

# Research on Expert Guidance System of Simulation Intelligent Training in Thermal Power Plant

Jin MA<sup>1</sup>, Qian Qian LIU<sup>2</sup>, Jian YANG and Ningning HAO

*Automation Department of North China Electric Power University, Bao Ding, China*

**Abstract.** With the rapid development of artificial intelligence, in order to solve the problems of many questions from trainees, poor efficiency of teachers in answering questions and high work intensity in the process of thermal power plant simulation practice, this paper proposes an intelligent training expert guidance system for thermal power plant simulation, and this paper uses Visual C++ software programming to implement an intelligent training expert guidance system with voice interaction and video presentation functions, using the speech development package Microsoft Speech SDK, in-depth study of the SAPI it provides, design and implementation of speech recognition function and speech synthesis function. The XML file is used to define the types of questions and the corresponding Q&A test statements for keyword extraction; the video presentation function is realized with the help of Windows Media Player. The test results show that, based on the successful speech recognition, the system can simulate the way of dialogue between teachers and learners and answer learners' questions in real time. The system has an open function, users can add questions independently, expand the database in real time, jump out of the traditional limitations of space, time and geography, and increase the interest of learning.

**Keywords.** Thermal power plant simulation, Microsoft speech engine, speech recognition, speech synthesis, video presentation, expert guidance system

## 1. Introduction

With the continuous improvement and development of current artificial intelligence technology, human-computer interaction has become a development trend. Artificial intelligence technology provides intelligent possibilities for simulating dialogue scenarios between teachers and trainees and database retrieval, and promotes the development of intelligent training expert guidance systems [1].

People have gradually started to study speech recognition technology since the 1950s, and it has gone through 70 years of history [2], which has been widely studied at home and abroad [3]. the emergence of Siri, the fourth generation of Apple's voice assistant in 2011, brought the peak of foreign voice interaction industry [4]; Amazon's smart speaker Echo in 2014 is an important human-computer interaction technology into

---

<sup>1</sup>Corresponding Author, Jin Ma, Automation department of North China Electric Power University, Bao Ding, China; Email: majinwj@sina.com

<sup>2</sup>Corresponding Author, Qian Qian Liu, Automation department of North China Electric Power University, Bao Ding, China; Email: 2592002601@qq.com

the home appliance industry node [5]. Speech recognition continues to make breakthroughs in the research field and is applied to various industries, and voice interaction function has become a development trend. At present, the leading domestic companies such as IFlytek and Baidu have speech recognition accuracy rates of more than 98% [6]. Microsoft provides a speech recognition engine with a more mature continuous speech recognition technology, which can accurately recognize customized recognition commands, which is necessary for the study of this paper.

This paper focuses on the intelligent training expert guidance system, combined with speech recognition technology, to extract some of the training expert's experience that cannot be described visually through the protocol document and cured into the screen video, so that the program partly replaces the instructor's work and reduces the impact of the training instructor on the simulation training due to staff changes such as sick leave, transfer or retirement. The system uses Microsoft design to implement the speech recognition function and speech synthesis function to simulate the scenario of the instructor communicating with the trainee for instruction, answering the trainee's questions through voice dialogue, and realizing the function of video playback through Windows Media Player, let trainees learn independently, solving the time and space constraints and reducing the problems of repetitive answers from instructors. It is also not necessary to retrain new instructors when old instructors have problems such as retirement or transfer, which has good engineering practical significance to improve the effectiveness of trainee simulation training and reduce the work intensity of instructors.

## **2. Expert System Module Design**

### *2.1. System Framework*

The process of thermal power plant simulation training is mainly composed of two parts: questions from the trainees and answers from the teacher. The teacher instructs the trainees to demonstrate the operation on the thermal power plant simulation machine, and the trainees perform unit start-up and shutdown, troubleshooting and other operations [7].

The thermal power plant simulation intelligent training expert guidance system is mainly divided into three layers, namely, application layer, logic layer and data layer. As shown in figure 1, the application layer includes trainee question and answer display, which provides a friendly human-computer interaction interface for trainees; the logic layer includes speech recognition, speech synthesis, question analysis, information retrieval, answer extraction, etc., which provides technical support for the Q&A system and realizes the Q&A function; the data layer is mainly for the storage of empirical knowledge and provides data support for the operation of the guidance system.

Figure 2 shows the workflow diagram of this system. Once the trainee wakes up the system, the system will respond to the trainee's voice input according to the customized voice command recognition pattern. If the system successfully identifies the voice input question, it will proceed directly to the next operation; if it does not, the system will prompt the most similar statement and the trainee will ask the question again. The recognized question will be processed for question analysis, information retrieval, and answer extraction. Finally, the extracted answers will be displayed in text form, voice reading and video presentation.

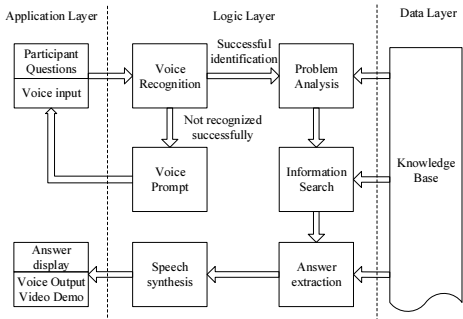


Figure 1. Intelligent expert system structure block diagram.

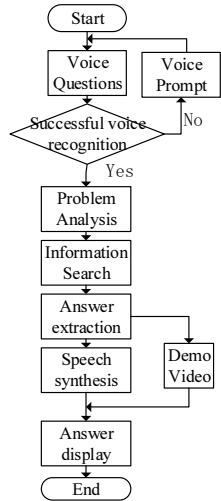


Figure 2. System flow chart.

2.2. Voice Interaction Module

A question and answer system with a voice interface, i.e., a question and answer system implemented with voice as the input and output interface. In this paper, the voice interaction function is applied to the expert guidance system of thermal power plant simulation intelligence training, so that the program can work instead of the instructor to answer difficult questions to the trainees and improve the work efficiency. The system can be generally divided into three processing links, i.e., speech recognition link, text question and answer link and speech synthesis link.

2.2.1. Speech Recognition Link.

Develop speech recognition functions using the speech recognition engine provided by Microsoft to recognize the participant's voice input questions into computer-recognized strings.

2.2.2. Text Q&A Session.

Based on the successful speech recognition, a question and answer test is conducted to do a series of analysis and processing on the recognized string, and the answer part is displayed in text.

2.2.3. Speech Synthesis Session.

Using the speech synthesis engine provided by Microsoft to develop the speech synthesis function, the text answers generated in the Q&A session are presented in the form of voice broadcast to the participants who ask the questions.

2.3. Video Demonstration Module

The operation demonstration video of the simulation training protocol was recorded on the web version of the simulation practice interface of thermal power plant, and the instructor's experiential knowledge was recorded by the simulation teaching and research

department of North China Electric Power University. The video was recorded using Bandicam screen recording software in AVI format; the free video player Windows Media Player provided by Microsoft was used to realize the function of video demonstration; the demonstration video corresponded to the questions and answers in the experience knowledge base, and the system extracted the answers while the corresponding demonstration video was extracted and played automatically.

2.4. Knowledge Quiz Module

Using XML files, we divide the question types, standardize the question forms, organize the types to which text questions belong, and realize question classification. Only one XML file needs to be designed and written, which saves workload and increases the speed of the system's operation. By organizing knowledge in the field of thermal power plant simulation practice, the types of interrogative sentences are roughly divided into three types: location-based questions, start-up questions, and tripping questions.

2.4.1. Location Class Problem.

Which screen is Device in? Where is Device located? How do I find Device? (see table 1)

Table 1. Example of location-based Q&A storage.

Number	Name	Answer Detail	Demo Video
Location-1	Air preheater A	Open the main screen, click on SCS and you will see "Air Prep A"	Video-L1.avi
Location-2	Induced draft fan A	Open the main screen and click on SCS to see "induced draft fan A"	Video-L2.avi
Location-3	Fan A	Open the main screen, click on SCS, and you will see "Fan A".	Video-L3.avi

Where Device indicates the name of each device on the thermal power plant simulation system, and can also refer to each subsystem or switch. The start-up and trip class problems are similar to the position class problems in terms of formulation, so they are not repeated here.

3. Key Technology Analysis

The two key technologies used in this paper are Microsoft Speech SDK 5.1, a speech recognition development kit introduced by Microsoft, and Speech SDK5.1LangPack, a Chinese language patching package, for speech recognition and speech synthesis technologies for application research and promotion; The Extensible Markup Language XML is used to mark up data, define data types [8][9], create syntax rules, have a small memory footprint, and allow users to define their own.

3.1. System Customization Program Implementation Of Microsoft Speech Recognition

The main implementation steps of Microsoft speech recognition development are as follows, and figure 3 shows the speech recognition processing.

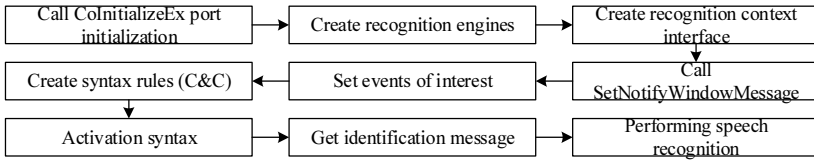


Figure 3. Speech recognition processing process.

### 3.2. Keyword Matching Algorithm

According to the way of searching in the text and the way of moving the search window, the keyword matching algorithm can be summarized into three basic search methods, which are: prefix-based search KMP algorithm, suffix-based search BM algorithm and substring-based search BDM algorithm [10]. BM algorithm has a simpler structure and faster speed compared to the other two algorithms, therefore, the BM algorithm is used in this paper to implement the information retrieval function in the research system.

#### 3.2.1. Bm Algorithm Based On Postfix Search.

The suffix-based search method is to search the longest common suffix of the text and pattern string  $p$  in the window from back to front (right to left along the reverse of the text) in a sliding window sliding along the text [10]. The main algorithms under this search category are the BM (Boyer-Moore) algorithm [11] and the CW (Commentz-Walter) algorithm [12].

The BM algorithm is a precise string matching algorithm that uses a right-to-left comparison method, and applies to two heuristic rules, namely the bad character rule and the good suffix rule, to determine the distance to jump to the right.

##### 1) Bad character rules.

According to the suffix search algorithm, the comparison is done sequentially from right to left until it stops when character  $x$  does not match, as discussed in the following two cases.

i. If the character  $x$  does not appear in the pattern string  $P$ , the  $m$  texts starting from the character  $x$  are not  $P[j](1 \leq j \leq m)$  matched with any of the pattern strings successfully, and all of them are skipped directly.

ii. If  $x$  appears in pattern  $P$ , it is aligned with that character. Let  $\text{Skip}(x)$  be the distance that  $P$  moves to the right, the length of the pattern string  $P$  is  $m$ , and  $\max(x)$  is the rightmost position of character  $x$  in  $P$ . The mathematical formula is expressed as

$$\text{skip}(x) = \begin{cases} m; x \neq P[j](1 \leq j \leq m), x \text{ does not appear in } P \\ m - \max(x); \{k \mid P[k] = x, 1 \leq k < m\}; x \text{ appears in } P \end{cases} \quad (3.1)$$

##### 2) Good Suffix Rule.

When a character is found to be mismatched while a portion of the substring has already been successfully matched, two cases are discussed.

i. During the move, if a substring in  $P[j+1, \dots, m](1 \leq j \leq m) = P[j-s+1, \dots, m-s]$   $P$ , let  $P$  shift right by  $s$  bits.

ii. If there is no character identical to it at any position in P, find the longest prefix that is identical to that suffix, i.e.  $P[s+1, \dots, m] = P[1, \dots, m-s]$ , where the left side of the equal sign is the suffix of  $P[j+1, \dots, m]$  the part that has been compared, and the right side of the equal sign is identical to the prefix of P. Then let P move s bits to the right.

$$\text{Shift}(j) = \min \left\{ \begin{array}{l} s \mid (P[j+1, \dots, m] = P[j-s+1, \dots, m-s]) \ \&\& (P[j] \neq P[j-s])(j > s), \\ P[s+1 \dots m] = P[1 \dots m](j \leq s) \end{array} \right\} \quad (3.2)$$

The best-case time complexity of the BM algorithm is  $O(n/m)$  and the worst-case time complexity is  $O(m*n)$ .

#### 4. Software Implementation

After waking up the system, the student can ask the system the questions he/she encountered. In the first case, if the system can recognize the question asked by the participant, the participant will go directly to the next step of question analysis; in the second case, if the participant's question does not conform to the grammatical rules, the system will prompt the most similar phrase or sentence, and the participant can ask the question again according to the prompt. In this paper, these two cases are introduced separately.

##### 4.1. Example Diagram of Partial Matching Pattern

When the question asked by the user does not conform to the grammar rules of the system, the system will prompt the most similar sentence, and the user can ask the question again according to the prompt. For example, if a student asks "What should I do if Air Prep A doesn't start?", the system answers "I didn't hear you clearly, you seem to be talking about how to start Air Prep A". The result chart of the Q&A test is shown in figure 4.

##### 4.2. Example Diagram of an Exact Match Pattern

When a user asks a question that conforms to the grammar rules, the system can immediately recognize it and perform the next operation, returning the corresponding answer to the user and automatically playing a demonstration video at the same time. As shown in figure 5, figure 6.

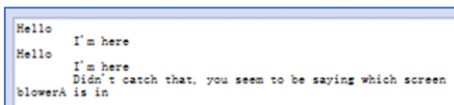


Figure 4. Partial Match mode Quiz Test.

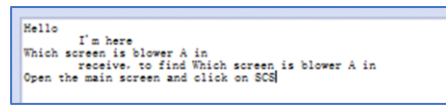


Figure 5. Exact Match mode Quiz Test.

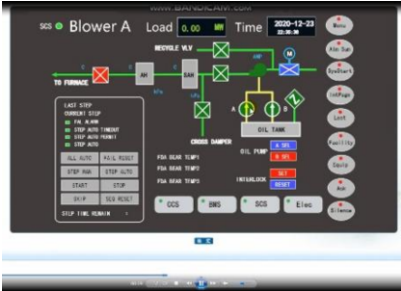


Figure 6. Operation demonstration video.

5. Conclusion

This paper applies speech recognition technology to the expert guidance system of thermal power plant simulation intelligence training, providing a new teaching means for simulation training teaching and a convenient way for trainees to acquire knowledge, which is a successful attempt of intelligent power plant idea in the field of simulation training.

In this paper, Visual C++ is used as the main programming software to implement the expected functions of the intelligent training expert guidance system, which makes full use of the experienced knowledge of training experts to guide the trainees' independent learning. The speech recognition engine and speech synthesis engine provided by Microsoft Speech SDK are used to design and implement the speech recognition function and speech synthesis function of the expert guidance system, and Windows Media Player is used to provide the function of video playback, simulate the dialogue mode between teachers and trainees, and conduct question and answer tests; XML files are used to design the storage Question templates with open function, users can add questions independently, expand the database in real time, jump out of the traditional limitations of space, time and location, increase the interest of students' learning, reduce teachers' work intensity and improve work efficiency.

The system is able to solidify the experience of training experts in on-screen videos that can be reused, solving the problem of training instructors having sick leave, changing jobs or retiring.

References

[1] Duan H Y 2018 Tianjin Normal University. Design and implementation of an intelligent question-answering system on the catechism platform. D

[2] Yao D and Katie S T 2020 Bridging the gap in mobile interaction design for children with disabilities: Perspectives from a pediatric speech language pathologist *J. International Journal of Child-Computer Interaction.* **5** 23-4

[3] Yang G P 2015 Design and implementation of embedded voice interaction technology for accusation system *J. Mechanics and Electronics.* (04):72-4

[4] Liao B Q, Luo P and Ma Y J 2019 Translation robot based on intelligent voice interaction system *J. Information and Computer (Theory Edition).* **31**(17):110-2

[5] Chen Xinyuan 2018 Intelligent voice interaction technology and its standardization *J. Electroacoustic Technology.* **42**(05):78-80

[6] Hao, Eurasia, Wu, Xuan, Liu and Rong kai 2020 Development status and application prospect of intelligent speech recognition technology *J. Electroacoustic Technology.* **44**(03):24-6

- [7] Wang Jing 2020 North China University of Electric Power. Research and development of intelligent question-answering system for simulation practice teaching in thermal power plants. D
- [8] Effantin B, Kheddouci H and Habi A 2017 *Search and Aggregation in XML Documents*. C. Int. Conf. on Database & Expert Systems Applications. springer, Cham
- [9] Lu Yang 2007 Beijing University of Posts and Telecommunications *Design and implementation of Voice XML-based voice service platform*. D
- [10] He Wei 2010 Fast exact string matching algorithm. D. Hefei University of Technology
- [11] Robert S. Boyer, J and Strother Moore 1977 A fast string searching algorithm *J. Communications of the ACM*. **20**(10)
- [12] Commentz-Walter Beate. 1980 A string matching algorithm fast on the average, TR 79.09.007 *J. Pergamon*. **5**(3)