Lecture Notes: Using Reference Texts and Pointers in AI Prompts

# 1. Provide Reference Text

Objective: Enhance the accuracy of AI model responses by providing reference texts that the model can consult.

Concept:

- Language models can confidently invent fake answers, especially when asked about esoteric topics or for citations and URLs.

- Providing a reference text helps the model answer questions more accurately and reduces the risk of fabrications, similar to how a student uses notes during a test.

Tactics:

- Instruct the model to answer using a reference text.

- Instruct the model to answer with citations from the reference text.

# 2. Use Reference Text in Prompts

Example:

- System Prompt: "Use the provided articles delimited by triple quotes to answer questions. If the answer cannot be found in the articles, write 'I could not find an answer.'"

- User Prompt: <insert articles, each delimited by triple quotes> Question: <insert question here>

# 3. Reference Pointers

Objective: Guide the AI to cite specific passages from a provided document, ensuring the answer is based on verifiable information.

Concept:

- When tasked with answering a question using only a provided document, the AI must cite the passage(s) from the document that supports its answer.

- If the required information is not in the document, the AI should acknowledge the lack of information rather than fabricate an answer.

Tactics:

- Answer the question using only the provided document.

- Cite the passage(s) of the document used to answer the question.

- If the document does not contain sufficient information, instruct the AI to indicate 'Insufficient information.'

- Use a structured format for citations within the prompt.

# 4. Use Reference Pointers in Prompts

Example:

- System Prompt: 'You will be provided with a document delimited by triple quotes and a question. Your task is to answer the question using only the provided document and to cite the passage(s) of the document used to answer the question. If the document does not contain the information needed to answer this question, then simply write: 'Insufficient information.' If an answer to the question is provided, it must be annotated with a citation. Use the following format to cite relevant passages: {'citation': ...}.'

- User Prompt: '"""<insert document here>""" Question: <insert question here>'