PROCEDURAL STORY GENERATION USING TRANSFORMERS

Niall Dillane

Computer Science and Information Systems

University of Limerick

Limerick, Ireland

`13132911@studentmail.ul.ie`

James Patten

Computer Science and Information Systems

University of Limerick

Limerick, Ireland

`james.patten@ul.ie`

ABSTRACT

Procedural content generation (PCG) – the process of generating data algorithmically – is a technique that has applications across a variety of domains. In the research outlined in this report, focus is directed to the use of PCG as a means to generate novel-like stories. The challenge is twofold: research into procedural generation methods, as well as the structure and language of stories, addressing questions such as: "What are the elements of a good story?". Finally, methods to codify these elements must be investigated, in such a way that they can be used by a procedural generation technique, effectively combining the two disciplines.

*K*eywords procedural generation · algorithms · AI · language · neural networks · transformers

1 Introduction

In order to be able to automatically generate stories we first need to understand the structure of a story. Indeed, not only the structure that would make one grammatically correct, but also those devices, structures and styles that help make a good story. This is probably the largest challenge in the project, a formula which is still unclear after millenia of storytelling.

Furthermore, there have been significant advancements over recent years in the Natural Language Processing (NLP) field, with new architectures and technologies appearing and outpacing each other. Choosing a direction which is sufficiently developed but still relatively state of the art will be difficult, especially since few of these have specifically tackled the topic of novel-like stories, instead focusing more on news article or script style content.
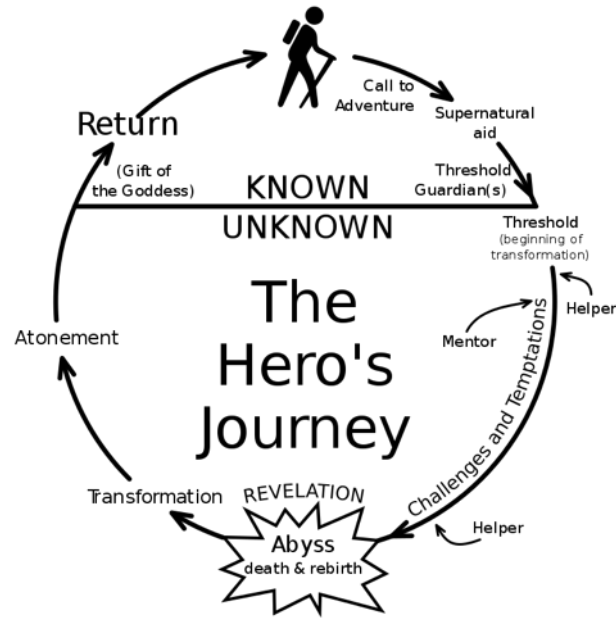
Figure 1: commons.wikimedia.org

However, it is this rapid progress and uncertain nature that make the topic so fascinating. There has been great progress in images, video, even games, but text generation has lagged behind somewhat. This signals that language is more nuanced, difficult to replicate, and still in its infancy.

2   Good Stories

2.1   Good Stories

Structure was the first element to be researched, as the highest level of a story planning process. Popular methods include the Hero's Journey (Campbell, 2008), which describes a cycle of sorts: a call to adventure, crossing from the known to the unknown, transformation etc. This is a more granular structure, most known for its application in Star Wars.

Another method is the three (or five) act structure (Trottier, 1998), wherein the story is broken down into distinct sections, typically: setup, confrontation, resolution. This is a bit more versatile, which was something to consider for the later steps of "filling in the gaps" with PCG. There was a balance to be struck: providing some structure so that the generation has some level of cogency, but also allowing flexibility so that not all stories are the same.

Following this, a tree-based approach became appealing, at the higher levels resembling more of a novel plan with plot points, and is eventually fleshed out below into sentences and phrases. This is a familiar concept in Computer Science, and it seems like the best way to maintain a coherent story. Some optional user interaction would be ideal – e.g. "Tell me a story about a dog with superpowers" – a kind of "seed". These could also be generated autonomously.

2.1.1 Headings: third level

Suspendisse vel felis. Ut lorem lorem, interdum eu, tincidunt sit amet, laoreet vitae, arcu. Aenean faucibus pede eu ante. Praesent enim elit, rutrum at, molestie non, nonummy vel, nisl. Ut lectus eros, malesuada sit amet, fermentum eu, sodales cursus, magna. Donec eu purus. Quisque vehicula, urna sed ultricies auctor, pede lorem egestas dui, et convallis elit erat sed nulla. Donec luctus. Curabitur et nunc. Aliquam dolor odio, commodo pretium, ultricies non, pharetra in, velit. Integer arcu est, nonummy in, fermentum faucibus, egestas vel, odio.

Paragraph The quick brown fox jumps over the lazy dog Sed commodo posuere pede. Mauris ut est. Ut quis purus. Sed ac odio. Sed vehicula hendrerit sem. Duis non odio. Morbi ut dui. Sed accumsan risus eget odio. In hac habitasse platea dictumst. Pellentesque non elit. Fusce sed justo eu urna porta tincidunt. Mauris felis odio, sollicitudin sed, volutpat a, ornare ac, erat. Morbi quis dolor. Donec pellentesque, erat ac sagittis semper, nunc dui lobortis purus, quis congue purus metus ultricies tellus. Proin et quam. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos hymenaeos. Praesent sapien turpis, fermentum vel, eleifend faucibus, vehicula eu, lacus.

3 Examples of citations, figures, tables, references

Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Donec odio elit, dictum in, hendrerit sit amet, egestas sed, leo. Praesent feugiat sapien aliquet odio. Integer vitae justo. Aliquam vestibulum fringilla lorem. Sed neque lectus, consectetuer at, consectetuer sed, eleifend ac, lectus. Nulla facilisi. Pellentesque eget lectus. Proin eu metus. Sed porttitor. In hac habitasse platea dictumst. Suspendisse eu lectus. Ut mi mi, lacinia sit amet, placerat et, mollis vitae, dui. Sed ante tellus, tristique ut, iaculis eu, malesuada ac, dui. Mauris nibh leo, facilisis non, adipiscing quis, ultrices a, dui.

The documentation for `natbib` may be found at

$$\texttt{http://mirrors.ctan.org/macros/latex/contrib/natbib/natnotes.pdf}$$

Of note is the command `\citet`, which produces citations appropriate for use in inline text. For example,

    \citet{hasselmo} investigated\dots

produces

   Hasselmo, et al. (1995) investigated...

$$\texttt{https://www.ctan.org/pkg/booktabs}$$

3.1 Figures

Suspendisse vitae elit. Aliquam arcu neque, ornare in, ullamcorper quis, commodo eu, libero. Fusce sagittis erat at erat tristique mollis. Maecenas sapien libero, molestie et, lobortis in, sodales eget, dui. Morbi ultrices
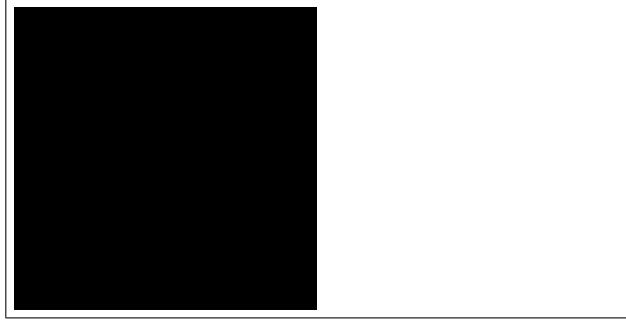
Figure 2: Sample figure caption.

Table 1: Sample table title

	Part		
Name	Description	Size (μm)
Dendrite	Input terminal	$\sim$100
Axon	Output terminal	$\sim$10
Soma	Cell body	up to 10^6

rutrum lorem. Nam elementum ullamcorper leo. Morbi dui. Aliquam sagittis. Nunc placerat. Pellentesque tristique sodales est. Maecenas imperdiet lacinia velit. Cras non urna. Morbi eros pede, suscipit ac, varius vel, egestas non, eros. Praesent malesuada, diam id pretium elementum, eros sem dictum tortor, vel consectetuer odio sem sed wisi. See Figure 2. Here is how you add footnotes. [1] Sed feugiat. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Ut pellentesque augue sed urna. Vestibulum diam eros, fringilla et, consectetuer eu, nonummy id, sapien. Nullam at lectus. In sagittis ultrices mauris. Curabitur malesuada erat sit amet massa. Fusce blandit. Aliquam erat volutpat. Aliquam euismod. Aenean vel lectus. Nunc imperdiet justo nec dolor.

3.2 Tables

Etiam euismod. Fusce facilisis lacinia dui. Suspendisse potenti. In mi erat, cursus id, nonummy sed, ullamcorper eget, sapien. Praesent pretium, magna in eleifend egestas, pede pede pretium lorem, quis consectetuer tortor sapien facilisis magna. Mauris quis magna varius nulla scelerisque imperdiet. Aliquam non quam. Aliquam porttitor quam a lacus. Praesent vel arcu ut tortor cursus volutpat. In vitae pede quis diam bibendum placerat. Fusce elementum convallis neque. Sed dolor orci, scelerisque ac, dapibus nec, ultricies ut, mi. Duis nec dui quis leo sagittis commodo. See awesome Table 1.

3.3 Lists

- Lorem ipsum dolor sit amet
- consectetur adipiscing elit.

---

[1]Sample of the first footnote.

- Aliquam dignissim blandit est, in dictum tortor gravida eget. In ac rutrum magna.

References

Campbell, J. (2008). *The hero with a thousand faces* (Vol. 17). New World Library.

Trottier, D. (1998). The screenwriter's bible: A complete guide to writing. *Formatting, and Selling Your Script*.