SUMMARISATION OF PLANNING RESPONSES WITH LLMS

Introduction

- Saves time; takes minutes rather than hours (or days)

- Reduces bias?

- All information can be considered equally.

- diverse forms of input; letters to brief comments, even handwritten text (gpt-4o).

- generate easy to understand summaries, removing any domain specific terminology. increased transparency etc.

- ensure the summaries are accurate, may require human oversight? For this we are using hallucination detection which works well.

Methodology

This project primarily considers the use of generative pre-trained transformer (GPT) large-language models (LLMs) for summarising planning responses. While *extractive* summarisation using encoder-transformer LLMs has been an established task for a number of years (e.g. with Google's BERT), only recently has this task been advanced through the use of larger scale GPT models (e.g. OpenAIs gpt-3/gpt-4 series). One benefit of these new models are their size; they are both *trained* on more human data, and have a larger number of *model parameters*. Both of these factors mean that such models are able to understand human text and semantic nuances to a greater degree. Additionally, their architectural differences mean that, while BERT-like models excel at extractive summarisation, GPT models are able to generate large amounts of human-like text.

Given these advances, a number of methods relating to document summarisation have been established in recent years (or months). In this project, we focus on the task of *map-reduce* summarisation; given a large set of documents, summarise each, then summarise those summaries to produce a final report.

For our use-case we established the following data-flow;
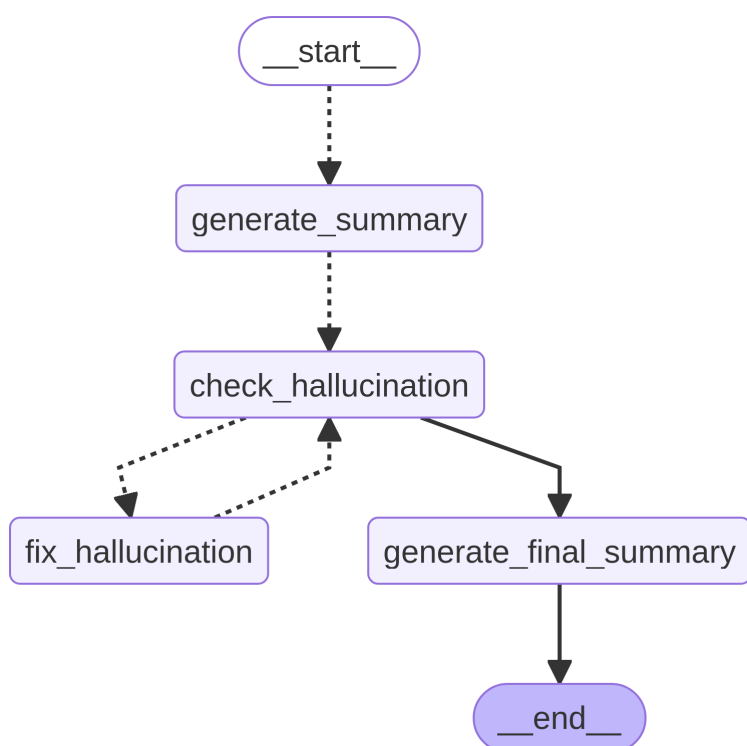
1. Summaries for each response are generated in parallel
2. Each summary is check to ensure there are no *hallucinations* (cyclically)

3. Summaries are used to form a final report

Hallucination removal

Hallucination detection and removal is a key part of this process. While summarisation typically accurately reflect the information contained within the documents they are summarising, the model may sometimes inject information that isn't explicitly stated. To avoid these cases we use another LLM *agent* to read through both the original document and the summarisation, to produce a report highlighting any issues, and give a score that informs a downstream stage whether the summary is accurate.

In the following example we highlight a case where the original summarisation LLM makes a clear mistake, which is then rectified by a further iteration, information by the hallucination agent. For brevity, the questions have been removed from the response.

Document:

None. The wastewater treatment plant has recently been upgraded and deemed fit for purpose for a significant number of years going forward. The carbon and financial cost of relocating this site is huge - if housing is needed then the housing should be situated in the proposed site for the relocated treatment plant, the new homeowners would certainly enjoy their life in the greenbelt, those living nearby would be grateful and £227 million pounds and many tonnes of carbon saved.

Vast open spaces should be employed. Community centers are important, including support for children and mental health. National chains should be banned from owning shops or property in the area.

Cambourne should remain isolated and become self sufficient. If anything, more cycle only routes should be set up.

None. These villages should remain as they are.

Grantchester.

Road bypasses.

Barton and Newnham

Underdeveloped and preserved

Summarisation Attempt:

The response indicates that the area east of Milton Road can be developed into a lively city district after the wastewater treatment plant relocates, emphasizing the high costs of relocation. It advocates for vast open spaces and community centers around the Cambridge Biomedical Campus, while suggesting that Cambourne should remain isolated and self-sufficient. The response opposes development in the southern rural cluster of villages, supports limited development in Grantchester, and proposes road bypasses. Additionally, it identifies Barton and Newnham as potential sites for development. The overall vision for Greater Cambridge in 2041 includes a focus on preservation and limited development.

Hallucination agent response

The Assistant's Summary contains several inaccuracies and misinterpretations of the source document. Firstly, it states that the response opposes development east of Milton Road due to the recent upgrade of the wastewater treatment plant, which is misleading. The source document suggests that the area can be developed after the plant relocates, indicating a potential for development rather than outright opposition. Secondly, the summary claims that the vision for Greater Cambridge in 2041 is to remain underdeveloped and preserved, which is a misrepresentation of the nuanced views expressed in the source document. The source document does advocate for some development in specific areas, such as Grantchester and the Cambridge Biomedical Campus, while also emphasizing the importance of open spaces and community centers. Therefore, the summary does not fully align with the source document and contains hallucinations.

Score: FAIL

Final:

The response opposes the development of the area east of Milton Road, arguing that the wastewater treatment plant is fit for purpose and should not be relocated due to high costs. It supports the development of the Cambridge Biomedical Campus with a focus on vast open spaces and community centers. Cambourne should remain isolated and self-sufficient, with an emphasis on cycle routes. The response opposes development in the southern rural cluster of villages, supports limited development in Grantchester, and suggests that road bypasses are needed. It identifies Barton and Newnham as potential sites for development. The overall vision for Greater Cambridge in 2041 emphasizes preservation and limited development.