Transactions of the Korean Nuclear Society Autumn Meeting
Changwon, Korea, October 24-25, 2024

Application of Large Language Model for Natural Language-based Human-System Interaction in Nuclear Power Plants

Seung Geun Kim [a*], Seungdon Yeom [a], Yonggyun Yu [a]
[a]*Applied Artificial Intelligence Section, Korea Atomic Energy Research Institute, Daedeok-daero 989beon-gil, Yuseong-gu, Daejeon 34057, Republic of Korea*
[*]*Corresponding author: sgkim92@kaeri.re.kr*

Keywords : nuclear power plant, natural language, human-system interaction, large language model

1. Introduction

Small modular reactors (SMRs) are gaining attention as they are expected to be beneficial for safety and economic efficiency. In terms of safety, SMRs are generally regarded as safer than conventional large nuclear power plants (NPPs), owing to its much lower power capacity and various passive safety systems. In terms of economic efficiency, SMRs are generally regarded as cheaper than conventional NPPs, owing to the efficient modular production and optimized deployment.

However, the advantages of SMRs could be diminished if the operation schemes of conventional NPPs are maintained. For most of conventional NPPs, one shift team of main control room (MCR) operators consists of five operators: senior reactor operator (SRO), shift technical advisor (STA), reactor operator (RO), turbine operator (TO), and electric operator (EO). This single team of operators operates one reactor at a time.

Owing to the low power capacity of SMRs, multiple reactors should be deployed to achieve similar total power output with large NPPs. If conventional operation scheme is maintained for SMRs, large number of shift teams are needed for the same total power output, which may greatly reduce the expected economic efficiency of SMRs. Accordingly, many SMR developers are considering operation schemes with increased number of reactors operated by one shift team and reduced number of operators per one shift team, to achieve higher economic efficiency.

To implement this new operation scheme, it is essential to verify whether a small number of operators can effectively monitor and control multiple reactors through enhancing the operational efficiency of individual operator. To address this challenge, various strategies are being considered to enable operators to perform their tasks more efficiently. These strategies include not only transferring tasks from operators to systems by automation but also enhancing operational efficiency by incorporating various operation support systems (OSSs) that assist operators in their tasks. These OSSs are designed with the primary objective of enhancing the efficiency of operators in performing their tasks while retaining the ultimate decision-making authority with the operators themselves. Therefore, to design an effective OSS, it is essential to consider not only the performances of algorithms that are required to perform specific tasks but also the issues related to the human-system interaction (HSI).

This study aims to develop a natural language-based HSI system based on large language models (LLMs) and speech recognition technology, to determine whether the interaction between operators and the OSS can be implemented similarly to traditional interactions or communication between operators. The experimental process utilized an integral pressurized water reactor (iPWR) simulator that simulates virtual SMR. The LLaMA3 LLM was utilized as a baseline model, and fine-tuned using nuclear-related documents. Additionally, the system was designed to implement OSS functions by utilizing the 'function call' function of the LLM, allowing it to select and execute the appropriate functions based on the given query.

The rest of paper is organized as follows. In chapter 2, brief backgrounds on LLMs, communication schemes in MCRs, and verbal HSI are provided. The processes of the experiments are described in chapter 3. The summary and the conclusion of the paper is provided in chapter 4.

2. Preliminaries

2.1 Large language models

Large language models (LLMs) are advanced machine learning models that developed for processing various natural language-based tasks. LLMs are still suffering from the biased or inaccurate information on certain topics, and they may sometimes generate content that does not exist in the input or training data, a phenomenon known as hallucination. Despite these limitations, LLMs exhibit the highest performance among the proposed natural language processing (NLP) models. They are widely used in applications such as chatbots, code generation, and language translation, and are also integrated with other generative artificial intelligence models to create content such as images (e.g. DALL-E, Midjourney, Stable Diffusion, Imagen), videos (e.g. Make-A-Video, Phenaki, Runway ML, Synthesia), and music (e.g. Jukebox, MusicLM, AIVA, Amper Music).

To effectively account for the variations and vast scope of natural language, machine learning models for NLP have become increasingly large and complex. Recent LLMs consist of tens or hundreds of billions of parameters and are typically designed using a transformer[1]-based architecture.

As the development of LLMs requires enormous amounts of data, most of them are developed by large corporations with the infrastructure capable of acquiring and processing such data effectively. Representative LLMs are GPT (OpenAI)[2], BERT (Google)[3], LaMDA (Google)[4], LLaMA (Meta AI)[5], and Claude (Anthropic).

Since LLMs are trained on huge amounts of data, they are capable of handling most general conversations and sentences. However, because much of the training data consists of documents that are available on the internet, LLMs may struggle with processing conversations or sentences that require a high level of expertise on domain knowledge. Therefore, when using LLMs for specific purposes, an additional fine-tuning process is necessary to adapt the model to the intended purpose. Many corporations provide baseline LLMs that can be applied across various fields through additional fine-tuning. This process allows for the development of purpose-specific models using significantly less data than the amount of data required for training the original LLMs.

In nuclear field, attempts have been made to introduce NLP and LLM technologies. However, owing to the scarcity of natural language data in nuclear field and the insufficient discussion on the adoption of AI technologies to the nuclear field, there is a lack of NLP application cases, and the introduction of LLMs remains in its early stages. Seong et al.[6] proposed a methodology that applies NLP technology specialized for the nuclear field to help operators easily find similar cases by accessing documents related to limiting operation conditions (LCOs) in NPPs. Kim et al.[7] introduced a new communication framework based on BERT aimed at reducing the number of operators in SMR MCR and tested whether the LLM can accurately recognize verbal queries.

2.2 Three-way communication and verbal HSI

In the current MCRs, operators communicate through a communication method known as 3-way communication. 3-way communication for utility workers, also known as 3-part communication or 3-way handshake, is a critical safety protocol used in high-risk industries such as nuclear industry.

Process of 3-way communication can be summarized as follows. First, the sender initiates the communication with a specific instruction or message (query). Then, the receiver repeats back the message to confirm understanding and answers for the query. Finally, the sender repeats back the answer if the receiver adequately acknowledged the query, or corrects any misunderstandings if the receiver does not.

As 3-way communication involves a query, a response, and a confirmation process, it helps to minimize communication errors based on repetitive cross-checking. In this regard, despite the fact that operators in a digitalized MCR can individually access all necessary information via monitors and display panels, they are still required to communicate based on 3-way communication method. Table I enlists the advantages of 3-way communication.

Table I: advantages of 3-way communication

Advantages	Description
Safety	Significantly reduces the risk of miscommunication
Clarity	Ensures that instructions are clearly understood before action is taken
Error prevention	Allows for immediate correction of any misunderstood information
Accountability	Creates a clear record about instructions and how they were understood
Standardization	Provides a consistent communication method across different teams and situations
Focus	Encourages active listening and engagement from both sender and receiver
Confidence	Builds trust between team members that instructions are being properly communicated and understood
Adaptability	Can be used in various situations, from routine operations to emergency responses

From the perspective of HSI, communication via voice and dialogue is considered to have several advantages. These advantages include not only improved accessibility for users and flexibility of application, but also many other advantages that are related to the nuclear industry. Potential advantages of HSI based on voice and dialogue in nuclear field can be summarized as follows: 1) Speech is a natural and intuitive way for humans to communicate, making it easier for operators to interact with systems. 2) Verbal commands allow operators to interact with systems remotely or while their hands are occupied, which is particularly useful in multitasking scenarios. 3) Speaking can often faster than typing or using physical controls, potentially increasing efficiency in certain tasks. 4) Speaking commands may require less cognitive effort than remembering specific keystrokes or navigating complex menu structures. 5) As auditory sensing is omni-directional while visual sensing is uni-directional, information can be effectively transferred even when the operator is not located in front of or focusing on the display.

There are many remaining problems to consider when applying verbal HSI systems, such as potential issues with reliability in noisy environments and the need for robust NLP capabilities. However, if such HSI systems are properly implemented, they can mitigate the drawbacks and merge the strengths of several interface method. For instance, if information derived from the OSS is transferred to operators through both visual and auditory channels, it could enable operators to quickly assess the overall situation through auditory feedbacks while allowing them to gradually obtain detailed information through visual displays.

3. Experiments

To check the feasibility of LLM application for natural language-based HSI in NPPs, lab-scale experiments were planned and conducted based on simulator. The experiments consist of several steps including: simulator preparation and LLM adaptation to the nuclear domain, establishing HSI system architecture, and case study based on test scenarios.

3.1 Simulator preparation and LLM domain adaptation

To assess the feasibility of a newly developed LLM-based HSI system, this study conducted experiments using an iPWR simulator developed by Spanish company Tecnatom and distributed by the International Atomic Energy Agency (IAEA). The iPWR simulator, designed for educational purposes, simulates a virtual SMR with most systems simplified in their modeling. It includes a variety of malfunction and control options. The virtual SMR has an electrical output of approximately 45 MWe and includes two steam loops. In addition to systems similar to those found in conventional PWRs, it also includes passive systems such as gravity injection system, passive decay heat removal system, and automatic depressurization system. Fig. 1 depicts an overview screen of the iPWR simulator.

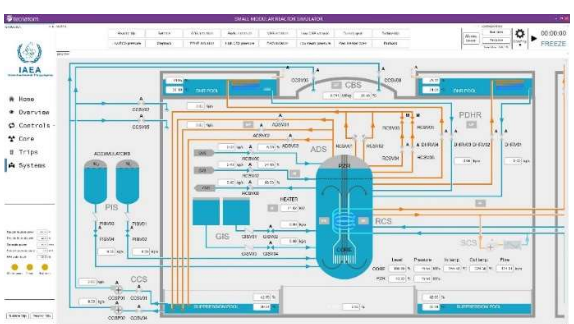

Fig. 1. Overview screen of the IAEA integral pressurized water reactor simulator

The iPWR simulator itself is not inherently designed for integration with other external programs, and therefore requires additional preparation steps to conduct experiments. For tasks related to monitoring, it is necessary to retrieve values of specific variables and device statuses from the iPWR simulator in real time. For tasks related to control, it is essential to issue control commands to the iPWR simulator without the use of manual mouse or keyboard inputs. To enable these functionalities, this study involved preparatory work that allowed direct access to the simulator's process memory by reverse engineering, enabling the reading of variable values and the issuance of control commands.

For the LLM domain adaptation, LLaMA3 LLM was selected as a baseline model and fine-tuned with nuclear field document data. The data used for fine-tuning was sourced exclusively from publicly available documents, such as nuclear terminology dictionaries (provided by Korea Hydro & Nuclear Power), with the majority of the data obtained through web crawling.

3.2 Establishing HSI system architecture

To implement a natural language-based HSI system, integration modules connecting the LLM, the simulator, and the program with specific functions should be developed. At the time of writing this paper, LLM with 'function call' function, simple monitoring and control function, and interaction module between LLM and simulator were developed.

The system operates in the following sequence: the system receives a voice query from the operator. This query is converted into a text query using a 'speech-to-text' module, which is then fed into the LLM trained on nuclear-related documents. The LLM processes the text query and generates an output, utilizing function call capabilities to invoke specific functions when necessary. The LLM interprets the content of the query, selects the appropriate function to handle it, and the selected function interacts with the simulator to perform the requested tasks. For example, when a monitoring function is executed, the system reads specific variable values from the simulator and may analyze trends if necessary. When a control function is executed, the system may issue commands to the simulator to alter the status of specific components.

Once a selected function interacts with the simulator, the corresponding output is generated. The output is firstly provided in text form, and subsequently transformed into sound (speech) using a 'text-to-speech' module. The operator receives feedback in both text and sound formats, allowing them to verify whether the system has correctly understood their query and executed the corresponding command appropriately. Fig. 2 depicts a schematic of the natural language-based HSI system operation scheme.

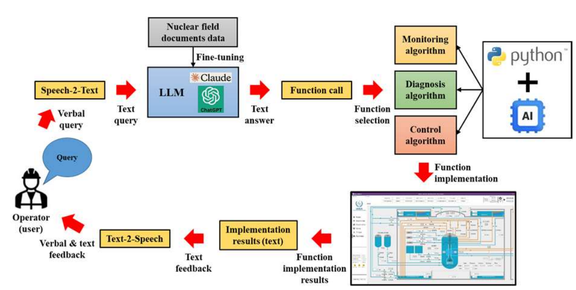

Fig. 2. Schematic of the natural language-based human-system interaction system operation scheme

3.3 Case study based on test scenarios

To test the natural language-based HSI system, scenarios were selected and classified into four levels based on their complexity and task type. At the time of writing this paper, the experiments were conducted on one-off monitoring tasks and one-off control tasks with refined text inputs that exclude any incomplete words or sentences. Table II enlists the selected scenarios.

Table II: list of selected scenarios

Scenario	Description
One-off monitoring	Reading the current value of specific variable
One-off control	Altering the status of specific component to the desired state
Continuous monitoring	Comprehending the changing trend of specific value (assisted with monitoring algorithm)
Continuous control	Controlling multiple components to achieve specific goal (assisted with control algorithm)

For the one-off monitoring tasks, the text inputs were given in a form of 'read me the {value/status} of {variable/component name}'. Totally 56 kinds of variables and 14 kinds of components were considered during the experiments.

For the one-off control tasks, 14 kinds of components considered, similar to that of one-off monitoring tasks. The text inputs were given in a form of 'open/close the {component name} to the {setpoint}' for the valves, and 'on/off the {component name}' for the pumps.

It was found that for both one-off monitoring tasks one-off control tasks, the LLM was capable of monitor/control the values/components adequately with given 'refined' text inputs, for every considered variables and components. However, the experiments conducted thus far are at a very early stage. Future work will involve the integration of speech-to-text and text-to-speech modules to enable speech input, as well as experiments incorporating incomplete inputs. These steps are necessary to verify whether the LLM model can successfully perform its intended role in real-world environments.

In addition, experiments for more complex continuous monitoring tasks and continuous control tasks will be conducted as further works. These tasks can be represented as a collection of multiple one-off monitoring and one-off control tasks, and thus the NLP performance are not expected to pose significant problems. However, these tasks may require algorithms that capable of making more complex judgments depending on the situation. For example, in the selected continuous monitoring scenario, the system should assess the trend of specific variable over a period and provide more nuanced responses, such as 'increasing rapidly', 'decreasing slowly', or 'maintained stable'. To achieve this, an algorithm must be implemented as a kind of function that considers the rate of change over time to accurately evaluate the trend.

4. Conclusions

This study aimed to implement a natural language-based HSI system using LLM and was designed to assess whether interaction between operators and the OSS could be conducted in a manner similar to communication between operators. Experiments were conducted using an iPWR simulator provided by IAEA, and the LLaMA3 LLM fine-tuned with nuclear-related documents. The system was further developed to function as an OSS, based on the 'function call' function of LLMs to facilitate interactions between the operator, LLM, and simulator.

Although only early-stage experiments were conducted based on highly refined cases, the results demonstrated that the application of fine-tuned LLM and utilization of 'function call' function could successfully perform simple monitoring/control tasks in response to natural language text queries and providing appropriate feedback to the operator.

Future work will include the integration of speech-to-text and text-to-speech modules to enable speech input to the system, as well as experiments with incomplete inputs, including those containing noise and wrong words/sentences, to evaluate performance in more realistic scenarios. Additionally, further experiments will be conducted on more complex continuous monitoring/control tasks to explore the potential of the proposed natural language-based HSI system.

ACKNOWLEDGMENT

This research was supported by a grant from Korea Atomic Energy Research Institute (KAERI) (No. KAERI-524540-24)

REFERENCES

[1] Vaswani, Ashish. "Attention is all you need." *arXiv preprint arXiv:1706.03762* (2017).
[2] Achiam, Josh, et al. "Gpt-4 technical report." *arXiv preprint arXiv:2303.08774* (2023).
[3] Devlin, Jacob. "Bert: Pre-training of deep bidirectional transformers for language understanding." *arXiv preprint arXiv:1810.04805* (2018).
[4] Thoppilan, Romal, et al. "Lamda: Language models for dialog applications." *arXiv preprint arXiv:2201.08239* (2022).
[5] Touvron, Hugo, et al. "Llama: Open and efficient foundation language models." *arXiv preprint arXiv:2302.13971* (2023).
[6] Seong, No Kyu, et al. "Retrieval methodology for similar NPP LCO cases based on domain specific NLP." *Nuclear Engineering and Technology* 55.2 (2023): 421-431.
[7] Kim, Taejin, et al. "A Novel Communication Framework for Reduction of Staffs in Advanced MCRs of SMRs: A Prototype Development of a NLP-Based System." *Nuclear Technology* (2024): 1-15.