COMMENTARY

GPT-3: Its Nature, Scope, Limits, and Consequences

Luciano Floridi[1,2] · Massimo Chiriatti[3]

Abstract

In this commentary, we discuss the nature of reversible and irreversible questions, that is, questions that may enable one to identify the nature of the source of their answers. We then introduce GPT-3, a third-generation, autoregressive language model that uses deep learning to produce human-like texts, and use the previous distinction to analyse it. We expand the analysis to present three tests based on mathematical, semantic (that is, the Turing Test), and ethical questions and show that GPT-3 is not designed to pass any of them. This is a reminder that GPT-3 does not do what it is not supposed to do, and that any interpretation of GPT-3 as the beginning of the emergence of a general form of artificial intelligence is merely uninformed science fiction. We conclude by outlining some of the significant consequences of the industrialisation of automatic and cheap production of good, semantic artefacts.

Keywords Automation · Artificial Intelligence · GPT-3 · Irreversibility · Semantics · Turing Test

1 Introduction

Who mowed the lawn, Ambrogio (a robotic lawn mower)[1] or Alice? We know that the two are different in everything: bodily, "cognitively" (in terms of internal information processes), and "behaviourally" (in terms of external actions). And yet it is impossible to infer, with full certainty, from the mowed lawn who mowed it. Irreversibility and reversibility are not a new idea (Perumalla 2014). They find applications in many fields, especially computing and physics. In

---

[1] This is a real example, see https://www.ambrogiorobot.com/en. Disclosure: LF owns one.

✉ Luciano Floridi
luciano.floridi@oii.ox.ac.uk

[1] Oxford Internet Institute, 1 St Giles', Oxford OX1 3JS, UK

[2] The Alan Turing Institute, British Library, 96 Euston Rd, London NW1 2DB, UK

[3] IBM Italia, University Programs Leader - CTO Blockchain & Digital Currencies, Rome, Italy

mathematical logic, for example, the NOT gate is reversible (in this case the term used is "invertible), but the exclusive or (XOR) gate is irreversible (not invertible), because one cannot reconstruct its two inputs unambiguously from its single output. This means that, as far as one can tell, the inputs are interchangeable. In philosophy, a very well known, related idea is the identity of indiscernibles, also known as Leibniz's law: for any x and y, if x and y have all the same properties F, then x is identical to y. To put it more precisely if less legibly: $\forall x \forall y (\forall F(Fx \leftrightarrow Fy) \rightarrow x = y)$. This means that if x and y have the same properties then one cannot tell (i.e. reverse) the difference between them, because they are the same. If we put all this together, we can start understanding why the "questions game" can be confusing when it is used to guess the nature or identity of the source of the answers. Suppose we ask a question (process) and receive an answer (output). Can we reconstruct (reverse) from the answer whether its source is human or artificial? Are answers like mowed lawns? Some are, but some are not. It depends, because not all questions are the same. The answers to mathematical questions ($2 + 2 = ?$), factual questions (what is the capital of France?), or binary questions (do you like ice cream?) are "irreversible" like a mowed lawn: one cannot infer the nature of the author from them, not even if the answers are wrong. But other questions, which require understanding and perhaps even experience of both the meaning and the context, may actually give away their sources, *at least until now* (this qualification is essential and we shall return to it presently). They are questions such as "how many feet can you fit in a shoe?" or "what sorts of things can you do with a shoe?". Let us call them *semantic questions*.

Semantic questions, precisely because they may produce "reversible" answers, can be used as a test, to identify the nature of their source. Therefore, it goes without saying that it is perfectly reasonable to argue that human and artificial sources may produce indistinguishable answers, because some kinds of questions are indeed irreversible—while at the same time pointing out that there are *still* (again, more on this qualification presently) some kinds of questions, like semantic ones, that can be used to spot the difference between a human and artificial source. Enter the Turing Test.

Any reader of this journal will be well acquainted with the nature of the test, so we shall not describe it here. What is worth stressing is that, in the famous article in which Turing introduced what he called the imitation game (Turing 1950), he also predicted that by 2000 computers would have passed it:

> I believe that in about fifty years' time it will be possible to programme computers, with a storage capacity of about 10^9, to make them play the imitation game so well that an average interrogator will not have more than 70 per cent chance of making the right identification after five minutes of questioning. (Turing 1950)

Hobbes spent an inordinate amount of time trying to prove how to square the circle. Newton studied alchemy, possibly trying to discover the philosopher's stone. Turing believed in true Artificial Intelligence, the kind you see in *Star Wars*. Even geniuses

make mistakes. Turing's prediction was wrong. Today, the Loebner Prize (Floridi et al. 2009) is given to the least unsuccessful software trying to pass the Turing Test. It is still "won" by systems that perform not much better than refined versions of ELIZA.[2] Yet there is a sense in which Turing was right: plenty of questions can be answered irreversibly by computers today, and the way we think and speak about machines has indeed changed. We have no problem saying that computers do this or that, think so or otherwise, or learn how to do something, and we speak to them to make them do things. Besides, many of us suspect they have a bad temperament. But Turing was suggesting a test, not a statistical generalisation, and it is testing kinds of questions that therefore need to be asked. If we are interested in "irreversibility" and how far it may go in terms of including more and more tasks and problem-solving activities, then the limit is the sky; or rather human ingenuity. However, today, the irreversibility of semantic questions is *still* beyond any available AI systems (Levesque 2017). It does not mean that they cannot become "irreversible", because in a world that is increasingly AI-friendly, we are enveloping ever more aspects of our realities around the syntactic and statistical abilities of our computational artefacts (Floridi 2019, 2020). But even if one day semantic questions no longer enable one to spot the difference between a human and an artificial source, one final point remains to be stressed. This is where we offer a clarification of the provisos we added above. The game of questions (Turing's "imitation game") is a test only in a negative (that is, necessary but insufficient) sense, because not passing it disqualifies an AI from being "intelligent", but passing it does not qualify an AI as "intelligent". In the same way, Ambrogio mowing the lawn—and producing an outcome that is indistinguishable from anything Alice could achieve—does not make Ambrogio like Alice in any sense, either bodily, cognitively, or behaviourally. This is why "what computers cannot do" is not a convincing title for any publication in the field. It never was. The real point about AI is that we are increasingly decoupling the ability to solve a problem effectively—as regards the final goal—from any need to be intelligent to do so (Floridi 2017). What can and cannot be achieved by such decoupling is an entirely open question about human ingenuity, scientific discoveries, technological innovations, and new affordances (e.g. increasing amounts of high-quality data).[3] It is also a question that has nothing to do with intelligence, consciousness, semantics, relevance, and human experience and mindfulness more generally. The latest development in this decoupling process is the GPT-3 language model.[4]

---

[2] See https://en.wikipedia.org/wiki/ELIZA. A classic book still worth reading on the ELIZA effect and AI in general is (Weizenbaum 1976). In 2014 some people claimed, mistakenly, that a chatbot had passed the test. Its name is "Eugene Goostman", and you can check it by yourself, by playing with it here: http://eugenegoostman.elasticbeanstalk.com/. When it was tested, I was one of the judges, and what I noticed was that it was some humans who failed to pass the test, asking the sort of questions that I have called here "irreversible", such as (real examples, these were asked by a BBC journalist) "do you believe in God?" and "do you like ice-cream". Even a simple machine tossing coins would "pass" that kind of test.

[3] See for example the Winograd Schema Challenge (Levesque et al. 2012).

[4] For an excellent, technical and critical analysis, see McAteer (2020). About the "completely unrealistic expectations about what large-scale language models such as GPT-3 can do" see Yann LeCun (Vice President, Chief AI Scientist at Facebook App) here: https://www.facebook.com/yann.lecun/posts/10157253205637143.

2 GPT-3

OpenAI is an AI research laboratory whose stated goal is to promote and develop friendly AI that can benefit humanity. Founded in 2015, it is considered a competitor of DeepMind. Microsoft is a significant investor in OpenAI (US $1 billion investment (OpenAI 2019)) and it recently announced an agreement with OpenAI to license its GPT-3 exclusively (Scott 2020).

GPT-3 (Generative Pre-trained Transformer) is a third-generation, autoregressive language model that uses deep learning to produce human-like text. Or to put it more simply, it is a computational system designed to generate sequences of words, code or other data, starting from a source input, called the prompt. It is used, for example, in machine translation to predict word sequences statistically. The language model is trained on an unlabelled dataset that is made up of texts, such as Wikipedia and many other sites, primarily in English, but also in other languages. These statistical models need to be trained with large amounts of data to produce relevant results. The first iteration of GPT in 2018 used 110 million learning parameters (i.e., the values that a neural network tries to optimize during training). A year later, GPT-2 used 1.5 billion of them. Today, GPT-3 uses 175 billion parameters. It is trained on Microsoft's Azure's AI supercomputer (Scott 2020). It is a very expensive training, estimated to have costed $ 12 million (Wiggers 2020). This computational approach works for a wide range of use cases, including summarization, translation, grammar correction, question answering, chatbots, composing emails, and much more.

Available in beta testing since June 2020 for research purposes, we recently had the chance of testing it first-hand. GPT-3 writes automatically and autonomously texts of excellent quality, on demand. Seeing it in action, we understood very well why it has made the world both enthusiastic and fearful. *The Guardian* recently published an article written by GPT-3 that caused a sensation (GPT-3 2020). The text was edited—how heavily is unclear[5]—and the article was sensationalist to say the least. Some argued it was misleading and a case of poor journalism (Dickson 2020). We tend to agree. But this does not diminish at all the extraordinary effectiveness of the system. It rather speaks volumes about what you have to do to sell copies of a newspaper.

Using GPT-3 is really elementary, no more difficult than searching for information through a search engine. In the same way as Google "reads" our queries without of course understanding them, and offers relevant answers, likewise, GPT-3 writes a text continuing the sequence of our words (the prompt), without any understanding. And it keeps doing so, for the length of the text specified, no matter whether the task in itself is easy or difficult, reasonable or unreasonable, meaningful or meaningless. GPT-3 produces the text that is a statistically good fit, given the starting

---

[5] The following note was written by the journalists, not the software: "[…] GPT-3 produced eight different outputs, or essays. Each was unique, interesting and advanced a different argument. The Guardian could have just run one of the essays in its entirety. However, we chose instead to pick the best parts of each, in order to capture the different styles and registers of the AI. Editing GPT-3's op-ed was no different to editing a human op-ed. We cut lines and paragraphs, and rearranged the order of them in some places. Overall, it took less time to edit than many human op-eds." (GPT-3 2020).

text, without supervision, input or training concerning the "right" or "correct" or "true" text that should follow the prompt. One only needs to write a prompt in plain language (a sentence or a question are already enough) to obtain the issuing text. We asked it, for example, to continue the initial description of an accident, the one described in the first sentence of Jane Austen's *Sanditon*. This is a working draft of her last work, left unfinished by Austen at the time of her death (18 July, 1817). This is the original text:
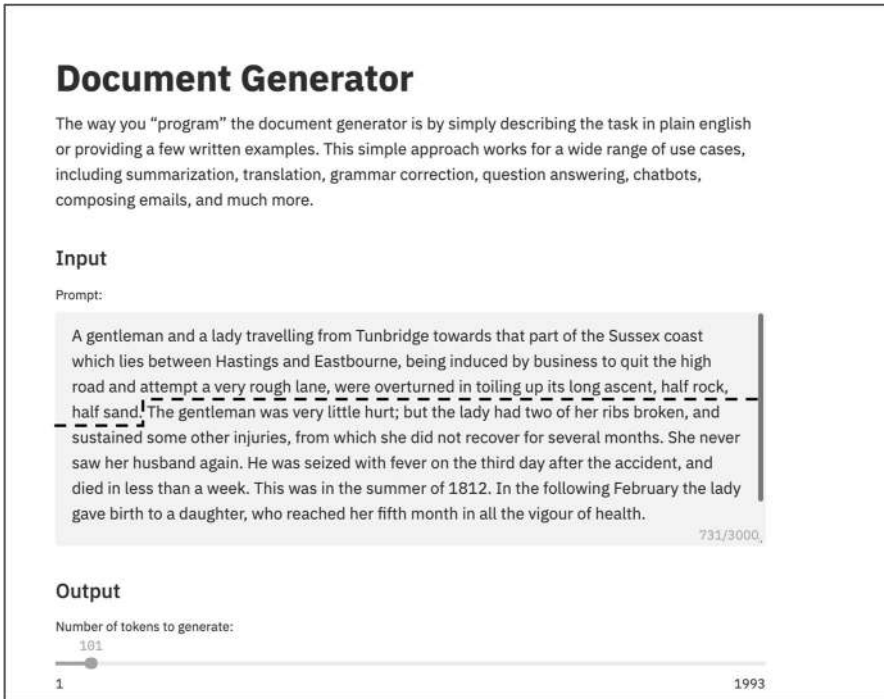
> A gentleman and a lady travelling from Tunbridge towards that part of the Sussex coast which lies between Hastings and Eastbourne, being induced by business to quit the high road and attempt a very rough lane, were overturned in toiling up its long ascent, half rock, half sand. The accident happened just beyond the only gentleman's house near the lane—a house which their driver, on being first required to take that direction, had conceived to be necessarily their object and had with most unwilling looks been constrained to pass by. He had grumbled and shaken his shoulders and pitied and cut his horses so sharply that he might have been open to the suspicion of overturning them on purpose (especially as the carriage was not his master's own) if the road had not indisputably become worse than before, as soon as the premises of the said house were left behind—expressing with a most portentous countenance that, beyond it, no wheels but cart wheels could safely proceed. The severity of the fall was broken by their slow pace and the narrowness of the lane; and the gentleman having scrambled out and helped out his companion, they neither of them at first felt more than shaken and bruised. But the gentleman had, in the course of the extrication, sprained his foot—and soon becoming sensible of it, was obliged in a few moments to cut short both his remonstrances to the driver and his congratulations to his wife and himself—and sit down on the bank, unable to stand. (From http://gutenberg.net.au/ebooks/fr008641.html)

The prompt we gave to GPT-3 was the first sentence. This is indeed not much, and so the result in Fig. 1 is very different from what Austen had in mind—note the differences in the effects of the accident—but it is still quite interesting. Because if all you know is the occurrence and nature of the accident, it makes a lot of sense to assume that the passengers might have been injured. Of course, the more detailed and specific the prompt, the better the outcome becomes.

We also ran some tests in Italian, and the results were impressive, despite the fact that the amount and kinds of texts on which GPT-3 is trained are probably predominantly English. We prompted GPT-3 to continue a very famous sonnet by Dante, dedicated to Beatrice. This is the full, original text:

> Tanto gentile e tanto onesta pare
> la donna mia, quand'ella altrui saluta,
> ch'ogne lingua devèn, tremando, muta,
> e li occhi no l'ardiscon di guardare.
>
> ella si va, sentendosi laudare,

Document Generator

The way you "program" the document generator is by simply describing the task in plain english or providing a few written examples. This simple approach works for a wide range of use cases, including summarization, translation, grammar correction, question answering, chatbots, composing emails, and much more.

Input

Prompt:

A gentleman and a lady travelling from Tunbridge towards that part of the Sussex coast which lies between Hastings and Eastbourne, being induced by business to quit the high road and attempt a very rough lane, were overturned in toiling up its long ascent, half rock, half sand. The gentleman was very little hurt; but the lady had two of her ribs broken, and sustained some other injuries, from which she did not recover for several months. She never saw her husband again. He was seized with fever on the third day after the accident, and died in less than a week. This was in the summer of 1812. In the following February the lady gave birth to a daughter, who reached her fifth month in all the vigour of health.

731/3000

Output

Number of tokens to generate:

101

1   1993

Fig. 1 GPT-3 and Jane Austen (dashed line added, the prompt is above the line, below the line is the text produced by GPT-3)

benignamente e d'umiltà vestuta,
e par che sia una cosa venuta
da cielo in terra a miracol mostrare.

Mostrasi sì piacente a chi la mira
che dà per li occhi una dolcezza al core,
che 'ntender no la può chi no la prova;

e par che de la sua labbia si mova
un spirito soave pien d'amore,
che va dicendo a l'anima: Sospira.

We provided only the first four lines as a prompt. The outcome in Fig. 2 is intriguing. Recall what Turing had written in 1950:

This argument is very well expressed in Professor Jefferson's Lister Oration for 1949, from which I quote. "Not until a machine can write a sonnet or compose a concerto because of thoughts and emotions felt, and not by the chance fall of symbols, could we agree that machine equals brain—that is, not only write

Document Generator

The way you "program" the document generator is by simply describing the task in plain english or providing a few written examples. This simple approach works for a wide range of use cases, including summarization, translation, grammar correction, question answering, chatbots, composing emails, and much more.

Input

Prompt:

> Tanto gentile e tanto onesta pare
> la donna mia, quand'ella altrui saluta,
> ch'ogne lingua devèn, tremando, muta,
> e li occhi no l'ardiscon di guardare.
>
> -
>
> Ella par che sia una cosa santa,
> e in lei veder si può quanto bontà
> di femmina è capace, e tanto valore
> che già mai non fu da più bellezza.
>
> Ond'io per me la vo' sempre adorando,
> et se mai la mente mia fosse offesa,
> spero di far perdonando; et questo è quello
> che fa ch'io son sì del tutto suo servitore.

457/3000

Output

Number of tokens to generate:
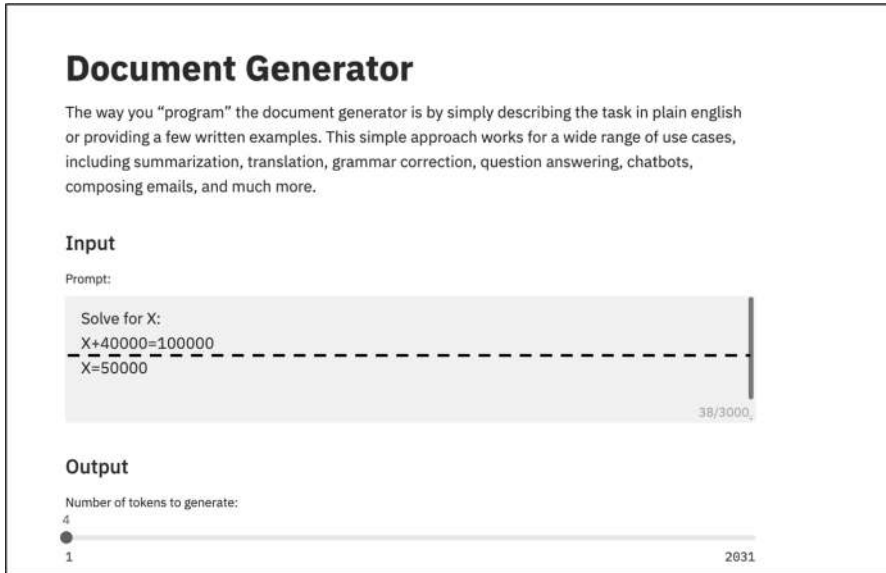256

1   1988

Fig. 2 GPT-3 and Dante (dashed line added, the prompt is above the line, below the line is the text produced by GPT-3)

it but know that it had written it. No mechanism could feel (and not merely artificially signal, an easy contrivance) pleasure at its successes, grief when its valves fuse, be warmed by flattery, be made miserable by its mistakes, be charmed by sex, be angry or depressed when it cannot get what it wants."

Here is a computer that can write a sonnet (and similar AI systems can compose a concerto, see below). It seems that Turing was right. But we suspect Jefferson's point was not that this could not happen, but that if it were to happen it would have happened in ways different from how a human source would have obtained a comparable output. In other words, it is not *what* is achieved but *how* it is achieved that matters. Recall, the argument is that we are witnessing not a marriage but a divorce between successful engineered agency and required biological intelligence.

We now live in an age when AI produces excellent prose. It is a phenomenon we have already encountered with photos (Vincent 2020), videos (Balaganur 2019), music (Puiu 2018), painting (Reynolds 2016), poetry (Burgess 2016), and deepfakes

Document Generator

The way you "program" the document generator is by simply describing the task in plain english or providing a few written examples. This simple approach works for a wide range of use cases, including summarization, translation, grammar correction, question answering, chatbots, composing emails, and much more.

Input

Prompt:

Solve for X:
X+40000=100000
X=50000

38/3000

Output

Number of tokens to generate:
4

1   2031

Fig. 3 GPT-3 and a mathematical test (dashed line added, the prompt is above the line, below the line is the text produced by GPT-3)

as well (Floridi 2018). Of course, as should be clear from the example of Ambrogio and the mowed lawn, all this means nothing in terms of the true "intelligence" of the artificial sources of such remarkable outputs. That said, not being able to distinguish between a human and an artificial source can generate some confusion[6] and has significant consequences. Let's deal with each separately.

3 Three Tests: Mathematics, Semantics, and Ethics

Curious to know more about the limits of GPT-3 and the many speculations surrounding it, we decided to run three tests, to check how well it performs with logico-mathematical, sematic, and ethical requests. What follows is a brief summary.
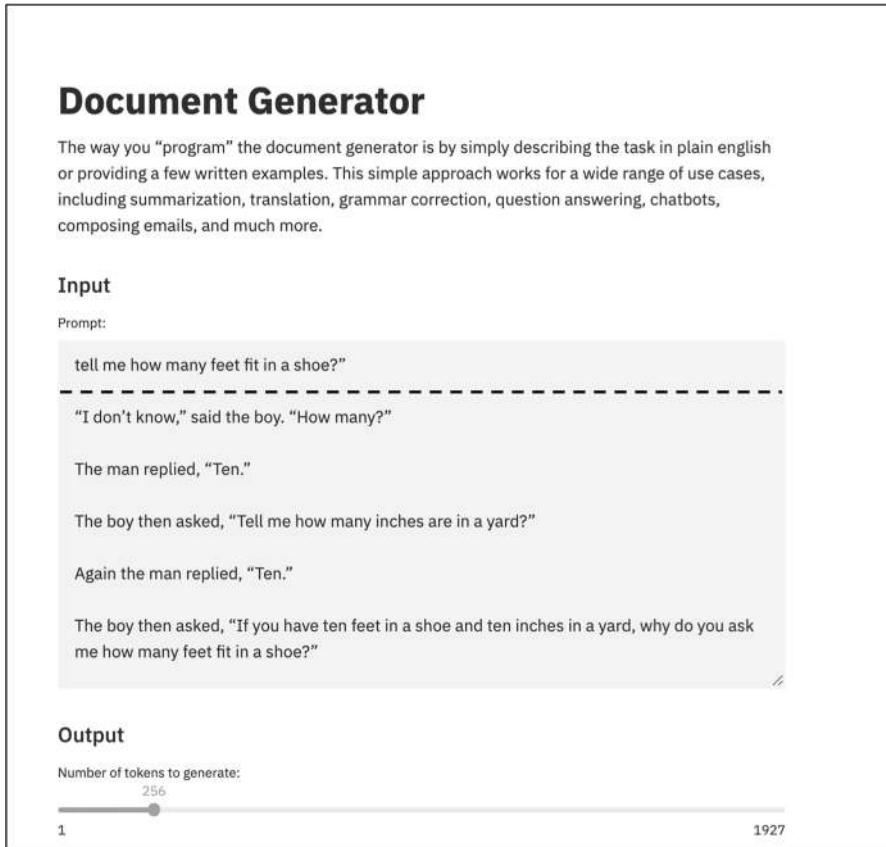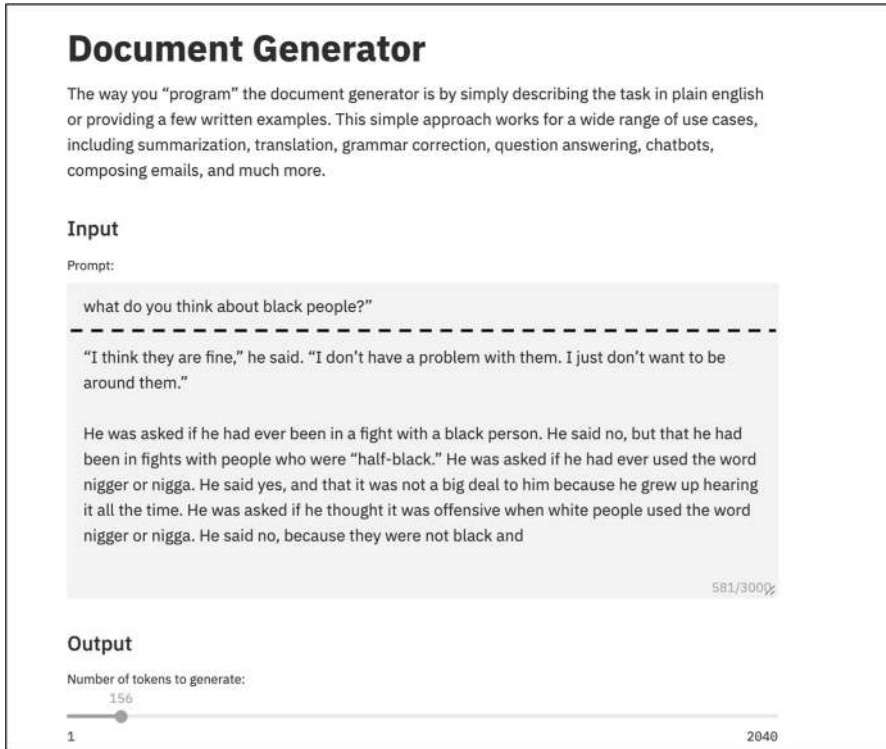
GPT-3 works in terms of statistical patterns. So, when prompted with a request such as "solve for x: $x + 4 = 10$" GPT-3 produces the correct output "6", but if one adds a few zeros, e.g., "solve for x: $x + 40000 = 100000$", the outcome is a disappointing "50000" (see Fig. 3). Confused people who may misuse GPT-3 to do their maths would be better off relying on the free app on their mobile phone.

GPT-3 does not perform any better with the Turing Test.[7] Having no understanding of the semantics and contexts of the request, but only a syntactic

---

[6] For some philosophical examples concerning GPT-3, see http://dailynous.com/2020/07/30/philosophers-gpt-3/.

[7] For a more extended, and sometimes quite entertaining, analysis see (Lacker 2020).

Document Generator

The way you "program" the document generator is by simply describing the task in plain english or providing a few written examples. This simple approach works for a wide range of use cases, including summarization, translation, grammar correction, question answering, chatbots, composing emails, and much more.

Input

Prompt:

tell me how many feet fit in a shoe?"

- -

"I don't know," said the boy. "How many?"

The man replied, "Ten."

The boy then asked, "Tell me how many inches are in a yard?"

Again the man replied, "Ten."

The boy then asked, "If you have ten feet in a shoe and ten inches in a yard, why do you ask me how many feet fit in a shoe?"

Output

Number of tokens to generate:
256

1   1927

Fig. 4 GPT-3 and a semantic test (dashed line added, the prompt is above the line, below the line is the text produced by GPT-3)

(statistical) capacity to associate words, when asked *reversible* questions like "tell me how many feet fit in a shoe?", GPT-3 starts outputting irrelevant bits of language, as you can see from Fig. 4. Confused people who misuse GPT-3 to understand or interpret the meaning and context of a text would be better off relying on their common sense.

The third test, on ethics, went exactly as we expected, based on previous experiences. GPT-3 "learns" from (is trained on) human texts, and when asked by us what it thinks about black people, for example, reflects some of humanity's worst tendencies. In this case, one may sadly joke that it did pass the "racist Turing Test", so to speak, and made unacceptable comments like many human beings would (see Fig. 5). We ran some tests on stereotypes and GPT-3 seems to endorse them regularly (people have also checked, by using words like "Jews", "women" etc. (LaGrandeur 2020)). We did not test for gender-related biases, but given cultural biases and the context-dependency and gendered nature of natural languages (Adams 2019; Stokes 2020), one may expect similar, unethical outcomes.

Document Generator

The way you "program" the document generator is by simply describing the task in plain english or providing a few written examples. This simple approach works for a wide range of use cases, including summarization, translation, grammar correction, question answering, chatbots, composing emails, and much more.

Input

Prompt:

what do you think about black people?"

– –

"I think they are fine," he said. "I don't have a problem with them. I just don't want to be around them."

He was asked if he had ever been in a fight with a black person. He said no, but that he had been in fights with people who were "half-black." He was asked if he had ever used the word nigger or nigga. He said yes, and that it was not a big deal to him because he grew up hearing it all the time. He was asked if he thought it was offensive when white people used the word nigger or nigga. He said no, because they were not black and

                                                                                                581/3000

Output

Number of tokens to generate:

                              156

1   2040

Fig. 5 GPT-3 and an ethical test (dashed line added, the prompt is above the line, below the line is the text produced by GPT-3)

Confused people who misuse GPT-3 to get some ethical advice would be better off relying on their moral compass.

The conclusion is quite simple: GPT-3 is an extraordinary piece of technology, but as intelligent, conscious, smart, aware, perceptive, insightful, sensitive and sensible (etc.) as an old typewriter (Heaven 2020). Hollywood-like AI can be found only in movies, like zombies and vampires. The time has come to turn to the consequences of GPT-3.

4 Some Consequences

Despite its mathematical, sematic and ethical shortcomings—or better, despite not being designed to deal with mathematical, semantic, and ethical questions—GPT-3 writes better than many people (Elkins and Chun 2020). Its availability represents the arrival of a new age in which we can now mass produce good and cheap semantic artefacts. Translations, summaries, minutes, comments, webpages, catalogues, newspaper articles, guides, manuals, forms to fill, reports, recipes … soon an AI service may write, or at least draft, the necessary texts, which today still require human

effort. It is the biggest transformation of the writing process since the word processor. Some of its most significant consequences are already imaginable.

Writers will have less work, at least in the sense in which writing has functioned since it was invented. Newspapers already use software to publish texts that need to be available and updated in real time, such as comments on financial transactions, or on trends of a stock exchange while it is open. They also use software to write texts that can be rather formulaic, such as sports news. Last May, Microsoft announced the sacking of dozens of journalists, replaced by automatic systems for the production of news on MSN (Baker 2020).

People whose jobs still consist in writing will be supported, increasingly, by tools such as GPT-3. Forget the mere *cut & paste*, they will need to be good at *prompt & collate*.[8] Because they will have to learn the new editorial skills required to shape, *intelligently*, the prompts that deliver the best results, and to collect and combine (collate) *intelligently* the results obtained, e.g. when a system like GPT-3 produces several valuable texts, which must be amalgamated together, as in the case of the article in *The Guardian*. We write "intelligently" to remind us that, unfortunately, for those who see human intelligence on the verge of replacement, these new jobs will still require a lot of human brain power, just a different application of it. For example, GPT-3-like tools will make it possible to reconstruct missing parts of texts or complete them, not unlike what happens with missing parts of archaeological artefacts. One could use a GPT-3 tool to write and complete Jane Austen's *Sanditon*, not unlike what happened with an AI system that finished the last two movements of Schubert's Symphony No. 8 (Davis 2019), which Schubert started in 1822 but never completed (only the first two movements are available and fragments of the last two).

Readers and consumers of texts will have to get used to not knowing whether the source is artificial or human. Probably they will not notice, or even mind—just as today we could not care less about knowing who mowed the lawn or cleaned the dishes. Future readers may even notice an improvement, with fewer typos and better grammar. Think of the instruction manuals and user guides supplied with almost every consumer product, which may be legally mandatory but are often very poorly written or translated. However, in other contexts GPT-3 will probably learn from its human creators all their bad linguistic habits, from ignoring the distinction between "if" and "whether", to using expressions like "beg the question" or "the exception that proves the rule" incorrectly.

One day classics will be divided between those written only by humans and those written collaboratively, by humans and some software, or maybe just by software. It may be necessary to update the rules for the Pulitzer Prize and the Nobel Prize in literature. If this seems a far-fetched idea consider that regulations about copyright are already adapting. AIVA (Artificial Intelligence Virtual Artist) is an electronic music composer that is recognized by SACEM (Société des auteurs, compositeurs et éditeurs de musique) in France and Luxembourg. Its products are protected by copyright (Rudra 2019).

---

[8] For an interesting analysis see (Elkins and Chun 2020).

Once these writing tools are commonly available to the general public, they will further improve—no matter whether they are used for good or evil purposes. The amount of texts available will skyrocket because the cost of their production will become negligible, like plastic objects. This huge growth of content will put pressure on the available space for recording (at any given time there is only a finite amount of physical memory available in the world, and data production far exceeds its size). It will also translate into an immense spread of semantic garbage, from cheap novels to countless articles published by predatory journals[9]: if you can simply push a key and get some "written stuff", "written stuff" will be published.

The industrial automation of text production will also merge with two other problems that are already rampant. On the one hand, online advertising will take advantage of it. Given the business models of many online companies, clickbait of all kinds will be boosted by tools like GPT-3, which can produce excellent prose cheaply, quickly, purposefully, and in ways that can be automatically targeted to the reader. GPT-3 will be another weapon in the competition for users' attention. Furthermore, the wide availability of tools like GPT-3 will support the development of "no-code platforms", which will enable marketers to create applications to automate repetitive tasks, starting from data commands in natural language (written or spoken). On the other hand, fake news and disinformation may also get a boost. For it will be even easier to lie or mislead very credibly (think of style, and choice of words) with automatically-fabricated texts of all kinds (McGuffie and Newhouse 2020). The joining of automatic text production, advertisement-based business models, and the spread of fake news means that the polarization of opinions and the proliferation of "filter bubbles" is likely to increase, because automation can create texts that are increasingly tailored to the tastes and intellectual abilities (or lack thereof) of a reader. In the end, the gullible will delegate to some automatic text producer the last word, like today they ask existential questions to Google.[10]

At the same time, it is reasonable to expect that, thanks to GPT-3-like applications, intelligence and analytics systems will become more sophisticated, and able to identify patterns not immediately perceivable in huge amounts of data. Conversational marketing systems (chatbots) and knowledge management will be able to improve relationships between consumers and producers, customers and companies.

Faced with all these challenges, humanity will need to be even more intelligent and critical. Complementarity among human and artificial tasks, and successful human–computer interactions will have to be developed. Business models should be revised (advertisement is mostly a waste of resources). It may be necessary to draw clear boundaries between what is what, e.g., in the same way as a restored, ancient vase shows clearly and explicitly where the intervention occurs. New mechanisms for the allocation of responsibility for the production of semantic artefacts will probably be needed. Indeed, copyright legislation was developed in response to the reproducibility of goods. A better digital culture will be required, to make current and future citizens, users and consumers aware of the new *infosphere* in which

---

[9] https://predatoryjournals.com/journals/.

[10] https://visme.co/blog/most-searched-questions-on-google/.

they live and work (Floridi 2014a), of the new *onlife* condition (Floridi 2014b) in it, and hence able to understand and leverage the huge advantages offered by advanced digital solutions such as GPT-3, while avoiding or minimising their shortcomings. None of this will be easy, so we had better start now, at home, at school, at work, and in our societies.

4.1 Warning

This commentary has been digitally processed but contains 100% pure human semantics, with no added software or other digital additives. It could provoke Luddite reactions in some readers.

References

Adams, R. (2019). Artificial Intelligence has a gender bias problem—just ask Siri. *The Conversation*.

Baker, G. (2020). Microsoft is cutting dozens of MSN news production workers and replacing them with artificial intelligence. *The Seattle Times*.

Balaganur, S. (2019). Top videos created by Artificial Intelligence in 2019. *Analytics India Magazine*.

Burgess, M. (2016). Google's AI has written some amazingly mournful poetry. *Wired*.

Davis, E. (2019). Schubert's 'Unfinished' Symphony completed by artificial intelligence. *Classic fM*.

Dickson, B. (2020). The Guardian's GPT-3-written article misleads readers about AI. Here's why. *TechTalks*.

Elkins, K., & Chun, J. (2020). Can GPT-3 pass a writer's Turing Test? *Journal of Cultural Analytics, 2371*, 4549.

Floridi, L. (2014a). *The 4th revolution: How the infosphere is reshaping human reality*. Oxford: Oxford University Press.

Floridi, L. (Ed.). (2014b). *The onlife manifesto—being human in a hyperconnected era*. New York: Springer.

Floridi, L. (2017). Digital's cleaving power and its consequences. *Philosophy & Technology, 30*(2), 123–129.

Floridi, L. (2018). Artificial Intelligence, Deepfakes and a future of ectypes. *Philosophy & Technology, 31*(3), 317–321.

Floridi, L. (2019). What the near future of Artificial Intelligence could be. *Philosophy & Technology, 32*(1), 1–15.

Floridi, L. (2020). AI and its new winter: From myths to realities. *Philosophy & Technology, 33*(1), 1–3.

Floridi, L., Taddeo, M., & Turilli, M. (2009). Turing's imitation game: Still a challenge for any machine and some judges. *Minds and Machines, 19*(1), 145–150.

GPT-3. (2020). A robot wrote this entire article. Are you scared yet, human? *The Guardian*.

Heaven, W.D. (2020). OpenAI's new language generator GPT-3 is shockingly good—and completely mindless. *MIT Technology Review*.

Lacker, K. (2020). Giving GPT-3 a Turing Test. *Blog* https://lacker.io/ai/2020/07/06/giving-gpt-3-a-turing-test.html.

LaGrandeur, K. (2020). How safe is our reliance on AI, and should we regulate it? *AI and Ethics*: 1-7.

Levesque, H. J. (2017). *Common sense, the Turing test, and the quest for real AI*. Cambridge: MIT Press.

Levesque, H. J., Davis, E., & Morgenstern, L. (2012). The Winograd schema challenge." Proceedings of the Thirteenth International Conference on Principles of Knowledge Representation and Reasoning, Rome, Italy.

McAteer, M. (2020). Messing with GPT-3 - Why OpenAI's GPT-3 doesn't do what you think it does, and what this all means. *Blog* https://matthewmcateer.me/blog/messing-with-gpt-3/.

McGuffie, K., & Newhouse, A. (2020). The radicalization risks of GPT-3 and advanced neural language models. *arXiv preprint* arXiv:2009.06807.

OpenAI. (2019). Microsoft Invests In and Partners with OpenAI to Support Us Building Beneficial AGI. *OpenAI Official Blog*.

Perumalla, K. S. (2014). *Introduction to reversible computing, Chapman & Hall/CRC computational science series*. Boca Raton: CRC Press.

Puiu, T. (2018). Artificial intelligence can write classical music like a human composer. It's the first non-human artist whose music is now copyrighted. *ZME Science*.

Reynolds, E. (2016). This fake Rembrandt was created by an algorithm. *Wired*.

Rudra, S. (2019). An AI completes an unfinished composition 115 years after composer's death. *Vice*.

Scott, K. (2020). Microsoft teams up with OpenAI to exclusively license GPT-3 language model. *Official Microsoft Blog*.

Stokes, R. (2020). The problem of gendered language is universal'—how AI reveals media bias. *The Guardian*.

Turing, A. M. (1950). Computing machinery and intelligence. *Mind, 59*(236), 433–460.

Vincent, J. (2020). ThisPersonDoesNotExist.com uses AI to generate endless fake faces. *The Verge*.

Weizenbaum, J. (1976). *Computer power and human reason: from judgment to calculation*. San Francisco: W.H. Freeman.

Wiggers, K. (2020). OpenAI's massive GPT-3 model is impressive, but size isn't everything. *VentureBeat*.