LLM Apps
Langchain: Memory

LLM Apps and Memory

- The memory of Foundation LLMs is limited by their context window.

- The best way to solve this problem is using the RAG technique.

- In addition to the RAG technique, LangChain offers several features to manage the memory of LLM Apps.

Examples

- Conversation Buffer Memory: stores past messages in the prompt.

- Conversation Buffer Window Memory: stores past messages in the prompt with a limit of stored messages.

- Conversation Token Buffer Memory: stores past messages in the prompt with a limit of stored tokens.

- Conversation Summary Buffer Memory: stores summaries of past messages in the prompt with a limit of stored tokens.

Other

- There are other advanced functionalities for managing the memory of Agents and Entities that you can see in the LangChain documentation.