Hochschule
Bonn-Rhein-Sieg
University of Applied Sciences

b-it Bonn-Aachen International Center for Information Technology

Human-Centered Interaction in Robotics

HCIR Assignment-3

Shrikar Nakhye
Trushar Ghanekar
Keerthan Shekarappa
Mohammad Ibrahim

May 2024

Contents

1

TASK 3

1.1 MAS Admission Assistance Chatbot Architecture

1.1.1 1. Processing Multimodal Input

Sensing ("receive")

- **Components:** User Interface (UI) for text input, Speech-to-Text Converter (if voice input is used).
- **Function:** Receives input from the user in text or speech form.

Recognition ("find")

- **Components:** Natural Language Understanding (NLU) module.
- **Function:** Identifies the intent and extracts relevant entities from the user input.

Interpretation ("understand")

- **Components:** Context Manager, Entity Resolver.
- **Function:** Understands the context of the conversation and resolves ambiguities in entities.

1.1.2 2. Mapping Responses

Deliberative Planning ("planning")

- **Components:** Dialogue Manager, Intent Handler.
- **Function:** Plans the conversation flow and determines the next action based on the user intent and conversation context.

Associative Selection ("selection")

- **Components:** Response Selector, Knowledge Base Access.

- **Function:** Selects the most appropriate response or information from a predefined set or from a database.

Reactive Rules ("rules")

- **Components:** Rule Engine, Error Handler.
- **Function:** Applies predefined rules to generate responses, handles unexpected inputs, and manages errors.

1.1.3 3. Generating Multimodal Output

Formulation ("compose")

- **Components:** Response Generator, Template Engine.
- **Function:** Composes the response message by filling in templates with appropriate data.

Realization ("create")

- **Components:** Natural Language Generation (NLG) module.
- **Function:** Converts the composed message into natural language text.

Execution ("perform")

- **Components:** User Interface (UI) for text output, Text-to-Speech Converter (if voice output is used).
- **Function:** Outputs the response to the user in text or speech form.

1.2 Detailed Components and Their Interactions

User Interface

- **Input:** Text input field, voice input via microphone.
- **Output:** Text display area, voice output via speaker.

Natural Language Understanding (NLU)

- **Functions:** Intent classification, entity recognition.
- **Tools:** Rasa NLU, spaCy, BERT.

Context Manager

- **Functions:** Tracks the conversation state, manages context switches.

- **Tools:** In-memory storage, Redis.

Entity Resolver

- **Functions:** Resolves entities to their canonical form, manages synonyms.
- **Tools:** Custom entity resolution logic.

Dialogue Manager

- **Functions:** Manages the conversation flow, decides the next action.
- **Tools:** Rasa Core, custom dialogue policies.

Intent Handler

- **Functions:** Handles specific intents, manages intent-specific logic.
- **Tools:** Custom logic, external API integrations.

Response Selector

- **Functions:** Selects predefined responses or retrieves information from a database.
- **Tools:** SQL/NoSQL databases, knowledge base systems.

Rule Engine

- **Functions:** Applies business rules, manages fallback and error responses.
- **Tools:** Drools, custom rule-based logic.

Response Generator

- **Functions:** Generates responses using templates, dynamic data insertion.
- **Tools:** Jinja2, custom template engines.

Natural Language Generation (NLG)

- **Functions:** Converts structured data into natural language text.
- **Tools:** Rasa NLG, GPT-based models.

Text-to-Speech Converter

- **Functions:** Converts text responses into speech (optional).
- **Tools:** Google Text-to-Speech, Amazon Polly.

1.3 Interaction Flow

1. User sends a message via the user interface.
2. NLU processes the message, identifying intent and entities.
3. Context Manager updates the conversation context.
4. Dialogue Manager determines the next action based on the current state and user intent.
5. Intent Handler or Response Selector generates the appropriate response.
6. Rule Engine applies any necessary business rules.
7. Response Generator composes the message using templates.
8. NLG (if needed) converts structured data into a natural language response.
9. UI displays the text response, or Text-to-Speech Converter outputs the voice

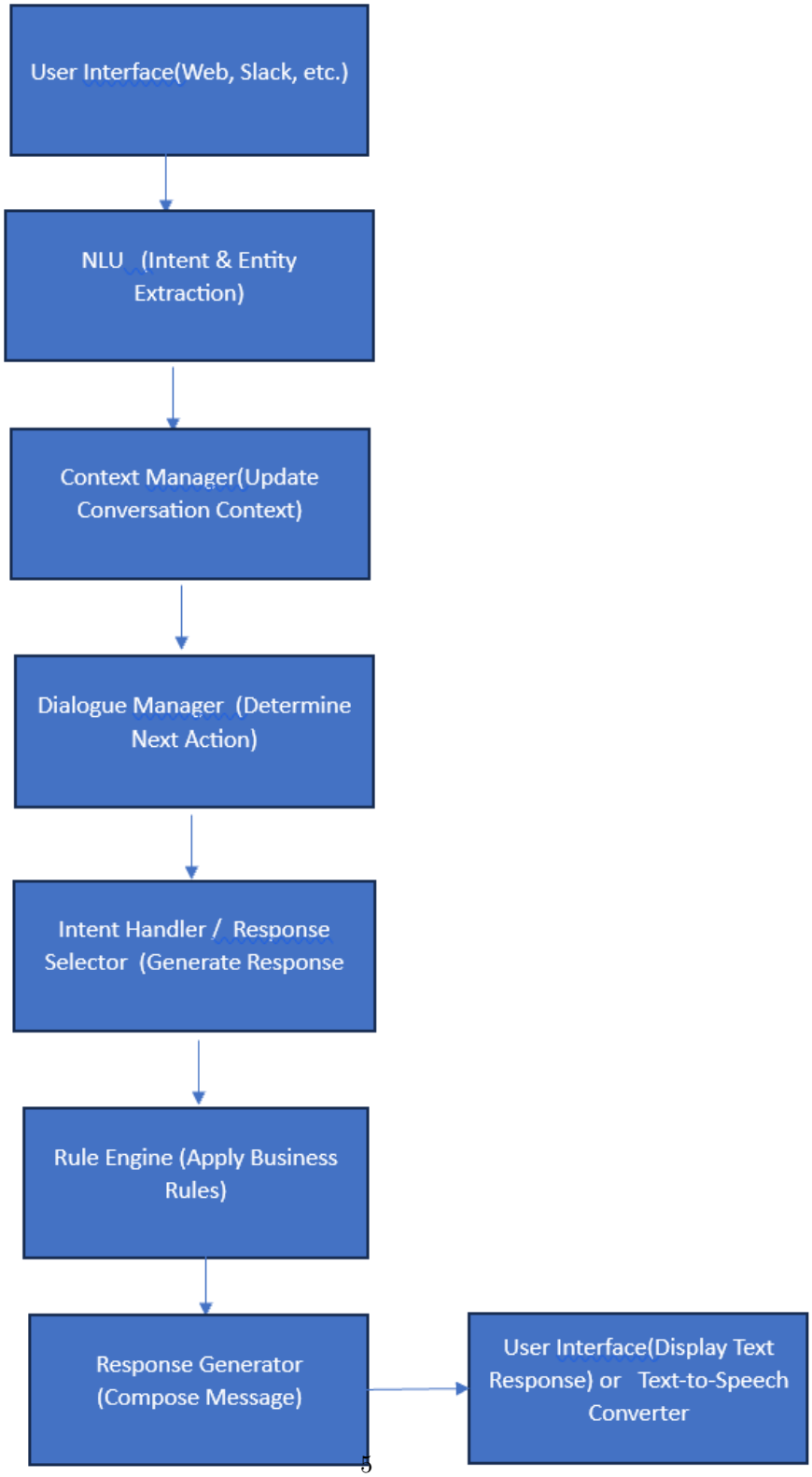

Figure 1.1: Interaction Architecture for MAS Admission Assistance Chatbot