LLM Application

LangGraph 세미나





다룰 내용

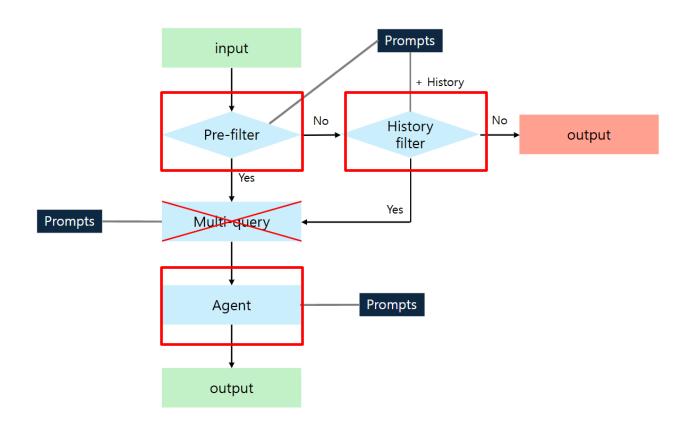
- MBC+ LangGraph 소개
- 강의 속 알짜배기 모음
 - 판다스 스튜디오
 - 케인
 - o AlSchool



MBC+ LangGraph 소개



이전 MBC+ LangChain 버전에서 문제



- 1. RAG 비용: 필터와 답변 생성을 위해 RAG를 2~3번 했어야 했습니다.
- 2. 단일 Agent: 하나의 Agent가 모든 기능을 수행 해야 했습니다.
- 3. 제어 불가: LangChain에서 제공하는 AgentExecutor를 사용해야 됐기 때문에 자체 튜닝이 어려웠습니다.



하지만 LangGraph를 사용한다면

1. RAG 비용: 필터와 답변 생성을 위해 RAG를 2~3번 했어야 했습니다.

: 최초 Retriever 수행 시 검색된 문서들을 LangGraph 상태로 저장하면 단 한 번의 RAG로 작업을 완료할 수 있습니다.

2. 단일 Agent: 하나의 Agent가 모든 기능을 수행 해야 했습니다.

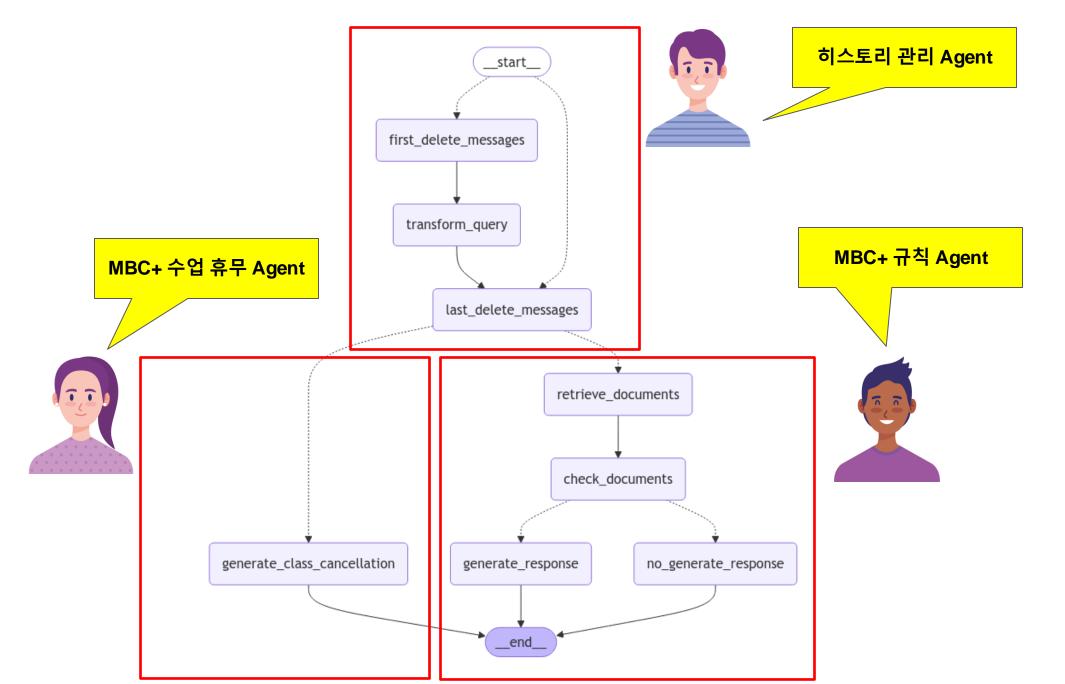
: 멀티 에이전트를 활용해 기능을 분배하면 답변의 품질을 향상시키고 응답 시간을 단축할 수 있습니다.

3. 제어 불가: LangChain에서 제공하는 AgentExecutor를 사용해야 됐기 때문에 자체 튜닝이 어려웠습니다.

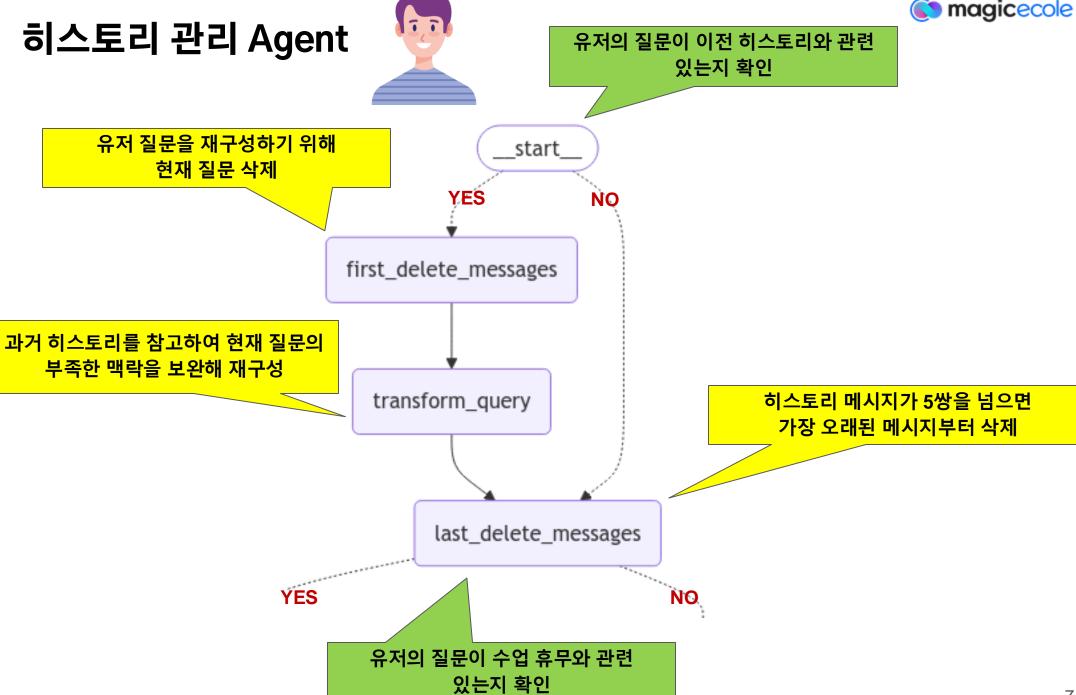
: LangGraph를 활용하면 원하는 아키텍처를 쉽게 구축할 수 있으며, 버그 발생 시 수정 작업도 간편합니다.



MBC+ 운영 챗봇 LangGraph 아키텍처

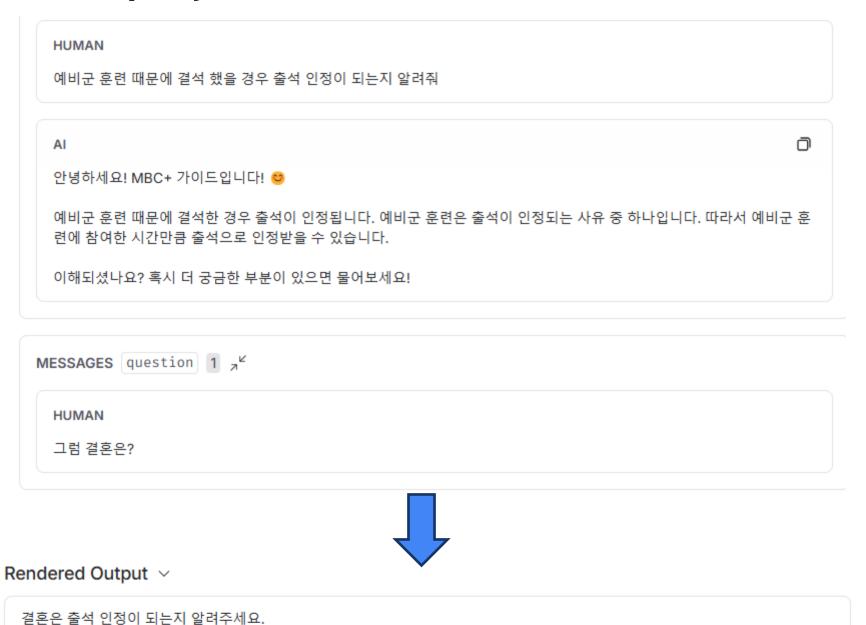






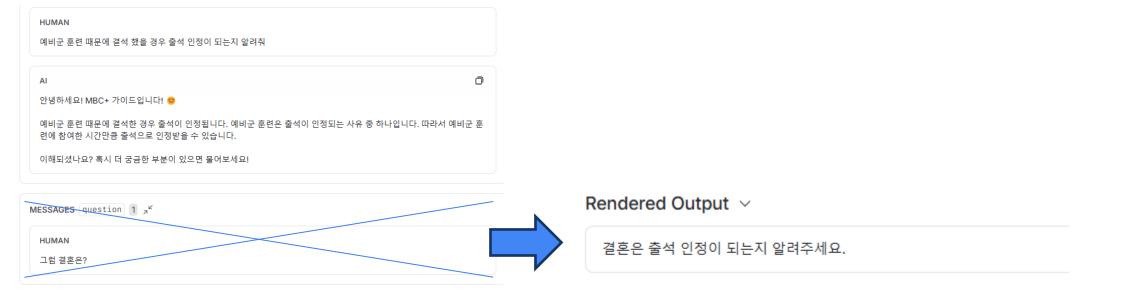


transform_query





first_delete_messages

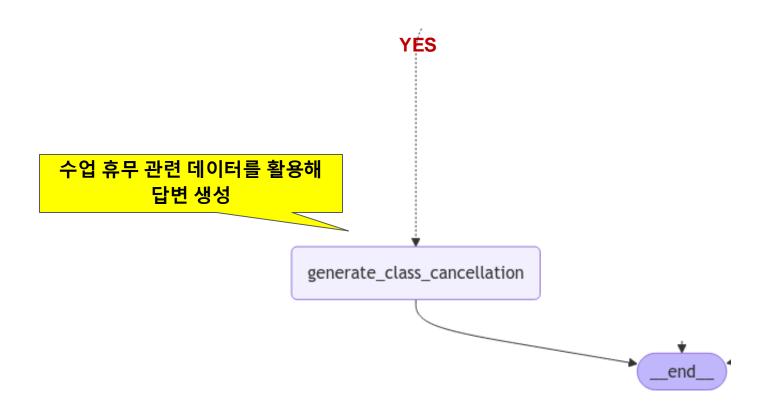




MBC+ 수업 휴무 Agent



질문이 수업 휴무와 관련이 있는 경우

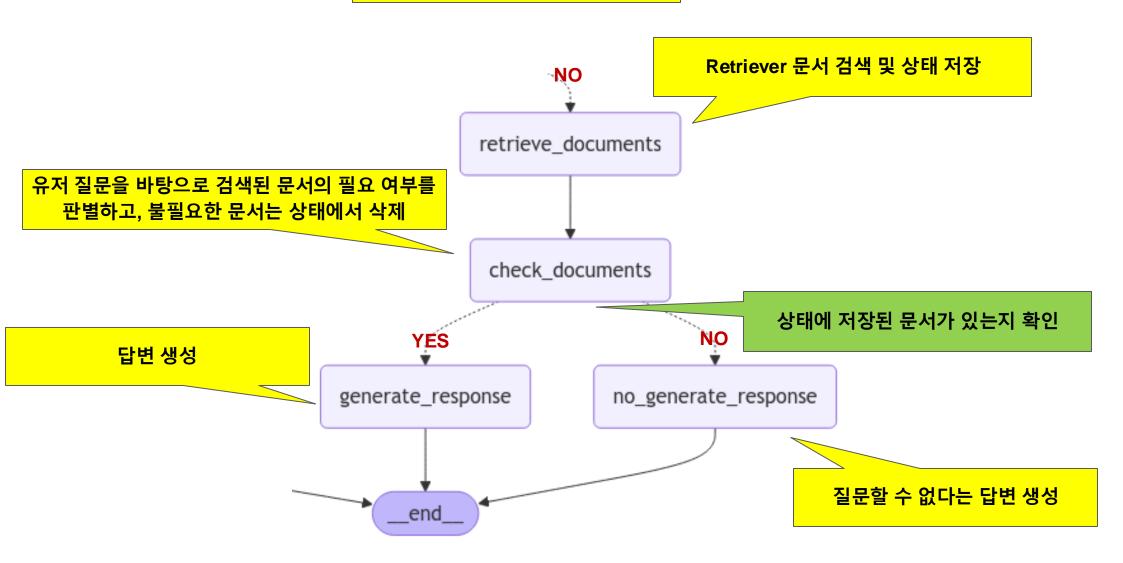








질문이 수업 휴무와 관련이 없는 경우





check_documents

불필요한 Document 제거!

DOCUMENTS documents 7 □ Q: 출석 일수를 산정할 때 쓰는 단위기간이 무슨 의미인지 A: 전체 훈련기간 중 (주말 및 공휴일 제외) 한달 ... □ 3. (스마트훈련과정(훈련유형과 관계없이 '스마트훈련단가'를 적용받는 훈련과정)) 스마트훈련과정 중 '집... □ ■ 내일배움카드 유형에 따른 훈련장려금 지급 1. "실업자"로 카드를 발급받은 경우 - 훈련장려금 지급 대상... □ 3. 휴가 사용 A. 휴가 신청 i. 휴가를 사용하기 전에 반드시 담당 강사와 운영 담당자에게 사전에 알리셔야 ... □ 2. (출석입력 요청) 훈련기관이 출석체크를 할 수 없는 경우 해당사유가 발생한 다음날까지... 직접 제작 문서 □ 제36조(출석률) ① 훈련생의 출석률은 다음 계산식에 따라 소수점 이하 첫째 자리까지 계... 직접 제작 문서 □ 모든 과제와 시험에 성실히 임하며, 부정행위를 하지 않습니다. 교·강사의 저작물을 비롯한... 직접 제작 문서





LangGraph 기술 공유 - Sqlite MemorySaver

from langgraph.checkpoint.sqlite import SqliteSaver

from langgraph.checkpoint.sqlite 모듈에서 SqliteSaver 클래스를 가져옵니다.

```
def __memory_init(self):
    db_path = "sqlite.db"
    conn = sqlite3.connect(db_path, check_same_thread=False)
    memory = SqliteSaver(conn)
    return memory
```

Memorysaver는 check same thread=False 로 생성해야합니다.

Number of requested results 20 is greater than number of elements in index 1, updating n_results = 1

BadRequestError: SQLite objects created in a thread can only be used in that same thread. The object was created in thread id 140079289106432 and this is thread id 140078655456832

- True 로 생성할 경우 발생하는 Error



LangGraph 기술 공유 - Sqlite MemorySaver

from langchain_core.messages import RemoveMessage

```
def last_delete_messages(self, state: operation_state):
    print("---- this node is last_delete_messages ----")

if len(state["messages"]) > 8:
    return {
        "messages": [RemoveMessage(id=m.id) for m in state["messages"][:-8]]
    }
```

히스토리를 삭제할 때는 별도의 노드를 생성해야 하며, 삭제할 메시지의 ID를 RemoveMessage에 담아 리스트 형태로 반환하면 됩니다.



LangGraph 기술 공유 - Sqlite MemorySaver

```
messages: Annotated[List, add messages]
from langgraph.graph.message import add messages
def add_messages(left: Messages, right: Messages) -> Messages:
Messages = Union[list[MessageLikeRepresentation], MessageLikeRepresentation]
MessageLikeRepresentation = Union[
    BaseMessage, list[str], tuple[str, str], str, dict[str, Any]
class BaseMessage(Serializable):
    id: Optional[str] = None
```



LangGraph 기술 공유 - 바뀐 pydantic v2 사용법

```
class ClassCancellationCheck(BaseModel):
    """Binary score indicating whether it is related to class cancellation."""
    relevance to class cancellation: str = Field(
        description="Binary relevance score: 'yes' if related to class cancellation, 'no' otherwise."
output_parser = PydanticOutputParser(pydantic_object=ClassCancellationCheck)
template_str = "{format_instructions}"
prompt_template = PromptTemplate(
  template=template_str,
  input_variables=["query"],
  partial_variables={"format_instructions": output_parser.get_format_instructions()},
model = ChatOpenAI(model="gpt-40", temperature=0)
query_filter_chain = prompt_template | model | output_parser
```



LangGraph 기술 공유 - 바뀐 pydantic v2 사용법

LangGraph를 사용하면서 langchain-core를 0.3 버전을 사용해야 합니다. Langchain-core를 사용하면 pydantic v2 버전을 사용해야 합니다.

차이점

특징	PydanticOutputParser	with_structured_output				
출력 생성 방식	LLM 출력 문자열을 받아서 Pydantic 모델로 파싱	LLM이 출력 자체를 Pydantic 모델 구조에 맞춰 생성				
에러 발생 가능성	출력이 구조화되지 않으면 파싱 실패 가능	LLM이 정확히 형식에 맞춰 출력하지 않으 면 에러 발생				
유연성	출력이 Pydantic 형식과 완전히 일치하지 않 아도 파싱 가능	LLM 출력이 Pydantic 형식과 반드시 일치해 야 함				
적합한 경 우	출력이 비교적 유연하고 후처리로 구조화가 필요한 경우	출력이 처음부터 완전히 구조화되어야 하는 경우				



LangGraph 기술 공유 - 바뀐 pydantic v2 사용법

Pydantic을 선언할 때는 필드명 , 설명 을 명확하게 선언해야 합니다!

```
class ClassCancellationCheck(BaseModel):
    """Binary score indicating whether it is related to class cancellation."""
    relevance_to_class_cancellation: str = Field(
         description="Binary relevance score: 'yes' if related to class cancellation, 'no' otherwise."
    )
```



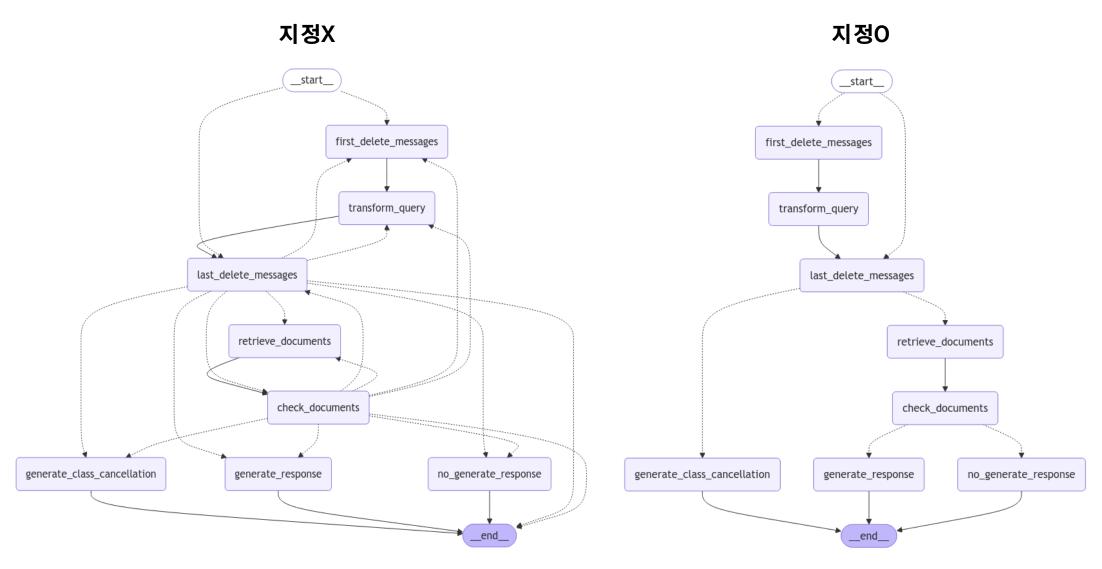
LangGraph 기술 공유 - Literal 사용 그래프 지정

add_conditonal_edge를 사용할 때 next 노드를 항상 명시 해줘야합니다!

```
def check class cancellation(self, state: operation state) -> Literal["Node A", "Node B"]:
     print("---- this node is check class cancellation ----")
     prompt = PromptTemplate(
         template=prompts.check class cancellation prompt,
         input variables=["question"],
     chain = prompt | self.model.with structured output(ClassCancellationCheck)
     response = chain.invoke({"question": state["question"]})
     if response.relevance to class cancellation == "yes":
         return "Node A"
     else:
         return "Node B"
```



LangGraph 기술 공유 - Literal 사용 그래프 지정

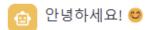




결과



불필요한 토큰을 사용하지 않기 때문에 속도 개선과 토큰 절감을 할 수 있습니다!



내일은 2024년 11월 29일 금요일입니다. 공휴일 정보에 따르면 내일은 공휴일이 아니므로 수업 이 정상적으로 진행됩니다. ■

수업에 참석해 주세요! 🧎

(☆) 내일 수업 취소하고 싶어

👍 안녕하세요! MBC+ 가이드입니다! 😊

내일 수업을 취소하고 싶으시다면, 우선적으로 운영진과 상담이 필요합니다. 디스코드 내 'MBC+ 과정운영팀'에게 과정 취소 상담과 관련한 내용을 DM으로 보내시면 빠르게 도움을 받으실 수 있습니다.

혹시 디스코드 사용이 어려우시면, 아래의 연락처로 문의해 주세요:

- 문의 전화: 02-730-4519
- 이메일: info@kdt-mbcplus.com

이해되셨나요? 혹시 더 궁금한 부분이 있으면 물어보세요! 😊

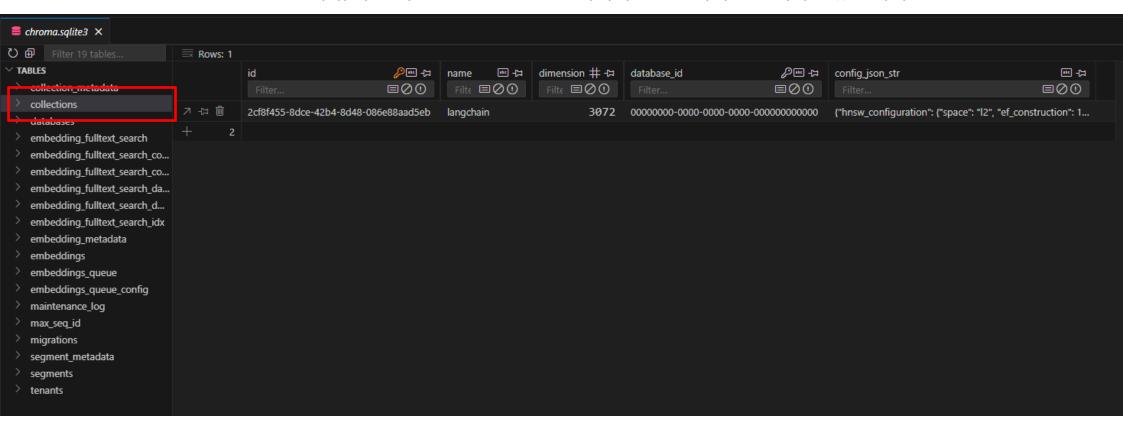
\odot	LangGraph	human: 내일 수업 취소	ai: 안녕하세요! M	2024. 11. 28. 오후 11	③ 6.32s		6,824	\$0.0187025
\odot	LangGraph	human: 내일 수업 취소	ai: 안녕하세요! 😊	2024. 11. 28. 오후 11	O 1.78s		1,407	\$0.00402



기술 공유 - chromaDB 원격 수정

서버가 실행 중일 때 변경된 데이터를 Chroma DB에 반영하려면, ChromaDB 파일을 삭제하고 다시 생성하는 방식은 사용하면 안 됩니다!

ChromaDB는 table과 유사한 개념인 collection으로 데이터를 그룹화하고 관리하고 있습니다.





기술 공유 - chromaDB 원격 수정

Collection 초기화

```
main_db = Chroma(embedo ng_function=embeddings, persist_directory=main_db_directory)
main_db.reset_collection()
print("현재 DB의 모든 컬렉션이 삭제되었습니다.")

recursive_splitted_document = load_documents(file_path)
main_db.add_documents(documents=recursive_splitted_document)
print("temp_DB의 무서들이 현재_DB에 추가되었습니다.")
```

Documents 업로드



강의 속 알짜배기 모음



판다스 스튜디오 - 병렬 노드 실행

add_conditonal_edge를 사용할 때 next node를 리스트 형태로 반환합니다.

from typing import Sequence part5.py def route datasources tool search(state: ToolSearchState) -> Sequence[str]: if set(state['datasources']) == {'search_menu'}: return ['search menu'] elif set(state['datasources']) == {'search_wine'}: return ['search_wine'] elif set(state['datasources']) == {'search_web'}: return ['search_web'] elif set(state['datasources']) == {'search_menu', 'search_wine'}: return ['search menu', 'search wine'] elif set(state['datasources']) == {'search_menu', 'search_web'}: return ['search_menu', 'search_web'] elif set(state['datasources']) == {'search_wine', 'search_web'}: return ['search wine', 'search web'] return ['search_web', 'search_menu', 'search_wine']



간단한 검색 tool 세팅

```
from langchain community.tools import TavilySearchResults
# Tool 정의
@tool
def search_web(query: str) -> List[str]:
   """데이터베이스에 존재하지 않는 정보 또는 최신 정보를 인터넷에서 검색합니다."""
   tavily search = TavilySearchResults(max results=3)
   docs = tavily search.invoke(query)
   formatted docs = "\n\n--\n\n".join(
          f'<Document href="{doc["url"]}"/>\n{doc["content"]}\n</Document>'
           for doc in docs
   if len(docs) > 0:
       return formatted docs
   return "관련 정보를 찾을 수 없습니다."
```



bind_tool을 사용하면 도구가 필요한 상황에서 바로 답변을 제공하지 않고, 대신 도구 호출을 수행합니다.

```
from langchain openai import ChatOpenAI
# LLM 모델
llm = ChatOpenAI(model="gpt-4o-mini", streaming=True)
# 도구 목록
tools = [search web]
# 모델에 도구를 바인딩
llm with tools = llm.bind tools(tools=tools)
# 도구 호출
tool call = llm with tools.invoke([HumanMessage(content=f"LangGraph는 무엇인가요?")])
# 결과 출력
print(tool call.additional kwargs)
```

```
('tool_calls': [{'index': 0, 'id': 'call_veljXtipNE0N007Yq9IKV2H6', 'function': {'arguments': '{"query":"LangGraph"}', 'name': 'search_web'}, 'type': 'function'}]}
```



도구 호출 결과를 ToolNode로 정의된 도구를 호출하면 해당 도구의 결과를 바로 얻을 수 있습니다.

```
from langgraph.prebuilt import ToolNode
# 도구 노드 정의
tool node = ToolNode(tools=tools)
# 도구 호출
tool_call = llm_with_tools.invoke([HumanMessage(content=f"LangGraph는 무엇인가요?")])
# 도구 호출 결과를 메시지로 추가하여 실행
results = tool node.invoke({"messages": [tool call]})
# 실행 결과 출력하여 확인
for result in results['messages']:
   print(result.content)
   print()
```

<Document href="https://github.com/langchain-ai/langgraph"/>

GitHub - langchain-ai/langgraph: Build resilient language agents as graphs. LangGraph Platform is a commercial solution for deploying agentic applications to production, built on the open-source LangGraph framework. Let's take a look at a simple example of an agent that can use a search tool. # Define the tools for the agent to use # This means that after `tools` is called, `agent` node is called next. workflow.add_edge("tools", 'agent') Define entry point and graph edges. First, we need to set the entry point for graph execution - agent node. Normal edge: after the tools are in voked, the graph should always return to the agent to decide what to do next LangGraph adds the input message to the internal state, then passes the state to the entrypoint node, "agent".

<a href="mailto:common:commons.comm

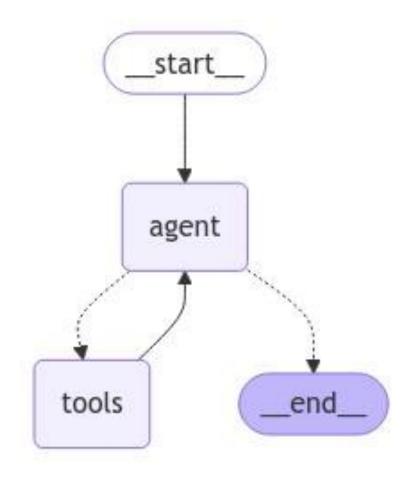
<Document href="https://www.langchain.com/langgraph"/>



```
def call model(state: GraphState):
   system message = SystemMessage(content=system prompt)
                                                          1. bind tool 된 IIm을 실행하는 노드
   messages = [system_message] + state['messages']
   response = llm with tools.invoke(messages)
   return {"messages": [response]}
def should continue(state: GraphState):
   last_message = state["messages"][-1]
   # 도구 호출이 있으면 도구 실행 노드로 이동
                                              2. Ilm이 tool을 호출했는지 판단하는 add conditional edge
   if last message.tool calls:
       return "execute tools"
   # 도구 호출이 없으면 답변 생성하고 종료
   return END
# 그래프 구성
builder = StateGraph(GraphState)
builder.add node("call model", call model)
builder.add node("execute tools", ToolNode(tools)) 3. tool node init
builder.add edge(START, "call model")
builder.add conditional edges(
    "call model",
   should continue,
       "execute tools": "execute tools",
       END: END
                                                 4. tool node에서 call model로 연결
builder.add edge("execute tools", "call model")
```



```
graph = builder.compile()
# 그래프 출력
display(Image(graph.get_graph().draw_mermaid_png()))
```



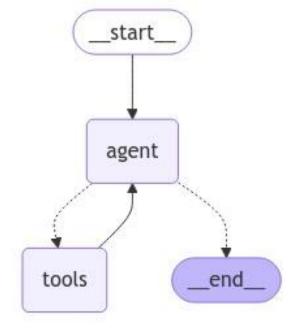


하지만 create_react_agent 를 사용하면 바로 결과를 얻을 수 있습니다!

```
from langgraph.prebuilt import create_react_agent

# 그래프 생성
graph = create_react_agent(
    llm,
    tools=tools,
    )

# 그래프 출력
display(Image(graph.get_graph().draw_mermaid_png()))
```





Create_react_agent 의 또 다른 장점 중 하나는 Tool System Prompt를 추가하여 에이전트의 동작 방식을 더욱 세부적으로 정의하고 조정할 수 있다는 점입니다.

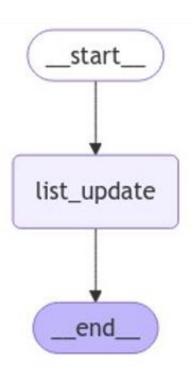
```
from langgraph.prebuilt import create react agent
from IPython.display import Image, display
from textwrap import dedent
# 시스템 프롬프트
system prompt = dedent("""
You are an AI assistant designed to answer human questions.
You can use the provided tools to help generate your responses.
Follow these steps to answer questions:
   1. Carefully read and understand the question.
   2. Use the provided tools to obtain necessary information.
   3. Immediately after using a tool, cite the source using the format below.
   4. Construct an accurate and helpful answer using the tool outputs and
citations.
   5. Provide the final answer when you determine it's complete.
# 그래프 생성
graph = create react agent(
   11m,
   tools=tools,
   state modifier=system prompt,
```



판다스 스튜디오 - 멀티 그래프 연결

간단한 num_list에 "2"를 추가하는 graph 입니다.

```
from langgraph.graph import StateGraph, START, END
from typing import TypedDict, List
from langgraph.graph.message import add messages
from typing import Annotated
import operator
class MyState(TypedDict):
    num list: Annotated[List, add messages]
graph builder = StateGraph(MyState)
def list update(state: MyState):
    return {"num list": "2"}
graph_builder.add_node("list_update", list_update)
graph_builder.add_edge(START, "list_update")
graph builder.add edge("list update", END)
graph = graph builder.compile()
```

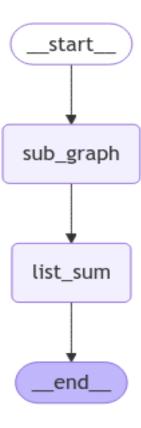




판다스 스튜디오 - 멀티 그래프 연결

State를 상속 받을 때 충돌을 주의해야 합니다!

```
class NewState(MyState):
    cnt : int
                                         state상속
new_graph_builder = StateGraph(NewState)
def list sum(state: NewState):
    tmp = 0
    for x in state["num_list"]:
        tmp += int(x.content)
   return {"cnt": tmp}
                                               Graph를 노드화
new_graph_builder.add_node("sub_graph", graph)
new graph builder.add node("list sum", list sum)
new_graph_builder.add_edge(START, "sub graph")
new_graph_builder.add_edge("sub_graph", "list_sum")
new_graph_builder.add_edge("list_sum", END)
new_graph = new_graph_builder.compile()
result = new graph.invoke({"num list" : "1"})
print(result)
```



Graph를 사용하면 코드 재사용성이 높아집니다.

['num_list'<mark>: [HumanMessage(content='1', additional_kwargs={}, response_metadata={}, id='542ec2a4-b78a-4284-ab25-4a15d70d395b')</mark> HumanMessage(content='2', additional_kwargs={}, response_metadata={}, id='6fdf9293-31c7-4503-9f06-a26a6160125b')], 'cnt': 3}



케인 - tools_condition을 사용한 다른 ToolNode 사용법

from langgraph.prebuilt import tools_condition

주의할 점! tools 라는 이름은 바꾸면 안 됩니다.

```
builder.add_conditional_edges(
    "call_model",
    should_continue,
    {
        "execute_tools": "execute_tools",
        END: END
    }
}
builder.add_conditional_edges(
        "call_model",
        tools_condition,
        {
             "tools": "execute_tools",
             END: END,
        }
        )
}
```

```
def tools_condition(
    state: Union[list[AnyMessage], dict[str, Any], BaseModel],
) -> Literal "tools", "__end__"]:
    """Use in the conditional_edge to route to the ToolNode if the last message
    has tool calls. Otherwise, route to the end.

Args:
    state (Union[list[AnyMessage], dict[str, Any], BaseModel]): The state to check for tool calls. Must have a list of messages (MessageGraph) or have the "messages" key (StateGraph).
Returns:
```



```
def call model(state: GraphState):
   system message = SystemMessage(content=system prompt)
                                                          1. bind tool 된 IIm을 실행하는 노드
   messages = [system_message] + state['messages']
   response = llm with tools.invoke(messages)
   return {"messages": [response]}
def should continue(state: GraphState):
   last_message = state["messages"][-1]
   # 도구 호출이 있으면 도구 실행 노드로 이동
                                              2. Ilm이 tool을 호출했는지 판단하는 add conditional edge
   if last message.tool calls:
       return "execute tools"
   # 도구 호출이 없으면 답변 생성하고 종료
   return END
# 그래프 구성
builder = StateGraph(GraphState)
builder.add node("call model", call model)
builder.add node("execute tools", ToolNode(tools)) 3. tool node init
builder.add edge(START, "call model")
builder.add conditional edges(
    "call model",
   should continue,
       "execute tools": "execute tools",
       END: END
                                                 4. tool node에서 call model로 연결
builder.add edge("execute tools", "call model")
```

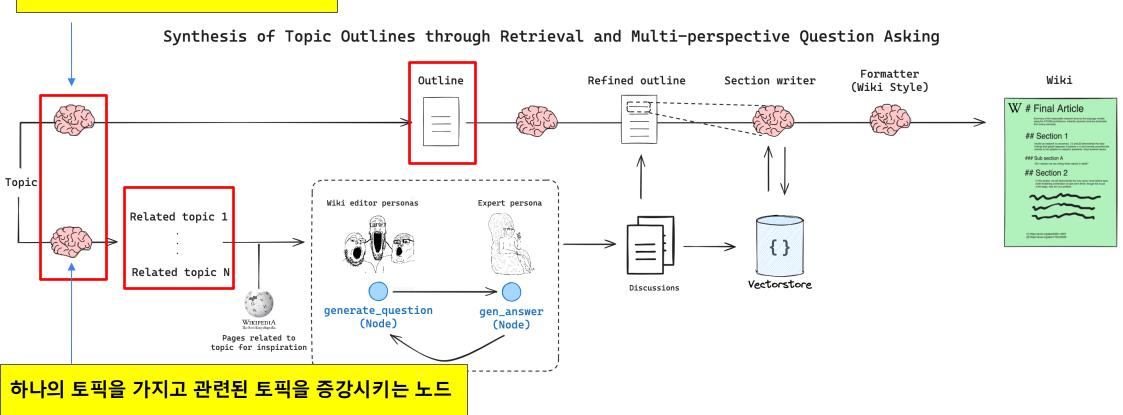


AlSchool - Web Storm Architecture

https://langchain-ai.github.io/langgraph/tutorials/storm/storm/

노트북 LM 아키텍처와 유사함!

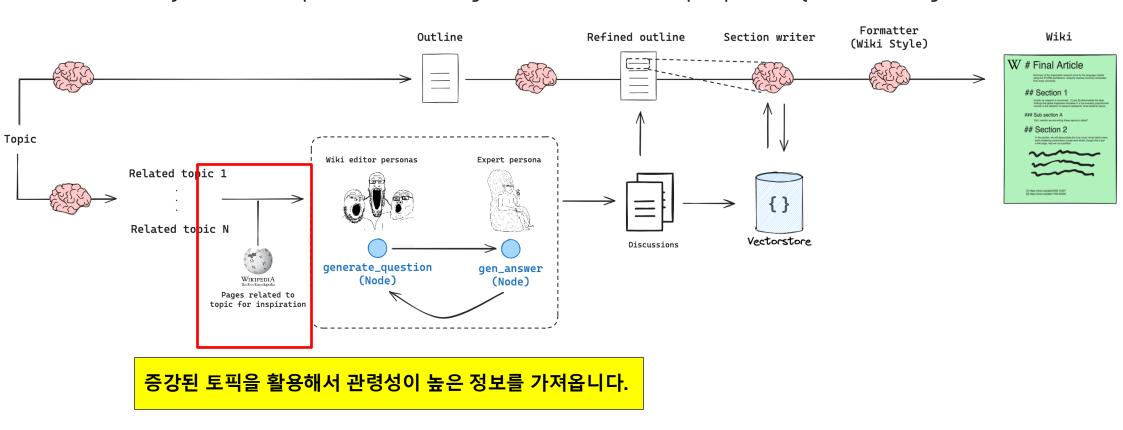
초기 outline(윤곽)을 생성하는 노드





AlSchool - Web Storm Architecture

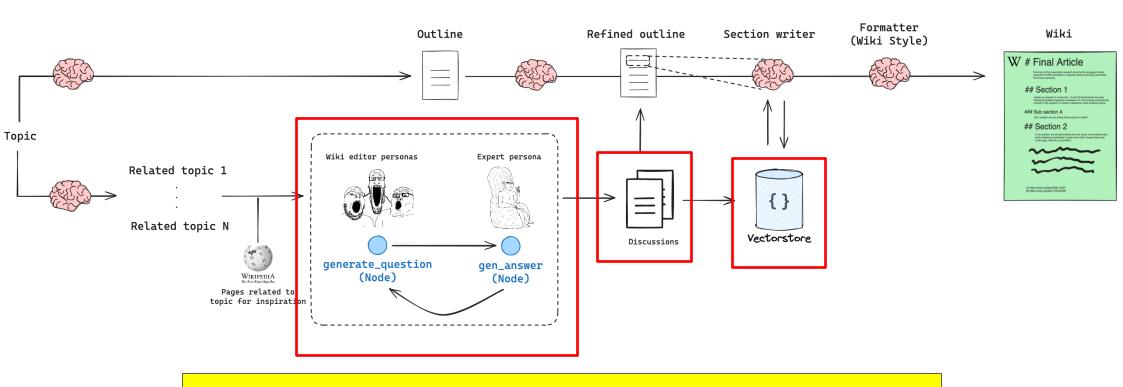
Synthesis of Topic Outlines through Retrieval and Multi-perspective Question Asking





AlSchool - Web Storm Architecture

Synthesis of Topic Outlines through Retrieval and Multi-perspective Question Asking

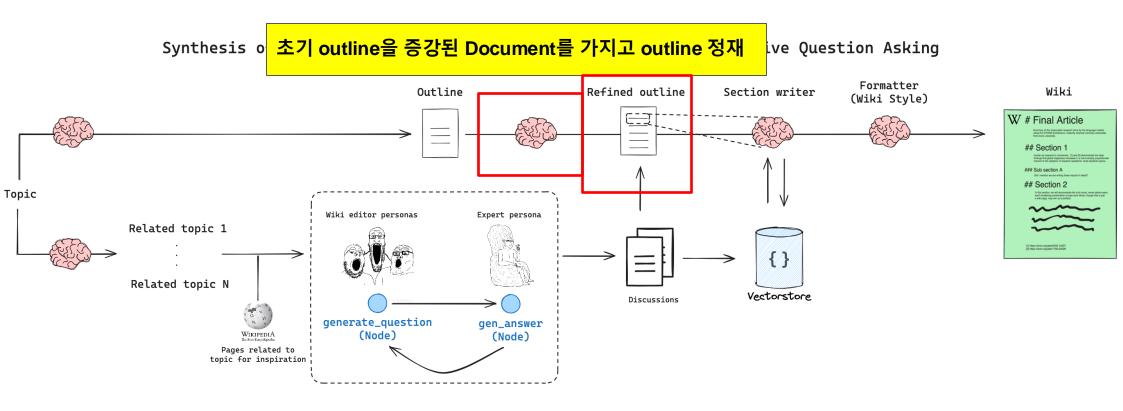


질문 노드와 답변 노드는 검색된 정보를 기반으로 상호작용하며 추가 정보를 생성합니다.

생성된 정보는 문서 형태로 변환되어 상태(state)에 업데이트합니다.

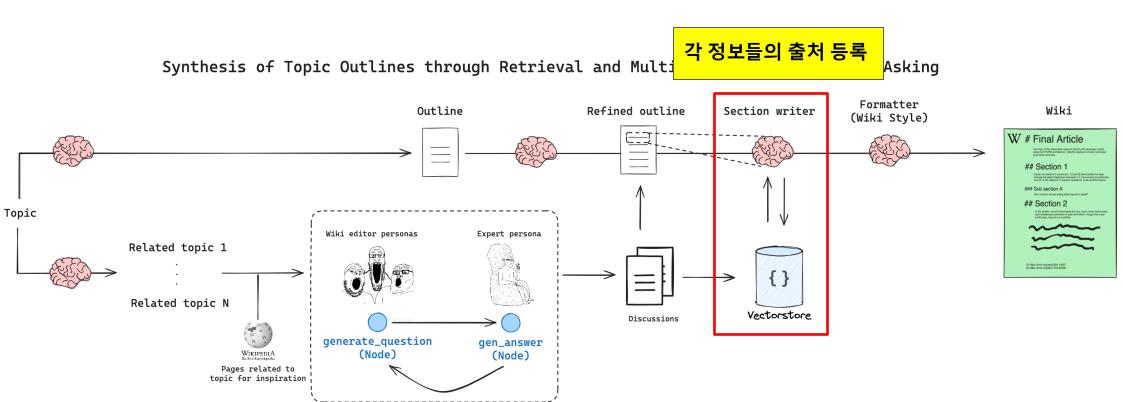


AISchool - Web Storm Architecture





AISchool - Web Storm Architecture

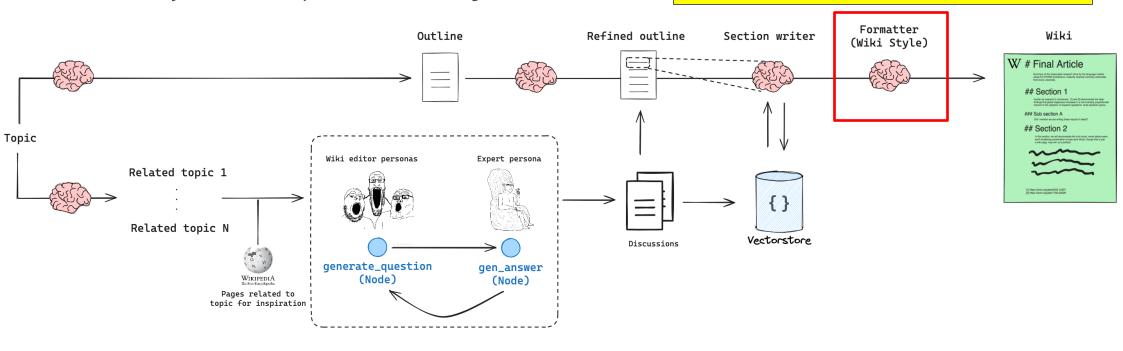




AISchool - Web Storm Architecture

Synthesis of Topic Outlines through Retrieval and Multi

증강된 모든 정보를 가지고 위키피디아 글 작성





감사합니다!