A Survey on Learning-Based Approaches for Modeling and Classification of Human–Machine Dialog Systems

Fuwei Cui, Qian Cui, and Yongduan Song, *Fellow, IEEE*

***Abstract*—With the rapid development from traditional machine learning (ML) to deep learning (DL) and reinforcement learning (RL), dialog system equipped with learning mechanism has become the most effective solution to address human–machine interaction problems. The purpose of this article is to provide a comprehensive survey on learning-based human–machine dialog systems with a focus on the various dialog models. More specifically, we first introduce the fundamental process of establishing a dialog model. Second, we examine the features and classifications of the system dialog model, expound some representative models, and also compare the advantages and disadvantages of different dialog models. Third, we comb the commonly used database and evaluation metrics of the dialog model. Furthermore, the evaluation metrics of these dialog models are analyzed in detail. Finally, we briefly analyze the existing issues and point out the potential future direction on the human–machine dialog systems.**

***Index Terms*—Artificial intelligence (AI), deep learning (DL), dialog model, machine learning (ML), reinforcement learning (RL), sequence to sequence (Seq2Seq) model.**

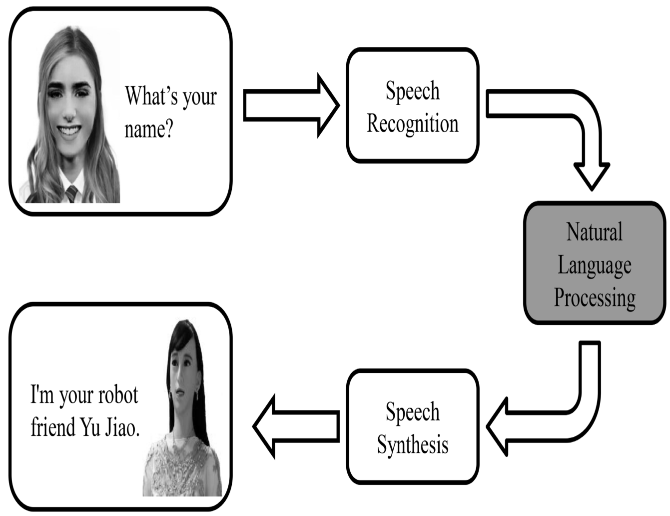

Fig. 1. Composition of dialog system.

I. Introduction

One of the original research purposes of the human–machine dialog system is to pass the Turing test. And human beings have been studying dialog systems for half a century. Early dialog systems were based on artificial rules, such as Eliza (1966) [1], Parry (1975) [2] which passed the Turing test, and Alice (2009) [3] which won the Loebprize three times recently. Although the rule-based dialog system has achieved good results, the establishment of rules is laborious, and its transferability is poor. Most importantly, many rules eventually lead to the software system either too costly or seldom practical.

Manuscript received July 27, 2019; revised January 18, 2020 and March 5, 2020; accepted March 28, 2020. This work was supported by the National Natural Science Foundation of China under Grant 61803053, Grant 61833013, Grant 61860206008, and Grant 61773081. *(Corresponding author: Yongduan Song.)*

Fuwei Cui is with the Institute of Advanced Control System, College of Electronic Information Engineering, Beijing Jiaotong University, Beijing 100044, China (e-mail: 17111051@bjtu.edu.cn).

Qian Cui and Yongduan Song are with the State Key Laboratory of Power Transmission Equipment and System Security and New Technology, and Chongqing Key Laboratory of Intelligent Unmanned Systems, School of Automation, Chongqing University, Chongqing 400044, China (e-mail: qiancui@cqu.edu.cn; ydsong@cqu.edu.cn).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNNLS.2020.2985588

With the development of speech recognition [4]–[6], speech synthesis [7], [8], natural language processing [9], [10], and information retrieval (IR) [11], [12], especially deep learning (DL) [13]–[15] and reinforcement learning (RL) [54], some data-driven-based models that use DL or RL have been proposed, such as IR models, generation models, RL models, and hybrid models. And so far, many human–machine dialog products have emerged, such as Cortana and Microsoft Xiaobing in 2014, Baidu Duer and Ali Xiaomi in 2015, Apple Siri and Google Assistant in 2016, Tencent tinkling in 2017 and so on. Although the existing dialog system can communicate with human beings in some occasions, the system itself is not intelligent enough, calling for the implementation of much strong artificial intelligence (AI). Therefore, it is crucial to carry out extensive research on the dialog system, which makes it necessary to have a general grasp of the current research situation of the dialog system.

The dialog system is generally composed of speech recognition, natural language processing, speech synthesis (early template-based dialog model without speech recognition and speech synthesis module), as conceptually shown in Fig. 1, where the function of the speech recognition module is to convert human speech signal into text signal for the natural language understanding module. Next, the natural language

TABLE I
TERMINOLOGIES AND ABBREVIATIONS

Abbr	Full Name
AI	artificial intelligence
AIML	artificial intelligence markup language
Alice	artificial linguistic internet computer entity
AWI	attention with intention
BLEU	bilingual evaluation understudy
BS	beam search
CNN	convolutional neural network
DL	deep learning
EA	embedding average
ECM	emotional chatting memory
GAN	generative adversarial networks
GM	greedy matching
GRU	gated recurrent unit
HRED	hierarchical recurrent encoder-decoder
IR	information retrieval
KB	knowledge based
LSTM	long short-term memory
MAP	mean average precision
MemNN	memory networks
ML	machine learning
MP	match pyramid
MRR	mean reciprocal rank
PPL	perplexity
RL	reinforcement learning
RNN	recurrent neural networks
ROUGE	recall-oriented understudy for gisting evaluation
Seq2Seq	sequence to sequence
VAEs	variational autoencoders
VE	vector extrema
VHCR	variational hierarchical conversation RNNs
VHRED	latent variable hierarchical recurrent encoder-decoder

understanding module first inputs the transformed text signal into the dialog model, then recognizes the human intention, and finally generates the corresponding reply. Lastly, the function of speech synthesis is to convert the text signals returned by the natural language understanding module into speech signals, and then output them.

The core of the dialog system is to build a dialog model, and building dialog model is the main task of this module. In this article, we review the current research status of dialog models, including the construction, classification, database, evaluation metrics, analysis of evaluation metrics, challenges and possible future research directions, and attempt to sketch a more comprehensive and clear outline for the study of the dialog model, in order to provide a useful reference for the related research in this field. For easy reference, terminologies and abbreviations that appear more than once in the article are listed in Table I.

II. Process of Establishing Dialog Model

There are two typical methods for building a dialog model, i.e., nondata-driven and data-driven. The general process for building a dialog model is illustrated in Fig. 2.

To build a dialog model with the nondata-driven method, we should first be familiar with the business scenarios of the application of the dialog model, then extract the corresponding rules through business analysis, and finally integrate all the rules to build a template (we call a template also a conversation model).

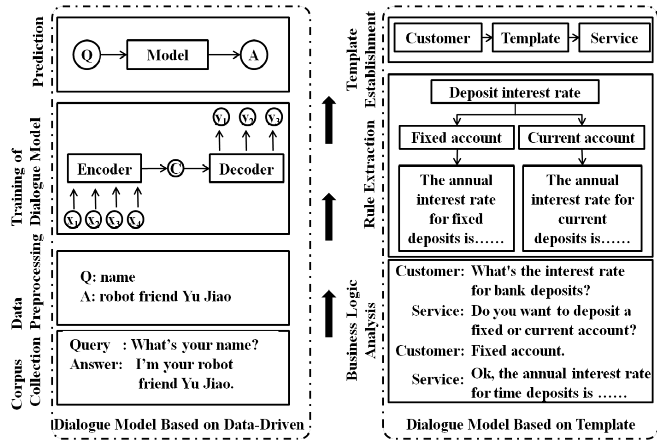

Fig. 2. Flow chart of building dialog model.

To build a data-driven dialog model, first we need to prepare a corpus, which is the basis of building the dialog model. Note that the quality of the corpus directly affects the training effect of the dialog model. There are two ways to develop a corpus. One is to use the open corpus on the Internet directly. In Section IV-A, some corpus commonly used for dialog model training are listed. The other is to crawl from the Internet. Second, we need to preprocess the data, the main operations include removing stop words, word segment (not involved in English though) and so on. Next, the processed data are put into the dialog model for training, and different dialog models are selected according to different business scenarios. Finally, the trained dialog model is used to predict, receive the user's input and generate the response.

With the comprehensive introduction of the process of building the dialog model in this section, we then are ready to discuss the core of the dialog model, analyze and compare the basic principles, research status, advantages and disadvantages, existing problems and future possible research directions of the existing dialog models in the sequel.

III. Classification of Dialog Models

As mentioned above, the dialog system has become an increasing active research area in natural language processing. According to the different implementation technologies, we can classify most existing models into the one based on an artificial template, and the one based on IR, generation, RL or combination of them. And the classification system of the dialog model is shown in Fig. 3, which is discussed and analyzed in what follows.

A. Dialog Model Based on Nondata-Driven

In the early works for the development of dialog system, computer and storage technology have its limit, and there is a lack of corresponding data sets. The corresponding dialog system in this period can only be constructed by the nondata-driven method, which generally uses traditional

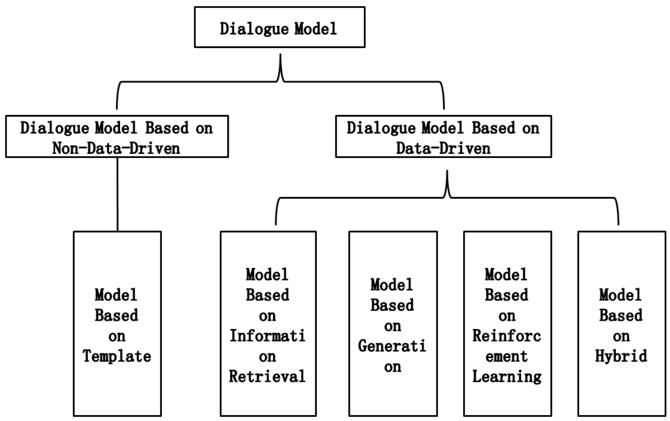

Fig. 3. Classification of dialog model.

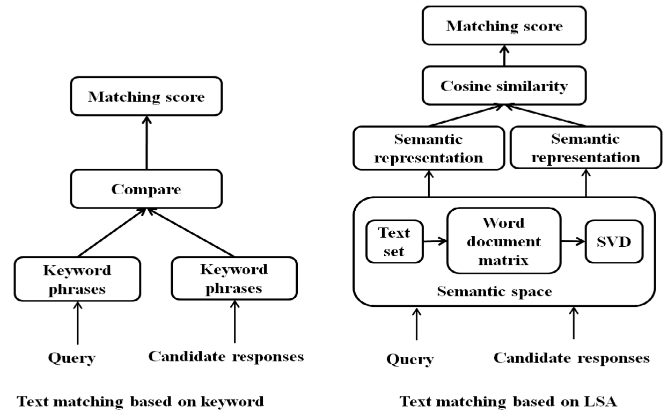

Fig. 4. Traditional ML algorithms used in dialog model (SVD is singular value decomposition).

machine learning (ML) algorithms, such as some pattern matching algorithms. Moreover, the commonly used single pattern matching algorithms are Boyer–Moore (BM) [81] algorithm and KMP (discovered by Knuth, Morris, and Pratt) [82], and Multipattern matching algorithms as well as Wu–Manber (WM) [83] algorithm and Aho–Corasick (AC) [84] algorithm.

The dialog model based on the nondata-driven method is mainly a template-based dialog model, which is realized by setting rules manually. When the user's question matches the preset rules, the corresponding response will be triggered. For example, when a user greets a dialog robot, the user may ask "hola, hello, hi, hey." The dialog robot will randomly select one of several preset greetings to respond to. And the main workflow of the template-based dialog model includes receiving user input, question normalization processing, question query reasoning, and template processing. Among them, receiving user input is mainly to get text signal; question normalization processing is mainly to replace strings that need to be replaced in question sentences, such as replacing "I'm" with "I am"; question query reasoning is to match the normalized question sentences with the rules in the rule base to get the best matching result; template processing is to process the special tags existing in the matching results to get the final response results, and then return the results to the user.

Early dialog models are mainly template-based dialog models, such as Eliza mentioned earlier, whose template was written by the script of the dialog, and the script is composed of keywords and corresponding transformation rules. When there is user input, first checking the keywords in the input statement and selecting the keywords with the highest ranking, then finding the corresponding conversion rules, and generating the response statement through the rules. In fact, Eliza is used in the field of psychological counseling, acting like a counselor. Furthermore, it demonstrates the possibility of template-based communication between human and machine.

Artificial linguistic internet computer entity (Alice), the recent template-based dialog system, is written in an AI markup language (AIML) [16], which is an XML language that can create rules for robot dialog. In reality, the language adopts the "Stimulus-Response" theory and is developed in JAVA language. More specifically, there are some basic processes. After receiving the user's input, first extracting keywords from the input statement, replacing them and removing noise. Then, matching keywords through rules, locating the position of questions in the template. Finally, getting the response through the template. Although Alice can understand the context and expand knowledge easily, it has limitations in dealing with synonyms.

In a word, the template-based model has the advantage of high controllability and reliability. But it needs to establish rules manually, which is time-consuming and costly, and has poor adaptability to changes in user input wording, and it also needs designers to know the real business scenarios very well. When facing new business scenarios, it needs to be redesigned and the portability is poor. Particularly, with the increasing amount of data, the whole knowledge base may conflict and lead to system collapse.

B. Dialog Model Based on Data-Driven

The data-driven dialog model needs a dialog data set, which is also called corpus, the data in which is processed to train the model. And the traditional ML, DL or RL algorithms are often used to train model, which can use some algorithms or neural networks to learn the information from the data set.

The traditional ML algorithms, as shown in Fig. 4, are mainly text matching algorithms, which match the similarity between two texts by extracting keywords or semantic information.

The DL algorithms, as shown in Fig. 5, can transform the initial "low-level" feature representation into "high-level" feature representation by constructing a multilayer artificial neural network to extract and filter the input information layer by layer. It's like the process that human neurons transmit information through neural networks, reflecting the ability of human abstract learning. And the commonly used artificial neural networks are convolutional neural network (CNN), recurrent neural network (RNN) and deep belief network (DBN).

The RL algorithm, as shown in Fig. 6, which is an algorithm that imitates the learning behavior of human or animal, originates from the utility rule of behavioral psychology.

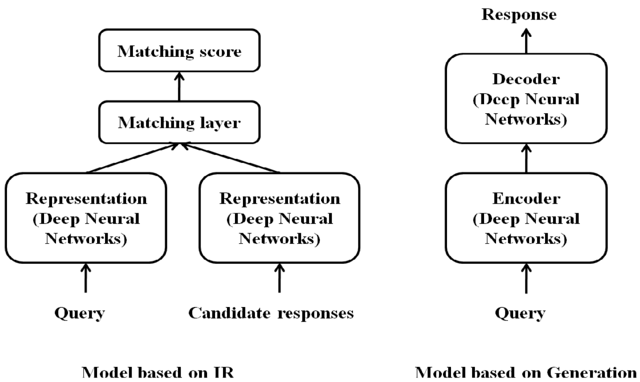

Fig. 5. DL algorithms used in dialog model (deep neural networks contain deep CNN, RNN, DBN, and other deep neural networks).

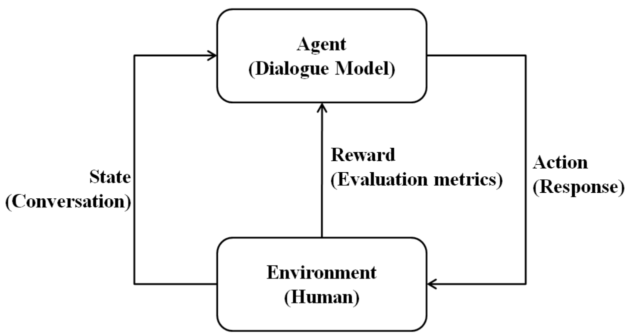

Fig. 6. RL algorithms used in dialog model.

By interacting with the environment through trial and error, it learns how to achieve the best state and action in order to get the greatest reward. The theoretical basis of RL is the Markov decision process (MDP) [46]. And the key elements of RL include action, state, strategy and reword.

1) Model Based on IR: Model based on IR is widely used in industrial production. More specifically, the principle of the model is: first, extracting the keywords or semantic representation of the question, then matching the similarity with the question in the corpus, and finally, selecting the response corresponding to the question with the highest similarity as the final output by sorting algorithm. Therefore, the core problem of IR-based model can be abstracted as a text-matching problem. And the commonly used methods of text matching include: 1) text matching based on keyword, the commonly used keywords contain term frequency-inverse document frequency (TF-IDF) and best matching (BM25); however, there are some limitations in keyword-based matching, which cannot use the semantic information of text; 2) text matching based on shallow semantic, latent semantic analysis (LSA) or latent semantic indexing (LSI), can solve the problem of synonymy at the semantic level, but cannot solve the problem of polysemy and ignore the order of words; and 3) text semantic matching based on DL mainly includes a representation-based matching method and interaction-based matching method. Next, we mainly introduce text semantic matching based on DL. The summary of the IR-based dialog models is shown in Table II and the summary of the use of evaluation metrics is shown in Fig. 7.

The representation-based matching method first generates the representation of the text and then calculates the matching degree. Huang *et al.* [17] propose a deep structured semantic model (DSSM), which maps high-dimensional sparse text features to low-dimensional dense features by using deep neural networks, i.e. replacing the traditional bag-of-words model with word hash, so as to achieve the goal of dimensionality reduction. Although the model can extract the semantic vector of sentence granularity better, the representation based on sentence granularity is slightly rough, and the temporal relationship between words is not considered. Wan *et al.* [18] propose a multiview long short-term memory (MV-LSTM) model, which first extracts the position representation of sentences by bidirectional LSTM. Then, calculating the degree of matching between two sentences by using interaction tensor and stores them in the matching matrix, Next, choosing the former k interactions of the matching matrix by k-max pooling. Finally, calculating the matching score by multilayer perceptron. The model can describe sentence information in fine granularity and extract temporal information of sentences. Furthermore, the similarity function with parameters in formula (1) is used to calculate similarity, which captures more diverse interactions information between two positional sentence representations, compared with cosine and bilinear similarity function. Because the result $s(\bullet)$ is an interaction tensor, which can represent more diverse information than the interaction matrices of the result of cosine or bilinear function. So it is more in line with the characteristics of language diversity

$$s(\mu, \nu) = f\left(\mu^T M^{[1:c]}\nu + W_{\mu\nu}\begin{bmatrix}\mu\\ \nu\end{bmatrix} + b\right) \tag{1}$$

where $s(\bullet)$ is the similarity function, $f(z) = \max(0, z)$, μ and ν are two vectors, M^i $(i \in [1{:}c])$ is one slice of the tensor parameters, $W_{\mu\nu}$ and b are the parameters of the linear part.

The interaction-based matching method calculates matching features directly and then extracts deep matching information on the basis of matching features. Pang *et al.* [19] propose a match pyramid (MP) model, which constructs a matching matrix in three ways and regards the matching matrix as a picture. And CNN is used to convolute and pool the matching matrix to extract features. Specifically, these three ways are indicator function, dot product, and cosine similarity respectively. By considering the relationship between words in sentences in various ways, the MP model can extract different granularity information and improve the matching ability of the model. Wu *et al.* [20] propose a sequential match network (SMN) model. Firstly, constructing a matching matrix from word and sentence granularity and extracting important matching information through convolution and pooling. Then, filtering noise through a gated recurrent unit (GRU) layer. Finally, obtaining a matching score through a hidden layer. Therefore, the model can take full account of dialog relations and important context information, extract features from word and sentence granularity, and achieve good results in the multi-turn dialog. Zhou *et al.* [21] propose a deep attention matching (DAM) model, which is inspired by Transformer [22]. Firstly, self-attention and cross-attention are used to extract matching

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE II
SUMMARY OF THE IR-BASED DIALOG MODELS

Reference	Model Name	Datasets
Huang et al.[17]	DSSM(Deep Structured Semantic Models)	the evaluation dataset
Wan et al.[18]	MV-LSTM(Multi-View Long Short-Term Memory)	collected form Yahoo community question answering system
Pang et al.[19]	MP(Match Pyramid)	MSRP dataset and a large academic dataset
Wu et al.[20]	SMN(Sequential Match Network)	Ubuntu corpus and Douban conversation corpus
Zhou et al.[21]	DAM(Deep Attention Matching)	Ubuntu corpus and Douban conversation corpus
Wang et al.[23]	IRGAN(Information Retrieval with GAN)	LETOR dataset, Movielens (100k) and Netflix, InsuranceQA dataset
Zhang et at.[29]	DUA(Deep Utterance Aggregation)	Ubuntu corpus, Douban conversation corpus, E-commerce dialogue corpus

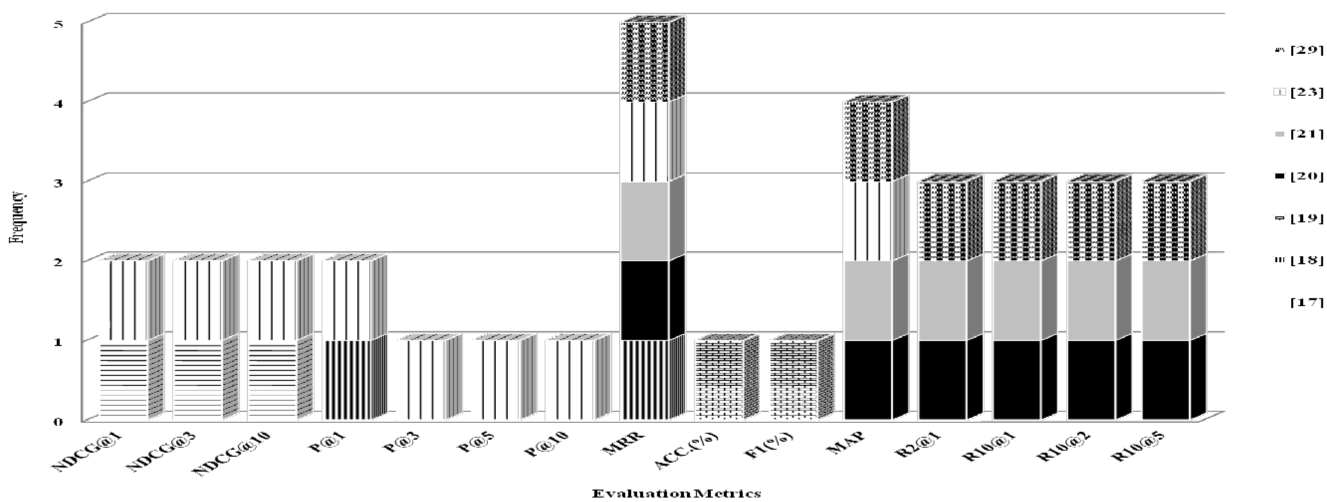

Fig. 7. Summary of the use of evaluation metrics in IR-based dialog models. The abscissa is the evaluation metric. The ordinate is occurrence frequencies of evaluation metrics appearing in articles. The legend on the right is the reference number.

information between context information and response from word level to sentence level. Then, matching information is aggregated into a 3D matching graph. After convolution and pooling, matching information is further extracted. Finally, calculating the matching score by a single-layer perceptron. In particular, this model breaks the structure of RNN and CNN before and achieves the best result in a multiturn dialog.

In a word, the method of text semantic matching based on DL, which is superior to the former two methods in performance, can not only extract semantic information of different granularity but also consider the temporal characteristics of sentences.

A few years ago, a model of applying generative adversarial networks (GAN) to the dialog area based on IR emerged [24]. Wang *et al.* [23] propose an information retrieval with GAN (IRGAN) model, which is inspired by GAN. IRGAN unifies the generative retrieval model and the discriminant retrieval model, overcoming the major shortcomings, and optimizing the two models iteratively by the minimax algorithm. It is worth mentioning that this model provides a new research idea for the development of the retrieval model, which surpasses the strong benchmark model in the application of web search, recommendation system and question answering system.

The main reason for the wide utilization of the model based on IR is that, on the one hand, it does not need to build template manually, and saves time and effort; and on the other hand, the model responding to the language is fluent and logical, with no grammatical errors. However, the model cannot deal with questions that do not exist in the corpus and cannot be applied in open domain dialogs.

2) Model Based on Generation: Model based on the generation normally adopts the structure of Seq2Seq [25], which generally includes encoder and decoder: the encoder mainly encodes the input questions and extracts the semantic information; the decoder uses the extracted semantic information to decode and generates replies. And the encoder and decoder are generally composed of RNN, which are LSTM [26] and GRU [27]. The summary of the generation-based dialog models is shown in Table III and the summary of the use of evaluation metrics is shown in Fig. 8.

Model based on the generation is usually trained by using short text or question–answer corpus, and the model is also the basic Seq2Seq model. Shang *et al.* [28] propose a new model of neural responding machine (NRM) for short text conversation, which is based on the neural machine translation (NMT) [29] model with a classical attention mechanism. At the decoder end of the Seq2Seq model, global strategy, local strategy, and hybrid strategy are proposed, and the improved model with three strategies is compared with the previous IR [30] and statistical machine translation (SMT) [31] on

TABLE III

SUMMARY OF THE GENERATION-BASED DIALOG MODEL

Style	Reference	Model Name	Datasets
Simple dialogue	Shang et al.[28]	NRM(Neural Responding Machine)	Conversations on Sina Weibo
Diversity dialogue	Li et al.[32]	Seq2Seq with penalized beam search	OpenSubtitles (OSDb) dataset and WMT'14 training dataset
	K Vijayakumar et al.[33]	Seq2Seq with diversity beam search	COCO and PASCAL-50S
	Shao et al.[34]	Seq2Seq with random beam search	Combines multiple dataset
meaningful dialogue	Mou et al.[35]	Seq2BF(Sequence to Backward and Forward Sequences)	Chinese dataset crawled from the Baidu Tieba
	Xing et al.[36]	TA-Seq2Seq (Topic Aware Sequence-to-Sequence)	Dataset from Baidu Tieba
Emotional dialogue	Zhou et al.[38]	ECM (Emotional Chatting Memory)	NLPCC emotion classification dataset and STC conversation dataset
	Asghar et al.[39]	Seq2Seq with Emotion	Cornell Movie Dialogues Corpus
Multi-turn dialogue	Serban et al.[41]	Extends the Hierarchical Recurrent Encoder-Decoder	MovieTriples dataset
	Vlad Serban et al.[43]	VHRED (Latent Variable Hierarchical Recurrent Encoder-Decoder)	Twitter Dialogue Corpus
	Yao, et al.[44]	AWI (Attention With Intention)	In-house dialogue dataset.
	Park et al.[48]	VHCR(Variational Hierarchical Conversation RNNs)	Cornell Movie Dialogues Corpus and Ubuntu Dialogue Corpus
	Shen et al.[49]	CSRR(Conversational Semantic Relationship RNN)	Cornell Movie Dialogues Corpus and Ubuntu Dialogue Corpus

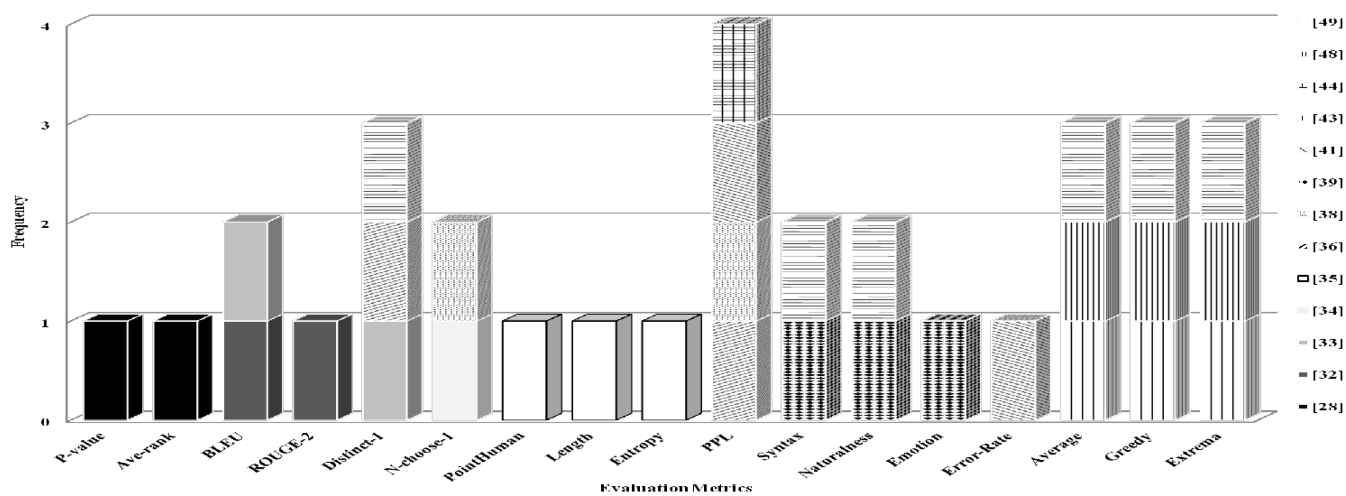

Fig. 8. Summary of the use of evaluation metrics in generation-based dialog models.

the Chinese data of Weibo. The evaluation results show that the results of the IR-based model are similar to those of the model with a global strategy, and the model with a hybrid strategy has the best results. But the model only carries out single-turn dialog, and without adding intention or emotional information.

Because the response generated by the dialog system based on the basic Seq2Seq model is too simple, contains too little information, and tends to generate general replies, some scholars begin to explore the direction of diversity and rich content of replies. As we all know, the traditional Seq2Seq model uses the beam search (BS) algorithm when choosing the response. Although this algorithm can reduce the space and time occupied by the search, the difference between the sentences output by the BS algorithm is very small, which cannot reflect the diversity of the language. Li *et al.* [32] introduced penalty factors into the BS of traditional Seq2Seq model to influence ranking results. At the same time, RL is used to automatically adjust the diversity rate of different inputs, which makes the output results more diverse. Vijayakumar *et al.* [33] propose a diversity BS algorithm, which first groups the beamwidth, then guarantees the difference between different groups by adding a diversity penalty, so that the generated replies remain diversity. Shao *et al.* [34] add a self-attention unit to ensure the length and coherence of the conversation and use a random BS algorithm to search candidate conversations in the solution space, then rearrange them to get the final results. The above three articles have achieved good results in resolving the diversity of replies, but replies not only need diversity but also need to make replies more meaningful and avoid universal replies. Mou *et al.* [35] propose a sequence to backward and

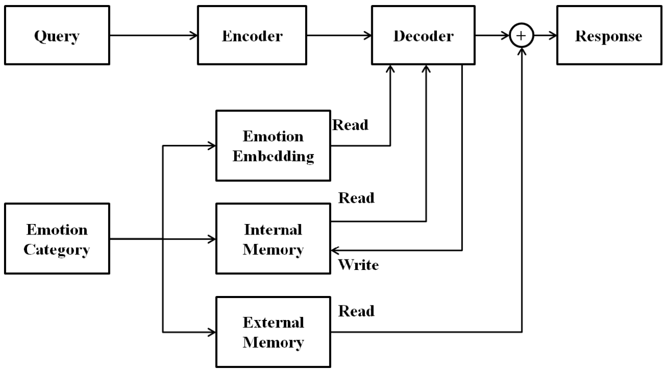

Fig. 9. Emotional chatting memory (ECM) model.

forward sequences (Seq2BF) model, which consists of three parts. The first part is the prediction keyword, using pointwise mutual information (PMI) to predict a noun as a keyword, the second part is the backward Seq2Seq model, which generates the responses of the first half of the sentence backward by using keywords, and the last part is the forward Seq2Seq model, which reverses the first half of the second part and inputs them into the forward model to generate the second half of the sentence reply. Therefore, the model can explicitly generate replies with keywords, and the reply content contains more accurate information. Xing *et al.* [36] propose a topic aware sequence-to-sequence (TA-Seq2Seq) model, which is an improvement based on the topic augmented joint attention-based Seq2Seq (TAJA-Seq2Seq) [37] model, which combines topic attention and information attention together to influence decoding output. Among them, topics are captured through the TwitterLDA model. The model uses prior subject information when decoding the first output tag, so the quality of the first tag will be higher, and the quality of the first tag will affect the decoding quality of the whole sentence, so the quality of the response generated by the model will be higher. And the TwitterLDA model is trained using data other than the training set, thus increasing the diversity of responses and the amount of information.

Human-to-human communication is emotional, so human-to-machine communication should also be emotional. Zhou *et al.* [38] propose an emotional chatting memory (ECM) model based on the memory network, which is the pioneering work of introducing emotional factors into a generative dialog system based on DL. This model generates responses with context information and emotional information through emotional embedding, internal memory, and external memory, as shown in Fig. 9. More specifically, the function of the emotional embedding part is to place the emotional category embedding vectors and the word embedding in decoder at the same level, and embedding them statically in the Seq2Seq model; the function of the internal memory part is to memorize the emotional state dynamically; the function of the external memory part is to introduce the external emotional memory mechanism so that the model can select words according to the external emotional dictionary or nonemotional dictionary. Asghar *et al.* [39] make three improvements on the basis of traditional Seq2Seq model to introduce emotion into dialog model: (1) adding emotional information to words embedding; (2) designing loss function with emotional factors; (3) designing emotional diversity decoding, mainly improving BS to affectively diverse beam search (ADBS). Asghar *et al.* [39] improve the basis of the ECM model and the Affect-LM model [40]. It not only improves the Affect-LM model which only considers the language model, but also explores the emotional factors in decoding, and improves the method of specifying emotions in the ECM model through emotional embedding to make it more realistic.

Most of the above articles are about single-turn dialogs. Normal dialogs often require multiturn dialogs. Therefore, the study of multiturn dialogs is also a research hotspot of dialogs system. Serban *et al.* [41] extend the hierarchical recurrent encoder–decoder (HRED) [42] model and applies the HRED model to the multiturn dialog. The model adds a context hiding layer between the encoding layer and decoding layer of the traditional Seq2Seq model, which stores and transmits context dialog information for use in the generation of a new turn of dialog. However, the response of the model is relatively single. Therefore, Vlad Serban *et al.* [43] improve the HRED model and proposes a latent variable hierarchical recurrent encoder–decoder (VHRED) model, which introduces a Gauss random variable into the context hidden layer to increase the diversity of responses. In addition, Yao, *et al.* [44] propose an attention with intention (AWI) model which is inspired by the thought of discourse structure [45], which includes language structure, intention structure, and attention state. And the AWI model also adopts the idea of layering. Unlike HRED model and VHRED model, AWI model adds attention mechanism to make the model pay more attention to important information, thus improving the effect of the model. Recently, latent variable models based on variational autoencoders (VAEs) [47] shows better performance for dialog generation. Some researchers use VAEs with hierarchical RNNs for multiturn dialog generation, which get better performance, but VAEs suffer from the degeneration problem. To solve the degeneration problem, Park *et al.* [48] propose a variational hierarchical conversation RNNs (VHCR) model, which alleviates the degradation problem to some extent by exploiting an utterance drop regularization. Shen *et al.* [49] make some improvements based on the VHCR and propose a conversational semantic relationship RNN (CSRR) model, which can generate the query and response that are consistent on the topic but different on the content.

The advantage of the generation model is that it does not need to label and extract features manually, and it can generate reasonable responses to questions that do not exist in the corpus. Because most of the generation models use neural network, and neural network has a certain learning ability. But the learning ability of neural network needs to be improved, so the fluency and logic of the generated replies are not good, and the generation model makes grammatical error easily and is difficult to train.

3) Model Based on RL: RL is widely used in robots [50], [51], games [52], [53] and network security [54]. In dialog system, action refers to generating dialog, state refers to human–machine conversation, strategy refers to deciding what kind of dialog to respond to according to the current state, and reward refers to evaluation metrics of the outcome of dialog.

RL is generally applied to dialog state tracking and dialog strategy selection, which belong to the natural language processing module in a dialog system. By combining with the traditional dialog model, the dialog effect of the model can be improved. Li *et al.* [55] use deep RL to solve the problem of multiturn dialog. By combining deep RL with the traditional Seq2Seq model, designing three kinds of returns to solve the problems of dull response, repetitive response and ungrammatical response. Avoiding the problem that the Seq2Seq model tends to generate general and security response. Making the response generated by the model richer and more diverse. As the training process of RL is a process of finding the optimal value through trial and error, the convergence speed of the model will be slow, so it is very important to improve the learning efficiency of RL. Lipton *et al.* [56] propose a bayes-by-backprop Q-network (BBQN) model, which is based on deep Q-network (DQN) [57], to solve the problem of convergence speed. The uncertainty of Q value is solved by combining Bayesian neural network, and BBQN model can significantly improve the exploration speed of deep Q-learning agents in the dialog system. In the same way, Zhao and Eskenazi [58] use the end-to-end model and deep RL to solve the problem of language understanding and dialog state strategy selection in a task-oriented dialog system and propose a hybrid algorithm combining RL and supervised learning to speed up learning efficiency.

The advantage of the RL model is that it can solve the dependence problem of multiple modules in the pipeline dialog model, has a good effect of multiturn dialog, and generates rich and logical responses. But the exploring efficiency of RL needs to be improved.

4) Model Based on Hybrid: The hybrid model integrates several parts of template-based model, IR-based model, generation-based model, RL-based model, and external knowledge, so as to exert their respective advantages and improve the overall effect of the model.

Qiu *et al.* [59] integrate the IR-based model and generation-based model into a dialog model, which includes three parts: IR-based model, generation-based model, and rerank model. First, the IR-based model is used to retrieve the set of candidate question–answer pairs, and the question Q and the candidate answer R are formed into a pair. Then, the confidence scores are calculated by the rerank model. And the score of the question–answer pair with the highest score is O. Finally comparing the size of O to the threshold value T, if $O > T$, the final answer is R. On the contrary, the final answer is R' generated by the generation-based model. This model can give full play to the fluency and logic of the IR-based model, and can also use the generate-based model to deal with questions that have not appeared in the database. Vougiouklis *et al.* [60] add external knowledge to the dialog generation and combines external knowledge with the generation-based model. The dialog model in this article includes two parts: one part is the sentence modeling, which first uses Wikipedia data set to pretrain the CNN, so that the CNN can extract external knowledge. After training, this part extracts local information corresponding to the input sequence. The other part is the sequence modeling, which is an RNN model. The last hidden layer state of this model will add the local information extracted from the previous part. In particular, this article is the first attempt to construct an end-to-end learning system by using the aligned data of two different data sources to automatically generate context-aware responses. Compared with the pure sequence model based on RNN, the model in this article has a 55% improvement in the degree of perplexity. The article [61] is also a dialog model combining external knowledge, but the difference between the model in [61] and that in [60] lies in the different way of combining external knowledge. Vougiouklis *et al.* [60] combine the historical information and external knowledge at the decoder end; however, Ghazvininejad *et al.* [61] combine the historical information and external knowledge at the encoder end, and the external knowledge is searched from the knowledge base through keyword-based retrieval. When dealing with entities that do not appear in training data, the model can also give appropriate responses based on external knowledge, so that the model can be enriched by external knowledge without retraining the model. In the same year, Madotto *et al.* [62] propose a memory-to-sequence (Mem2Seq) model, which consists of end-to-end memory networks (MemNN) [63] encoder and memory decoder. More specifically, the function of MemNN is to represent the historical information of dialog in vector; memory decoder generates replies by reading, writing, and copying memory. And the content of memory is made up of dialog history information and knowledge-based (KB) information. If the expected word is in the KB information, it is selected from the KB information, otherwise it is selected from the dialog history information.

The advantage of the hybrid model is that it can give full play to the advantages of multiple models and make the whole model perform well in fluency, logicality, controllability, and handling of unappeared questions. But the design of the model is complex, which requires designers to be familiar with the above models, and the training time and prediction time of the model will increase.

IV. Common Data sets and Evaluation Metrics

A. Data sets

Data sets are the basic condition of the dialog system, and high quality data sets can make the training model more effective. By consulting a large number of documents, we summarize some data sets commonly used in single-turn or multiturn dialogs in English or Chinese, as shown in Table IV, and give a brief introduction to the composition of the data sets.

The following is a more detailed description of each data set.

Cornell Movie-Dialogues Corpus [64] is a dialog of 9035 roles extracted from 617 movies, with 10 292 movie roles, having 220 579 conversations. The corpus also contains the title information, role information, the actual text of each dialog, the structure of the dialog, and the original source of the dialog.

TABLE IV
SUMMARY OF THE COMMON DATA SETS

Dataset name	Download address	Single turn or multi-turn dialogue	Data sources
Cornell Movie-Dialogs Corpus	http://www.cs.cornell.edu/~cristian/Cornell_Movie-Dialogs_Corpus.html	multi-turn	Extract fictional dialogue from movie scripts
Ubuntu Dialogue Corpus	1)http://dataset.cs.mcgill.ca/ubuntu-corpus-1.0/ 2)https://github.com/rkadlec/ubuntu-ranking-dataset-creator	multi-turn	Dialogue extracted from Ubuntu dialogue log
Douban Dialogue Corpus	https://github.com/MarkWuNLP/MultiTurnResponseSelection	multi-turn	Dialogue from Douban comments
Short-Text Conversation Dataset	http://data.noahlab.com.hk/conversation/	single turn	Posts and comments extracted from Sina Weibo
Aligning Reddit and Wikipedia Dataset	https://github.com/pvougiou/Aligning-Reddit-and-Wikipedia	multi-turn	Sentences extracted from Reddit and Wikipedia
Papaya Conversational Dataset	https://github.com/bshao001/ChatLearner	single turn	Manual dialogue data and clean online data

Ubuntu Dialog Corpus [65] is a dialog extracted from the Ubuntu chat log. The data set consists of 930 000 dialogs and 7 100 000 utterances. And the main data sets include training set, validation set, and test set. More specifically, there are 1 000 000 examples in the training set, 50% positive (label 1) and 50% negative (label 0), and the data of training set is composed of contexts, utterances, and labels; there are 19 560 data in the validation set and 18 920 data in the test set, and the data in the validation set and the test set are composed of contexts, ground truth utterances, and nine interference responses.

Douban Dialog Corpus is the dialog data collected by Wu *et al.* [20] on the Douban. The training set consists of 1-M dialog data; the validation set consists of 50-K dialog data; the test set consists of 10-K dialog data. In addition, the minimum number of turns per task is 3, and the average number of turns per task is about 6.

Short-Text Conversation Data set [66], [67] is a short text content extracted from posts and comments under posts on Sina Weibo. There are 4.8 M "postresponse" pairs in the training set and 422 posts in the test set, each of which has about 30 responses.

Aligning Reddit and Wikipedia Data set [60] is composed of conversational sequence in Reddit and aligned sentences in Wikipedia. It contains a total of 15 K comments sequence and 75 K aligned Wikipedia sentences. And the data set is mainly used as the external knowledge base of the dialog system.

Papaya Conversational Data set consists of two parts: core data and peripheral data. More specifically, the core data are manufactured to maintain the consistent personality of the chat robot, who can be trained as a polite, patient, and humorous role by using core data, and users can modify the role information according to their own needs; the peripheral data are a collection of online resources, including scene dialogs designed to train robots, Cornell movie dialogs and 170 000 pairs of Reddit data cleaned up. In particular, the data set website has a program for automatically generating Reddit data, which can generate millions of pairs of Reddit data.

B. Evaluation Metrics

The evaluation metrics of the dialog model are divided into objective evaluation metrics and subjective evaluation metrics. The objective evaluation metrics mainly includes the evaluation metrics of the IR-based model and the generation-based model, and the subjective evaluation metrics mainly includes the human metrics. The classification of evaluation metrics is shown in Fig. 10.

1) Evaluation Metrics of IR Model: The evaluation metrics commonly used in retrieval models are recall@k (R@k), mean average precision (MAP) and mean reciprocal rank (MRR).

R@k refers to the choice of k most likely responses to a given question to see if the correct response is in the k responses. Usually, people use R@1 to make k equal to 1, because when building a data set, there is usually only one correct answer for each question in the test set.

MAP [68] is a commonly used evaluation metrics in the field of object detection and text classification. Precision refers to the proportion of positive classes in classification results. And AP refers to the average of all precision maxima when recall values are between 0 and 1. Because the AP we obtained above is the AP corresponding to a topic, the average of AP for all topics is the MAP, which can be calculated by the following formula:

$$\mathrm{MAP} = \frac{\sum_{q=1}^{Q} \mathrm{AP}}{Q} \tag{2}$$

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

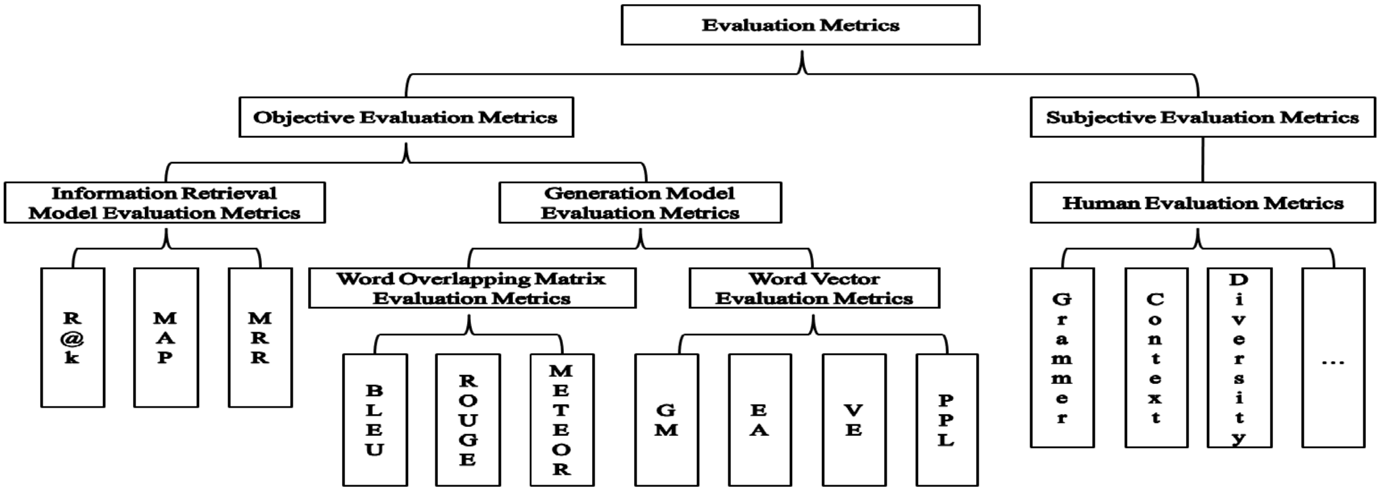

Fig. 10. Classification of evaluation metrics.

where Q is a collection of sample queries, $|Q|$ represents the number of queries.

MRR [69] evaluates the performance of the retrieval model by ranking the correct retrieval results in retrieval results. When calculating, the reciprocal ranking of the correct answers of a query in the search results is taken as its accuracy, and then averaging the accuracy of all queries. The MRR is defined as

$$\text{MRR} = \frac{1}{|Q|}\sum_{I=1}^{|Q|}\frac{1}{\text{rank}_i} \tag{3}$$

where rank_i is the position of the first relevant result in the ranking of the query i.

2) Evaluation Metrics of Generating Model: The evaluation metrics of the generated model are divided into word overlap matrix metrics and word vector metrics. More specifically, word overlap matrix mainly reflects the quality of generating responses by counting the number of occurrences of some phrases in sentences, and word vector calculates the similarity between words at the semantic level to express the similarity between sentences.

a) Word Overlap Matrix Metrics: Bilingual evaluation understudy (BLEU) [70] was originally designed to evaluate the quality of machine translation results and is now used in the field of dialog systems. The BLEU is represented by

$$P_n\left(r,\hat{r}\right) = \frac{\sum_k \min\left(h\left(k,r_i\right), h\left(k,\hat{r}_i\right)\right)}{\sum_k h\left(k,r_i\right)} \tag{4}$$

$$\text{BLEU-}N = b\left(r,\hat{r}\right)\exp\left(\sum_{n=1}^{N}\beta_n \log P_n\left(r,\hat{r}\right)\right) \tag{5}$$

where $h(k,r_i)$, $h(k,\hat{r}_i)$ represent the number of times each phrase appears in true and generated responses respectively. $P_n(r,\hat{r})$ reflects the accuracy of each n-gram phrase in the whole data set. Since the value of n is generally 1–4, it is necessary to the weighted sum P_n under different values of n, and β_n is the weight coefficient. In order to avoid short response, a penalty factor $b(r,\hat{r})$ is added to the evaluation metrics to improve the effect of the model.

Recall-oriented understudy for gisting evaluation (ROUGE) [71] is an automatic summary evaluation method, which evaluates by calculating the n-gram phrase recall rate between a candidate output and a reference output set. ROUGE, which is an improved version of BLEU, concentrates on recall rate, while BLEU concentrates on accuracy. Both of them can reflect word order, but the words reflected by ROUGE can be discontinuous, and the words reflected by BLEU must be continuous. The calculation formula of ROUGE is shown as follows:

$$\text{ROUGE} = \frac{\sum_{S\in\hat{E}\{\text{Ref Summaries}\}}\sum_{g_n\in\hat{E}S}\text{Count}_{\text{match}}\left(g_n\right)}{\sum_{S\in\hat{E}\{\text{Ref Summaries}\}}\sum_{g_n\in\hat{E}S}\text{Count}\left(g_n\right)} \tag{6}$$

where Ref is the abbreviation for reference, g is the abbreviation for gram, n is the length of n-gram phrases, and $\text{Count}_{\text{match}}(g_n)$ is the maximum number of n-gram phrases appearing in both candidate output and reference output sets.

METEOR [72] improves on the basis of BLEU by adding the alignment relationship between the generated response and the true response to make it more relevant to manual discrimination. And the METEOR can be described as

$$\text{METEOR} = \text{(1-Pen)}\, F_{\text{mean}} \tag{7}$$

where

$$\text{Pen} = \gamma\,(\text{frag})^{\beta} \tag{8}$$

$$F_{\text{mean}} = \frac{P_m R_m}{\alpha P_m + (1-\alpha) R_m} \tag{9}$$

where P_m and R_m refer to the precision and recall rate, respectively, frag is the fragment fraction, α, β, and γ are the weight coefficients, Pen is the penalty term, and F_{mean} is a statistic used to evaluate the classification model.

b) Word Vector Metrics: Greedy matching (GM) expresses the similarity of two sentences by calculating the similarity of word embedding vectors in two sentences. The GM is defined as follows:

$$G\left(r,\hat{r}\right) = \frac{\sum_{\omega\in r}\max_{\hat{\omega}\in\hat{r}}\text{cos\_sim}\left(e_\omega, e_{\hat{\omega}}\right)}{|r|} \tag{10}$$

$$\text{GM}\left(r,\hat{r}\right) = \frac{G\left(r,\hat{r}\right) + G\left(\hat{r},r\right)}{2} \tag{11}$$

where r and $\hat{r}$ are true response and generating response, respectively. ω and $\hat{\omega}$ are words in r and $\hat{r}$, respectively. e_ω and $e_{\hat{\omega}}$ are embedded vectors corresponding to ω and $\hat{\omega}$, respectively. First, each word in the true responses is computed its maximum cosine similarity in the generated responses, and mean value is calculated. Then, the same calculation is done for each word in the generated response. Finally, the final matching value is obtained by averaging the two words.

Embedding average (EA) is a method of calculating sentence vectors by word vectors in sentences. It has been applied to dialog systems and text similarity tasks. And it can be obtained by the following formula:

$$\bar{e_r} = \frac{\sum_{\omega \in r} e_\omega}{\sum_{\omega' \in r} e_{\omega'}} \tag{12}$$

$$\mathrm{EA} = \cos\left(\bar{e_r}, \bar{e r}\right) \tag{13}$$

where $\bar{e_r}$ denotes the mean of word vectors of all phrases in a sentence and taking the cosine similarity of the mean of word vectors of two sentences as the metrics of evaluating sentence similarity.

Vector extrema (VE) [73] is a method of calculating the similarity between two sentences on the sentence level vector. In this method, the maximum 1-D vector in a sentence is selected as the sentence vector. The formula is as follows:

$$e_{rd} = \begin{cases} \max_{\omega \in r} e_{\omega d}, & \text{if } e_{\omega d} > |\min_{\omega' \in r} e_{\omega' d}| \\ \min_{\omega \in r} e_{\omega d}, & \text{otherwise} \end{cases} \tag{14}$$

$$\mathrm{VE} = \cos\left(e_{rd}, e_{\hat{r}d}\right) \tag{15}$$

where d is the dimension of word vector and $e_{\omega d}$ is the first d dimension of ω embedded vector.

Perplexity (PPL) is a metrics based on information theory to measure the quality of a probability model to predict samples. The calculation formula is as follows:

$$\mathrm{PPL} = b^{-\frac{1}{N}\sum_{i=1}^{N} \log_b p(x_i)} \tag{16}$$

where $p(x_i)$ refers to the probability of the occurrence of the word x. N is the number of all words. The value of b is usually 2 or e. And the smaller the value of PPL, the better the prediction effect of the language model.

3) Human Evaluation Metrics: The existing objective evaluation metrics reflect the relationship between questions and responses to a certain extent, but there is no metrics that can solve the evaluation problem of the dialog system very well. Therefore, many researchers and enterprises prefer to use human evaluation metrics to evaluate the results of the dialog system. Human evaluation metrics are generally from grammar, contextual relevance, diversity, and other aspects. The higher the score, the higher the quality of the response. On the contrary, the lower the quality of the response.

The advantage of human evaluation is that it can fully reflect the real feelings of human beings and make the design of the dialog system develop in line with the daily conversation habits of human beings. But human evaluation needs to find volunteers and design questionnaires, which is time-consuming and laborious.

V. Analysis of the Evaluation Metrics

A. Analysis

The evaluation metrics of the IR model, generating model, and human are shown in Figs. 7 and 8. If we know the specific evaluation metric values of each model, we can compare the models. But we just count the frequency of each metric in the corresponding literature. There are two reasons why we do this, on the one hand, because the data sets used in these articles are generally different, the evaluation metrics of different data sets are not comparable; on the other hand, because there is no unified evaluation metrics adopted by all articles, the evaluation metrics of each article are very different. So it is not feasible to compare different models by the value of evaluation metrics.

Next, we give the analysis of the evaluation metrics for each type of model as follows.

1) Model Based on IR: As shown in Fig. 7, first, we can catch sight of MRR being widely used in the articles using IR-based model, and we can infer that MRR is more suitable for different data and tasks in IR-based model. Second, we can see that literature [23] has the most metrics, because this article has three tasks, but other articles just have one or two tasks. Then, we also can find that the literature [20], [21], and [29] have the same metrics, the reason is that these three articles use the same data set as shown in Table II and face the same task. So, we can find that the evaluation metric is related to the task, and completing different task requires different data sets. Finally, the human evaluation metrics are not found in Fig. 7, the IR-based model does not need it, because the responses of this model are usually logical and fluent.
2) Model Based on Generation: As shown in Fig. 8, we can find PPL occurs more frequently, which may be more applicable for more data sets and tasks. There are many human evaluation metrics used in these articles. Because the responses of generation-based model are not fluent and logical, there are many problems needed to solve which just like the style column shows in Table III and other problems not shown. It is a challenge to solve these problems. And it is even more challenging to design an metric to evaluate the model of solving these problems in a unified way. So this leads to the phenomenon of sparse metrics in Fig. 8.

In summary, the distribution of evaluation metrics of these two kinds of models is sparse, and no unified metric can evaluate these two kinds of models very well. The design of evaluation metrics of the IR-based model is relatively easy; however, the design of evaluation metrics based on the generation model is relatively more complex and most of the evaluation metrics are manually evaluated.

B. Concluding Remarks

From the above analysis, some conclusions can be drawn as follows.

1) The evaluation metrics of both IR-based model and generation-based model can reflect the quality of the

response to a certain extent. But so many metrics are confused that a unified evaluation of the dialog model cannot be formed.

2) It is one-sided to evaluate the dialog system by referring to the evaluation metrics of other tasks of natural language processing, which may mislead the training of the dialog model. Therefore, it is necessary to design automatic evaluation metrics highly related to human evaluation metrics.
3) Through comparison, it is found that the human evaluation metric is more often used in the generated-based dialog model.
4) Data set is also an important factor affecting the evaluation results. And so many data sets with different quality are easy to interfere with the evaluation results of dialog models. So different tasks require a standard data set to facilitate performance comparison between models.

VI. Existing Issues and Future Research Directions

Although dialog models have made considerable progress in recent years, several typical issues still remain to be better addressed due to the complexity of natural language processing. In particular, we list the following typical issues and outline the potential research directions toward solving them.

A. Existing Typical Issues

1) On the High Quality Large Data set: The size and quality of the data set determine the response effect of the dialog model. And the larger the data set and the higher the quality, the more useful information it contains, and the better the response effect of the dialog system will be. But, most of the existing data sets are generally of low quality and contain little useful information.
2) On the Unified Automatic Evaluation Metrics: The existing objective evaluation metrics have no clear correlation with human evaluation [74], which cannot meet the needs of dialog model evaluation, and human evaluation metrics are time-consuming. Therefore, it is necessary to design automatic evaluation metrics highly related to human evaluation.
3) On the Generalization Ability of the Model: When one trained dialog model faces the new business scenarios and new data set, the generalization ability of the model may be poor. Especially, with the continuous updating and enlargement of data sets, it is time-consuming and computational cost-consuming to continuously retrain the dialog model. And how to integrate multidomain knowledge for the training and use of dialog system or how to easily transfer the knowledge learned from the dialog model to new business scenarios is a problem needed to solve, which is a key problem to improve the generalization ability of the model.
4) On the Better Personification: Everyone has his own unique personality, which is a combination of emotions, values, personality, and other factors. Existing dialog robots can generate personalized responses with emotions, but they still give people a sense of inauthenticity, which is related to the lack of better personification.
5) On the Reasoning Ability: The products related to a dialog system that can be seen on the market give people a feeling of not smart. The main reason is that the existing dialog system cannot reasoning like human beings, which is the most critical factor restricting the development of the dialog system to a higher level of intelligence.

B. Future Research Directions

In the face of the above problems, combined with the development trend and actual needs of human–machine interaction system, the following aspects of research will be the potential directions in this field in the future.

1) Construction of High Quality Large Data set: The existing models are all based on data-driven, and the quality and scale of data are very important to the models. Therefore, how to build a large data set, especially high quality data set is an important research direction.
2) Establishment of Dialog Evaluation System: The evaluation metrics designed in the future should be able to evaluate the reasonableness of response from different granularity of characters, words, and sentences. Designers need to fully consider the influence of different granularity factors and design more relevant evaluation metrics to human beings, which can improve the effect of dialog system. Kannan and Vinyals [75] try to use GAN to evaluate the dialog system. Lowe *et al.* [76] design an automatic dialog evaluation model (ADEM) on the basis of hierarchical RNN. These two works provide some references for the design of new evaluation metrics.
3) Improving the Generalization Ability of the Model: In order to improve the generalization ability of the dialog model, there are two directions worth studying. One is to dynamically integrate the dialog system with the knowledge graph so that the dialog system can continuously learn new knowledge in the process of communicating with human beings and store the learned knowledge in the knowledge graph. The dialog system can call the knowledge in the knowledge graph at any time. And adding knowledge in the knowledge graph as prior knowledge into the dialog model can improve the response effect of the dialog model. The other is to integrate the multidomain knowledge which can reduce the errors caused by task switching and improve the generalization ability of the model and use transfer learning which can save time and labor costs and be more in line with the human learning process.
4) Design of Anthropomorphic Model: The anthropomorphic design of the dialog system is also a direction worthy of study. Researchers need to take into account the various factors that make up human personality and integrate them into the dialog system so that the dialog robot can highly personify the human personality and

give people a more real feeling. We also need to maintain the consistency of personality. Some of the existing methods are controlled by training data sets and rules. We can learn from the thought of control theory and control the output of the neural network by designing appropriate controllers [77].

5) Research on Inductive Reasoning Theory: The reasoning ability of the dialog model has always been a challenging research problem. In recent years, graph neural network [78]–[80] can combine end-to-end learning with inductive reasoning, which is expected to solve the problem that DL cannot reasoning. The research of inductive reasoning theory is applied in the field of a dialog system, so that the dialog system can also reasoning is the key to achieve strong AI.

VII. CONCLUSION

Dialog model is crucial to building dialog systems. In this article, a comprehensive survey on the learning-based dialog models is presented. In particular, the current development status and construction process of human–machine dialog system is reviewed, and the typical dialog models and methods as well as the corresponding advantages and disadvantages are compared and analyzed. Also, the challenging issues on the high quality large data set, the unified automatic evaluation metrics, the generalization ability of the model, the better personification and the reasoning ability, together with the potential research directions are briefly discussed.

All-round research on learning-based dialog model is not only of great theoretical significance but also of important social value. On the one hand, the research in this field will help to understand the nature of human dialog and promote the development of a dialog system and its related technologies. On the other hand, the research will bring bright prospects for the collaborative development of linguistics and AI.

REFERENCES

[1] J. Weizenbaum, "ELIZA—A computer program for the study of natural language communication between man and machine," *Commun. ACM*, vol. 9, no. 1, pp. 36–45, 1983.

[2] K. M. Colby, *Artificial Paranoia; a Computer Simulation of Paranoid Processes*. New York, NY, USA: Pergamon, 1975.

[3] S. R. Wallace, "The anatomy of A.L.I.C.E.," in *Parsing the Turing Test*. Berlin, Germany: Springer, 2009, pp. 181–210.

[4] G. Hinton *et al.*, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, Nov. 2012.

[5] L. Deng *et al.*, "Recent advances in deep learning for speech research at microsoft," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Vancouver, BC, Canada, May 2013, pp. 8604–8608.

[6] W. Xiong *et al.*, "Toward human parity in conversational speech recognition," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, vol. 25, no. 12, pp. 2410–2423, Dec. 2017.

[7] Y. Qian, Y. Fan, W. Hu, and F. K. Soong, "On the training aspects of deep neural network (DNN) for parametric TTS synthesis," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Florence, Italy, May 2014, pp. 3829–3833.

[8] A. van den Oord *et al.*, "WaveNet: A generative model for raw audio," 2016, *arXiv:1609.03499*. [Online]. Available: http://arxiv.org/abs/1609.03499

[9] Y. Bengio, R. Ducharme, P. Vincent, and C. Janvin, "A neural probabilistic language model," *J. Mach. Learn. Res.*, vol. 3, pp. 1137–1155, Feb. 2003.

[10] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, Montreal, QC, Canada, 2014, pp. 3104–3112.

[11] G. Salton, and M. J. McGill, *Introduction to Modern Information Retrieval*. New York, NY, USA: McGraw-Hill, 1986.

[12] A. M. Elkahky, Y. Song, and X. He, "A multi-view deep learning approach for cross domain user modeling in recommendation systems," in *Proc. 24th Int. Conf. World Wide Web (WWW)*, Florence, Italy, 2015, pp. 278–288.

[13] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, 2014.

[14] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," 2014, *arXiv:1409.0473*. [Online]. Available: http://arxiv.org/abs/1409.0473

[15] G. Mesnil, X. He, L. Deng, and Y. Bengio, "August. Investigation of recurrent-neural-network architectures and learning methods for spoken language understanding," in *Proc. Interspeech*, 2013, pp. 3771–3775.

[16] H. Zhang, R. Kishore, R. Sharman, and R. Ramesh, "Agile integration modeling language (AIML): A conceptual modeling grammar for agile integrative business information systems," *Decis. Support Syst.*, vol. 44, no. 1, pp. 266–284, Nov. 2007.

[17] P.-S. Huang, X. He, J. Gao, L. Deng, A. Acero, and L. Heck, "Learning deep structured semantic models for Web search using clickthrough data," in *Proc. 22nd ACM Int. Conf. Conf. Inf. Knowl. Manage. (CIKM)*, San Francisco, CA, USA, 2013, pp. 2333–2338.

[18] S. Wan *et al.*, "A deep architecture for semantic matching with multiple positional sentence representations," in *Proc. 30th AAAI Conf. Artif. Intell.*, Phoenix, AZ, USA, 2016, pp. 2835–2841.

[19] L. Pang *et al.*, "Text matching as image recognition," in *Proc. 30th AAAI Conf. Artif. Intell.*, Phoenix, AZ, USA, 2016, pp. 2793–2799.

[20] Y. Wu, W. Wu, C. Xing, M. Zhou, and Z. Li, "Sequential matching network: A new architecture for multi-turn response selection in retrieval-based chatbots," in *Proc. 55th Annu. Meeting Assoc. Comput. Linguistics*, Vancouver, BC, Canada, 2017, pp. 496–505.

[21] X. Zhou *et al.*, "Multi-turn response selection for chatbots with deep attention matching network," in *Proc. 56th Annu. Meeting Assoc. Comput. Linguistics*, Melbourne, VIC, Australia, 2018, pp. 1118–1127.

[22] A. Vaswani *et al.*, "Attention is all you need," 2017, *arXiv:1706.03762*. [Online]. Available: http://arxiv.org/abs/1706.03762

[23] J. Wang *et al.*, "IRGAN: A minimax game for unifying generative and discriminative information retrieval models," in *Proc. 40th Int. ACM SIGIR Conf. Res. Develop. Inf. Retr. (SIGIR)*, Shinjuku City, Japan, 2017, pp. 515–524.

[24] I. J. Goodfellow *et al.*, "Generative adversarial networks," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 3, 2014, pp. 2672–2680.

[25] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," 2014, *arXiv:1409.3215*. [Online]. Available: http://arxiv.org/abs/1409.3215

[26] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[27] J. Chung, C. Gulcehre, K. Cho, and Y. Bengio, "Empirical evaluation of gated recurrent neural networks on sequence modeling," in *Proc. Deep Learn. Represent. Learn. Workshop (NIPS)*, Montreal, QC, Canada, 2014, pp. 1–9. [Online]. Available: http://arxiv.org/abs/1412.3555

[28] L. Shang, Z. Lu, and H. Li, "Neural responding machine for short-text conversation," in *Proc. 53rd Annu. Meeting Assoc. Comput. Linguistics 7th Int. Joint Conf. Natural Lang. Process.*, Beijing, China, 2015, pp. 1577–1586.

[29] Z. Zhang *et al.*, "Modeling multi-turn conversation with deep utterance aggregation," in *Proc. 27th Int. Conf. Comput. Linguistics*, Santa Fe, NM, USA, 2018, pp. 3740–3752.

[30] Z. Ji, Z. Lu, and H. Li, "An information retrieval approach to short text conversation," 2014, *arXiv:1408.6988*. [Online]. Available: http://arxiv.org/abs/1408.6988

[31] A. Ritter, C. Cherry, and W. B. Dolan, "Data-driven response generation in social media," in *Proc. EMNLP*, Edinburgh, Scotland, 2011, pp. 27–31.

[32] J. Li, W. Monroe, and D. Jurafsky, "A simple, fast diverse decoding algorithm for neural generation," 2016, *arXiv:1611.08562*. [Online]. Available: http://arxiv.org/abs/1611.08562

[33] A. K Vijayakumar *et al.*, "Diverse beam search: Decoding diverse solutions from neural sequence models," 2016, *arXiv:1610.02424*. [Online]. Available: http://arxiv.org/abs/1610.02424

[34] Y. Shao, S. Gouws, D. Britz, A. Goldie, B. Strope, and R. Kurzweil, "Generating high-quality and informative conversation responses with sequence-to-sequence models," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, Copenhagen, Denmark, 2017, pp. 2200–2209.

[35] L. Mou *et al.*, "Sequence to backward and forward sequences: A content-introducing approach to generative short-text conversation," in *Proc. COLING*, Osaka, Japan, 2016, pp. 11–17.

[36] C. Xing *et al.*, "Topic aware neural response generation," in *Proc. AAAI*, San Francisco, CA, USA, 2017, pp. 3351–3357.

[37] C. Xing *et al.*, "Topic aware neural response generation," 2016, *arXiv:1606.08340*. [Online]. Available: http://arxiv.org/abs/1606.08340

[38] H. Zhou *et al.*, "Emotional chatting machine: Emotional conversation generation with internal and external memory," in *Proc. AAAI*, New Orleans, LA, USA, 2018, pp. 730–738.

[39] N. Asghar, P. Poupart, J. Hoey, X. Jiang, and L. Mou, "Affective neural response generation," 2017, *arXiv:1709.03968*. [Online]. Available: http://arxiv.org/abs/1709.03968

[40] S. Ghosh, M. Chollet, E. Laksana, L.-P. Morency, and S. Scherer, "Affect-LM: A neural language model for customizable affective text generation," in *Proc. 55th Annu. Meeting Assoc. Comput. Linguistics*, Vancouver, BC, Canada, 2017, pp. 634–642.

[41] I. V. Serban, A. Sordoni, Y. Bengio, A. Courville, and J. Pineau, "Building end-to-end dialogue systems using generative hierarchical neural network models," in *Proc. 30th AAAI Conf. Artif. Intell.*, Phoenix, AZ, USA, 2016, pp. 3776–3783.

[42] A. Sordoni, Y. Bengio, H. Vahabi, C. Lioma, J. G. Simonsen, and J.-Y. Nie, "A hierarchical recurrent encoder-decoder for generative context-aware query suggestion," in *Proc. 24th ACM Int. Conf. Inf. Knowl. Manage. (CIKM)*, Melbourne, VIC, Australia, 2015, pp. 553–562.

[43] I. V. Serban *et al.*, "A hierarchical latent variable encoder-decoder model for generating dialogues," in *Proc. AAAI*, San Francisco, CA, USA, 2017, pp. 3295–3301.

[44] K. Yao, G. Zweig, and B. Peng, "Attention with intention for a neural network conversation model," 2015, *arXiv:1510.08565*. [Online]. Available: http://arxiv.org/abs/1510.08565

[45] B. J. Grosz, "Attention, intentions, and the structure of discourse," *Comput. Linguistics*, vol. 12, pp. 175–204, Jul. 1986.

[46] E. Levin, R. Pieraccini, and W. Eckert, "Using Markov decision process for learning dialogue strategies," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Seattle, WA, USA, May 1998, pp. 201–204.

[47] D. Kingma and M. Welling, "Auto-encoding variational Bayes," in *Proc. 2nd Int. Conf. Learn. Represent.*, Banff, AB, Canada, 2014, pp. 1–14.

[48] Y. Park, J. Cho, and G. Kim, "A hierarchical latent structure for variational conversation modeling," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol.*, Melbourne, VIC, Australia, 2018, pp. 1792–1801.

[49] L. Shen, Y. Feng, and H. Zhan, "Modeling semantic relationship in multi-turn conversations with hierarchical latent variables," in *Proc. 57th Annu. Meeting Assoc. Comput. Linguistics*, Florence, Italy, 2019, pp. 5497–5502.

[50] Z. Yang, K. Merrick, L. Jin, and H. A. Abbass, "Hierarchical deep reinforcement learning for continuous action control," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 11, pp. 5174–5184, Nov. 2018.

[51] B. Gokce and H. L. Akin, "Implementation of reinforcement learning by transfering sub-goal policies in robot navigation," in *Proc. 21st Signal Process. Commun. Appl. Conf. (SIU)*, Haspolat, Turkey, Apr. 2013, pp. 1–4.

[52] A. Jeerige, D. Bein, and A. Verma, "Comparison of deep reinforcement learning approaches for intelligent game playing," in *Proc. IEEE 9th Annu. Comput. Commun. Workshop Conf. (CCWC)*, Las Vegas, NV, USA, Jan. 2019, pp. 0366–0371.

[53] M. Xu, H. Shi, and Y. Wang, "Play games using reinforcement learning and artificial neural networks with experience replay," in *Proc. IEEE/ACIS 17th Int. Conf. Comput. Inf. Sci. (ICIS)*, Singapore, Jun. 2018, pp. 855–859.

[54] Z. Ni and S. Paul, "A multistage game in smart grid security: A reinforcement learning solution," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 9, pp. 2684–2695, Sep. 2019.

[55] J. Li, W. Monroe, A. Ritter, M. Galley, J. Gao, and D. Jurafsky, "Deep reinforcement learning for dialogue generation," 2016, *arXiv:1606.01541*. [Online]. Available: http://arxiv.org/abs/1606.01541

[56] Z. C. Lipton, X. Li, J. Gao, L. Li, F. Ahmed, and L. Deng, "BBQ-networks: Efficient exploration in deep reinforcement learning for task-oriented dialogue systems," 2016, *arXiv:1608.05081*. [Online]. Available: http://arxiv.org/abs/1608.05081

[57] V. Mnih *et al.*, "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, Feb. 2015.

[58] T. Zhao and M. Eskenazi, "Towards end-to-end learning for dialog state tracking and management using deep reinforcement learning," in *Proc. 17th Annu. Meeting Special Interest Group Discourse Dialogue*, Los Angeles, CA, USA, 2016, pp. 1–10.

[59] M. Qiu *et al.*, "AliMe chat: A sequence to sequence and rerank based chatbot engine," in *Proc. 55th Annu. Meeting Assoc. Comput. Linguistics*, Vancouver, BC, Canada, 2017, pp. 498–503.

[60] P. Vougiouklis, J. Hare, and E. Simperl, "A neural network approach for knowledge-driven response generation," in *Proc. COLING*, Osaka, Japan, 2016, pp. 3370–3380.

[61] M. Ghazvininejad *et al.*, "A knowledge-grounded neural conversation model," in *Proc. 32nd AAAI Conf. Artif. Intell.*, New Orleans, LA, USA, 2018, pp. 5110–5117.

[62] A. Madotto, C.-S. Wu, and P. Fung, "Mem2Seq: Effectively incorporating knowledge bases into end-to-end task-oriented dialog systems," in *Proc. 56th Annu. Meeting Assoc. Comput. Linguistics*, Melbourne, VIC, Australia, 2018, pp. 1468–1478. [Online]. Available: http://arxiv.org/abs/1804.08217

[63] S. Sukhbaatar, J. Weston, and R. Fergus, "End-to-end memory networks," in *Proc. Adv. Neural Inf. Process. Syst. (NIPS)*, Montreal, QC, Canada, 2015, pp. 2440–2448.

[64] C. Danescu-Niculescu-Mizil, M. Gamon, and S. Dumais, "Mark my words!: Linguistic style accommodation in social media," in *Proc. 20th Int. Conf. World Wide Web (WWW)*. Hyderabad, India: ACM, 2011, pp. 745–754.

[65] R. Lowe, N. Pow, I. Serban, and J. Pineau, "The ubuntu dialogue corpus: A large dataset for research in unstructured multi-turn dialogue systems," 2015, *arXiv:1506.08909*. [Online]. Available: http://arxiv.org/abs/1506.08909

[66] H. Wang, Z. D. Lu, H. Li, and E. H. Chen, "A dataset for research on short-text conversation," in *Proc. Empirical Methods Natural Lang. Process. (EMNLP)*, Seattle, WA, USA, 2013, pp. 935–945.

[67] M. Wang, Z. Lu, H. Li, and Q. Liu, "Syntax-based deep matching of short texts," in *Proc. 24th Int. Joint Conf. Artif. Intell.*, Buenos Aires, Argentina, 2015, pp. 1354–1361.

[68] R. Baeza-Yates, *Modern Information Retrieval*. Beijing, China: Machinery Industry Press, 2004.

[69] E. Voorhees, "The TREC-8 question answering track report," in *Proc. 8th Text Retr. Conf. (TREC)*, 1999, pp. 79–82.

[70] K. Papineni, S. Roukos, T. Ward, and W.-J. Zhu, "BLEU: A method for automatic evaluation of machine translation," in *Proc. 40th Annu. Meeting Assoc. Comput. Linguistics (ACL)*, Philadelphia, PA, USA, 2001, pp. 311–318.

[71] C.-Y. Lin and E. Hovy, "Manual and automatic evaluation of summaries," in *Proc. ACL Workshop Text Summarization Branches (WAS)*, Barcelona, Spain, 2004, pp. 74–81.

[72] A. Lavie and A. Agarwal, "Meteor: An automatic metric for MT evaluation with high levels of correlation with human judgments," in *Proc. 2nd Workshop Stat. Mach. Transl. (StatMT)*, Prague, Czech Republic, 2007, pp. 228–231.

[73] G. Forgues, J. Pineau, J.-M. Larchevêque, and R. Tremblay, "Bootstrapping dialog systems with word embeddings," in *Proc. Mod. Mach. Learn. Natural Lang. Process. Workshop (NIPS)*, 2014, pp. 1–5.

[74] C.-W. Liu, R. Lowe, I. Serban, M. Noseworthy, L. Charlin, and J. Pineau, "How NOT to evaluate your dialogue system: An empirical study of unsupervised evaluation metrics for dialogue response generation," in *Proc. Conf. Empirical Methods Natural Lang. Process.*, Austin, TX, USA, 2016, pp. 2122–2132.

[75] A. Kannan and O. Vinyals, "Adversarial evaluation of dialogue models," 2017, *arXiv:1701.08198*. [Online]. Available: http://arxiv.org/abs/1701.08198

[76] R. Lowe, M. Noseworthy, I. V. Serban, N. Angelard-Gontier, Y. Bengio, and J. Pineau, "Towards an automatic turing test: Learning to evaluate dialogue responses," in *Proc. 55th Annu. Meeting Assoc. Comput. Linguistics*, Vancouver, BC, Canada, 2017, pp. 1116–1126.

[77] Y.-D. Song, X. Huang, and Z.-J. Jia, "Dealing with the issues crucially related to the functionality and reliability of NN-associated control for nonlinear uncertain systems," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 11, pp. 2614–2625, Nov. 2017.

[78] F. Scarselli, M. Gori, A. Chung Tsoi, M. Hagenbuchner, and G. Monfardini, "The graph neural network model," *IEEE Trans. Neural Netw.*, vol. 20, no. 1, pp. 61–80, Jan. 2009.

[79] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. 5th Int. Conf. Learn. Represent. (ICLR)*, Toulon, France, 2016, pp. 1–14. [Online]. Available: http://arxiv.org/abs/1609.02907

[80] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *Proc. 6th Int. Conf. Learn. Represent. (ICLR)*, Vancouver, BC, Canada, 2017, pp. 1–12. [Online]. Available: http://arxiv.org/abs/1710.10903

[81] R. S. Boyer and J. S. Moore, "A fast string searching algorithm," *Commun. ACM*, vol. 20, no. 10, pp. 762–772, Oct. 1977.

[82] D. E. Knuth, J. H. Morris, Jr., and V. R. Pratt, "Fast pattern matching in strings," *SIAM J. Comput.*, vol. 6, no. 2, pp. 323–350, Jun. 1977.

[83] S. Wu and U. Manber, "Fast text searching: Allowing errors," *Commun. ACM*, vol. 35, no. 10, pp. 83–91, Oct. 1992.

[84] A. V. Aho and M. J. Corasick, "Efficient string matching: An aid to bibliographic search," *Commun. ACM*, vol. 18, no. 6, pp. 333–340, Jun. 1975.

Fuwei Cui received the M.S. degree from the North China University of Technology, Beijing, China, in 2017. He is currently pursuing the Ph.D. degree in control science and engineering with Beijing Jiaotong University, Beijing, China.

His specific areas of research interest mainly focus on deep learning, natural language processing, and human–computer interaction systems.

Qian Cui received the B.S. degree from the Wuhan Institute of Technology, Wuhan, China, in 2017. She is currently pursuing the Ph.D. degree with the School of Automation, Chongqing University, Chongqing, China.

Her research interests include robust adaptive control, neural networks, and learning systems.

Yongduan Song (Fellow, IEEE) received the Ph.D. degree in electrical and computer engineering from Tennessee Technological University, Cookeville, TN, USA, in 1992.

He held a tenured Full Professor with North Carolina A&T State University, Greensboro, NC, USA, from 1993 to 2008, and a Langley Distinguished Professor with the National Institute of Aerospace, Hampton, VA, USA, from 2005 to 2008. He is currently the Dean of the School of Automation, Chongqing University, Chongqing, China. He was one of the six Langley Distinguished Professors with the National Institute of Aerospace (NIA), Hampton, VA, USA, and the Founding Director of Cooperative Systems with NIA. His current research interests include intelligent systems, guidance navigation and control, and bio-inspired adaptive and cooperative systems.

Prof. Song was a recipient of several competitive research awards from the National Science Foundation, the National Aeronautics and Space Administration, the U.S. Air Force Office, the U.S. Army Research Office, and the U.S. Naval Research Office. He is an IEEE Fellow and has served/been serving as an Associate Editor for several prestigious international journals, including the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, the IEEE TRANSACTIONS ON SYSTEMS, MAN AND CYBERNETICS, and the IEEE TRANSACTIONS ON DEVELOPMENTAL AND COGNITIVE SYSTEMS, etc.