The Book of Endless History: Authorial use of GPT2 for Interactive Storytelling

John Austin

A Stranger Gravity, San Francisco, CA 94107
http://astrangergravity.com
http://johnaustin.io :: @kleptine

Abstract. We present *The Book of Endless History*, an infinite Wikipedia of fantasy stories written in the style of Borges and Calvino, exploring the use of structural conditioning on GPT2 to generate text with explicit subjects and embedded web-links. Users are presented with a Wikipedia-like interface, containing a short fantasy description of the topic and containing embedded web-links to other related subjects. Users may click on links to learn more about different topics, following an endless trail of generated pages. The GPT2 architecture is a text completion model – it has no explicit understanding of structure and it can be a challenge to integrate it with authorial intent. Nevertheless, through this work we show that it can be conditioned to learn to write *about* a subject and additionally to generate the topology for an encyclopedia. We refer to this technique as subject conditioning, or more generally, *structural conditioning*.

Keywords: Structural Conditioning · GPT2 · Interactive Storytelling

1 Introduction

The GPT2 deep learning architecture, released in February 2019 [3], has made massive strides in the fields of general text generation and analysis. GPT2 is text-completion model, able to generate long-form, complex blocks of text given only a sample prefix. Recently, news articles generated from the largest GPT2 models were found to be at least as believable as equivalent human-written articles, and in some cases, *more* believable [4]. Further, and of interest in the narrative generation domain, the model architecture is able to recall information about the dataset it is trained on. As the primary dataset was taken from a random walk of Reddit, the model is able to generate stories about figures such as Luke Skywalker, Hillary Clinton, Gandalf, and even pull in related concepts such as The Shire without additional prompting.

A primary strength of GPT2 is the generality of the approach: the model is trained simply to produce the most probable next word, given a sequence of tokens. Because of this, GPT2 can even be formulated to perform *non-generative* text analysis tasks such as question-answering, translation, and summarization. The trick is to formulate these tasks as text completion tasks, providing them in "Question: Answer" format. Although the performance of GPT2 on these tasks

compared to other specialized approaches is not state of the art, the generality of the system is staggering.

In the narrative generative domain, traditionally the properties of GPT2 are difficult to control. It generates incredibly realistic text, but because it is primarily a text-completion system, it lacks any degree of authorship beyond allowing the user to provide a prefix. Even when prompted with a clear prefix, it can be challenging to compel the model to talk *about* the intended subject, rather than completing valid but unrelated text.

With *The Book of Endless History*, we present an approach we refer to as subject conditioning, or more generally *structural conditioning*, making use of the unique non-generative flexibility noted above. By transforming our dataset into a "{subject}\n text" representation, we are able to condition the model to expect all text following a line containing the "{subject}" syntax to discuss that subject. We use an off-the-shelf POS tagger and Python to perform this transformation.

Additionally, we layer another structure into our generator: links out. By surrounding *all* named entities in our dataset with square brackets, GPT2 learns generally to place brackets around subjects. After training, all returned text contains a variety of phrases and words 'linked' by GPT2 (see Fig. 2). In the book, we show these as clickable web-links, which query the book recursively for the linked page.

With the combination of these two techniques we are able to create an infinite fantasy Wikipedia of sorts: one where every page is generated, and the topology of the encyclopedia is generated as well. Clicking on a link within the encyclopedia takes a user to a generated page for that topic, and presents them with another set of links to choose from. The result of a random walk through the pages of the book is a bit like a deep-dream[2] experience, where the result of the generator is fed back into the generator itself.

2   Approach

Our approach consists of two steps: link conditioning and subject conditioning. To create a model that generates text with links, we first collect the complete works of Jorge Luis Borges and Italo Calvino in text format and transform this dataset using the off-the-shelf tool, SpaCy[1], to identify named entities and surround them with square brackets. We fine-tune the 345 million parameter version of GPT2 on this data[1]. This results in a model that generates fairly interesting prose in the style of Borges and Calvino, and additionally one that surrounds subjects within the text with brackets.

At this point we perform the subject transformation and training, but on a much smaller selection of books, selected for their focus on short, descriptive stories: *Invisible Cities*, *Book of Imaginary Beings*, *A Universal History of Infamy*, and *Labyrinths*. Again, we use SpaCy, selecting the first named entity in

---

[1] Approximately 4-6 hours on an NVidia T4, provided for free by Google Colab.

each line, and inserting it in curly braces on a new line above.[2] A sample of this final transformation can be see in Fig. 1.

We take this two pronged approach, because while we would like to copy the style of Borges and Calvino from their full body of works, many of their books are first person stories. Performing the subject conditioning on the full collection results in the model generating quite a bit of dialogue (which is undesirable for an encyclopedia), because the primary place of named entities in these stories occur in lines such as:

```
"The next time I kill you", replied [Scharlach].
```

On the other hand, four books is not nearly enough data to create a compelling and varied generator. The two-phase training allows us the best of both worlds. By training first on a large corpus we get improved style and generality. Then, by confining our subject conditioning to just the works that are written in a non-fiction or third person style, we push the generator to generate non-fiction as well.

A sample query on the final generated text can be seen below in Fig. 2

```
{Zora}
BEYOND SIX RIVERS and three mountain ranges rises [Zora], a city that ...
```

Fig. 1. A sample line of text in the final training dataset.

```
Query:      {Einstein}
Completes:  The most significant incident in Einstein's life occurred on
            a summer's day in 1869. He was riding his bicycle along the
            avenues of [Munich] when suddenly, in front of him, was a
            figure which he could never forget.  ...
```

Fig. 2. A sample result querying the book for "Einstein".

The Book of Endless History is presented online[3] in the format of an encyclopedia of the present day found by a hypothetical future society - one which ruminates on the nature of truth and history. As we move into an age where false text becomes easier to create than the real, there is somewhat of a dark irony to the nature of the book.

3   Conclusion

We hope that this project will inspire further research into the usage and authorial control of GPT2 and other deep learning architectures for storytelling.

---

[2] As you may notice, this causes our text to be double spaced, with subjects in between real lines. However, GPT2 manages to learn to ignore this structure incredibly quickly.

[3] http://bookofendlesshistory.com

While GPT2 may lack in consistency and structure, it makes up for it in creativity and prose. Correspondingly, structured approaches such as grammars and planners can be powerful tools for authorship, but are limited in their ability to generate eloquent text (or require unreasonable amounts of labor to do so). We believe that there is an ideal hybrid system: one in which an author may provide structural 'hints' at the level of detail they require and have a deep learning model fill in the detail. While this project applies this concept simplistically, we are eager to see similar results with more provided structure and larger models. With the recent acceleration of research on GPT2, we are excited to see the how the landscape of tools available to interactive story generators transforms over the next few years.

References

1. Honnibal, M., Montani, I.: spaCy 2: Natural language understanding with Bloom embeddings, convolutional neural networks and incremental parsing (2017), to appear
2. Mordvintsev, A., Olah, C., Tyka, M.: Inceptionism: Going deeper into neural networks. Blog (2015), `https://ai.googleblog.com/2015/06/inceptionism-going-deeper-into-neural.html`
3. Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I.: Language models are unsupervised multitask learners. OpenAI (2019)
4. Zellers, R., Holtzman, A., Rashkin, H., Bisk, Y., Farhadi, A., Roesner, F., Choi, Y.: Defending against neural fake news. CoRR **abs/1905.12616** (2019), `http://arxiv.org/abs/1905.12616`