Prototype of Natural Language Interface for Nuclear Reactor Facility-wide Intelligence using AI and Robot Technology

Sungmoon Joo* and Jonghwan Lee
Korea Atomic Energy Research Institute, Daedeok-daero 989-111, Yuseong-gu, Daejeon, Korea
**Corresponding author: smjoo@kaeri.re.kr*

1. Introduction

This study introduces a prototype of NURI, **NU**clear **R**eactor facility-wide **I**ntelligence. NURI is a chatter-bot-style artificial intelligence with natural language interface.

2. System Overview

The following block diagram (Figure 1) shows how NURI communicates and interacts with human operators using either speech or text,

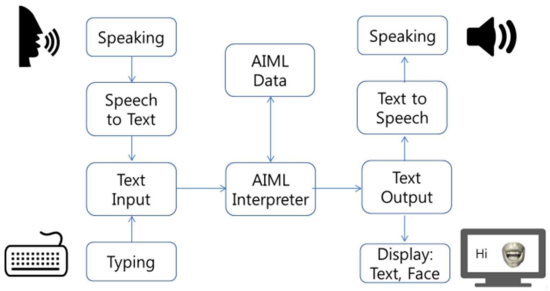

Figure 1. NURI System Overview

NURI has both speech and text interfaces; NURI can convert human speech to texts using the speech recognition (speech-to-text) system, and can convert text data to speech using speech synthesis (text-to-speech).

After receiving text data from a speech-to-text system or direct text input from a keyboard, the text input is sent to the AIML interpreter which will be described later. The AIML interpreter retrieves the most meaningful reply from the AIML dataset. The reply from AIML dataset will be converted to speech using a text-to-speech system. The speech output can also be displayed on a screen either in texts or in the form of an animated face.

3. Speech Recognition and Speech Generation

3.1 Speech Recognition

Figure 2 shows a block diagram of a typical speech recognition system, which illustrates how a speech signal is converted to texts [2]. The speech signal is received through a microphone, and will be converted to a digital format. This digital format can be processed by a series of software modules which make up the speech recognition system.

3.2 Speech Synthesis

Speech synthesis is the process of converting text data to speech. The block diagram in Figure 3 illustrates the process involved in converting texts to speech [2].

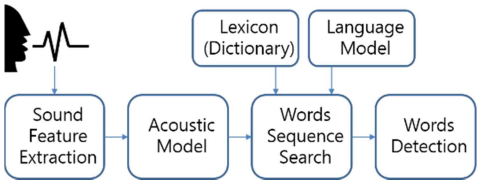

Figure 2. Speech Recognition System

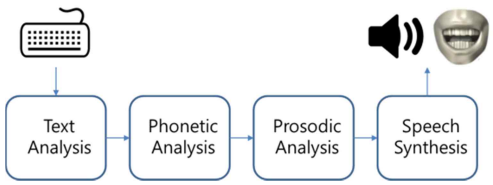

Figure 3. Text-to-Speech System

4. Speech Animation

Speech animation is the process of adopting the mouth movements of animated speaker to the phonemes of a speech sound. Among many variations of phoneme mouth shapes, this study adopts the basic ten shapes of the Preston Blair phoneme series [1]. Figure 4 shows the phoneme mouth shapes. These ten basic phoneme shapes can match almost any sound of speech with the in-between frames moving from one to the other.

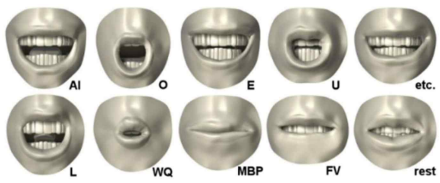

Figure 4. Preston Blair Standard Phoneme Mouth Shapes

5. Natural Language Understanding and AIML

Natural Language Understanding (NLU) is the process of translating natural language into a language interpretable by a computer. There are several approaches for NLU including statistical, pattern matching, syntactically driven parsing, semantic grammars, case frame instantiation, etc. This study follows the line of pattern matching methods.

The pattern matching approach for natural language analysis is based on interpreting the syntax as a whole rather than constructing an interpretation through a combination of structures and semantics of words or other constituents. The interpretation is obtained by pattern matching from words [2].

AIML (Artificial Intelligence Markup Language) is an XML (eXtensible Markup Language)-based language that contains knowledge inside XML tags. Knowledge is stored in a structured way in AIML data so that knowledge is easily accessed whenever required. The AIML interpreter feeds human operator's input to the AILM data, and retrieves the best possible reply from it.

The AIML elements contain AIML objects which are composed of topics and categories. The data are elements that can be interpreted by the AIML interpreter. The following AIML example illustrates a category where the input matching "CHECK SECONDARY ALARM" with output template "I will check secondary alarm".

```
<category>
<pattern>CHECK SECONDARY ALARM</pattern>
<template>I will check secondary alarm</template>
</category>
```

6. Implementation of NURI system using ROS

6.1 ROS

Robot Operating System is a middleware framework that is widely used in robotics. The philosophy is to create functionalities that can be shared and used in other robots easily by making a piece of software that could work in other robots with minimal or no changes in the code.

Figure 5 shows the concept of ROS. ROS master manages the communication between nodes. ROS nodes are single-purpose, executable program. They are individually compiled, executed, and managed without disturbing other nodes. Nodes communicates over topics; nodes can publish or subscribe to a topic; Topic is a name for a stream of messages. A ROS message is a data structure defining the type of a topic [3].

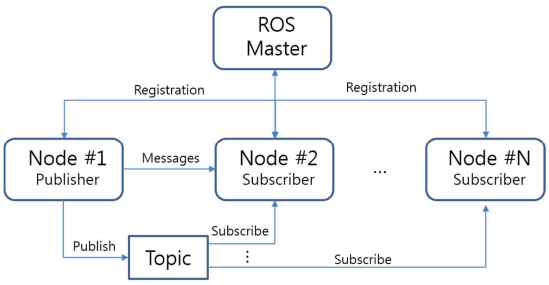

Figure 5. ROS Concept

6.2 Integration of AIML-based AI and Natural Language Interface into ROS Package

ROS Python nodes were developed to handle AIML files. The functionalities described previously were tied into a package in ROS. Figure 6 shows the structure of the ROS package. AIML clients can receive NURI's reponse (from the server) through /text_out topic. The response from AIML server can be either displayed as texts and converted to speech through speech synthesis module along with an animated face. AIML clients can also send human operator's input taken from the keyboard or microphone to the AIML server through /text_in topic.

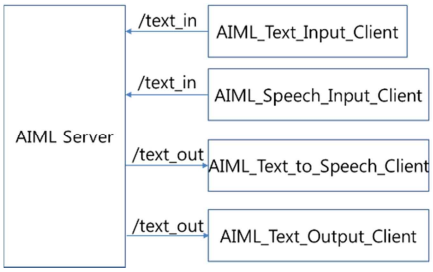

Figure 6. ROS implementation of NURI

7. Conclusion

This study introduces a prototype of NURI, NUclear Reactor facility-wise Intelligence. NURI can react to human operator's voice or text input, through speaking with an animated face or texts on screen.

Future work includes the addition of reactor operation dialogue scenario, and semantic interpretation proposed in [4].

ACKNOWLEDGEMENTS

This work was supported by the nuclear research and development program through the national research foundation of Korea funded by the Ministry of Science and ICT.

REFERENCES

[1] www.garycmartin.com/mouth_shapes.html, last access on Aug. 28 2017.
[2]R. S. Wilkens, "A study of the use of Natural Language Processing for Conversational Agents", M.S. Thesis, 2016
[3]M. Quigley, GB. Gerkey, and W. D. Smart, Programming Robots with ROS, O'Reilly,2016
[4]S. Joo, S. Seo, Y. Suh, and C. Park, "Linguistic Formalism for Semi-Autonomous Reactor Operation", Transactions of the Korean Nuclear Society Spring Meeting, Jeju, Korea, May 18-19, 2017