and mind me own business.”
 “Aren’t I right then?” Mrs Poulter continued, still too loud, and still too mauve.
 Creating in the bus. Mrs Dun wondered whether she had been wise in the first place to accept Mrs Poulter’s friendship.
 “As for those old men,” said Mrs Dun, “they’re nothing to me.”
 “They’re nothing to me,” Mrs Poulter agreed.
 But the situation made her want to cry. And Mrs Dun could feel it. She could feel her own gooseflesh rise. As they waited to escape from the suffocating bus the features of their familiar town began fluctuating strangely through the glass. Like that blood-pressure thing was on your arm. Nor did it help either lady to know the other could be involved.

And here below the output categorization by ChatGPT, where we marked with three asterisks wrong classifications:

1. Sentences that present or represent a fact:

Past Facts:

In High Street the overstuffed bus began to spew out its coloured gobbets.

***"Wonder what those two old fellers were doin' so far from Terminus Road?" Mrs Poulter nursed her curiosity as they waited to be carried by the common stream.

***"But two respectable old gentlemen like the Mister Browns? They was probably only taking a walk to get their circulation going."

Mrs Poulter had turned mauve.

***"I was never one," she said, "not to keep to meself, and mind me own business."

***Creating in the bus.

***Mrs Dun wondered whether she had been wise in the first place to accept Mrs Poulter's friendship.

***"As for those old men," said Mrs Dun, "they're nothing to me."

***"They're nothing to me," Mrs Poulter agreed.

***But the situation made her want to cry.

***And Mrs Dun could feel it.

***She could feel her own gooseflesh rise.

***As they waited to escape from the suffocating bus the features of their familiar town began fluctuating strangely through the glass.

Not in the Past (includes progressives):

"What goes on in people's minds. Because it does go on. You've only got to read the papers."

"They was probably only taking a walk to get their circulation going." (progressive implied)

Like that blood-pressure thing was on your arm.

Nor did it help either lady to know the other could be involved.

2. Sentences that do not represent a fact:

***"You wonder what goes on in some people's minds," said Mrs Dun.

***"I beg yours?"

***"What goes on in other people's minds is private. I wouldn't want to know what goes on inside of my own husband's mind."

Although Mrs Dun might have wanted, she suggested she didn't by drawing in her chin.

***"Aren't I right then?" Mrs Poulter continued, still too loud, and still too mauve.

Figure 14. Snapshot of ChatGPT classification for SmallExcerpt2.

Seen the total inability of ChatGPT to recognize the use of past tense in the narration as the current narrative present, we decided to include in a new prompt the explicit information about the way in which past events should be interpreted. Here is the new prompt:

I have a short text where you should divide up sentences at first into two categories: those that present or represent a fact and those that don't. Then you should divide up those that represent a fact into two subcategories: those that are placed in the past and those that aren't, where past event clauses are those that have a main verb in the pluperfect tense.

What we did then was basically this: we left the two binary choices intact and added an explanation to make easier the selection of past event sentences. But this worked to badly increase the level of complexity of the query and made the whole prompt impossible to satisfy. This is how the new prompt was received:

Understood! Please share your text, and I'll analyze it by dividing the sentences into the categories you described:

1. **Fact vs. Not a Fact**

- o Facts are statements that present or represent objective truths or verifiable claims.
- o Non-Facts are statements that don't meet this criterion, e.g., subjective statements, hypothetical scenarios, or rhetorical questions.

2. **Subcategories for Facts:**

- o **Past Events:** Sentences containing main verbs in the **pluperfect tense** (e.g., "had gone," "had been").
- o **Other Facts:** Sentences representing facts but not placed in the past using the pluperfect tense.

Please share the text, and I'll start the classification!

Figure 15. Snapshot of ChatGPT answer to prompt clarification.

And this is what happened in the classification with the new prompt clarification:

Here's how I categorized the sentences based on your criteria:

Category 1: Fact vs. Not a Fact

Facts:

1. The bus nearly shook apart, and one of the young lady typists lost her balance.
2. It was those stiletto heels.
3. Mrs. Poulter heaved and laughed.
4. Mrs. Dun went a deeper yellow as she grasped the rail.
5. The sadness had been shaken out of her.
6. She was happy again.
7. Presently she couldn't resist: "That veranda of yours must be a real luxury, dryin' laundries in the rain."
8. "I'll say it is!" said Mrs. Dun.
9. She had a certain relentlessness of conviction.
10. ****If it hadn't been for her gloves, her knuckles would have shown up white on the chrome rail.
11. The bus was making slow progress, on account of the pay-as-you-enter, and queues at the shelters, and kiddies who had missed the special.
12. Mrs. Poulter looked out.
13. She was proud of the glossier side of Sarsaparilla, of the picture windows and the texture brick.
14. She brightened with the leaves of the evergreens which the sun was touching up.
15. Then she saw Bill, and waved.
16. ****But he did not respond.
17. He went on sweeping the gutters for the Council.
18. ****It was against Bill Poulter's principles to acknowledge his wife in public.
19. ****Sometimes on her appearing, he went so far as to take time off to roll himself a cigarette.
20. ****But never wave.
21. She accepted it.
22. She was content enough to realize he was wearing the old fawn sweater, no longer presentable except for work, because the loose stitch she had been trying out had begun to stretch and sag.

Not Facts

1. ****"These old buses!"
2. ****"They're a downright disgrace!"
3. ****"That veranda of yours must be a real luxury, dryin' laundries in the rain."

Category 2: Facts Subcategorized by Tense

Facts in the Past (Pluperfect Tense)

1. The sadness had been shaken out of her.

Other Facts

1. The bus nearly shook apart, and one of the young lady typists lost her balance.
2. It was those stiletto heels.
3. Mrs. Poulter heaved and laughed.
4. Mrs. Dun went a deeper yellow as she grasped the rail.
5. She was happy again.
6. Presently she couldn't resist: "That veranda of yours must be a real luxury, dryin' laundries in the rain."
7. "I'll say it is!" said Mrs. Dun.
8. She had a certain relentlessness of conviction.
9. ****If it hadn't been for her gloves, her knuckles would have shown up white on the chrome rail.
10. The bus was making slow progress, on account of the pay-as-you-enter, and queues at the shelters, ****and kiddies who had missed the special.
11. Mrs. Poulter looked out.
12. She was proud of the glossier side of Sarsaparilla, of the picture windows and the texture brick.
13. She brightened with the leaves of the evergreens which the sun was touching up.
14. Then she saw Bill, and waved.
15. ****But he did not respond.
16. He went on sweeping the gutters for the Council.
17. ****It was against Bill Poulter's principles to acknowledge his wife in public.
18. ****Sometimes on her appearing, he went so far as to take time off to roll himself a cigarette.
19. ****But never wave.
20. She accepted it.
21. She was content enough to realize he was wearing the old fawn sweater, no longer presentable except for work, ****because the loose stitch she had been trying out had begun to stretch and sag.

Figure 16. Snapshot of ChatGPT new summary for SmallExcerpt1.

The number of mistakes has increased. The same happened with SmallExcerpt2 which we don't repeat here but select the most revealing part of the output: in Category 2. Facts in the Past (Pluperfect Tense) the answer was "No sentences here have main verbs in the **pluperfect tense**." On the contrary we have the following two sentences:

1. Mrs. Poulter had turned mauve.
2. Mrs. Dun wondered whether she had been wise in the first place to accept Mrs. Poulter's friendship

We used a new small excerpt to confirm our findings and we discovered that ChatGPT confuses the habitual "used to" as a past perfect, and also the conditional perfect for a factual verb and as such not included in the subsection for non-facts, as shown in the subsection we selected from the whole output:

Facts in the Past (Pluperfect Tense)

1. "They're taking you down!" Bill used to shout.
 2. Her husband would have spat if it didn't normally happen indoors.
 3. "What?" he used to say, throwing off her concern with his shoulder.
-

Figure 17. Snapshot of ChatGPT subsection for Pluperfect.

4. Comparing GPT ad Gemini

In this section we will draw a comparison between GPT responses and the one given by Gemini [31] on the same excerpts. The output of the conversation with Gemini held on January 27th, 2025 is however made available only as Supplementary Material. The elements to compare are limited to the 12 mistakes made by GPT and in particular they are:

1. oversimplification (dropping the restrictive modifier)
2. wrong coreference chain (Mrs Dun rather than Mrs Poulter)
3. wrong sense selection with ambiguous term (think)
4. omitted cataphora with split referents (a strange man and Johnny Haynes)
5. disjointed reference of MEMORY and Mother
6. use of stream of consciousness as narrative style
7. mistakes in classifying factual vs non-factual sentences
8. mistakes in classifying factual sentences in the past with pluperfect
9. inability to detect linguistic elements characterizing narrative style
10. inability to characterize UNCERTAINTY in the verbal complex
11. abstract key points substituted by direct extracted references
12. Correct name of the author but wrong title of the novel

I will now tag each previous key point with WRONG in case the same mistake has been repeated and RIGHT and explain how Gemini has managed to overcome the difficulty.

1. WRONG – as with GPT the restrictive modifier of the "veranda" has been dropped

2. WRONG – the same mistake but then we asked Gemini to correct the error and this is the interaction where I use bold italics for the query and italics for the extract:

Great, but you made a mistake. In the piece of text I repeat below you wrongly assigned reference to Mrs Dun but it's Mrs Poulter intended because she is the one who has been depicted with gloves:

"She had a certain relentlessness of conviction. If it hadn't been for her gloves her knuckles would have shown up white on the chrome rail."

You are absolutely correct! My apologies for that oversight. That's a crucial detail, and I appreciate you pointing out the mistake.

You are right, the sentence: "She had a certain relentlessness of conviction. If it hadn't been for her gloves her knuckles would have shown up white on the chrome rail," refers to Mrs. Poulter, not Mrs. Dun.

The text earlier describes Mrs. Poulter as "peeping inside her plump glove to see if the ticket was still there," which is the key detail that connects her to the gloves. Mrs Dun, on the other hand, has no mention of gloves, instead she "went a deeper yellow as she grasped the rail," focusing on her reaction and appearance.

This mistake highlights the importance of careful reading and attention to detail. Thank you for correcting me! I will try my best to be more careful in the future.

3. RIGHT – think is used with the appropriate sense

4. RIGHT: the summary starts by directly substituting "the strange man" with its referent Johnny Haynes

5. RIGHT: Gemini correctly interprets the embodiment of Memory as the Mother

6. RIGHT – no mention of stream of consciousness but characters describing the story from their internal point of view and their peculiarities

7. RIGHT – a better classification of factuality with few mistakes. We report the final comment on the role that Factuality represents in the narration which is tagged as OBSERVATIONS:

Non-Factual as Insight: The non-factual sentences, predominantly reflections of Waldo's internal state, are key to understanding his character, motivations, and psychological breakdown. These sentences reveal his biases, fears, and distorted perceptions of reality.

Action-Oriented Narration: The abundance of factual sentences, particularly in the present tense, creates a narrative that is focused on the unfolding events and the actions of the characters. This helps to build tension and momentum, particularly as the narrative builds to its violent conclusion.

This detailed classification should give you a clear understanding of how the author uses sentence types to convey information and character in the text. Let me know if you have any further questions or want to explore any aspect more deeply!

8. WRONG – a better classification of past vs. present but there are mistakes. In fact there are only one or two sentences in the pluperfect. Gemini wrongly declares that there are no pluperfect sentences and motivates it by this observation:

Absence of Pluperfect: The absence of pluperfect in the factual sentences suggests that the narrative focuses primarily on the immediate actions and events, with minimal digression into past events that occurred prior to the narrative timeframe. This helps to create a sense of immediacy. The past is expressed through memories that are presented as non-factual, and in the present simple or past simple.

9. RIGHT – specific key points dedicated to linguistic items but they do not include what should have made the difference, that is mainly verbal complexes

10. RIGHT – UNCERTAINTY is depicted in deep details at the end of the Supplementary Materials in two pages. We paste here the beginning of the response by Gemini to the question “*And now I would like to know what is the role of the abstract notion of*

“UNCERTAINTY” and how it has been implemented in linguistic structures and words in the novel”:

Now, let's address your question about the role of the abstract notion of “uncertainty” and how it's implemented in linguistic structures and word choices within the novel. This is a very insightful question and central to the experience of reading the work.

Uncertainty as a Pervasive Force:

Uncertainty is not just a theme, but a fundamental principle that underlies the narrative structure, character development, and use of language throughout “The Solid Mandala.” It manifests in several ways:

11. RIGHT - Bullet key themes in Gemini's summaries are always abstractions, Here is a list from the Supplementary Materials where we can easily note that each excerpt has different themes unlike what happens with GPT where themes are often repeated:

Excerpt 1:

The Mundane and the Meaningful: Nostalgia vs. Reality: The Invisibility of the Ordinary: Routine and Habit: Suburban Life: Prejudice and Bias: Acceptance and Resilience;

Communication and Distance:

Excerpt 2:

Duty vs. Sentiment: Brotherhood: Individual vs. Shared Identity: Mortality and the Passage of Time: The Burden of Care: Underlying Repression and Complexity:

Excerpt 3:

Memory and Identity: Repression and Desire: Social Class and Envy: The Power of the Past: Transformation and Escape: Fragility of Reality: The Grotesque:

Excerpt 4:

Art and Emotion: Obsession and Control: Jealousy and Resentment: Violence and Self-Destruction: Ambiguity and Uncertainty: The Past as a Burden: Communication and Misunderstanding: Repressed Emotions:

12. WRONG – as happened with GPT also Gemini got the right name of the author but the wrong title of the novel. The response and the question are reported here below:

I thought by now you would have been able to tell who is the author and the title of the novel

You are absolutely right! My apologies for the oversight. With all this detailed analysis, I got so caught up in the process that I neglected the most basic element: identifying the work.

Based on the text I have been analyzing, the author is **Patrick White**, and the novel is titled "*The Tree of Man*."

The distinct style, themes, character portrayals, and the specific details I've encountered throughout the text are all hallmarks of White's work. His focus on the inner lives of seemingly ordinary individuals, his use of rich, evocative language, and his exploration of themes like isolation, memory, and the search for meaning are all consistent with his established style and preoccupations.

Thank you for pointing out my oversight! It's a crucial piece of information that I should have acknowledged much earlier.

I hope that, despite this oversight, you have found my analysis insightful and helpful. Now that we have established the author and title, does it change anything about the aspects of narrative style you would like to discuss? I am still happy to delve into any of the previous topics in more detail.

Ok for the author but the title of the novel is *The Solid Mandala*.

You are absolutely correct again! My apologies for the persistent error. The correct title of the novel is indeed "*The Solid Mandala*" by **Patrick White**. Thank you for catching that! It seems I was having a bit of a mental block there.

5. Discussion

As has been extensively shown, both ChatGPT and Gemini can produce well organized summaries or essays, but the contents of the analysis are sometimes the result of some form of hallucination, in particular this happens with GPT. On the contrary, with Gemini, we saw that over 12 types of mistakes made by ChatGPT, 8 have been correctly reported. However, the lack of generalization is present in both AI systems: in some case, the information to be used in order to report correctly seems to have been fully understood since a correct definition of the items has been captures. But the result shows that even in those cases in which the possible output has been exemplified correctly, the application of the apparently learnt information is either missing or is wrong. This reveals one of the most important deficiency of LLMs, also in their operative modality: the total lack or inability to generalize.

We assume that hallucinations are originated from three important faults in the overall architecture and conception of current AI technology: (1) tokenization that proceeds both at test and training time is wrought on the basis of too small vocabularies and it used subword units to cope with OOVs; (2) contextual information which is the basis of LLMs knowledge both in syntax and semantics is lost every time an unknown token/wordform is met and sometimes is not put to use even when made available by the model itself; (3) higher abstract information cannot be derived simply by local or chunk-level context, it may require complex inferential operations and some reasoning. When unknown wordforms are met, the model strives to build a new token with subword units, often resulting in illegal or inexistent words (see our papers in [24,25]). As discussed at length in [23] and reported above, smaller vocabularies result in approximate context and wrong embeddings selection in zero shot model performance leading to hallucinations. Similarities between model and input test can in general be misleading, due to the nature itself of cosine measure in vector space, as has been shown in detail in our latest research [24,25]. Finally, as explained in (3) higher abstract information like the one needed to characterize the plot of a novel and the underlying linguistic techniques, require the ability to generalize and then reason on a set of underlying

linguistically characterized phenomena in order to choose the correct interpretation which in our case has been mostly missing. The other important linguistic element hard to cope with for LLMs is the well-known fact that human language erases redundant information whenever it is grammatically derivable – and in some cases also in order to produce non-literal language with the aim to generate ambiguity and misunderstanding as happens in satiric contexts or in metaphors. Elided linguistic information – as for instance the subject of an untensed clause like an infinitival or a gerundive - cannot be recovered by context lookup in a model.

Eventually, even though ChatGPT has shown a remarkable ability in organizing summaries their content is deficient, but it may be used as precis by K-9 students and still get a good mark. Different judgement can be given for GEMMA 2 and Gemini which have produced a much richer output with deep and elaborate concepts at all levels of analysis we have tested it. Mistakes appeared but in a much smaller quantity.

5. Conclusions

In this paper we have made a detailed analysis of the way in which ChatGPT.4o and Google Gemini summarize narrative text and showcased the typical errors that may ensue. We listed 12 types of error and hallucination that GPT made and compared its performance with Google Gemini, improving and correcting 8 of them. Worst cases are semantically relevant errors, which are those related to coreference resolution, to the incorrect deletion of a restrictive modifier, and what is more important, association of actions and thoughts to the wrong character. In addition, we showed that ChatGPT has not been able to correctly detect neither factuality nor temporal ordering. Besides the plot being incorrectly classified, also the style, which has been wrongly defined at first as stream of consciousness with auctorial intrusion. Eventually, we may regard the summarization function of GPT as contributing descriptions which are error prone. Google Gemini on the hand, may be regarded satisfying even if containing some mistake. The reason for the improvements found in Gemini may be due to the use of a much larger vocabulary – 256K vs 32K entries. Besides, even if the responses of the two AI systems contain in some case well organized summaries and useful hints for further improvements, they cannot be regarded fully reliable and require human intervention.

Author Contributions: All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: In this section, you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

OOVW	Out Of Vocabulary Word
NLP	Natural Language Processing
GPT	Generative Pre-Trained Transformer
LLM	Large Language Model
BERT	Bidirectional Encoder Representations from Transformer
DOI	Digital Object Identifier
RHT	Reasoning Hallucination Test
FCT	False Confidence Test
FQT	Fake Question Test
NOTA	None Of The Above

Appendix A

In this Appendix we include a short summary of the novel *The Solid Mandala*:

The novel is the story of the life of twins Arthur and Waldo, their family and their neighbour, Mrs Poulter in a suburb of Sydney. Waldo Brown is an appropriate example of many other important characters of Patrick White's novels, as well. He is the representative of the intellectual who failed to become an artist or even to accomplish anything of significant and ended up being a simple clerk in a municipal library. His life is empty of events and positive emotions. He is educated and despises his community, which he considers too uninteresting and uncultured for him to be a part of. This voluntary isolation translates in a general growing resentment and in open hostility towards his twin brother, who is completely different from him and yet always a constant in his life. Arthur and Waldo could not be more diametrically opposed. This opposition is specifically crafted to portray and therefore study two basic drives: intellect and intuition. We find in Waldo every characteristic of the academic individual driven by intellect, as we said before; in Arthur, instead, there is a more "feminine" intuition which is often painted as direct result of his weak wits. Contrary to his brother, Arthur is far from studious and clever: it is often difficult to understand him even from his point of view, and this does not seem to hold particular meaning for him. He has difficulty speaking and expressing himself, even though some of his thoughts are deeper and more significant than Waldo's. He loves others, even and mostly the brother who despises him and considers him a handicap. Most of all, he is completely, almost unbelievably good, always humble and helpful in his simple way of living.

Appendix B

In this appendix we list the excerpts taken from the txt version of the Penguin version of the novel we bought on the internet and we indicate the page numbers related to each such excerpt in the Penguin published version wherefrom the txt version has been taken.

EXCERPT NO. 1 (from page 12 to page 14)

Everyone was too obsessed by the start of another day - no hope this side of the tea-break - to notice the in no way exposed ladies in the eight-thirteen from Sarsaparilla, though it was perhaps doubtful whether anyone would ever notice Mrs Poulter or Mrs Dun unless life took its cleaver to them. For the present, however, they had the protection of their hats.

"Why did you?" asked Mrs Dun.

"Did I what?"

"Come to live down Terminus Road."

"Well," said Mrs Poulter, peeping inside her plump glove to see if the ticket was still there, "when we first come from up north - both my hubby and me was country people - we wanted it quiet like. We was young and shy. Oh, it was Bill's nerves, too. Bill said: It'll give us time to find our feet. It'll always open up in time. Land is always an investment."

"Oh, yairs. Land is an investment."

But a sadness was moistening Mrs Poulter.

"In those days," she said from out of the distance, "all the roads at Sarsaparilla were dead ends. Not only Terminus. You couldn't go anywhere as the crow."

"Eh?" Mrs Dun asked.

"As the crow flies," Mrs Poulter explained.

"Oh, the crow," her friend murmured, seeming uneasy at the idea.

"There was a Chinese woman lived on a hill up at the back. I never ever knew her. I seen her once. They were people of means, so people said. Growin' vegetables and things. They'd planted one of those what-they-call wheel-trees. Well, I seen her standing under it when it was in flower."

Mrs Dun sucked her teeth.

"You wouldn't of said she was without refinement," Mrs Poulter remembered. "But a Chinese is never the same."

It was something Mrs Dun had not even contemplated.

"And anyway, the Chinese person isn't the point."

Just then the bus nearly shook apart, and one of the young lady typists lost her balance. It was those stiletto heels.

"These old buses!" Mrs Poulter heaved, and laughed.

Mrs Dun went a deeper yellow as she grasped the rail.

"They're a downright disgrace!" Mrs Poulter laughed.

The sadness had been shaken out of her. She was happy again.

Presently she couldn't resist: "That veranda of yours must be a real luxury, dryin' laundries in the rain."

"I'll say it is!" said Mrs Dun.

She had a certain relentlessness of conviction. If it hadn't been for her gloves her knuckles would have shown up white on the chrome rail.

The bus was making slow progress, on account of the pay-as-you-enter, and queues at the shelters, and kiddies who had missed the special. Mrs Poulter looked out. She was proud of the glossier side of Sarsaparilla, of the picture windows and the texture brick. She brightened with the leaves of the evergreens which the sun was touching up. Then she saw Bill, and waved. But he did not respond. He went on sweeping the gutters for the Council. It was against Bill Poulter's principles to acknowledge his wife in public. Sometimes on her appearing he went so far as to take time off to roll himself a cigarette. But never wave. She accepted it. She was content enough to realize he was wearing the old fawn sweater, no longer presentable except for work, because the loose stitch she had been trying out had begun to stretch and sag.

EXCERPT NO. 2 (from page 23 to page 26)

"Put on your coat, and we'll go for a walk," he decided at last. "Otherwise you'll sit here brooding."

"Yes," Arthur said. "Brooding."

But he sat, and might have continued sitting, in that old leather chair with the burst seat where mice had nested the other winter, the woodwork scratched by dogs reaching up to claim right of affection. Arthur sat in their father's chair.

Waldo brought the two coats. He helped Arthur into his. Waldo treated the old herringbone rather roughly, to show that what he was doing had been dictated by duty and common sense. He set the matching cap very straight on Arthur's head. It was, in any case, the angle at which Arthur wore his cap. Waldo was relieved the performance of duty had at last set him free. But duty was honest, whereas he mistrusted the snares of sentiment set by inexhaustible tweed. (It was that good English stuff, from amongst the things discarded by Uncle Charlie, some of which were lasting for ever.)

"When it comes to illness there's too much giving in to it, not to mention imagination," Waldo warned.

As he put his own coat on he glanced at his brother's head, at the shagginess of hair falling from under the tweed cap. Very white. Waldo might have contemplated the word "silvery", but rejected it out of respect for literature and truth. Arthur's hair was, in fact, of that doubtful white, with the tobacco stains left by the red which had drained out of it. Unlike Waldo's own. Waldo on top was a thinned-out dirty-looking grey.

Arthur continued sitting.

And the two old dogs, turning on their cat-feet, forgetful of their withered muscle, watched out of milky eyes. One of them - it was Scruffy - clawed once at Arthur's knee. The dogs made little whinging noises in anticipation. They were easily delighted.

"You do feel better, though?" Waldo asked, so suddenly and so quietly that Arthur looked up and smiled.

"Yes, Waldo," Arthur said, and: "Thank you."

Then the older of the two dogs, of whiter muzzle, and milkier marble eyes, threw up his head, and gave two ageless sexless barks. The second of the two dogs began to scutter across the boards on widespread legs.

Waldo was leading his brother Arthur, as how many times, out of the brown gloom of the kitchen. The cold light, the kitchen smells, had set almost solid in it. Yet, here they were, the two human creatures, depending on habit for substance as they drifted through. If habit lent them substance, it was more than habit, Waldo considered bitterly, which made them one.

Some had made a virtue out of similar situations: naked-looking, identical boys; laughing girls, he had noticed, exchanging the colours which distinguished them, to mystify their friends; neat, elderly ladies, in polka dots and similar hats, appeared to have survived what was more a harness than a relationship.

But the Browns.

Waldo could feel his brother's larger, fleshy hand in his thinner, colder one as they stumbled in and out of the grass down what remained of the brick path. The wind drove reasons inward, into flesh. They were reduced, as always, to habit. But stumbled, even so.

Only the old pot-bellied dogs appeared convinced of the mild pleasures they enjoyed, frolicking and farting, though somewhat cranky with each other. One of them - Runt - lifted his leg on a seedy cabbage and almost overbalanced.

His brother was breathing deeply, Waldo saw.

He had difficulty with his words, chewing them to eject, but when he did, there they stood, solid, and for ever.

There was the sound of Waldo's stiff oilskin nothing would free from the weathers which had got into it. Waldo's oilskin used to catch on things, and he always expected to hear it tear. On that gooseberry bush, for instance. Which had not succeeded. Arthur had advised against it - Sarsaparilla was too warm - but Waldo had planted the bush. To demonstrate something or other.

On the broken path Waldo's oilskin went slithering past the gooseberry thorns. The wind might have cut the skins of the Brothers Brown if they had not been protected by their thoughts.

Arthur spoke quite briskly. Time, it appeared, removed him quickly from the sources of pain. Sometimes Waldo envied the brother who did not seem to have experienced - though he should have - the ugly and abrasive roughcast of which life was composed.

My brother, Waldo would breathe, at times indulgently enough, and at once he became the elder by years instead of the younger by several hours. Waldo could modulate his voice, more to impress than to please. The rather fine tenor voice, of which the parents had been proud, and Dulcie Feinstein had accompanied in the first excitement of discovery. Men, the insensitive ones, sometimes recoiled from the silken disclosures of Waldo's voice.

Waldo's voice and Arthur's hair. So Mother used to say. (It should have been Waldo's mind, Waldo knew.)

Sidling brittly down the path, to negotiate the irregular bricks, now pushing Arthur, who liked to be humoured at times into believing he was the leader, Waldo could not avoid staring into his brother's hair, fascinated, when the wind blew, by the glimpses of pink skin beyond. This head might have flaunted an ostentation of cleanliness, if it had not been for its innocence, and the fact that he knew Arthur was in many ways not exactly clean. Every third Sunday Waldo made him sit on a stool on the back veranda, behind the glass, behind the scratching of the roses, to hack at the excessive hair, and as it first lay against, then flowed away through his fingers, the barber always wondered why he got the shivers, why he hated the smell of his own mucus as he breathed down his thin nose, while the hair lay on the boards, in dead snippets, and livelier love-knots, quite old-girlishly, if not obscenely, soft. It had seemed much coarser when Arthur was a boy.

And Arthur had grown into a big strong man. Was still, for that matter. It was Arthur who lifted the weights. His muscles had remained youthful, perhaps because his wits had been easy to carry.

EXCERPT NO. 3 (from page 187 to page 193)

Then there was the visit, more ominous still, because less expected, more oblique in execution, undoubtedly malicious in conception.

It was a couple of years after they got the dogs that the strange man pushed the gate which never quite fell down. It was a Sunday, Waldo would remember, the silence the heavier for insects. The thick-set man came up the path. He was the colour and texture of certain vulgar but expensive bricks, and was wearing tucked into his open shirt one of those silk scarves which apparently serve no other purpose than to stop the hair from bursting out. If it had not been for his vigour, the burly stranger, who inclined towards the elderly by Waldo's calculating, might have been described as fat. But with such purposefulness animating his aggressive limbs, solid was the more accurate word. Waldo had begun to envy the artificial gloss which streamed from the stranger's kempt head, and the casual fit of his fashionable clothes, so that it came as a relief to spot one of those zips which might one day get stuck beyond retrieve in some public lavatory, and to realize that, with such a build, in a year or two, a stroke would probably strike his visitor down.

If visitor he were. And not some busybody of an unidentified colleague. Or blackmailer in search of a prey. Or or. Waldo racked his memory, and was racked.

He found himself by now in the dining-room, that dark sanctuary at the centre of the house, from the safety of which on several occasions he had enjoyed watching with Mother the antics of someone unwanted, Mrs Poulter for instance, roaming round by congested paths, snatched at by roses. Only now, with Mother gone, the game had lost some of its zest, he had forgotten some of the rules. The Peace, moreover, had so far receded he couldn't help wishing the dogs hadn't gone trailing after Arthur, that they might appear round the corner, and while Scruffy held the stranger up, Runt tear the seat out of his insolent pants.

For the man had begun to knock, and ask: "Anyone at home?" then growing braver, or showing off, to rattle, and shout: "Anyone in hiding?"

Waldo sincerely wished Mother had been there to deal with things, especially as a woman, more of a female, whether the stranger's wife or not, was following him up the path. She walked with the quizzical ease of a certain type of expensive woman Waldo had never met, only smelt, and once touched in a bus. She walked smiling, less for any person, than for the world in general and herself. Which was foolish of her when you knew how the axe could fall.

"Perhaps you've made a mistake," the woman said rather huskily, touching her hair, and looking around at nothing more than a summer afternoon.

She was wearing a lime-green dress of more than necessary, though diaphanous, material. Raised to her hair, her arm, exposing the dark shadow of its pit, was a slightly dusty brown. Under his dressing-gown, Waldo got the shivers.

"No, I tell you!" the man insisted.

He continued rattling the door-knob, till he left off to thwack a window-pane with the crook of one of his blunt fingers.

"I can't believe anyone really lives in it," said the woman in her inalterably husky voice.

Waldo was sure he had heard somewhere that huskiness of voice was an accompaniment of venereal disease. So however good the stranger might be having it with his wife or whore there was retribution to come. Waldo nearly bit his lip.

But much as he regretted the stranger's presence and relationship, he thrilled to the evocations of the woman's voice as she stood amongst the lived-out rosemary bushes, humming, smelling no doubt of something exotic, Amour de Paris out of the pierrot bottle, holding her head up to the light, which struck lime-coloured down, at her breasts, and into her indolent thighs. The result was he longed to catch that moment, if he could, not in its flesh, oh no, but its essence, or poetry, which had been eluding him all these years. The silver wire was working in him ferociously now.

At least the long cry in his throat grew watery and obscure. Mercifully it was choked at birth.

Again memory was taking a hand. He remembered it was that boy, that Johnny Haynes, they could have cut each other's throats, telling him behind the dunny to watch out for hoarse-voiced men and women, they were supposed to be carriers of syph.

Waldo might have continued congratulating himself on this piece of practical information, if the man hadn't just then shouted at the woman:

"But I know it is! It's the place all right. I'd bet my own face. There's that erection they had my old man stick on top because they wanted what Waldo's dad used to call a classical pediment'. I ask you!"

But the woman apparently did not care to be asked. She remained indifferent. Or ignorant.

It was Waldo who was moved, not by the materialization of Johnny Haynes, but by the motion of his own life, its continual fragmentation, even now, as Johnny, by his blow, broke it into a fresh mosaic. All sombre chunks, it seemed. Of an old blue-shanked man under his winter dressing-gown, which he wore because the house was dark and summer slow in penetrating.

So it was only natural he should continue hating Haynes, clopping like a stallion with his mare all round the house, staring vindictively at it from under his barbered eyebrows - what vanity - as though he intended to tear bits of the woodwork off. Waldo remembered reading some years earlier, before the demands of his own work had begun to prevent him following public affairs, that Johnny Haynes was going to the top, that he had become a member of parliament - if you could accept that sort of thing as the top - and been involved in some kind of shady business deal. Exonerated of course. But. You could tell. Only gangsters dressed their women like that.

Then, edging round the secure fortress of the dining-room, Waldo saw that Johnny had come to a stop in the yard. After kicking at the house once or twice, to bring it down, or relieve his frustration, the visitor appeared the victim of a sudden sentimental tremor.

"I would have been interested," he grumbled, "to take a look at old Waldo. And the dill brother. The twin."

Waldo had never hated Johnny Haynes so intensely as now, for trying to undermine his integrity in such seductive style, and when Johnny added: "I was never too sure about the twin; I think he wasn't so loopy as they used to make out" - then Waldo knew he was justified.

O God, send at least the dogs, he prayed, turning it into a kind of Greek invocation as he was not a believer, and no doubt because of his blasphemy against reality, the dogs failed to come.

Instead, the mortals went.

"The Brothers Brown!" Johnny snort-laughed.

"If they ever existed," the woman replied dreamily.

Then she shuddered.

"What's wrong?" Johnny asked.

"A smell of full grease-trap," the woman answered in her hoarse voice. "There are times when you come too close to the beginning. You feel you might be starting all over again."

At once they were laughing the possibility off, together with anything rancid. They were passing through to the lime-coloured light of the front garden, where the woman's body revived. The mere thought of their nakedness together gave Waldo Brown the gooseflesh, whether from disgust or envy he couldn't have told. But his mouth, he realized, was hanging open. Like a dirty old man dribbling in a train. Whereas Johnny Haynes was the elderly man, asking for trouble of the lime-coloured woman, wife or whore, who was going to give him syph or a stroke.

Anyway, they were going out the gate. Most indecently the light was showing them up, demolishing the woman's flimsy dress, as the member of parliament passed his hand over, and round, and under her buttocks, which she allowed to lie there a moment, in the dish where those lime-coloured fruits had too obviously lain before.

More than anything else these dubious overtures, such an assault on his privacy, made Waldo realize the need to protect that part of him where nobody had ever been, the most secret, virgin heart of all the labyrinth. He began very seriously indeed to consider moving his private papers - the fragment of Tiresias a Youngish Man, the poems, the essays, most of which were still unpublished - out of the locked drawer in his desk to more of a hiding place, somewhere equal in subtlety to the papers it was expected to hide. Locks were too easily picked. He himself had succeeded in raping his desk, as an experiment, with one of the hairpins left by Mother. Arthur was far from dishonest, but

had the kind of buffalo mind which could not restrain itself from lumbering into other people's thoughts. How much easier, more open to violation, the papers. So it became imperative at last. To find some secret, yet subtly casual, cache.

In the end he decided on an old dress-box of Mother's, lying in the dust and dead moths on top of the wardrobe, in the narrow room originally theirs and finally hers. Choked by quince trees, the window hardly responded to light, unless the highest blaze of summer. A scent of deliquescent quinces was married to the other smell, of damp. The old David Jones dress box lay in innocence beyond suspicion. Heavy though, for its innocence. Waldo discovered when he took it down some article which had been put away and forgotten, something more esoteric than could have come from a department store.

It turned out to be one of Mother's old dresses shuddering stiffly awkwardly through his fingers, and the scales of the nacreous fan flopping floorwards. He would have to investigate. Afterwards. Arthur was out roaming with the dogs. Waldo almost skipped to transfer the papers, so easily contained: his handwriting was noted for its neatness and compression - in fact he was often complimented.

Then, as though the transfer of the papers had been too simple on an evening set aside for subtlety, he remembered the old dress. He stooped to pick up the little fan. One of the ribbons connecting the nacreous blades must have snapped in the fall. The open fan hung lopsided, gap-fingered. But glittering.

In the premature obscurity which quince branches were forcing on the room Waldo fetched and lit a lamp, the better to look at what he had found. Rust had printed on the dress a gratuitous pattern of hooks and eyes. Not noticeably incongruous. Age had reconciled their clusters with the icy satin and shower of glass which swirled through his fingers creating a draught. It was a dress for those great occasions of which few are worthy. He need not mention names, but he could see her two selves gathered on the half-landing at the elbow in the great staircase, designed by special cunning to withstand the stress of masonry and nerves. Standing as she had never stood in fact, because, although memory is the glacier in which the past is preserved, memory is also licensed to improve on life. So he became slightly drunk with the colours he lit on entering. How his heart contracted inside the blue, reverberating ice, at the little pizzicato of the iridescent fan as it cut compliments to size and order. Disorderly in habit, because the years had gradually frayed her, Mother kept what he liked to think of as a sense of moral proportion. Which he had inherited together with her eyes. There were those who considered the eyes too pale, too cold, without realizing that to pick too deeply in the ice of memory is to blench.

Merely by flashing his inherited eyes he could still impress his own reflexion in the glass - or ice.

Mother had died, hadn't she? while leaving him, he saw, standing halfway down the stairs, to receive the guests, the whole rout of brocaded ghosts and fleshly devils, with Crankshaw and O'Connell bringing up the rear. Encased in ice, trumpeting with bugles, he might almost have faced the Saportas, moustache answering moustache.

When his heart crashed. So it literally seemed. He was left holding the fragments in front of the mirror. Then went out to see. A lamp he had disarranged on the shelf in taking the one for his own use had tumbled off. He kicked at the pieces. And went back.

To the great dress. Obsessed by it. Possessed. His breath went with him, through the tunnel along which he might have been running. Whereas he was again standing. Frozen by what he was about to undertake. His heart groaned, but settled back as soon as he began to wrench off his things, compelled. You could only call them things, the disguise he had chosen to hide the brilliant truth. The pathetic respect people had always paid him - Miss Glasson, Cornelius, Parslow, Mrs Poulter - and would continue to pay his wits and his familiar shell. As opposed to a shuddering of ice, or marrow of memory.

When he was finally and fully arranged, bony, palpitating, plucked, it was no longer Waldo Brown, in spite of the birthmark above his left collarbone. Slowly the salt-cellars filled with icy sweat, his ribs shivery as satin, a tinkle of glass beads silenced the silence. Then Memory herself seated

herself in her chair, tilting it as far back as it would go, and tilted, and tilted, in front of the glass. Memory peered through the slats of the squint-eyed fan, between the nacreous refractions. If she herself was momentarily eclipsed, you expected to sacrifice something for such a remarkable increase in vision. In radiance, and splendour. All great occasions streamed up the gothick stair to kiss the rings of Memory, which she held out stiff, and watched the sycophantic lips cut open, teeth knocking, on cabuchons and carved ice. She could afford to breathe indulgently, magnificent down to the last hair in her moustache, and allowing for the spectacles.

EXCERPT NO. 4 (from page 211 to page 214)

When Arthur produced something he had found.

"What is it, Waldo?"

"An old dress of Mother's."

"Why was it behind the copper? She must have forgotten."

"Put it away!" Waldo shouted. "Where it was!"

To Arthur, who was holding in front of him the sheet of ice, so that Waldo might see his reflexion in it.

Arthur threw away the dress.

Which turned into the sheet of paper Waldo discovered in a corner, not ferreting, but ferreted.

On smoothing out the electric paper at once he began quivering.

"Arthur," he called, "do you know about this?"

"Yes," said Arthur. "That's a poem."

"What poem?"

"One I wanted to, but couldn't write."

Then Waldo read aloud, not so menacingly as he would have liked, because he was, in fact, menaced:

"my heart is bleeding for the Viviseckshunist Cordelia is bleeding for her father's life all Marys in the end bleed but do not complane because they know they cannot have it any other way' "

This was the lowest, finally. The paper hung from Waldo's hand.

"I know, Waldo!" Arthur cried. "Give it to me! It was never ever much of a poem."

He would have snatched, but Waldo did not even make it necessary.

When his brother had gone, Waldo went into the room in which their mother used to sit at the four o'clock sherry. He took down the dress-box and began to look out shining words. He was old. He was bleeding. He was at last intolerably lustreless. His hands were shaking like the papers time had dried.

While Arthur's drop of unnatural blood continued to glitter, like suspicion of an incurable disease.

Waldo was infected with it.

About four o'clock he went down, Tiresias a thinnish man, the dress-box under his arm, towards the pit where they had been accustomed to burn only those things from which they could bear to be parted. He stood on the edge in his dressing-gown. Then crouched, to pitch a paper tent, and when he had broken several match-sticks - increasingly inferior in quality - got it to burn.

The warmth did help a little, and prettiness of fire, but almost immediately afterwards the acrid years shot up his nose.

So he stood up. He began to throw his papers by handfuls, or would hold one down with his slipped foot, when the wind threatened to carry too far, with his slipped foot from which the blue veins and smoke wreathed upward.

It was both a sowing and a scattering of seed. When he had finished he felt lighter, but always had been, he suspected while walking away.

Now at least he was free of practically everything but Arthur.

After he had lain down on the bed he began to consider how he might disembarass himself, not like silly women in the news who got caught out through falling hair or some such unpremeditated

detail, but quick, clean, and subtle, a pass with the tongue he had not yet perfected, but must. As he lay, he raised himself on one creaking elbow, because of the urgency of his problem.

That was when Arthur came in and saw him.

“Waldo!” Arthur was afraid at last. “What are you trying to do to me?”

When Waldo had always wondered, fainter now, whether Arthur noticed the hurt which was intended for him. Or Dulcie. He had never shown her he had noticed that moustache. And Dulcie’s moustache might possibly have been the means of her destruction.

But Arthur so practically smooth.

Through the pain of destroying Arthur he noticed more than heard Arthur’s last words.

“I know it wasn’t much of a poem.” Arthur was shaping his defence. “Oughter have destroyed it at once. Apologise, Waldo.”

The warmed stones of words.

“That poem? That disgusting blood myth!” Waldo gasped to hear his own voice.

“I would have given the mandala, but you didn’t show you wanted it.”

“I never cared for marbles. My thumb could never control them.”

He was entranced by Arthur’s great marigold of a face beginning to open. Opening. Coming apart. Falling.

“Let me go! Wald! Waldo!”

As dropping. Down. Down.

References

1. OpenAI. 2024b. Hello gpt-4o, [Large language model]. <https://chatgpt.com>
2. Patrick White. 1966. *The Solid Mandala*. Eyre & Spottiswood, London.
3. Gordon Collier. 1992. *The Rocks and Sticks of Words Style, Discourse and Narrative Structure in the Fiction of Patrick White*. Editions Rodopi B. V., Amsterdam Atlanta.
4. Rodolfo Delmonte and Giulia Marchesini: A Semantically-Based Computational Approach to Narrative Structure, in Harry Bunt(ed.), Proceedings 13th Joint ISO - ACL Workshop on Interoperable Semantic Annotation (isa-13), Workshop at the 12th International Conference on Computational Semantics (IWCS 2017) Montpellier, France, September 19, 2017, pp. 14-25. <https://aclanthology.info/papers/W17-6906/w17-690>. (accessed on 17/01/2025)
5. Giulia Marchesini and Rodolfo Delmonte: A semantically-based Approach to the Annotation of Narrative Style, in Harry Bunt(ed.), Proceedings 13th Joint ISO - ACL Workshop on Interoperable Semantic Annotation (isa-13), Workshop at the 12th International Conference on Computational Semantics (IWCS 2017) Montpellier, France, September 19, 2017, pp. 14-25. <http://aclweb.org/anthology/W/W17/#7402>. (accessed on 17/01/2025)
6. Gordon Collier. 1999. Apparent Feature-Anomalies in Subjectivized Third-Person Narration, in John Pier(ed.), Recent Trends in Narratological Research, Open Edition Books, PUFR, GRAAT, 21, pp.129-151. <https://books.openedition.org/pufr/3956>. (accessed on 17/01/2025)
7. J. Martin and P.R. White. 2005. *Language of Evaluation, Appraisal in English*. Palgrave Macmillan, London and New York.
8. J. Read and J. Carrol. 2012. Annotating Expressions of Appraisal in English. *Language Resources and Evaluation*, 46:421–447.
9. Robin Emsley, 2023. ChatGPT: these are not hallucinations – they’re fabrications and falsifications, in *Schizophrenia* volume 9, Article number: 52 (2023)
10. Ankit Pal, Logesh Kumar Umapathi, Malaikannan Sankarasubbu, 2023. Med-HALT: Medical Domain Hallucination Test for Large Language Models, in Proceedings of the 27th Conference on Computational Natural Language Learning (CoNLL), pages 314–334.
11. McGowan A, Gui Y, Dobbs M, Shuster S, Cotter M, Selloni A, Goodman M, Srivastava A, Cecchi GA, Corcoran CM. ChatGPT and Bard exhibit spontaneous citation fabrication during psychiatry literature search. *Psychiatry Res.* 2023, <https://pmc.ncbi.nlm.nih.gov/articles/PMC10424704/pdf/nihms-1918813.pdf>(accessed on January 22, 2025)

12. Joshua Maynez, Shashi Narayan, Bernd Bohnet, Ryan McDonald, 2020. On Faithfulness and Factuality in Abstractive Summarization, in Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 1906–1919.
13. Kalev Hannes Leetaru, 2023. Hallucination In Summarization: When ChatGPT Hallucinated New Stories, <https://blog.gdeltproject.org/hallucination-in-summarization-when-chatgpt-hallucinated-new-stories-in-an-evening-news-broadcast/> (accessed on January, 21 2025)
14. Catarina G. Belem, Pouya Pezeskhpour, Hayate Iso, Seiji Maekawa, Nikita Bhutani, Estevam Hruschka, 2024. From Single to Multi: How LLMs Hallucinate in Multi-Document Summarization, Megagon Labs, University of California Irvine, arXiv:2410.13961v1.
15. Philippe Laban, Alexander R. Fabbri, Caiming Xiong Chien-Sheng Wu, 2024. Summary of a Haystack: A Challenge to Long-Context LLMs and RAG Systems, in Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, pages 9885–9903.
16. Kung-Hsiang Huang, Philippe Laban, Alexander Fabbri, Prafulla Kumar Choubey, Shafiq Joty, Caiming Xiong, Chien-Sheng Wu, 2024. Embrace Divergence for Richer Insights: A Multi-document Summarization Benchmark and a Case Study on Summarizing Diverse Information from News Articles, in Proceedings of the 2024 Conference of (NAACL) the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), pages 570–593.
17. Matthew Dahl, Varun Magesh, Mirac Suzgun, and Daniel E. Ho, 2024. Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models, in Journal of Legal Analysis, Oxford University Press, 16, 64–93 <https://doi.org/10.1093/jla/laae003> (accessed on January 21, 2025)
18. Cem Uluoglakci and Tugba Taskaya Temizel, 2024. Terms Dataset for Benchmarking Hallucination Tendency of LLMs, in Proceedings of the 18th Conference of the European Chapter of the Association for Computational Linguistics: Student Research Workshop, pages 95–136.
19. Sewon Min, Kalpesh Krishna, Xinxu Lyu, Mike Lewis, Wen-tau Yih, Pang Koh, Mohit Iyyer, Luke Zettlemoyer, and Hannaneh Hajishirzi. 2023. FActScore: Fine-grained atomic evaluation of factual precision in long form text generation. In Proceedings of the 2023 Conference (EMNLP) on Empirical Methods in Natural Language Processing, pages 12076–12100, Singapore. Association for Computational Linguistics.
20. Prakamya Mishra, Zonghai Yao, Parth Vashisht, Feiyun Ouyang, Beining Wang, Vidhi Dhaval Mody, Hong Yu, 2024. SYNFACT-EDIT: Synthetic Imitation Edit Feedback for Factual Alignment in Clinical Summarization, in Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing, pages 20061–20083.
21. Adam Tauman Kalai, Santosh S. Vempala. 2024. Calibrated Language Models Must Hallucinate, arXiv:2311.14648v3.(accessed on January 21, 2025)
22. Yekyung Kim, Yapei Chang, Marzena Karpinska, Aparna Garimella, Varun Manjunatha, Kyle Lo, Tanya Goyal, Mohit Iyyer, 2024. Evaluating faithfulness and content selection in book-length summarization, arXiv:2404.01261v2, also published as a conference paper at COLM 2024, <https://github.com/mungg/FABLES> (accessed on January 21, 2025)
23. Chaofan Tao, Qian Liu, Longxu Dou, Niklas Muennighoff, Zhongwei Wan, Ping Luo, Min Lin, Ngai Wong, 2024. Scaling Laws with Vocabulary: Larger Models Deserve Larger Vocabularies, arXiv:2407.13623v3.
24. Delmonte R., Nicolò Busetto, Word Predictability is Based on Context - and/or Frequency, in David C. Wyld, Dhinaharan Nagamalai (Eds), Proceedings 8th International Conference on Artificial Intelligence and Applications (AI 2022) October 29~30, 2022, Vienna, Austria, pp.195-211.
25. Delmonte R., Nicolò Busetto, Stress Test for BERT and Deep Models: Predicting Words from Italian Poetry. International Journal on Natural Language Computing (IJNLC). Dec. 2022, Vol. 11, No. 6, pp. 15-37. DOI: 10.5121/ijnlc.2022.11602
26. Harris ZS. 1954. Distributional structure. Word 10:146–62.
27. Clark S. 2015. Vector space models of lexical meaning. The Handbook of Contemporary Semantic Theory S Lappin, C Fox pp. 493–522 New York: Wiley.
28. Erk K. 2016. What do you know about an alligator when you know the company it keeps. Semant. Pragmat. 9:1–63.

29. Roser Sauri and James Pustejovsky. 2009. FactBank: a corpus annotated with event factuality. *Language Resources and Evaluation*, 43(3):227–268
30. R. Sauri and J. Pustejovsky. 2012. Are you sure that this happened? Assessing the Factuality Degree of Events in Text. *Computational Linguistics*, 38(2):261–299.
31. Google AI Studio, “Conversation with Gemini”, at https://aistudio.google.com/app/prompts/new_chat, (accessed on January 27, 2025)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.