Computer Science and Information Systems 19(3):1533–1548 https://doi.org/10.2298/CSIS210820022S

The Research and Implementation Feasibility Analysis of an Intelligent Robot for Simulating Navigational English Dialogue under the Background of Artificial Intelligence

Wei Sun

Foreign Language College, Anhui Xinhua University,
Hefei 230000, Anhui, China
sunwei@axhu.edu.cn

Abstract. The rapid development of artificial intelligence and robots, the research and development of intelligent dialogue robots are very necessary for today's society. However, a robotic algorithm system that simulates navigational English conversation has not yet been developed. In order to find a suitable algorithm system for dialogue robots, this paper uses the test data set to test the analytical model of navigational English dialogue instructions. The experimental results show that the conditional random field (CRF) + domain dictionary + ambiguity resolution method has the highest segmentation effect. The calculated percentages of the analytical model are correct rate: 76.85%; recall rate: 80.36%; F-value: 88.46%. This paper implements a robot teaching and reproduction method based on simulated navigational English conversation and human-computer interaction under the background of artificial intelligence, and designs robot motion realization experiments and speech recognition experiments. The three-dimensional error after fine-tuning the voice is between 1.6798mm and 2.9968mm. This article constructs a simulation navigational English dialogue robot system. The FAQ component has up to 79.2%; others have a lower accuracy rate of only 59.03%.

Keywords: Artificial Intelligence, Simulated Navigational English Conversation, Dialogue Robot, Teaching and Reproduction, Speech Recognition, Two-Dimensional Error, Conditional Random Field (CRF).

1. Introduction

As AI advances, robots have become the mainstream direction of robot development, and they have increasingly received the attention and attention of researchers. Therefore, the research and development of intelligent dialogue robots are very necessary for today's society.

At present, most of the simulation dialogue robots are still pre-set voice instructions, and then perform specific tasks according to the set instructions, which does not achieve the real human-computer intelligent interaction that people desire [1]. When people express natural language, some words will be omitted because of the context environment, which leads to the text information after speech recognition cannot be correctly converted into the control instructions that can be executed by the simulation dialogue robot. It is unrealistic to simply use the knowledge stored in the knowledge base itself to include all possible situations. This requires the simulation dialogue robot to use a dialogue management technology combined with knowledge base to judge and inquire these unknown or missing information, to complete the missing information by reasoning, and to store the unknown words. As an unstructured way of expression, natural language can be understood by human beings in line with human habits, but simulation dialogue robots tend to understand structured language [2]. How to take efficient methods to process navigational English instructions and convert them into executable instructions of simulation dialogue robots is a subject that needs further study.

In this paper, the test data set is used to test the established model of navigational English dialogue instruction analysis. In this paper, the robot teaching and playback method based on the simulation of navigational English conversation and human-computer interaction is realized, and the robot action realization experiment and speech recognition experiment are designed to verify the method proposed in this paper. This paper constructs a simulated navigational English dialogue robot system which can use knowledge to understand problems on a large scale.

2. Related Work

Many research teams at home and abroad have conducted in-depth research on the scheme research of simulated navigational English conversation under the background of artificial intelligence. In [3], the authors evaluated whether playing a dialogue game could motivate participants to participate in an advance care plan (ACP). The study found that people who played dialogue games had a higher ACP behavioral execution rate within 3 months. In [4], the author used a linear regression model to analyze the random intercept of the speaker. The results also show that the VOT of Hawaiian words ($\beta = -0.01$, $t(289) = -2.0$ and $p < 0.05$) is significantly shorter. In [5], an experienced language teacher is tested in a weekly one-to-one navigational English conversation under the background of artificial intelligence. The results show that all three recasting forms can effectively help learners improve the accuracy of navigational English past tense. In [6], the results showed that the intervention reduced the auditory recognition of the random number formula. The accessible, universal and multilingual nature of popular SNSs such as Facebook has inspired many scholars of second language teaching [7]. In [8-9], the research results show that the interaction mode of navigational English teacher educators is both heterogeneous and homogeneous. In [10], the author studied conversation excerpts from 5 academic conversations, exploring the different ways of metaphorical ideation and the reactions they elicited in conversation partners. In [11], the author checks the acoustic changes of vowels in the speaker during the whole speech task. The overall goal of this study is to understand the differences among speakers as an indicator of the range of normal vowel movement in American navigational English. In [12-13], the author studies the extended use of multiple phone corpora based on three languages. In [14-15], the author used dialogue analysis to study the repair sequence corpus related to pronunciation among Chinese and Japanese students in a Japanese university. The research claims that three segmental repair strategies are used in the interaction process to maintain mutually understandable pronunciation. In [15], as many as 93% of the endings were released by Putonghua speakers, 41% of which were not released by Cantonese speakers. In addition, it is found that Mandarin speakers do not consume human voice, but suck the end with 58% strong voice, which is contrary to most previous studies. In [16-17], the author explores how learners can take collaborative planning tasks as local emergency activities in class. The analysis shows that group planning is essentially a non-linear, social and practical activity, in which students' management participants work together to achieve effective task completion.

Aiming at the algorithm research and system development of intelligent robot, many research teams at home and abroad have carried out in-depth research. In [18], the author evaluates the commercial robots that have been deployed in suburban houses. In [19], the author considers the application of artificial agent in the navigation of semi-automatic mobile robot in the environment with obstacles. In [20], the author puts forward a new type of articulated cantilever sensor structure, analyzes and compares the simple cantilever sensor and articulated sensor. In [21], the author also analyzes the "kidnapping robot problem" and puts forward practical solutions. In [22-23], the author puts forward a new strategy of people to people cooperation and interaction. The proposed strategy is based on data fusion between inertial measurement unit and laser rangefinder to obtain human gait parameters. Finally, the controller shows how the walkers' orientation follows the human orientation in the actual experiment [24]. The unique performance of intelligent materials greatly improves the performance of soft robots [25]. In [26-27], the author proved the feasibility of rotating sensors and smart phone brain for mobile robots. In [28], from the system point of view, the author reviewed the latest work of underwater robots supporting IPMC from the perspective of modeling, manufacturing and biologically inspired design. In [29-31], the author's research results demonstrate that cultural genetic algorithms can effectively solve delays and problems, avoid falling into a peak and guide the mobile robot more effectively.

3. Method

3.1. Navigational English Word Segmentation Algorithm Model Based on Conditional Random Field

Set G = (V, E) as undefined, V is a node in the graph, and E is an undirected edge in the graph. $Y=\left\{Y_v \mid v\in V\right\}$. Let us define an observation sequence X. If each random variable Yv satisfies the following formula, then:

$$p\left(Y_v \mid X, Y_o, \omega \neq v\right) = p\left(Y_v \mid X, Y_\omega, \omega \approx v\right) \quad (1)$$

(X, Y) is called a conditional random field, where the parameter ω≈v represents two adjacent nodes in the graph. Its advantage is that it only needs to take into account the characteristics of the current observed state, and there are no special requirements for independence. In natural language processing, there is:

$$p\left(Y_i \mid X, Y_1, Y_2, \ldots, Y_n\right) = p\left(Y_i \mid X, Y_{i-1}, Y_{i+1}\right) \quad (2)$$

Given a given observation sequence X with a value of x and a random variable Y with a value of y, the parameter expression is shown as:

$$P(y \mid x) = \frac{1}{Z(x)} \exp\left(\sum_{i,k} \lambda_k t_k \left(y_{i-1}, y_i, x, i\right) + \sum_{i,l} \mu_l s_l \left(y_i, x, i\right) \right) \quad (3)$$

Where Z (x) is the normalization factor, λ K and μ L are the corresponding weights, TK and SL are the characteristic functions, K and l represent the number of characteristic functions, and I represent all possible values of Y. The conditional probability defined by CRF is expressed by formula:

$$p(Y \mid X, \lambda) = \frac{1}{Z(X)} \exp\left(\lambda_1 F_j(Y, X)\right), \quad Z(X) = \sum_Y \exp\left(\lambda_j F_j(Y, X)\right) \quad (4)$$

The conditional random field model, like other probability graph models, also needs to solve three standard problems: feature selection, parameter training and decoding. Viterbi algorithm is used in the process of label prediction.

Probabilistic combinatorial category grammar is to apply probabilistic knowledge to combinatorial category grammar, and use the distribution of probability to carry out the optimal solution, so as to eliminate the ambiguity of analytical results, and to eliminate the ambiguity by introducing probability and finding the maximum value of probability. If x is a natural language sentence, Z is a semantic form, and Y is a syntax structure that guides generation, the model representation is as shown in the formula:

$$P(y, z \mid x, w, \Phi) = \frac{\exp^{w\varphi(x,y,z)}}{\sum_{\left(y',z\right)} \exp^{w\varphi(x,y,z)}} \quad (5)$$

Where, Φ is the dictionary, $\varphi(x, y, z) = \left\{ f_1(x, y, z), \ldots f_n(x, y, z) \right\}$ is the eigenvector of N dimension, $w \in R^n$ is the parameter vector, corresponding to the eigenvector, $w \cdot \varphi(x, y, z) = \sum_{j=1}^{n} w \cdot f_i(x, y, z)$. If there are N training samples $\left\{\left(x_i, z_i\right), i = 1 \ldots n\right\}$, the parameter estimation is the optimal w that maximizes the value of the corresponding log likelihood function by solving, as shown in $L(w) = \sum_{i=1}^{n} \log\left(\sum_y P\left(z_i, y \mid x_i; w, \Phi\right) \right)$. For the method of solving the log likelihood estimation function, we can use gradient descent algorithm or inward outward algorithm. Relatively speaking, the efficiency of inward outward algorithm is higher.

3.2. Answer Sorting Algorithm based on Robot Conversation System

(1) Listwise sort learning algorithm. Create a feature vector $x_j^i = \psi\left(q^{(i)}, d_j^{(i)}\right), i = 1, 2, \ldots, m; j = 1, 2, \ldots, n^i$ a for each (query document) pair. Each eigenvector list $x^{(i)} = \left(x_1^i, \ldots, x_n^i(i)\right)$ and its score set $y^{(\mathrm{i})} = \left(y_1^i y_2^i, \ldots, y_n^i\right)$ constitute a training example.

Given a eigenvector list xi, we can get a fraction list $z^{(i)} = \left(f\left(x_1^{(i)}\right), f\left(x_2^{(i)}\right), \ldots, f\left(x_n^{(i)}\right)\right)$.

$$\sum_{i=1}^{m} L\left(y^{(i)}, z^{(i)}\right) \tag{6}$$

When a document list needs to be sorted, given a new document, we can build a feature vector x(i') and adopt sort parameter to give a score to this document. Finally, we sort the candidate documents from high to low based on their scores.

(2) Permutation probability calculation. We denote the arrangement as $\pi = \langle \pi(1), \pi(2), \ldots, \pi(n) \rangle$, where π (j) represents the object at the j-th position in the arrangement.

Suppose there is a sorting function that can score each object. Use s to represent the score set $s = \left(s_1, s_2, \ldots, s_n\right)$, we believe that there is uncertainty when using a sort function to predict a sorted list. In other words, any permutation and combination is possible, but the probability of different permutations and combinations is different for a given sorting function. It has some desired properties in the process of representing the likelihood of permutations and combinations. At this time, given the score list s, the probability of permutation and combination π is:

$$P_{\mathrm{S}}(\pi) = \prod_{j=1}^{n} \frac{\phi\left(s_\pi(\mathrm{j})\right)}{\sum_{k=j}^{n} \phi\left(s_\pi(\mathrm{k})\right)} \tag{7}$$

$s_{\pi(j)}$ is the value of the object at the j-th position of the permutation combination π.

(3) Top k probability. Its formula is shown as:

$$\wp_k\left(j_1, j_2, \ldots, j_k\right) = \left\{\pi \in \Omega_n \mid \pi(t) = j_t, \forall t = 1, 2, \ldots, k\right\} \tag{8}$$

It can be seen that the number of elements here is much less than the number of Ωn elements. The constraints of these permutations and combinations are: the object list $\left(j_1, j_2, \ldots, j_k\right)$ must be at the top k position of the permutation.

3.3. Teaching Reproduction Technology of Dialogue Robot

(1) Robot teaching reproduction based on gesture. The basic principle of the Kalman is to use the state equation of the linear system, the input observation data of the current system and the last-minute estimated system state, and finally obtain the optimal system state estimation function through iteration. The iterative model of state estimation of standard Kalman filter includes system state equation and observation equation:

$$\begin{aligned} x_k &= F_k x_{k-1} + B_k u_{k-1} + w_{k-1} \\ z_k &= H_k x_k + v_k \end{aligned} \tag{9}$$

Each iteration of the Kalman filter is divided into two steps: the prediction process and the update process. The observation data and prediction are fused by the product of independent Gaussian distributions. The idea of maximum likelihood is used to suppress noise, thereby obtaining the optimal estimation of the system state. The iterative process is as follows:

Forecasting process:

1) Forecast system status: $\tilde{x}_k = F_k \tilde{x}_{k-1} + B_k u_{k-1}$

2) Prediction system error covariance matrix: $\tilde{P}_k = F_k \tilde{P}_{k-1} F_k^T + Q_{k-1}$

Update process:

3) Update the Kalman gain matrix: $K_k = \tilde{P}_k \tilde{H}_k^T \left[H_k \tilde{P}_k \tilde{H}_k^T + R_k \right]^{-1}$

4) Estimate system covariance: $\hat{P}_k = \left[I - K_k H_k \right] \tilde{P}_k$

5) Estimate system status: $\hat{x}_k = \tilde{x}_k + K_k \left(z_k - H_k \tilde{x}_k \right)$

(2) Speech-based robot teaching and reproduction. Assuming that the acquired position data of the gesture is represented as $X = (x, y, z)^T$, the model of the adaptive double-exponential smoothing filter can be expressed as:

$$\begin{aligned} b_n &= \beta \left(\hat{X}_n - \hat{X}_{n-1} \right) + (1-\beta) b_{n-1} \\ \hat{X}_n &= \alpha X_n + (1-\alpha) \left(\hat{X}_{n-1} + b_{n-1} \right) \end{aligned} \tag{10}$$

In the formula, X_n represents the position measurement value of the gesture at time n, $\hat{X}_n$ is the gesture position value output by the filter at the current moment, $\hat{X}_{n-1}$ is the gesture position value output by the filter at the previous moment, α and β are adaptive weight factors, The value range is α, β ∈ (0,1), b_n represents the trend of input data at the current moment, and b_{n-1} represents the trend of the previous moment. From the two equations of the filter model, it can be seen that the trend b_n is calculated by exponential filtering of the difference between the first two outputs of the filter, and then the current trend b_{n-1} and the previous output $\hat{X}_{n-1}$ of the filter are used to calculate the output of the filter [32]. The trend b_n reduces the delay of the filter. The weighting factor β that weights the input data is used to calculate the trend b_n. Therefore, the weighting factor β controls the sensitivity of the trend b_n when the input data changes.

4. Experiment

4.1. Data Source

In order to make the results objective, a mature annotated corpus, people's daily, was selected as the training set. The corpus was developed by the Institute of linguistics of Peking University and provided to the researchers. Because it is a simulated navigational English dialogue oriented to the professional field, 150 sentences were collected as the test set for the evaluation experiment.

This experiment uses three different ways to simulate navigational English conversation: conditional random field, conditional random field and domain dictionary, conditional random field, domain dictionary and ambiguity elimination. We use the Python extension package provided by Harbin University of technology. The language model of the extension package includes: Modules such as unknown word recognition, navigation English, etc. In this experiment, only the Chinese word segmentation module is used. First, the module is trained through the corpus of people's daily to establish a language model, and then tested in different ways. The format of the domain dictionary used in the experiment is shown in Table 1.

Table 1. Example of domain dictionary

Terms	Candidate part of speech	Terms	Candidate part of speech
Coffee table	nz	Refrigerator	nz
Wash basin	n	Fruit knife	nz
Color TV	ns	Wooden dining table	nz

4.2. Description of Experimental Scene

The real scene of teaching reproduction experiment based on simulated navigational English dialogue and human-computer interaction is as follows: the operator stands at the table, wears Hololens glasses, and observes the movement of the robot. Before the teaching, the robot has completed three-dimensional registration and overlapped on the real robot. Kinect placed on the table is used to capture the operator's voice commands, and the standard camera beside the tool board is used to observe the movement of the robot's end axis. The hardware involved in the teaching experiment is shown in Table 2:

Table 2. Hardware equipment of analog navigational English conversation

Serial number	Name	Number	Features
1	Kinect sensor	2	Gesture tracking and speech recognition
2	Camera	1	Observe the end axis movement of the robot
3	Dialogue robot	1	Simulation dialogue experiment tool
4	HoloLens glasses	1	Augmented reality display
5	Experiment Tool Board	1	Simulation dialogue experiment tool
6	Computer	2	navigational English dialogue reproduction program
7	Wireless Router	1	HoloLens communicates with a PC

4.3. Experimental Verification

(1) Speech recognition. During the experiment, the tip of the left index finger of the operator moves along the center line of the S-shaped groove on the tool board, the Kinect of the hand eye position tracks the gesture posture of the operator, which is converted into the joint angle required for the movement of the dialogue robot, and sent to Hololens through the wireless network. The operator wears Hololens glasses to observe the situation of the dialogue robot moving with the fingertip. When the operator speaks the voice command "follow the fingertips," an navigational English conversation is started and the action is realized. When the action is completed, the operator only needs to say the voice command "dialogue completed" to stop the dialogue and movement, and then the dialogue robot moves to the initial position. This experiment mainly verifies the performance of speech recognition and action realization.

(2) Action realization. During the experiment, the operator simulated the navigational English conversation of the robot through two natural human-computer interaction modes, gesture and voice. The operator first uses gestures to guide the end of the dialog robot to the target hole. Since the diameter of the robot end axis is 16mm, the diameter of the target hole is 20mm, and the end axis of the robot is 2mm away from the center of the target hole, the jack cannot be completed. The voice instructions are fine-tuned and inserted into the target hole.

4.4. Evaluation Criteria

According to the characteristics of the word segmentation system, the performance of each system is mainly measured by using three indicators of accuracy (P), recall (R), and

F-measure as evaluation indicators. The calculation formulas are as shown in the formulas:

$$Accuracy = \frac{Number\ of\ correct\ results}{Total\ number\ of\ results}$$
$$Recall = \frac{Correct\ results\ in\ word\ output}{Correct\ results\ in\ the\ test\ set} \tag{11}$$
$$F = \frac{2 \times P \times R}{P + R}$$

5. Results and Discussions

5.1. Analysis of Navigational English Word Segmentation and Speech Based on CRF

Word segmentation based on different combinations of CRF models yields different results, and calculations of various performance indicators are performed according to evaluation criteria. This article summarizes and compares the data obtained from the evaluation of various models. The results are shown in Table 3 and the corresponding histogram 1. It is found through the chart that combined with navigational English literature knowledge, a dictionary of navigational English conversation has been added, and the word segmentation after conditional random field combined with the field dictionary has significantly improved accuracy, recall, and F value compared with the word segmentation using conditional random field. Combined with the ambiguity resolution method, 6.0%, 6.2%, and 5.3%, the corresponding evaluation index after the experiment is the highest among the three methods, which are 9.4%, 8.4%, and 10.2% higher than using the CRF model alone.

Table 3. Comparison of segmentation results of different combinations

Word segmentation model	Accuracy	Recall	F value
CRF model	0.804	0.799	0.802
CRF Model + Domain Dictionary	0.864	0.861	0.855
CRF model + domain dictionary + ambiguity resolution	0.898	0.883	0.904

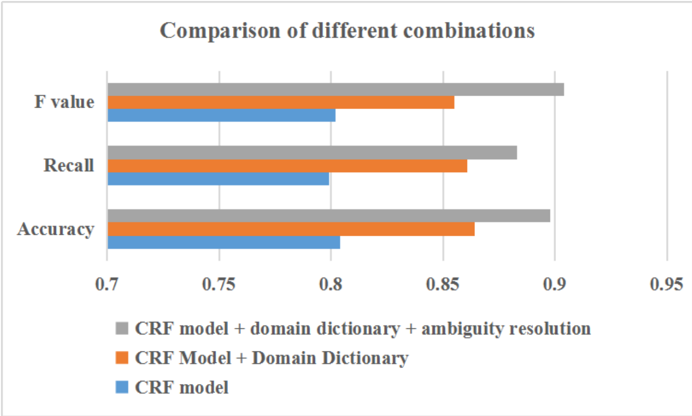

Fig 1. Comparison results of different combinations

Through the above experimental steps, a semantic analysis model for dialog robots is initially established. The analytical model is tested through the test set. Three major indicators are evaluated based on the test. The percentages of the analytical model are calculated as correct rate: 76.85%; Recall: 80.36%; F-value: 88.46%. Experiments show that the semantic parsing model can parse navigational English natural language instructions after word segmentation, and nearly 77.95% of sentences can be parsed and output correctly, which means that the semantic parsing model can convert navigational English into an intermediate representation. Finally, internal instructions that the robot can recognize are generated.

It is known from experiments that the CRF model can significantly improve the accuracy of word segmentation by combining with the domain dictionary and ambiguity resolution. The main reason is that the domain dictionary can identify proper nouns and unregistered words in the domain, and can correct the word segmentation results in the domain, and the use of ambiguity removal technology can further solve the problem of result ambiguity on this basis.

5.2. Analysis of Teaching Results of Dialogue Robots

In the simulated navigational English dialogue teaching and reproduction, the teaching accuracy is completely dependent on the accuracy of the gesture pose data acquisition, so the corresponding filtering algorithm is used to promote the precision. Gesture and posture data is filtered by Kalman filter. The filtering effect is shown in Figure 2. The blue dotted line in the figure is the gesture posture information (observed value) directly obtained by the gesture tracking system. The green solid line is after Kalman filtering. It can be seen from the figure that the attitude data directly acquired by the sensor has a lot of random noise, which fluctuates greatly. After Kalman filtering, the noise is effectively filtered, making the data smoother and more stable, and improving the stability of the teaching system. Because the tool board is placed horizontally and both are within two degrees.

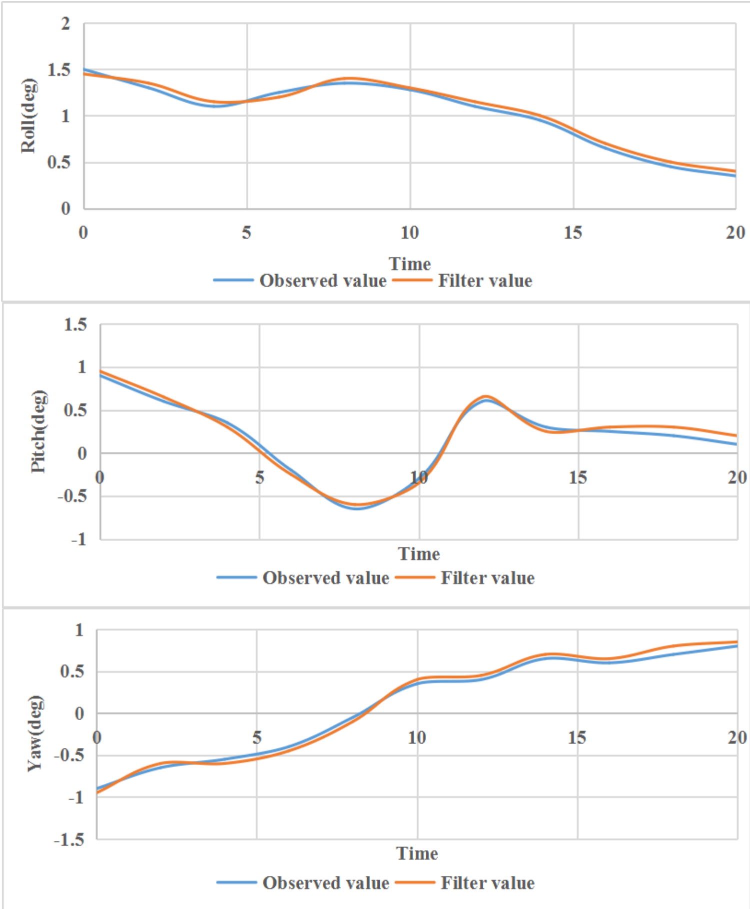

Fig 2. Filter comparison of gesture and pose data

The comparison is shown in Figure 3. The blue dotted line in the figure is the gesture position information (observation value) obtained directly by the gesture tracking system, and the green solid line is the estimated value (filter value) after the adaptive double exponential smoothing filter. Because the robot end moves perpendicular to the tool plate placed horizontally, only the X and Y direction motion data are considered. It can obtain a stable and smooth trajectory.

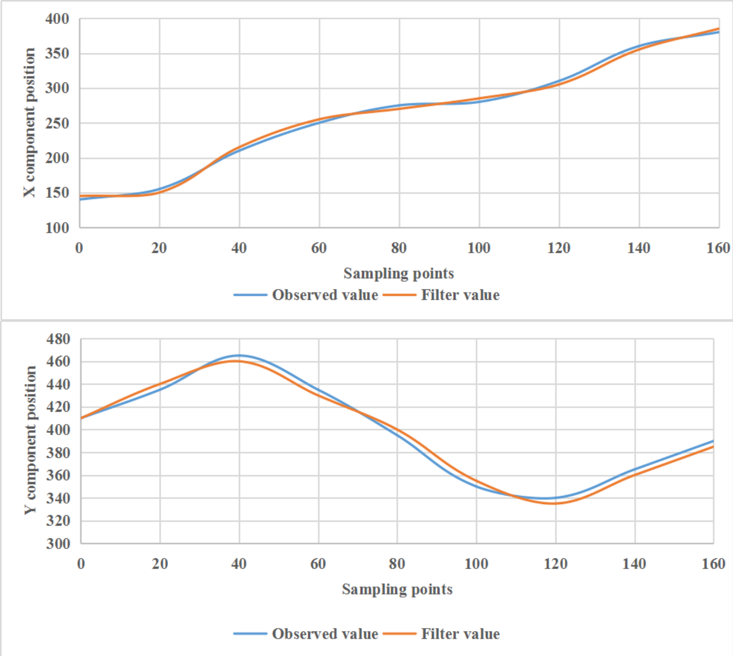

Fig 3. Filtering comparison of gesture position data

In the navigational English conversation and action realization reproduction experiment, first guide the robot end axis to the top of the target hole, because the accuracy of gesture teaching is 5mm, and the robot end axis can not complete the jack as long as it is 2mm away from the hole center, so gesture teaching can not meet the requirements of the jack, and can not carry out the actual jack, so only analyze the two-dimensional error, as shown in Table 4. Then use the voice command to fine tune the robot end axis after hand potential teaching to complete the jack. Table 5 shows the experimental error data after speech fine teaching, including two-dimensional error and three-dimensional error. The change of two-dimensional error after gesture teaching and speech fine teaching is shown in Fig. 4.

Table 4. Two dimensional error of gesture Teaching

Test number	1	2	3	4	5	6
X-axis error	2.6513	-1.0006	0.9913	3.6438	2.9956	1.1124
Y-axis error	-1.5688	3.2511	1.2293	3.5551	-1.8864	1.6643
Two-dimensional error	2.6781	3.4437	1.6606	5.0007	3.8511	1.8888

As shown in Figure 4, the error range of gesture teaching is relatively large. and those with small error can be directly inserted into the hole without the need of voice teaching for adjustment, and those with large error can be successfully inserted into the target hole within 2mm after voice fine-tuning. The three-dimensional error of speech fine tuning is between 1.6798mm and 2.9968mm. The errors of gesture teaching and speech teaching are in millimeter level. There are two factors that affect the robot's end trajectory. One is that there is a small error in the modeling process of dialogue robot.

The simulation results show that the error is 0.1684mm; The second is that the accuracy of the registration of the dialogue robot on the real robot is not high, which is limited by the 3D registration technology based on vision, and its registration accuracy is within 1.25 mm, thus affecting the accuracy of the simulation dialogue of the robot.

Table 5. Experimental error after fine speech teaching

Test number	1	2	3	4	5	6
X-axis error	0.3316	-1.0058	0.6687	0.2175	0.6799	0.7361
Y-axis error	-1.4431	1.2221	.11633	0.1999	1.3467	1.3491
Z-axis error	2.2655	1.4389	3.2451	2.1137	2.8134	2.6643
Two-dimensional error	1.9934	1.2964	1.2473	0.2946	1.0673	1.3450
Three-dimensional error	2.8115	1.8873	3.0521	2.0455	2.9431	2.9910

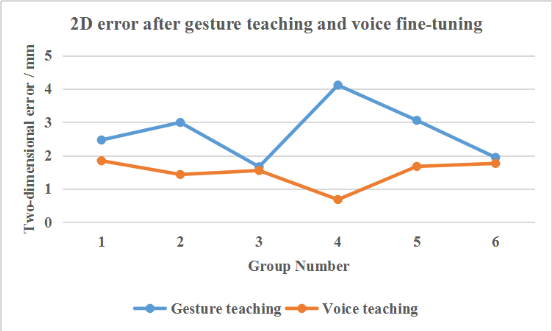

Fig 4. Two dimensional error after gesture teaching and speech fine tuning

5.3. Result Analysis of Answer Sorting of Dialogue Robot

(1) Sorting accuracy analysis. In order to compare the advantages and disadvantages of different sorting methods, three sorting methods were used in the experiment: (1) sorting purely according to the correlation degree of BM25, without considering any other sorting factors, this method is called rawrank; (2) Listnet sorting learning method we use; (3) Ranknet is the representative of pairwise sorting learning method and the most effective method in pairwise, so ranknet method is selected as the contrast object in this paper. We divide the training data set into five sub sets on average and carry out 5 fold cross validation. In each round of test, 3 data are used as training data, 1 data is used for verification, and the remaining 1 data is used for test. For Ranknet and Listnet, verifying which data is used to determine the number of iterations. Figure 5 and Table 6 show the accuracy.

Table 6. Ranking accuracy (map metrics)

Algorithm	ListNet	RawRank	RankNet
Result	0.394	0.267	0.202

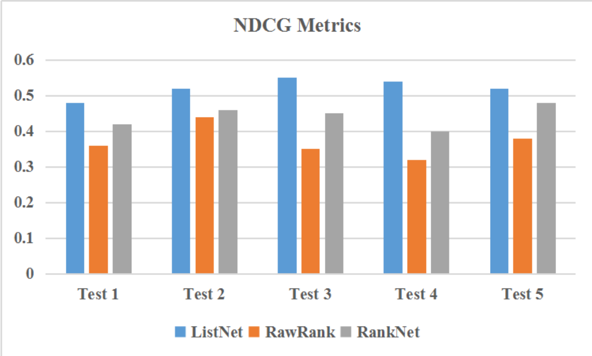

Fig 5. Sorting accuracy (NDCG metrics)

(2) Performance analysis of dialogue robot prototype system. The NLP part can solve 430 questions. In the FAQ section, 312 input problems can be successfully found, But in general, there were still no answers to the 219 input questions, so the recall rate was 78.1%. As can be seen from the second column of Table 7, the FAQ component has up to 78.13%.

Table 7. Evaluation results of FAQ and NLP components

	Return times	Precision
FAQ	312	78.13%
No-wiki-NLP	399	72.43%
wiki-NLP	31	60.98%
ALL	802	73.62%

Analyzing the experimental data, we found that in the first two months of the system's launch, the regular matcher component solved 1.1% of the problems raised by users. Later, in the first four months of the system's launch, we found that 2.01% of the problems had been resolved in the regular matcher. This shows that the regular matcher component works as we wish.

6. Conclusions

In this paper, a test data set is used to test the established navigational English dialogue instruction analysis model. The results of this work proves to be efficient, and the CRF + domain dictionary + ambiguity analysis method has the highest word segmentation results. The calculated percentages of the analytic model are correct rate: 76.85%; recall rate: 80.36%; F-value: 88.46%. Nearly 77.95% of sentences can be parsed and output correctly.

In this paper, a robotic teaching and reproduction method based on simulated navigational English conversation and human-computer interaction is implemented, and robot motion realization experiments and speech recognition are used to test methods proposed in this paper. The gesture teaching with large error can be smoothly inserted into the target hole after the two-dimensional error can be within 2mm after fine-tuning the voice. The three-dimensional error after fine-tuned speech is between 1.6798mm and 2.9968mm.

This paper constructs a simulated navigational English dialogue robot system that can use knowledge to understand problems on a large scale. It consists of a three-tier system. The FAQ component has up to 79.2%.

References

1. Šoić, R., Vuković, M., Ježić, G. (2021). Spoken Notifications in Smart Environments Using Croatian Language. Computer Science and Information Systems, Vol. 18, No. 1, 231–250.
2. Rakesh kumar S, Muthramalingam S, F. Al-Turjman, (2021). "Multimodal News Feed Evaluation System with Deep Reinforcement Learning Approaches", ACM Transactions on Asian and Low-Resource Language Information Processing, vol. 20, no. 1, pp. 8 – 20.
3. Yin, S., & Yuschenko, A. S. (2019). Dialogue system of controlling robot based on the theory of finite-state automata. Mekhatronika Avtomatizatsiya Upravlenie, 20(11), 686-695.
4. Drager, K., Grama, J., Gonzalez, L., & Copeland, C. (2016). Voice onset time and closure duration of word-initial voiceless plosives in pidgin conversation. Acoustical Society of America Journal, 140(4), 3052-3052.
5. Jain, D., Jakhalekar, I. R., & Deshmukh, S. S. (2017). Navigational strategies and their neural correlates. Journal of the Indian Institute of Science, 97(4), 511-525.
6. Sara, D., Steve, S., Ha Phuong, V., Lan-Xi, D., & Ana, J. (2017). Rendering website traffic data into interactive taste graph visualizations. Big Data and Information Analytics, 2(2), 107-118.
7. C. Warner, & H.-I. Chen. (2017). Designing talk in social networks: what facebook teaches about conversation. Language Learning & Technology,21(2), 121-138.
8. Edgar Lucero, & Jeesica Scalante-Morales. (2018). navigational English language teacher educator interactional styles: heterogeneity and homogeneity in the elte classroom. HOW, 25(1), 11-30.
9. James F. navigational English, & Ted Underwood. (2016). Shifting scales: between literature and social science. Modern Language Quarterly, 77(3), 277-295.
10. Azmuddin, R. A., Nor, N., & Hamat, A. (2017). Metacognitive online reading and navigational strategies by science and technology university students. Gema Online Journal of Language Studies, 17(3), 18-36.
11. Christina Kuo, & Gary Weismer. (2016). Vowel reduction across tasks for male speakers of american navigational English. Journal of the Acoustical Society of America, 140(1), 369-383.
12. Lee, Heeju, Su, Danjie, & Tao, Hongyin. (2018). A crosslinguistic study of some extended uses of what-based interrogative expressions in chinese, navigational English, and korean. Chinese Language & Discourse, 8(2), 137-173.
13. Pearson, M. (2016). Staging power in tudor and stuart navigational English history plays: history, political thought, and the redefinition of sovereignty by kirstin m. s. bezio. Theatre Journal, 68(4), 698-699.
14. Brenda Henry-Offor. (2017). Early modern women in conversation by katherine r. larson. Esc navigational English Studies in Canada, 43(1), 110-112.
15. Cicenia, J., Avasarala, S. K., & Gildea, T. R. (2020). Navigational bronchoscopy: a guide through history, current use, and developing technology. Journal of Thoracic Disease, 12(6), 3263-3271.
16. Lu Wanling. (2016). The examination of acoustic feature of navigational English obstruent coda by mandarin and cantonese speakers. Acoustical Society of America Journal, 140(4), 3336-3336.

17. Shu, W., Dong, Z. Y., F Luo, Ke, M., & Zhang, Y. (2018). Stochastic collaborative planning of electric vehicle charging stations and power distribution system. IEEE Transactions on Industrial Informatics, 14(99), 321-331.
18. Paul Gilmore. (2017). Charles brockden brown's romance and the limits of science and history. ELH navigational English literary history, 84(1), 117-142.
19. Oh, Y. H., Kim, J., & Ju, D. Y. (2019). Investigating the preferences of older adults concerning the design elements of a companion robot: analysis on type, weight and material of companion robot. Interaction Studies, 20(3), 426-454.
20. C. Ton, Z. Kan, & S. S. Mehta. (2017). Obstacle avoidance control of a human-in-the-loop mobile robot system using harmonic potential fields. Robotica, 36(4), 1-21.
21. Lisheng Kuang, Yunjiang Lou, & Shuang Song. (2017). Design and fabrication of a novel force sensor for robot grippers. IEEE Sensors Journal, 18(4), 1.
22. Cottrill, C. ., Pereira, F. C., Fang, Z., Dias, I. F., Lim, H. B., & Ben-Akiva, M. E., et al. (2018). The future mobility survey: experiences in developing a smartphone-based travel survey in singapore. Transportation Research Record, 2354(1), 59-67.
23. Cifuentes, C. A., Rodriguez, C., Frizeraneto, A., Bastosfilho, T. F., & Carelli, R. (2017). Multimodal human-robot interaction for walker-assisted gait. IEEE Systems Journal, 10(3), 933-943.
24. H Su, Lallo, A. D., Murphy, R. R., Taylor, R. H., & Krieger, A. (2021). Physical human–robot interaction for clinical care in infectious environments. Nature Machine Intelligence, 3(3), 184-186.
25. Wang, Q., Yang, X., Huang, Z., Ma, S., Li, Q., & Gao, D. W., et al. (2018). A novel design framework for smart operating robot in power system. IEEE/CAA Journal of Automatica Sinica, 5(2), 531-538.
26. T. Li, G. Li, Y. Liang, T. Cheng, & Z. Huang. (2016). Review of materials and structures in soft robotics. Chinese Journal of Theoretical & Applied Mechanics, 48(4), 756-766.
27. K Wang, Yang, C., & Wang, T. (2020). A smart robot training data acquisition and learning process recording system based on blockchain. Open Access Library Journal, 07(9), 1-5.
28. Amir Firouzeh, & Jamie Paik. (2017). An under-actuated origami gripper with adjustable stiffness joints for multiple grasp modes. Smart Material Structures, 26(5), 1-17.
29. Zheng Chen. (2017). A review on robotic fish enabled by ionic polymer–metal composite artificial muscles. Robotics & Biomimetics, 4(1), 24.
30. Yong-feng Dong, Hong-mei Xia, & Yan-cong Zhou. (2016). Disordered and multiple destinations path planning methods for mobile robot in dynamic environment. Journal of Electrical and Computer Engineering,2016(6), 1-10.
31. Luo, D., Chen, D., Wang, J., Zhu, G., & Xu, W. (2020). The smart robot crafting approach to computing materials. Construction Robotics, 4(2), 1-11.
32. Li, H., Han, D. (2021). Multimodal Encoders and Decoders with Gate Attention for Visual Question Answering. Computer Science and Information Systems, Vol. 18, No. 3, 1023–1040.

Wei Sun, was born in Chuzhou, Anhui, P.R. China, in 1982. She received the master's degree from South Central University for Nationalities, P.R. China. Now, she studies in Lyceum of the Philippines University for doctor's degree. Her research interests include linguistics, application of computational intelligence into translation. E-mail: sunwei@axhu.edu.cn

Received: August 20, 2021; Accepted: May 10, 2022.