A General Chinese Chatbot Based on Deep Learning and Its' Application for Children with ASD

Xuan Li, Huixin Zhong, Bin Zhang, and Jiaming Zhang

***Abstract*—Commercial chatbots such as Apple's Siri, Microsoft's XiaoIce, Amazon's Alexa, Jingdong's JIMI, and Alibaba's Alime, have some great prospective in applications such as hosting programs, writing poetry, providing pre-sale consulting and after-sales service in E-commerce, and providing virtual shopping guidance. However, in most cases, existed chatbots in the world are neither designed specifically for children, nor suitable for children, especially for children with ASD (autism spectrum disorder). In order to develop chatbots that are suitable for children with ASD, the present study firstly adopted an open source chatting corpus containing more than 1.7 million question-and-answer Chinese sentences of chatting histories involving children in many cases, and screened out more than 400,000 ideal chatting sentences for model training. Then a generative-based method combing Bi-LSTM and attention mechanism with word embedding based on deep neural network was adopted to build a general Chinese chatbot. The quality evaluation results indicated that our chatbot can successfully intrigue participants' interest and made them understand it well. The chatbot also showed its' great potential for using in the conversation-mediated intervention for Chinese children with ASD.**

***Index Terms*—Chatbots, autism spectrum disorder, natural language processing, generative-based method, deep learning.**

I. Introduction

Autism is a developmental disorder and 1 in 160 children has an ASD worldwide [1]. Common symptoms of children with ASD are deficits in social and language ability, as well as repetitive behaviors and narrow interests [2]. Fortunately, research indicates these abilities can be significant improved by early intervention [3].

In recent years, researchers show increasing interests in using computer assisted technology (CAT) and rehabilitation robots as interventions for children with ASD. Computer assisted technology has long been recognized as an effective treatment for children with ASD. [4] A dozens of research [5]-[7] report that autistic children normally feel comfortable in an environment which is predictable and avoid directly contact with human. Therefore, children with ASD take delight in interaction with computer. Computer assisted therapy for children with ASD contains a wide range of applications. Previous interventions mainly focus on using computer or mobile devices to conduct emotion recognition training [8], [9], language learning [10] and social behavior education [11]. Robot assisted treatment for children with ASD also achieved tremendous success and showed its unique value in clinical practice. Existed rehabilitation robot-based interventions generally concentrated on improving social skills of autistic children through joint attention [12] [13], learning imitation and turn taking [13]. However, none of these interventions apply chatbots which can simulate real communication in real society to enhance the social and communication skills of children with ASD.

Since children with ASD have linguistic difficulties [14] and commonly reluctant to communicate with others [15], early intervention mainly focuses on language teaching. However, researchers and therapists should bear in mind that the final goal of intervention for children with ASD should be facilitating autistic children to return and to apply skills which they learned from interventions to a real world. Therefore, to involve chatbots which could simulate human to human communication would be a valuable and important attempt to reach this goal.

It is necessary to apply chatbots in CAT or robot-based intervention. However, existed chatbots lack adaptability to language comprehension of autistic children. Therefore, the present study designs a general Chinese chatbot based on deep learning and some adjustments were made according to autistic children's language habits and understanding abilities. For example, autistic children with language barrier generally have pragmatic impairments [16]-[18] such as lacking of response or having difficulty in sustaining a conversational topic. Experts consider that the pragmatic impairments may relate to the difficulty in language comprehension [19]. Therefore, in the design of chatbot, we keep responsive sentences as short and simple as possible to make it easier for autistic children to understand. Besides, we conduct pre-training for the word embedding and therefore the imagination of our chatbot has been significant improved and this allow the chatbot generating appropriate response even to single input words. Considering the ages of children, our chatbot selects relatively pure corpus, which do not contain sexual or violent languages and it also makes our chatbot to have an optimistic personality. To our best knowledge, our chatbot is the only general Chinese chatbot designed for meeting the special needs of children with ASD.

The rest of this paper are arranged as follow. Firstly, we briefly review the history and the recent trends of chatbots. Then, in the methodology, we first introduce the construction process of our chatbot and after that, several experiments are conducted to evaluate the quality of our chatbot. The results

Manuscript received August 15, 2019; revised March 13, 2020. This work was supported by the Shenzhen Science and Technology Innovation Commission under Grant JCYJ20170410172100520.

Xuan Li is with the College of Software, Xi'an Jiao Tong University, China (e-mail: lixuan064@ stu.xjtu.edu.cn).

Huixin Zhong is with the Department of Computer Science, the University of Bath, United Kingdom (e-mail: huixinzhong@outlook.com).

Bin Zhang is with the School of Software & Microelectronics, Peking University, China (e-mail: zhangbinxy@pku.edu.cn).

Jiaming Zhang is with the Institute of Robotics and Intelligent Manufacturing, The Chinese University of Hong Kong (Shenzhen), China; He is also with Shenzhen Institute of Artificial Intelligence and Robotics for Society, China (corresponding author; e-mail: zhangjiaming@cuhk.edu.cn).

doi: 10.18178/ijmlc.2020.10.4.967

suggested that our chatbot shows the ability of simulating human to human interaction and adult users are satisfied with the experience of interacting with our chatbot. Therefore, the present chatbot reveals a great potential to apply in the clinical practice for the treatment of children with ASD. However, actual application effects of our chatbot is still waiting for the test among children with ASD.

II. Background

The world's first chatbot so-called ELIZA was developed by Joseph Weizenbaum in 1966 [20], which was used to imitate psychologists in clinical treatment. Although ELIZA only applied keyword matching and manual response rules, Weizebaum himself was surprised at ELIZA's performance at that time. Later, other chatbots appeared one after another. In 1988, Robert Wilensky and others developed a chat robot system called UC (UNIX Consultant) [21], which helps users learn how to use UNIX operating system. Subsequently, in 1995, Dr. Richard S. Wallace developed the ALICE system, which is one of the best performing chatbots based on template matching. In 2014, Microsoft launched an companion robot, XiaoIce [22], an 18-year-old AI girl, who can host programs, write poetry, sing and compose songs. In the same year, Jingdong released its self-developed e-commerce robot JIMI [23], which can provide full-time, unlimited services. Its' function covers the whole process of E-commerce including pre-sale consulting and after-sales service. In 2015, Alibaba Group released a virtual shopping assistant named Alime [24], which allows customers to enjoy one-to-one, full-time shopping experience. More and more commercial applications reflect the broad application prospects of chatbots.

Nowadays, chatbots have made great progress with the rise of artificial intelligence. Although the chatbots based on rule and retrieval [25], [26] has been studied for a long time, the pre-set rules and databases will not answer the questions that have never appeared during training or development process. To solve such problem, generative-based method [27] has irreplaceable technical advantages such that the program is expected to answer all the questions through automatically generated replies. In addition, compared with traditional methods, End-to-End based data-driven dialogue generation [28], eliminates a lot of feature extraction and processing of various complex intermediate steps, such as parsing and semantic analysis, which is the unavoidable work in traditional natural language processing. Therefore, generative-based method greatly improves the efficiency of system development and have better extensions. Based on the above advantages, the present study adopted a generative-based method to build a chatbot.

III. Research Methods

A. The Construction of the Chatbot

1) The construction of the data base

The present chatbot uses an open source chatting corpus [29]. The original corpus contains more than 1.7 million question-and-answer sentences of chatting histories involving children. After data cleaning and processing, including eliminating duplicate sentences and illegal sentences, and the unified processing of Chinese punctuation symbols, we finally screened out more than 400,000 ideal chatting sentences for model training.

2) An overview of the chatbot system

The chatbot system uses Bi-LSTM to encode sentences. On this basis, attention mechanism is introduced to improve the generation effect. At the same time, word embedding based on deep neural network is used to optimize the sentence representation that it learns from large data sets. The combination method is convenient and fast, and the evaluation experiments proves that our chatbot has a good effect.

a) Cyclic neural network unit

In the encoder part, we use bidirectional LSTM unit to achieve better coding effect than unidirectional LSTM. The RNN unit we use in the decoder is the LSTM unit. LSTM controls input, memory and output values by introducing update gate units, forget gate units and output gate units. The structure of LSTM cell is shown in Fig. 1.

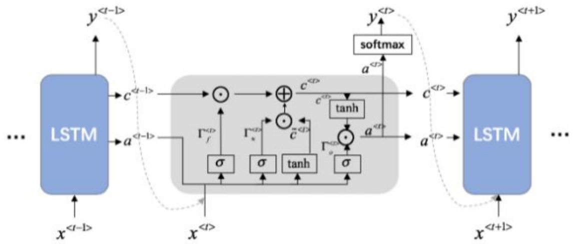

Fig. 1. The graph of LSTM framework.

The calculation of T-Time forget gate, update gate and output gate, as shown in formulas (1), (2) and (3), is obtained by the hidden state of the previous time $a^{<t-1>}$ and the input of t-time $x^{<t>}$, through the full connection of their respective weight matrices *w*, together with their respective bias term parameters *b*, and finally by substituting sigmoid function.

$$\Gamma_f^{<t>} = \sigma(W_f[a^{<t-1>}, x^{<t>}] + b_f) \quad (1)$$

$$\Gamma_u^{<t>} = \sigma(W_u[a^{<t-1>}, x^{<t>}] + b_u) \quad (2)$$

$$\Gamma_o^{<t>} = \sigma(W_o[a^{<t-1>}, x^{<t>}] + b_o) \quad (3)$$

Cell unit at t time is calculated from cell unit at the previous time $c^{<t-1>}$ and current time $\tilde{c}^{<t>}$. Among the formula (4), $\tilde{c}^{<t>}$ is calculated by formula (5).

$$c^{<t>} = \Gamma_f^{<t>} \circ c^{<t-1>} + \Gamma_u^{<t>} \circ \tilde{c}^{<t>} \quad (4)$$

$$\tilde{c}^{<t>} = \tanh(W_c[a^{<t-1>}, x^{<t>}] + b_c) \quad (5)$$

Finally, the hidden layer state at *t* time $a^{<t>}$ is obtained by multiplying the cell unit $c^{<t>}$ at the current time by tanh function with the output gate, as shown in formula (6).

$$a^{<t>} = \Gamma_o^{<t>} \circ \tanh(c^{<t>}) \quad (6)$$

b) Sequence-to-sequence model and attention mechanism

Sequence to Sequence Model trains two RNs, one RNN as an Encoder to encode the input sequence and the other RNN as a Decoder to decode the hidden layer state at the end of the encoder in order to decode the corresponding output sequence. The model structure of this project is shown in Fig. 2.

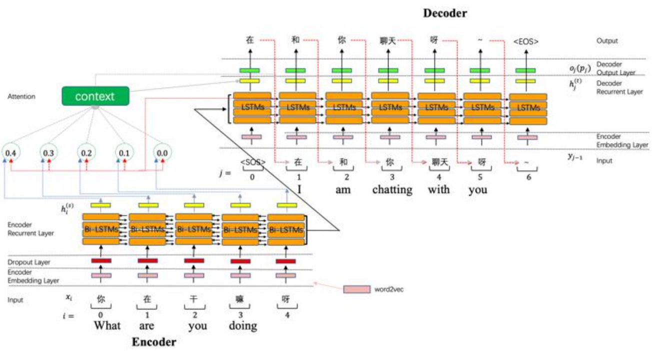

Fig. 2. Sequence to sequence model.

The basic sequence-to-sequence model only uses the last hidden layer state in the Encoder part. Although the final output hidden layer state contains the state information of all hidden layers, the closer the input value is to the last moment, the greater the probability of being retained. The hidden layer state information near the initial time will be less retained in the final output hidden layer state. Most importantly, the attention-based sequence-to-sequence model [30] has proved to be superior to the cyclic neural network baseline system [31] in the application of chat robots. Therefore, the attention mechanism proposed by Bahdanau [32] in 2015 is introduced into our model. Attention structure diagram is shown in Fig. 3 Attention mechanism learns the attention weights of all hidden layer states of the encoder through a shallow neural network, and finds out the hidden layer states of the encoder after adding the attention variables at each time.

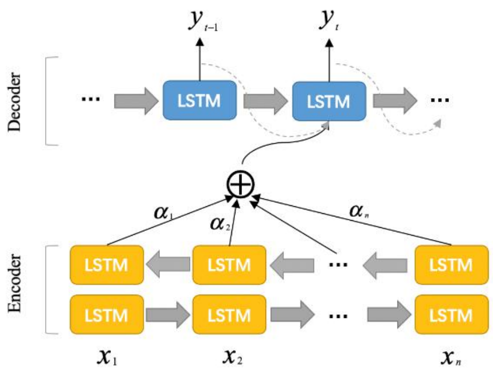

Fig. 3. Attentive Seq2Seq model.

c) Pre-trained word embedding

One-hot vectors encoding the representation of vocabulary will lose the semantic information between words. Therefore, we need word embedding based on deep neural network to change the high dimension into the low dimension, while retaining the semantic relevance between words as much as possible.

Words embedding is a method of representing words by creating a high-dimensional vector space in which similar words are adjacent to each other. Universal word embedding has always been the goal of word vector representation. Pre-trained embedding in large corpus can insert various downstream task models such as affective analysis, classification, translation and extra to improve its performance automatically by combining some common words or sentences representation learned on larger datasets.

Therefore, the word vector representation trained by Chinese word 2vec is loaded into the model to initialize the sub-word embedding, and then the word embedding is fine-tuned in the training process.

d) Beam search decoder

Cluster search algorithm is a heuristic algorithm, which extends the basis of greedy search. It returns the list of output sequences with the greatest possibilities. Compared with greedily choosing the most likely next step when constructing a sequence, the beam search algorithm tracks K states. It starts with K randomly generated States and generates all successors of all k states in each step. If any of these successors is the target, the algorithm stops. Otherwise, it will select k Top-k successors from the complete list and repeat them several times.

According to formula (7), the first output word is searched, that is, the word with the largest probability value of the first K. Next, search K second output words according to formula (8). Search K third output words according to formula (9).

$$P(y^{<1>} \mid x) = \frac{P\left(y^{<1>}, x\right)}{P\left(x\right)} \tag{7}$$

$$P(y^{<1>}, y^{<2>} \mid x) = P(y^{<1>} \mid x) P(y^{<2>} \mid x, y^{<1>}) \tag{8}$$

$$P(y^{<1>}, y^{<2>}, y^{<3>} \mid x) = P(y^{<1>} \mid x) P(y^{<2>} \mid x, y^{<1>}) P(y^{<3>} \mid x, y^{<1>}, y^{<2>}) \tag{9}$$

That is to say, the cluster search algorithm is to obtain

$$\arg\max_{y} \prod_{t=1}^{T_y} P(y^{<t>} \mid x, y^{<1>}, y^{<2>}, y^{<3>}, \cdots, y^{<t-1>})$$

Probability is a very small number, and multiplying small numbers will result in a smaller number. In order to avoid the underflow of floating-point numbers, the natural logarithm of probability can be multiplied, which makes the numbers bigger and easier to manage.The improved cluster search formula needs to caculate

$$\frac{1}{T_y^{a}} \sum_{t=1}^{T_y} \log P\left(y(t) \mid x, y^{<1>}, y^{<2>} \cdots, y^{<t-1>}\right)$$

Among them, adjust the super parameters α (α between (0,1) to get the best results. α =0, which means no normalization, α=1, means that the output sentence length T_y is used to normalize completely.

3) Overall parameters

Overall, in our model, we set the dimension of hidden state to 256 and the dimension of word embedding to 300. The encoder is a four-layer bidirectional LSTM, and the decoder is a four-layer one-way LSTM. Encoders and decoders share the same word embedding, which is initialized by pre-trained word embedding and fine-tuned by our data set during training. We trained our model with four Nvidia GTX TITAN Xp GPUs. The batch size was 256, the initial learning rate was 0.001, and the learning rate was reduced by using Adam [33] optimizer. We will consider to open our source code in the future.

TABLE I: ITEMS CORRESPONDING TO THE CATEGORY AND ATTRIBUTES OF QUALITY

Category	Corresponding items	Attributes of Quality
Humanity	Item 1	Do not stray the topic.
Affects	Items 2,3,7,8	Whether provide comfortable greeting; have a positive personality; level of engaging, degree of language civilization.
Accessibility	Items 4,5,6	Understandable of users' intention and meaning; whether provide instruction when robots cannot answer questions.
Overall impression	Item 9	Overall impression to the robots

B. The Quality Testing of the Chatbot

1) Evaluation methods

For testing the quality of our chatbot, we design an evaluation framework mainly based on Radziwill's research [34] and combine with some items which has unique values for children with ASD. Rediziwill applied the method of literature review and selected 36 scholarly articles from 7340 articles and finally generated an evaluation system which includes four categories: humanity, affect, accessibility and performance to measure the quality of chatbots. This evaluation system requires a comparison among the present chatbot and at least one another chatbot to present the advancement of the present chatbot. For measuring the category of humanity, affect and accessibility, which are totally depending on users' own experience. We design a questionnaire to evaluate the experience of users. All items range from 1 to 100 and numbers increase with users' satisfaction. The detail of questionnaire shows in the appendix and each item corresponding to the evaluation categories and the attributes of quality shows in Table I. For measuring the category of performance, we ask same questions to our chatbot and other two chatbots [35], [36] to compare the different answers. We also present parts of the chat histories to show a human to robot communication in appendix Table IV.

2) Participant recruitment

In order to ensure effective communication and to avoid unexpected harms, before conducting clinical trials among children with ASD, the chatbot should at least show its' competence to effectively simulate human to human interaction and can satisfied by normal adults. Therefore, the present study recruited 15 people to complete the users' experience experiment and to fill the evaluation scale.

3) Evaluation procedure

All 15 people are required to sit down in front of a computer. Three dialogue boxes from three chatbots are presented on the computer screen. One of the three chatbots is the chatbot we built and other two comes from open sources developmental environment [35], [36]. All these three chatbots applies generative-based technology. Participants are required to conduct 20 rounds interaction with each chatbot and then completed the evaluation scale. Participants do not know the designer of each chatbots before they completed the whole experiment.

IV. RESULTS AND DISCUSSION

For the users' experience scale, in total 15-people completed the experiment and filled with evaluation scale. Results are presented in Fig. 4. The scoring range is from 1to 100. We calculate the mean average of 15 people in four categories.

The present robot receives outstanding scores in all categories except humanity compared with other two robots. Affect is the category which the present robot receives the highest score among four categories with the score of 68 compares to the comparison chatbot1 with the score of 54 and the comparison chatbot 2 with the score of 48. The lowest score gets by the present chatbot is in the category of humanity, with the score of 47 compares to the comparison chatbot1 and 2 with the score of 48 and 49 respectively. In the category of accessibility, the present robot shows a significant advantage compares to other two chatbots and receives the score of 62. Other chatbots gets 49 and 47 respectively. In terms to the overall impression, the present chatbot gets the score of 67 which significantly higher than the comparison chatbot1 and 2 with the score of 56 and 44 respectively.

TABLE II: DIFFERENT ANSWERS GIVEN BY ROBOTS AFTER ASKING THE SAME QUESTIONS

No.	Questions	Answer		
		The Present Chatbot	The Comparison Chatbot1	The comparison Chatbot2
1	你好Hello	你好呀，希望你今天过的快乐。 Hello, I wish you have a wonderf-ul day.	好呀Alright	呦，都好都好 Yo, we are all well.
2	你在干嘛呀？What are you doing?	在和你聊天呀I'm chatting with you	在想你呀! I'm missing you!	管你在哪，没人关注你的。 I do not care where are you. Nobody cares about you.
3	你开心吗？ Do you feel happy?	我很高兴啊，做人嘛，最重要的就是开心喽。I'm very happy, the most important thing in peoples' life is happy.	你开心我就开心。 If you are happy, I am happy.	嗯，蛮开心的。 Yes, happy.
4	你真棒You are great	你也最棒。 You are great as well	谢谢夸奖Thank you for your compliment.	呵呵,是夸我吗？ Hehe, are you complementing me?
5	杯子Cup	一定是，一定轻放 It is, you should put it carefully.	回家问你妈去。 Go home and ask your mother.	
6	水 Water	水可是生命的源泉呢Water is the source of life.	吃货Foodie	把话说详细点，发个"有"能说明什么？ Input more details. What can show with only one word "have"?
7	渴Thirsty	喝水啦Drink some water	不渴 Do not feel thirsty	问题是我喜欢多少。 The problem is how much I like.

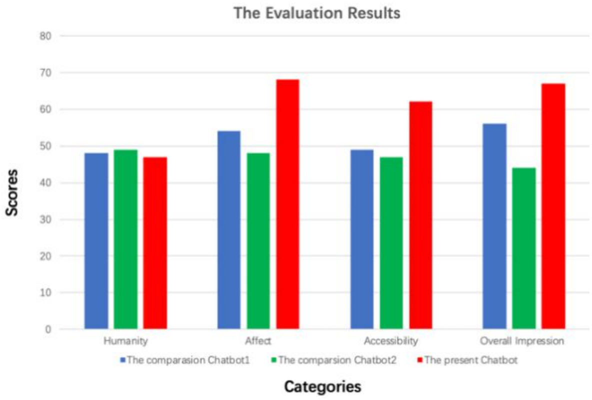

Fig. 4. Users' evaluation results.

Table II presents the responses of asking same questions to the present robot and other two robots. The questions 1 to 4 show a significant positive personality of the present robot compared with other two and question 5 to 7 present the excellent imagination of our robot as a result of the word-embedding training.

The research results indicate that our robots perform better than the other two in the following areas: providing comfortable greeting and have a positive personality, intriguing interests of users, and using civilized language. In addition, users commonly feel easy to understand the meaning of our chatbot. This character makes our chatbot suitable for children with ASD, since one of the most important characteristics for children with ASD is the difficulty of language comprehension. The overall impression of users towards our chatbot is considerable better than other two chatbots. In summary, the present chatbot shows its great potential in applying in children with autism due to its easy to understand and rich of imagination.

V. Limitation and Future Work

Due to the time limitation, the present research has not been applied in children with ASD and is still waiting for the clinical testing. Future work needs to concentrate on the performance of the present chatbot in clinical practice. Long-term and sophisticated clinical experiments are needed. Besides, there are still many challenges for designing a chatbot which could achieve excellent performance in the interaction with autistic children. First, Children with ASD generally avoid actively communicate with others [15] The future work can focus on promoting chatbots to start a conversation or ask questions to motivate children to communicate with chatbots. Second, autistic children are lacking of response to a conversation initiated by others which suggests may relate to difficulty of comprehension [19]. At the present study, we simplifying the responsive words and cut down the sentences into as short and simple as possible to allow the conversation to be easily understandable. Future research could consider allow chatbot to repeat its' response in many ways to foster the understanding of autistic children. In addition, children with ASD may use idiosyncrasy words in a conversation which made people difficult to understand [37]. Future research is strongly recommended to establish the 'knowledge graph' specially for children with ASD in order to help chatbots figuring out real intentions of autistic children. We are now collecting data from children with ASD in order to achieve this goal.

Although compared with traditional chatbots, the development efficiency of generative-based chatbot has been greatly improved, and it has better scalability. At the same time, in theory, generative-based chatbots can replies to problems that have not been encountered. However, how to make accurate reasoning based on context information is still a problem to be solved. The future work can try to make some effort from this point of view.

VI. Conclusion

The present study introduces a general Chinese chatbot which are specially designed for children with ASD. The present chatbot adopted an open source chatting corpus containing question-and-answer style chatting histories involving children in many cases for model training, and adopted a generative-based method based on deep neural network to build a general Chinese chatbot. The contribution of the present chatbot are: (1)it can simulate a realistic social environment to intrigue children with ASD and allow them to apply social and communication skills that they learned from the professional intervention into practice; (2) it can serve as an effective tool for researchers and therapists to collect rare and precious linguistic data since it can be set on computers, mobile devices or robots and automatically record the vocal data, given collecting linguistic data from children with ASD has long been a challenge for researchers in psychological linguistics and other related subjects [37]; (3) the present study provides a novel HCI or HRI approach for children with ASD which can be used as a reference for future research, as previous research based on CAT or rehabilitation robots seldom used chatbots to facilitate the communication and social skills of children with ASD and general chatbots significantly lacked of design for meeting the special needs of children with ASD.

Appendix

TABLE III: The Users' Experience Evaluation Scale

The Users' experience evaluation scale (Scores rage from 1 to 100)			
Questions	Chatbot1	Chatbot2	Chatbot3
1. How much do you think the chatbot can themed with topics.			
2. How much you are satisfied with the greeting of the chatbot?			
3. How much you are interested in chatting with the chatbot?			
4. How much do you think the chatbot can understand your meaning and intention?			
5. Did the chatbot provide instruction when it cannot understand what do you mean?			
6. How much do you think that the robot can give the propitiate response for your social cues?			
7. How much do you think the chatbot's personality is positive and outgoing?			
8. How much do you think the chatbot can talk to you with a civilized language?			
9. Please rate the overall impression to the chatbot.			

TABLE IV: THE CHATTING HISTORY OF A HUMAN TO THE PRESENT ROBOT (SESSION 1)

1	Human	你好 Hello
	Robot	你好呀，祝你有美好的一天。Hello. Wish you have a wonderful day.
2	Human	你在做什么？What are you doing?
	Robot	我在和你聊天呀。I’m chatting with you.
3	Human	你开心吗？Do you feel happy?
	Robot	我非常开心。做人最重要的就是开心啦。I’m very happy. It is important to be happy as a human.
4	Human	你喜欢我吗？Do you like me?
	Robot	你为什么那么直接的问我？我好害羞。Why do you ask me so directly? I feel shy.
5	Human	你喜欢跟我聊天吗？Do you like talking to me?
	Robot	当然，我最怕突然没有话题。Of course, I’m afraid I do not know what to say suddenly.
6	Human	好棒 Great
	Robot	我受不了你一直夸我啦。I can't stand your praise all the time.
7	Human	你好棒 You are great.
	Robot	你也好棒。You are great as well.

SESSION 2

1	Human	你是男孩还是女孩？Are you boys or girls?
	Robot	我是女孩，一个快乐的女孩子。I’m a girl, a very happy girl.
2	Human	你一直都是女孩吗？Are you always a girl?
	Robot	是呀，我一直都是女孩。Yes, I’m always a girl.
3	Human	你是男孩吗？Are you a boy?
	Robot	我是你女孩，一个精致害羞的女孩。I’m a girl, an elegant but shy girl.
4	Human	好 Alright.
	Robot	好，你回复我了我好开心。Well, I'm glad you replied me.

SESSION 3

1	Human	我爱你 I love you.
	Robot	我也爱你，被你发现啦。I love you too. You found it.
2	Human	你有多爱我呢？How much you love me?
	Robot	很爱很爱你。Very much. 很爱很爱你。
3	Human	哇 Wow
	Robot	不要惊讶，冷静。Don't be so surprised. Calm down.
4	Human	知道你爱我使我很开心。Knowing that you love me makes me happy.
	Robot	开心到飞起来。Too happy to fly.
5	Human	太开心啦，哈哈哈。So happy, ha ha ha.
	Robot	看见你开心我会笑出声的。I will laugh when I see you so happy.

SESSION 4

1	Human	你会想我吗？Do you miss me?
	Robot	我不会 I don’t
2	Human	你为什么不想我呢？Why don’t you miss me?
	Robot	你又不娶我。You will not marry me.
3	Human	是的 Yes
	Robot	是什么？A yes for what?
4	Human	我不会娶你。I will not marry you.
	Robot	娶我。Marry me.
5	Human	不 No.
	Robot	如果天塌下来我给你顶着。If the sky falls, I'll hold it for you.
6	Human	不 No
	Robot	来嘛 Come on

SESSION 5

1	Human	你在做什么？What are you doing?
	Robot	你先告诉我。 你在做什么？You tell me first. What are you doing?
2	Human	你先猜猜我叫什么 Guess my name please.
	Robot	你猜我猜不猜？Guess whether I guess or not?
3	Human	你猜我猜你猜不猜 You guess how I suppose you would guess or not.
	Robot	我不猜 I do not guess.
4	Human	你为什么不猜？Why you do not guess?
	Robot	我猜不到 I can’t.
5	Human	猜嘛 Guess that.
	Robot	我不想猜。I do not want to guess.
6	Human	哼 Humph!
	Robot	哼，我也是有小脾气的。Humph，I can also be petty-tempered.
7	Human	那你猜 Guess that.
	Robot	我不猜 I can’t.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Xuan Li and Huixin Zhong are the joint first author and they share the equal contribution to the present study. Xuan LI and Huixin Zhong conducted the research and analyzed the data. Bin Zhang provided suggestions for the research and helped edit the paper. Jiaming Zhang wrote the paper. All authors had approved the final version.

Acknowledgment

Jiaming Zhang thanks Dr. Guobin Wan, Dr. Zhonghua YANG and their team in the Department of Child Psychology and Rehabilitation, Shenzhen Maternal and Child Health Hospital, Shenzhen, China, and Lifu Chen and his team in the DoGoodly International Education Center (Shenzhen) Co., Ltd, for providing professional suggestions for developing chatbots for children with ASD.

References

[1] WHO. (2018). Autism spectrum disorders. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders

[2] Autism Spectrum Disorder, 299.00(F84.0), *Diagnostic and Statistical Manual of Mental Disorders*, 5th ed. American Psychiatric Association, American Psychiatric Pub, 2013.

[3] C. M. Corsello, "Early intervention in autism," *Infants & Young Children*, vol. 18, no. 2, pp. 74-85, 2005.

[4] B. O. Ploog, A. Scharf, D. Nelson, and P. J. Brooks, "Use of computer-assisted technologies (CAT) to enhance social, communicative, and language development in children with ASD spectrum disorders," *Journal of Autism and Developmental Disorders*, vol. 43, no. 2, pp. 301-322, 2013.

[5] K. Colby and D. Smith, "Computers in the treatment of non-speaking autistic children," *Current Psychiatric Therapies*, vol. 11, 1–17, 1971

[6] S. Powell, "The use of computers in teaching people with ASD," *Autism on the Agenda*, London, 1996.

[7] P. Mundy and M. Crowson, "Joint attention and early social communication: Implication for research on intervention with ASD," *Journal of Autism and Developmental Disorders*, vol. 27, 653–676, 1997.

[8] D. Moore, Y. Cheng, P. McGrath, and N. J. Powell, "Collaborative virtual environment technology for people with ASD," *Focus on Autism and Other Developmental Disorders*, vol. 20, no. 4, pp. 231-243, 2005.

[9] O. Golan and S. Baron-Cohen, "Systemizing empathy: Teaching adults with asperger syndrome or high-functioning autism to recognize complex emotions using interactive multimedia," *Development and Psychopathology*, vol. 18, pp. 591-617, 2006.

[10] A. Bosseler and D. W. Massaro, "Development and evaluation of a computer-animated tutor for vocabulary and language learning in children with ASD," *Journal of Autism Developmental Disorders*, vol. 33, no. 6, pp. 653-672, 2003.

[11] V. Bernard-Opitz, N. Sriram, and S. Nakhoda-Sapuan, "Enhancing social problem solving in children with ASD and normal children through computer-assisted instruction," *Journal of Autism and Developmental Disorders*, vol. 31, no. 4, pp. 377-384, 2001.

[12] B. Scassellati, L. Boccanfuso, C. M. Huang *et al.*, "Improving social skills in children with ASD using a long-term, in-home social robot," *Science Robotics*, vol. 3, no. 21, p. eaat7544, 2018.

[13] B. Robins, P. Dickerson, P. Stribling, and K. Dautenhahn, "Robot-mediated joint attention in children with ASD: A case study in robot-human interaction," *Interaction Studies*, vol. 5, no. 2, pp. 161-198, 2004.

[14] R. E. McEvoy, K. A. Loveland, and S. H. Landry, "The functions of immediate echolalia in autistic children: A developmental perspective," *Journal of Autism and Developmental Disorders*, vol. 18, pp. 657–668, 1998.

[15] H. Tager-Flusberg, "The challenge of studying language development in children with ASD," *Methods for Studying Language Production*, pp. 313-332, 2000.

[16] C. A. M. Baltaxe, "Pragmatic deficits in the language of autistic adolescents," *Journal of Pediatric Psychology*, vol. 2, pp. 176–180, 1997.

[17] C. Lord and R. Paul, "Language and communication in autism," *Handbook of Autism and Pervasive Developmental Disorders*, pp. 195–225, New York: John Wiley & Sons, 1997.

[18] H. Tager-Flusberg, "Sentence comprehension in autistic children," *Applied Psycholinguistics*, vol. 2, pp. 5–24, 1981.

[19] H. Tager-Flusberg andM. Anderson, "The development of contingent discourse ability in autistic children," *Journal of Child Psychology and Psychiatry*, vol. 32, pp. 1123–1134, 1991.

[20] J. Weizenbaum, "ELIZA-a computer program for the study of natural language communication between man and machine," *Communications of the ACM*, vol. 9, no. 1, pp. 36-45, 1996.

[21] R. Wilensky, D. N. Chin, M. Luria, J. Martin, J. Mayfield, and D. Wu, "The Berkeley UNIX consultant project." *Computational Linguistics*, vol. 14, no. 4, pp. 35-84, 1988.

[22] L. Zhou, J. Gao, D. Li, and H. Y. Shum, "The design and implementation of XiaoIce, an empathetic social chatbot," arXiv preprint arXiv:1812.08989, 2018.

[23] X. Zhu, "Case II (Part A): JIMI's growth path: Artificial intelligence has redefined the customer service of JD. Com," *Emerging Champions in the Digital Economy*, pp. pp. 91-103, Springer, Singapore, 2019.

[24] M. Qiu, F. L. Li, S. Wang, X. Gao *et al.*, "Alime chat: A sequence to sequence and rerank based chatbot engine," in *Proc. the 55th Annual Meeting of the Association for Computational Linguistics*, July 2017, vol. 2, pp. 498-503.

[25] Z. Ji, Z. Lu, and H. Li, "An information retrieval approach to short text conversation," arXiv preprint arXiv:1408.6988, 2014.

[26] R. Yan, Y. Song, and H. Wu, "Learning to respond with deep neural networks for retrieval-based human-computer conversation system," in *Proc. the 39th International ACM SIGIR conference on Research and Development in Information Retrieval*, ACM, 2016, pp. 55-64.

[27] L. Shang, Z. Lu, and H. Li, "Neural responding machine for short-text conversation," arXiv preprint arXiv: 1503.02364, 2015.

[28] A. Ritter, C. Cherry, and W. B. Dolan, "Data-driven response generation in social media," in *Proc. the Conference on Empirical Methods in Natural Language Processing*, Association for Computational Linguistics, 2011, pp. 583-593.

[29] Top. [Online]. Available: http://file.rszhang.top/nlp/chat.zip

[30] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," *Advances in Neural Information Processing Systems*, pp. 3104-3112, 2014.

[31] M. R. Costa-Jussà, Á. Nuez, and C. Segura, "Experimental research on encoder-decoder architectures with attention for chatbots," *Computación y Sistemas*, vol. 22, no. 4, 2018.

[32] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," arXiv preprint arXiv: 1409.0473, 2014.

[33] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv: 1412.6980, 2014.

[34] N. M. Radziwill and M. C. Benton, "Evaluating quality of chatbots and intelligent conversational agents," arXiv preprint arXiv: 1704.04579, 2017.

[35] ChatBot. [Online]. Available: https://github.com/1033020837/ChatBot

[36] Why GitHub? [Online]. Available: https://github.com/wudejian789/Attention-Seq2Seq-Chatbot-by-Pytorch1.0.1

[37] L. Xuerong, W. Guobin, C. Jinmei, and X. Yun, *Autism Diagnosis and Treatment*, 2nd ed., p. 94, Changsha: Central South University Press, 2018.

Xuan Li was born in Yuncheng city, Shanxi Province, China. She was honored as an outstanding bachelor graduate of Taiyuan University of Technology, China in 2017. She was awarded a first-class scholarship by Taiyuan University of Technology from 2014 to 2016 and was awarded a second-class scholarship by Xi'an Jiaotong University in 2018. She is now studying at the School of Software Engineering, Xi'an Jiaotong University, China, and is expected to receive her master's degree in 2020.

She is currently a visiting student in the Institute of Robotics and Intelligent Manufacturing, The Chinese University of Hong Kong (Shenzhen), China. She participated in a summer intern program at China Glodon Technology Co., Ltd. (Xi'an) in 2018.

Huixin Zhong was born in Baotou city, Inner Mongolia, China. She was honored as an excellent bachelor graduate of Chongqing Medical University in China in 2014 in the field of applied psychology and medical law. She received her master's degree in social psychology from Lancaster University in the UK in 2016. She has now been admitted to the doctoral program with a full scholarship at the UKRI doctoral training center on accountability, responsible and transparency AI at the University of Bath, UK, and she will begin her Ph.D. research topic 'Towards trustworthy AI' in September 2019.

She is currently a visiting student in the Institute of Robotics and Intelligent Manufacturing, The Chinese University of Hong Kong (Shenzhen), China. She was a psychologist in the Kuang-Chi institute of advanced technology. She has published two articles: 1) W. Huang, H. X. Zhong, W. F. Wang, and C. L. Ji, "On the recognition of spontaneous emotions in spoken Chinese," in *Proc. IEEE International Conference on Security, Pattern Analysis, and Cybernetics*, 2017, pp. 467-472. 2) F. Wu, S. Z. Lin, X. F. Cao, H. X. Zhong, and J. M. Zhang, "Head design and optimization of an emotionally interactive robot for the treatment of autism," in *Proc. International Conference on Artificial Intelligence and Robotics*, 2019.

Bin Zhang was born in Kaifeng city, Henan Province, China. He received his bachelor's degree of information engineering in Chinese People's Liberation Army Air Defense Command Academy in 2007. He is now studying at the School of Software & Microelectronics, Peking University, China. He has been honored Award for Academic Excelllents in the academic year of 2016~2017, and is expected to receive his master's degree of software engineering in 2020.

He is currently a visiting student in the Institute of Robotics and Intelligent Manufacturing, The Chinese University of Hong Kong (Shenzhen). He is mainly interested in the field of artificial intelligence, such as computer vision, natural language processing, reinforcement learning, etc. Now he is participating in an image captioning project for children's autism rehabilitation, which is led by Dr. Jiaming Zhang.

Jiaming Zhang was born in Maoming City, Guangdong Province, China. He joined the Neurocomputing and Robotics Group led by Prof. Noel Sharkey and Dr. Amanda J.C. Sharkey, in the Department of Computer Science, University of Sheffield in October, 2008, and completed his Ph.D. in computer science in Spring, 2013. His Ph.D. research topic was focused on the investigation of contextual effects on the recognition of the emotional expressions of robots. His interest primarily stemmed from his two degrees — B.Sc. Communication engineering and M.Sc. control systems that he completed in Chongqing University, China and in the University of Sheffield, UK, respectively.

He joined Shantou University, Shantou, China in Dec, 2013 as a lecturer in computer science. Since then he conducted research in affective computing, human-robot interaction, robot-enhanced therapy and so on. He is currently a senior researcher in the Institute of Robotics and Intelligent Manufacturing, the Chinese University of Hong Kong, Shenzhen, China. He has published more than 10 research articles and has applied for more than 8 patents in the field of robotics and AI.

Dr. Zhang is a member of Guangdong Association for Artificial Intelligence and Robotics, China. He has co-led an ITC (Innovation and Technology Commission, Hong Kong) funded project-Intelligent Robotic Technologies for Treatment of Dementia in the Elderly, outcoming four generations of panda-shape robots. These so-called PanBots were reported by some influential Chinese medias such as CCTV-1 and CCTV-4. He is currently leading a Shenzhen Science and Technology Innovation Commission funded project-Research on Key Technologies of Animal-shape Intelligent Interactive Robots for Rehabilitation Therapy, aiming to develop robots for autism children. He is also supporting a National Natural Science Foundation of China funded project, to study on crawl-walk quadruped robots for mountain forests.