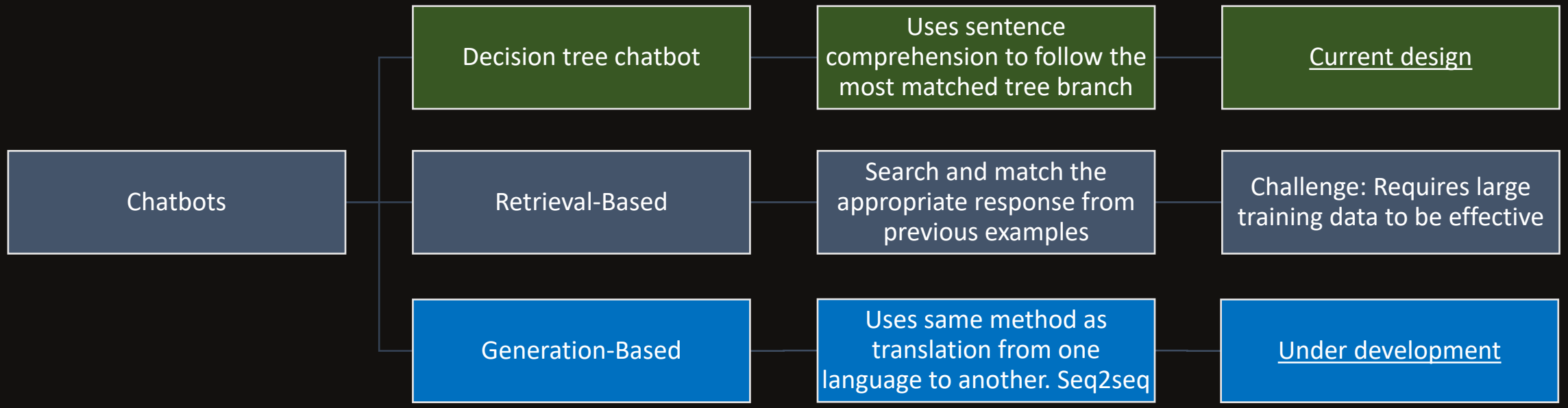
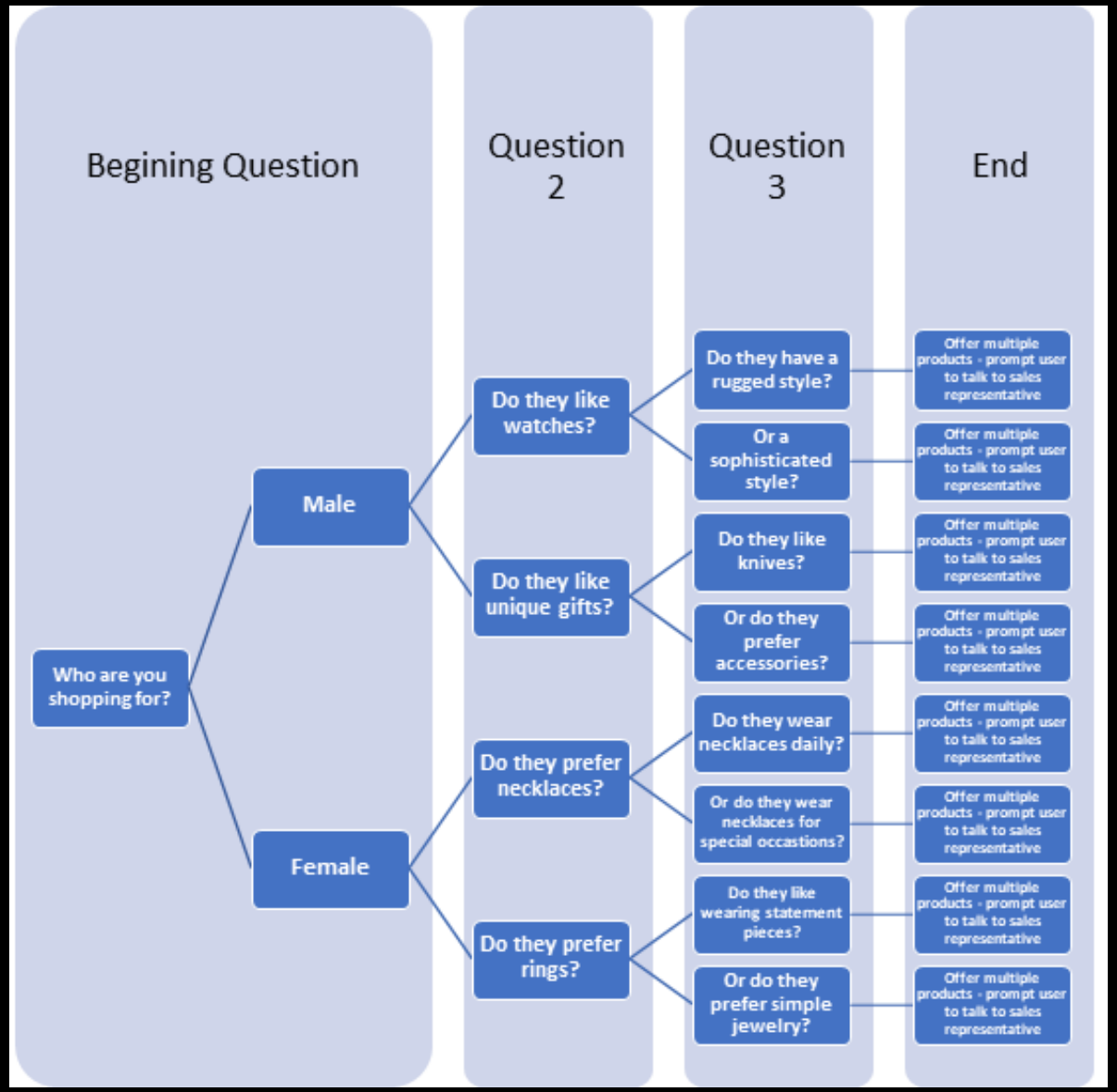
Intel-PA Progress Report 2022-04-06

Abdul Rehman

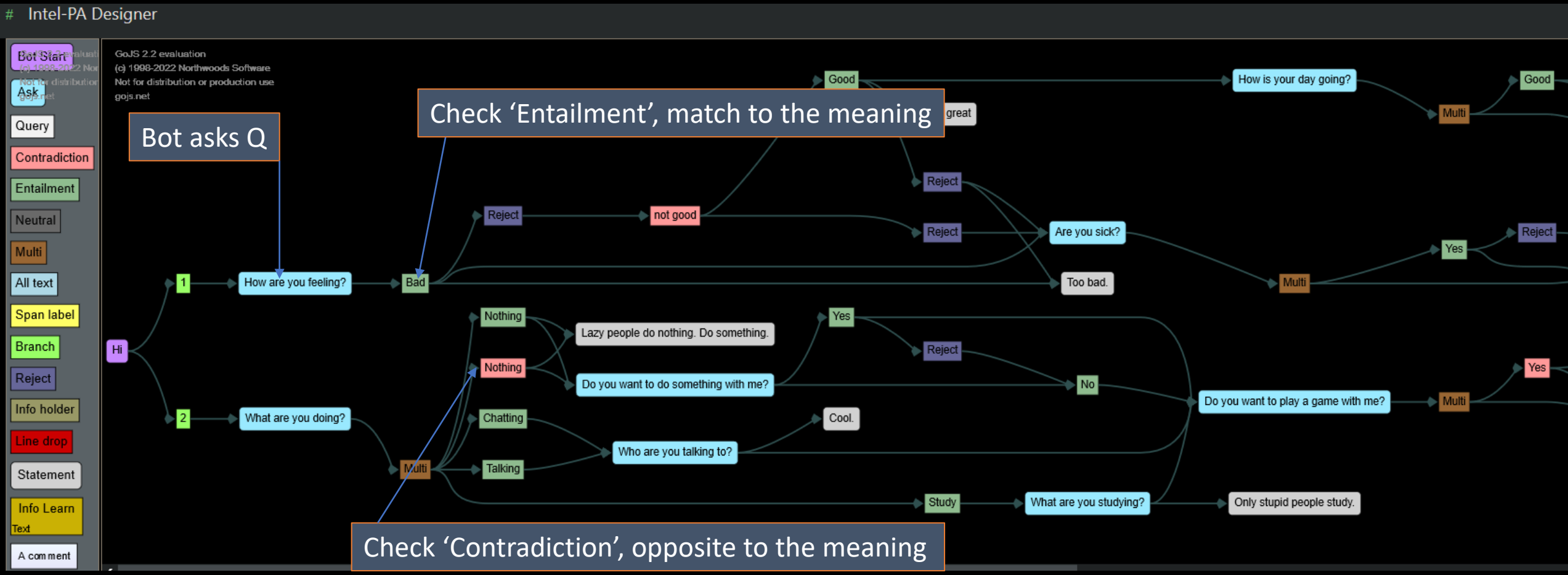

Chatbots
Decision tree chatbot
Uses sentence comprehension to follow the most matched tree branch
Current design
Retrieval-Based
Search and match the appropriate response from previous examples
Challenge: Requires large training data to be effective
Generation-Based
Uses same method as translation from one language to another. Seq2seq
Under development

Decision tree chatbot

- Uses the sentence comprehension
  to understand which of the branches
  match best to the response



Decision tree chatbot



Decision tree chatbot

- Currently using: ELMo
  - Deep Contextualized Word Embeddings to quantify/vectorize matches by meaning.
  - Context sensitive.
  - Same word in different context changes the embedding vector.

Decision tree chatbot

- Also using: BERT
  - For internal query.
  - For example, if doctor wants to know if “someone has cancer”, then BERT summarizes the answer from all the previous responses as

    Dr to Bot: “Did anyone has cancer?”
    BERT: Patient said “mom has cancer”

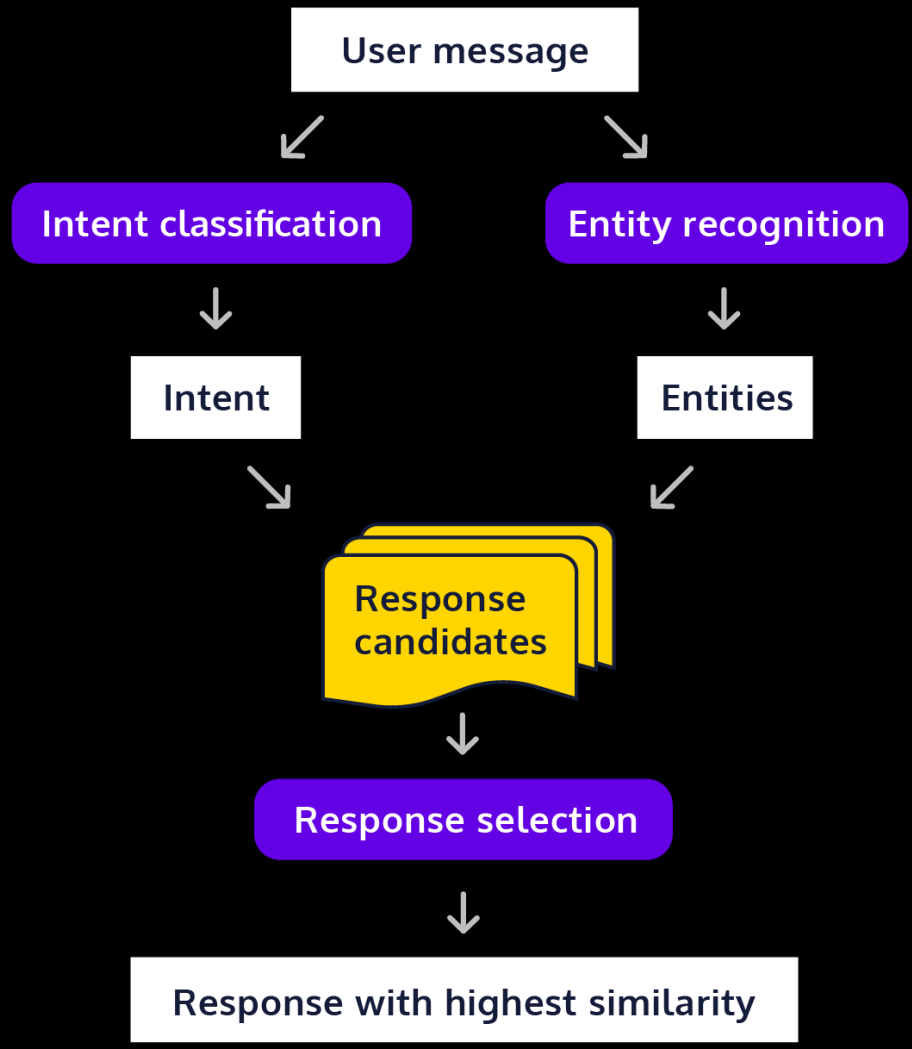
  - If that “internal query” fails (patient hasn’t told us yet), then a new question tree is launched.

Retrieval-Based

- Usually used for QA bots.
- Helpdesk bots. Wiki bots.

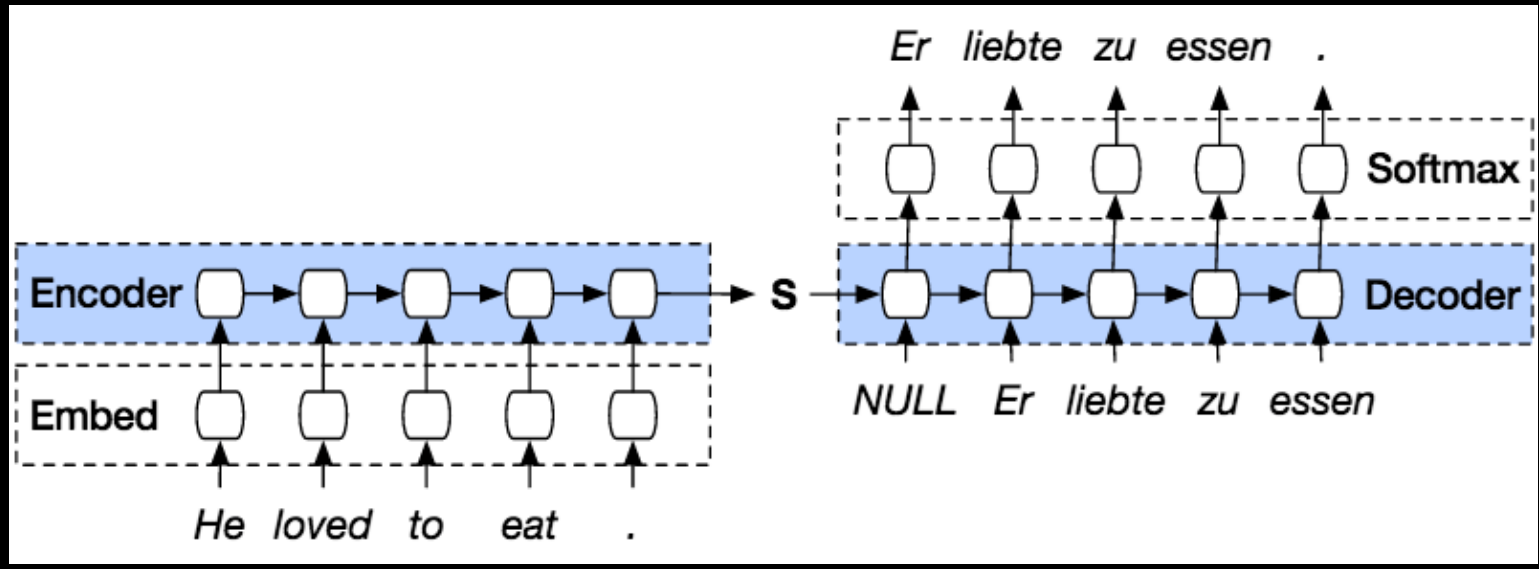
- They are trained on QA datasets.
- Microsoft’s Xiaoice uses retrievals from previous chats 70% of the time.

- Not suitable for our project due to lack of data.



Generation-Based

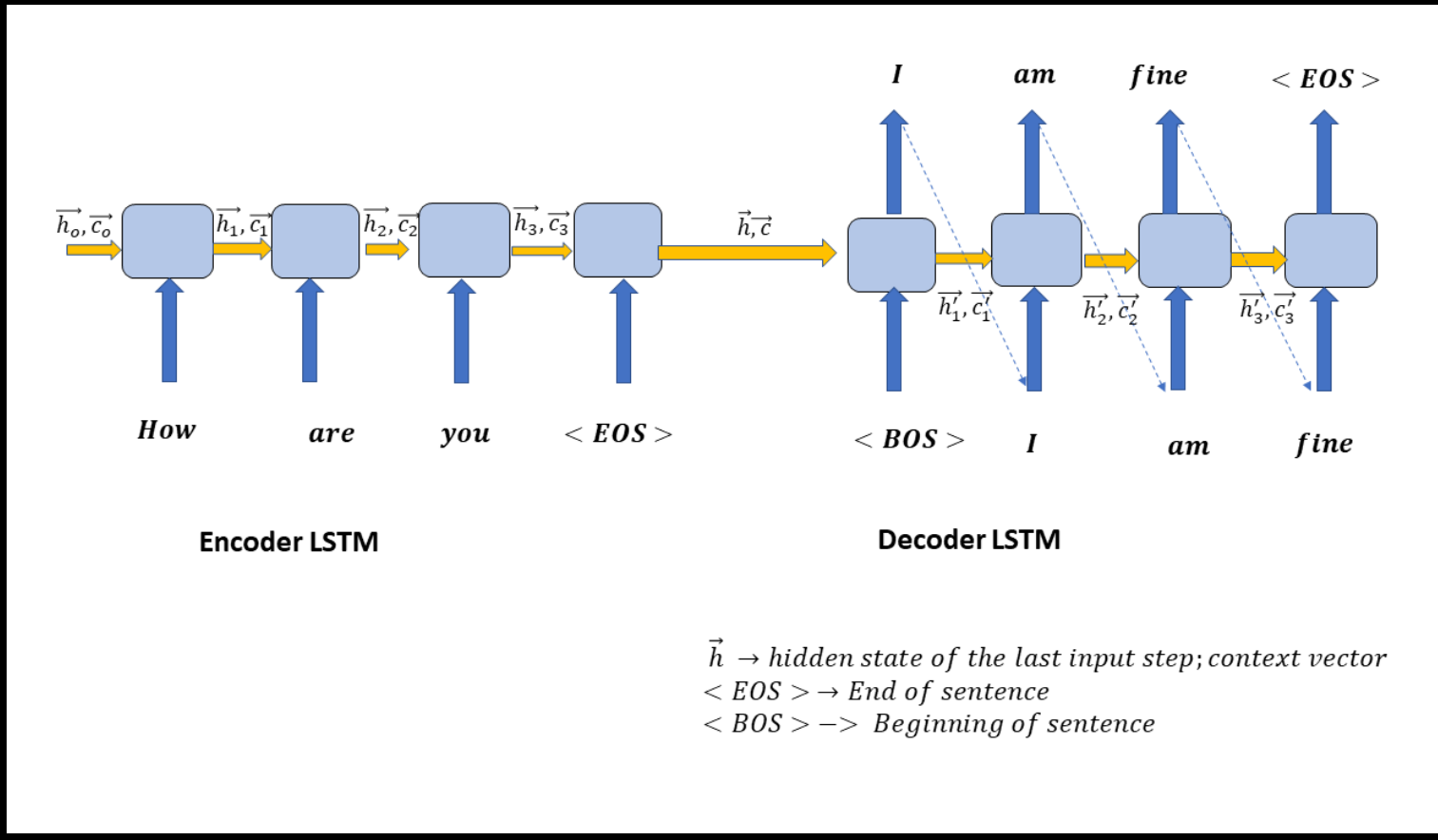
- Using Sequence-to-sequence models, same as translation models



- Instead of language-to-language, we use question-to-answer sequence learning

Generation-Based

- Most modern chatbots use this method as last resort when there is lack of data.
- But they have goldfish memory.



- https://parl.ai/projects/msc/
- https://arxiv.org/pdf/2107.07567.pdf

Current plan

- Loosely incorporate *Generation-based* response generator with the decision tree chat-bot.

Current plan

- Collect example chat by simulating the chat ourself.
- 1k example responses would be enough to train.