A Survey of the Chatbots system

Mohammed A. Jaleel
Department of Computer Science, Al-Turath University College

Alia K. Abdul Hassa
Department of Computer Science, University of Technology

Sarah S. Jasim
Department of IT, Middle Technical University

ABSTRACT. A chatbot is an artificial intelligence (AI) system programmed to simulate human conversation via text or natural language interaction. The majority of chatbots make use of AI algorithms to offer an acceptable response. In their interactions with humans, older chatbots just gave the impression of intelligence by using rule-based and generative-based models in place of today's more sophisticated pattern-matching and string-processing design methodologies withthe need for the conversation to be secure. But as new technologies have arisen, more intelligent systems have emerged, employing intricate knowledge-based models. This study aims to examine the many types of chatbots, how the two main types deal with conversational context, and the design strategies utilized in creating both older and more recent chatbots.

KEYWORDS: Conversational agents · Chatbots · Generative based · Classifications · Design techniques conversational context · Neural network · Algorithms · Rule based ·Retrieval based · Chat script · AIML · Pattern matching ·Machine learning

1 INTRODUCTION

Natural language interfaces replace graphical user interfaces as the preferred method of interaction between humans and computers. Recent years have seen a rise in interest in natural language processing as an appealing personalization technique. That's also because it enables users to express themselves in their own words rather than learning and memorizing a set of standard responses. As a result, it's possible to dialogue with the machine [1]. For example, the standard format for a text- or voice-based interaction begins with the user asking a question in natural language and the chatbot responding similarly [2]. They are typically programmed to mimic human speech or text to facilitate conversation between humans and chatbots.

Although they first emerged as simple conversation and entertainment systems, chatbots have since advanced to new levels of sophistication [2]. However, while sophisticated chatbots employ AI to respond to user queries, current state-of-the-art systems still struggle to hold conversations with people in a way that is coherent, contextual, and natural [3]. The use of chatbots is not novel. The period officially began when Joseph Weizenbaum released his ELIZA program in.

There was a strong impression that the application might make the user believe they were having a conversation with a live person [4]. ELIZA and other early chatbots relied on keyword matching and shallow contextual recognition, making it impossible for them to sustain a

conversation. When the ELIZA program detects a user-supplied keyword, it applies the rules connected with that word to the phrase and reformats it accordingly. The chosen decomposition rule has a connection to a group of assembly and reassembly rules responsible for producing the replies. His SCRIPT (data, not part of the program itself) for a particular conversation type consists of keywords and their accompanying transformation rules [5]. In addition to being one of the most advanced chatbots, ALICE (Artificial Linguistic Internet Computer Entity) was developed between 2001 and 2004 to process natural language. [6]. The extensible Markup Language (XML), created during the past decade, serves as the foundation for ALICE [7]. The vast majority of current chatbot frameworks are AIML-compatible. Those built with AIML are now the most widely used among chatbots due to their low resource requirements and simple configuration. Having to take into account the proper context when generating predefined responses to a particular set of queries also places restrictions on conversation bots. As a result of developers' and researchers' efforts, new, more functional components have been added to preexisting techniques, and entirely new design architectures have been proposed in some situations. Utilizing ontologies and recalling conversational context was key to most of these enhancements [8]. There are two main categories of chatbots: those meant to carry out specified tasks (called "task-oriented chatbots") and those built for introductory chit-chat (called "non-task-oriented chatbots"). I'm here. After introducing the various categories into which chatbots fall, this study analyzes the most frequent methods of chatbot design within each of those groups.

2 CLASSIfiCATION OF CHATBOTS

Due to the rapid evolution of the chatbot industry in recent years brought on by technological advances, chatbots are currently challenging to categorize precisely into distinct groups as necessary.

Some examples of what goes into creating these chatbots are the types of interactions they're designed for, the areas of knowledge they draw on, the purposes serve, and the design processes they employ (response generation methods). The philosophy of the chatbot's basic design, the need to preserve and consider the context for conversational understanding, and the nature and goal of the discussion are all examples of such criteria. [8]. The following criteria can use as a rough classification: (Fig. 1).

1. Preferred method of communication (typing or talking).
2. A Chatbot Application for Communicating with Robots (Task-Oriented or Non-Task-Oriented).
3. Intelligence based on rules or machine learning, deep learning etc.
4. Dedicated or Public Domains.

Chatbots can be classified into two broad categories [9–11]. (Fig. 2):

1. Chatbots that focus on tasks
2. Chatbots that don't focus on tasks

Fig. 1. Broad classification of Chatbots

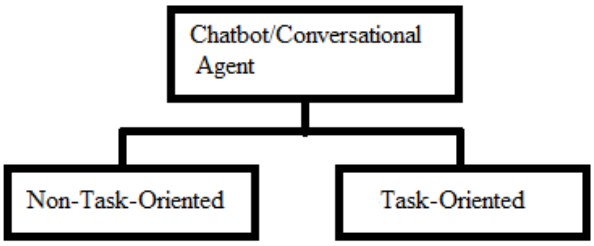

Fig. 2. Main classification of Chatbots

Conversational task-oriented chatbots are programmed for limited interaction within a narrow domain and limited conversation length. Non-task-oriented chatbots, on the other hand, can imitate human communication and carry idle chit-chat in an open domain for amusement purposes [10].

3 CHATBOTS DESIGN TECHNIQUES

Chatbots can be developed in a task-oriented or non-task-oriented fashion using various methods. Both types of chatbots may share features from each other's design philosophies. Since the very first chatbot ELIZA, numerous methods have been developed for creating chatbots in various frameworks, and we'll look at them all here. There are primarily three groups into which these methods fall [8, 10, 12-15].

1. Rules-based methods
2. Methods that Emphasize Access and Retrieval
3. Inherently generative methods

These three primary strategies can each make use of the following tools and methods:

1. Parsing Method
2. Matching Patterns
3. AIML

4. Chat script
5. Ontologies
6. Markov Chain Model,
7. Models Using Neural Networks

7.1. Recurrent Neural Networks (RNNs)

7.2. The Sequence-to-Sequence Neural Model

7.3. The Structure of the Networks That Support Long-Term and Short-Term Memories (LSTMs)

1. Parsing. Using a process called "parsing," by breaking it to a string of tokens, information can be extracted from a piece of text (lexical analysis), for example, discovering the grammatical of the sentence. ELIZA, an early chatbot, employed a rudimentary parsing mechanism to identify keywords in sentences. The keywords are then used to retrieve relevant documents from the corpus [16]. Semantic analysis is a more refined form of analysis. It takes a sentence typed in by the user and turns it into a form that a computer can understand. Today, commercial chatbot technologies like Dialogflow rely on it to detect user intent [17].
2. They were recognizing and resembling a pattern. The most popular method used in chatbots is pattern matching to categorize user input as a pattern> and generate a suitable answer from a template>. These patterns> template> sets are lovingly made by hand. While both primitive and advanced chatbots use pattern-matching techniques, the underlying algorithms' complexity varies widely. ALICE is a more sophisticated chatbot that employs complicated pattern-matching methods and assigns some conversational context to its searches [7]. Earlier chatbots, such as ELIZA, employed simple pattern-matching principles. Typically, QA chatbots will utilize a pattern-matching technique. The ability to adapt the discussions you make with this method is a big plus. Scalability issues arise because of the manual effort to construct all conceivable patterns. The scale problem constrains Chatbots' information extraction skills and the bots' responses may be predictable and routine.
3. AIM. Artificial Intelligence Markup Language (AIML) is an offshoot of Extensible Markup Language (XML) (Extensible Markup Language). The Artificial Intelligence Markup Language (AIML) was designed for coding chatbot conversation flows. The AIML language specification is freely available. AIML objects are data used in the language—these so-called AIML objects.

   Themes and categories are used to categorize the items. The subject element is an optional root node containing a name attribute and a collection of child categories. The smallest unit of information storage in an AIML file is a category. At least two items with the names "Patterns" and "Templates" are in a category. Each category is a rule for matching inputs and turning them into outputs. User input is matched to "patterns," and "templates" are used to make chatbot responses [18]. AIML is powerful, flexible, and easy to use to shape how chatbots talk to each other, but you need to know a little bit about natural language processing and programming to use it.
4. Chat script. Chat script is a free and open-source program for making interactive virtual assistants. For continuous user state preservation during talks, Chat script combines a natural language engine with a dialogue management system [19]. Rules are developed for this rule-based engine using a dialogue flow scripting technique. Scripts like these can be stored in a text file. Conversation logs from users can be analyzed with machine learning methods to fine-

tune the everyday experience. Concepts are used in the chat script. For example, you can create any exact word or adverb. There are currently around 2000 predefined concepts and scripts in the database, and new ones can be made with little effort [20].

5. Ontologies. When a chatbot needs to replace human-created domain knowledge with ontology-based domain knowledge, it turns into a domain ontology. Ontologies are not new; some components of conversational systems already use them. For example, it's been utilized as a foundation for systematic grammatical techniques in fields like language generation [21]. The primary benefit of utilizing domain ontologies in chatbots is that they can investigate concept nodes in the ontology to build links between concepts used in conversation. Still, they are also able to suggest new inferences [22]. However, ontologies' broader application to chatbots is limited.
6. Modelling random processes using Markov chains. Because it is based on probabilities, the Markov chain process is considered a probabilistic model. Markov chains are an attempt to model the likelihood of transitions between states. Markov chain models are based on the central premise that, given the current state, the chance of going to one or more states is constant. The chatbot will produce the output associated with the state change when this technique is used. Its paves the way for the chatbot to construct an improved, novel, and largely reliable set of responses. Because the user determines the starting point, responses are more likely to be pertinent. Markov chains are widely used to program recreational chatbots that can hold basic conversations as humans. Markov chain models are wrong at simulating complex talks because they simplify more involved decision-making processes [8]. A matrix can be used to implement a simple Markov chain model, and the model can be easily programmed.
   Models for Artificial Neural Networks. Improvements in chatbot intelligence have been made possible by recent developments in machine learning, in particular, the use of artificial neural networks. Search-based and generative approaches to response generation are both possible for chatbots based on artificial neural networks [23, 24], with the latter being the current focus of attention in the field of study. Distinctive features of artificial neural network chatbots versus rule-based chatbots Learning algorithms are present in the artificial neural network building method. One learning method implemented in machines is the artificial neural network model [25]. Machine learning algorithms can be either supervised or unsupervised. Deep learning is a subfield of machine learning that permits unsupervised training on unstructured or unlabeled data. Data is processed, and patterns are generated like the human brain does to arrive at conclusions and choices. Artificial neural networks have several uses, including computer vision, decision making, speech recognition, machine translation, social network filtering, and medical diagnosis. Artificial neural networks in various forms have been applied to NLP [26]. The discipline of conversation modelling is increasingly utilizing deep learning neural networks, with recurrent neural networks (RNNs), sequence-to-sequence networks, and extended short-term memory networks (LSTMs) dominating the space [24].

7:1 Recurrent Neural Network (RNN). In artificial intelligence, recurrent neural networks (RNNs) are a subset of the broader category of recurrent artificial neural networks. The system predicts the following output by storing the result of the previous layer's prediction and feeding this result to the current input. So recurrent neural networks may recall their past calculations and

apply that knowledge to the data they're currently analyzing. This one small step forever altered the landscape of chatbots. To depict an RNN, many copies of a neural network are created, and their combined outputs are fed into the input of a next network [27]. Traditional neural networks lack this capacity for looking back at data. Recurrent neural networks are excellent for processing sequential input because of their capacity to remember prior computations. RNNs can replicate natural language's inherent sequential nature, in which words acquire semantic meaning concerning their immediate context. Therefore, the RNN can remember the context and base its output on the words that came before the statement. Because of this, recurrent neural networks are ideally suited for usage in chatbots since they can learn from previous conversations and utilize this knowledge to comprehend user input better and provide relevant responses [28]. Chatbots are more likely to produce contextually accurate responses when they have access to information from previous talks.

7:2 Seq2Seq Neural Networks are a type of Deep Learning that directly model the relationship between two sequences. Two layers of a recurrent neural network (RNN) are used to create the Seq2Seq model. Two recurrent neural networks—an encoder that takes in data and an output decoder—form the Seq2Seq model based on RNN architecture [28]. The input set length to the model can vary by the encoder or the decoder. It was first proposed in 2014 and is a refinement of Ritter's generative model that takes advantage of current advances in deep learning to improve accuracy. Encoders take an input sentence and encode it, whereas decoders take an encoded sentence and decode it to get the intended result. The primary application of this approach is in translation (including statistical machine translation), where a sentence written in one language serves as input for a translation into another language. Chatbots can use this paradigm to map user input to a predetermined response [29]. To produce answers, nothing beats the Sequence-to-Sequence (Seq2Seq) approach. [29].

7:3 Long Short-Term Memory Networks (LSTMs). One subset of recurrent neural networks is called an extended short-term memory network (LSTM) [8]. Long short-term memories (LSTMs) are intended to solve the issue of reliance on RNNs in the long run. Using memory cells and gates, LSTM can teach cells to retain information for a considerable time. Similarly to how computer memory works, these memory cells can be read from, written to, and accessed by other devices. Input, memory, and output gates are used to manage the flow of data [30]. When compared to conventional RNNs, LSTM networks excel at learning from past examples. Since then, LSTM networks have become the de facto standard, surpassing RNNs in this application. Furthermore, when classifying, processing, and predicting time series with gaps of uncertain magnitude between key occurrences over extended periods, well-trained LSTM networks perform exceptionally well. These characteristics show LSTM's superiority to other sequence learning approaches, such as RNNs, Hidden Markov Models, and others. That's why LSTMs are such a boon to chatbot development. LSTM can make frequent allusions to information far removed in time [8, 30].

3 TASK-ORIENTED CHATBOTS

A task-oriented chatbot's primary function is facilitating the completion of defined user tasks. As an example, they are made to deal with situations like Booking flights and hotels, reserving lodgings, placing goods orders, organizing events, and facilitating access to specific data are all examples. Voice-based tasks include personal assistants like Cortana, Alexa, and Siri. Goal-

driven Conversational Agents/ Chatbots that are eager to complete the assignments given to them. Chatbots that are task-oriented excel in specific settings. These chatbots aren't trivia experts and can't answer your queries on current events. As an alternative, some chatbots are purpose-built to assist you in meeting a particular objective.

4 NON-TASK-ORIENTED CHATBOTS

Non-task-focused chatbots are computer programs for in-depth discussions instead of short-term interactions like booking a vacation. They are designed to simulate the casual nature of in-person conversations, or "chats." A lot of the time, these gadgets are fun to use. Architectures for chatbots that aren't task-oriented come in two forms:

-Sequence-to-sequence models use a generative approach to generate natural speech at the appropriate moments in a discussion.

-Develop the skill of retrieving relevant conversational responses from a stockpile on the go. Answers that never appeared in the corpus can be generated with greater accuracy using generative models. Data from human conversations are mined using corpus-based technologies. Machine translation paradigms, such as Using B. A neural network sequence-to-sequence system for learning to link user utterances with system replies can be used for this purpose as information retrieval (IR-based systems essentially duplicate human responses from past talks). With a database of response selection algorithms, search-based models may quickly and accurately choose the most pertinent one for any conversation [31]. Many different heuristics can be used to help choose the best course of action. This can be accomplished with anything as straightforward as rule-based expression matching or sophisticated as a suite of machine learning classifiers. Instead of coming up with a whole new solution, a system based on this idea would choose one from a set of possible ones. The rule-based system can be traced back to the pioneering and influential ELIZA and PARRY systems.

5 DIALOGUE CONTEXT IN TASK-ORIENTED CHATBOTS

As was mentioned before, a task-oriented chatbot's primary objective is to carry out a designated function. B. Place a pizza order, organize a gathering or troubleshoot your gadget from the inside [32]. The purpose of this chatbot is to make using the device easier for the user. In this scenario, the interaction context is a text chat or one using natural language. Different algorithms are used by chatbots depending on the scenario. Pattern-matching frameworks are commonly used for textual discussions. The ability to recognize patterns in the text is built into most chatbots. That information is used to formulate a reply to the user. Many users can generate the same work in different ways [33]. In this situation, the chatbot's ability to comprehend the conversational context is critical to the mission's success. For instance, a user may interpret "I would like a large vegetarian pizza, please" and "I would like a large veggie pizza, please" as having the same meaning. The response may fall short of users' expectations, however, if it is handled by a generic chatbot lacking pattern-matching skills [34]. Bots with pattern-matching abilities can learn from their interactions with the world and adapt their answers accordingly.

6 DIALOGUE CONTEXT IN NON-TASK-ORIENTED CHATBOTS

Chatbots that aren't task-oriented aren't designed to carry out specific actions. And yet, these chatbots have many uses and are helpful in many contexts. B. Strengthening the understanding of those learning a second language while keeping in touch with seniors. Each application

might not have to keep tabs on the same activities as task-oriented chatbots. Instead, it needs to interact with the user to keep the conversation going [35].

Therefore, task-oriented chatbots are more accessible to create algorithms than non-task-oriented ones. These chatbots are sometimes used in tandem with task-oriented chatbots to give users the best of both worlds and make the transition as painless as possible. Due to the dialogue context being entirely arbitrary in this scenario, it is feasible that people will avoid taking part in talks that are directly on the topic [36]. Instead, users can inquire about the weather, horoscopes, and current events, among other topics, with the chatbot. Chatbots need access to data stored in the cloud to have meaningful conversations with humans.

CONCLUSION

In this survey, we dove into the famous taxonomy of chatbots and outlined the many design approaches employed when creating these automated conversational assistants. To sum up, we can say that natural language has become a significant input for modern chatbots and represents a modernization of the traditional connection between humans and computer systems. In recent years, natural language processing has emerged as a promising personalization tool due to its flexibility in accommodating users' unique linguistic styles. Can engage with Chatbots are developed with various tactics and algorithms to process such natural language inquiries and provide valuable responses to users. So far, chatbots have been primarily classified as either task-oriented or non-task-oriented in the existing literature on the subject. Chatbots that aren't task-oriented can be used in various contexts; task-oriented bots, on the other hand, are designed to perform a specific action in response to user input. As a result, these chatbots have various conversational contexts. A chatbot designed to do a specific task, for instance, may be able to process a group of related questions using elementary algorithms based on pattern matching. On the other hand, task-agnostic chatbots must use many algorithms and strategies to respond correctly to a wide range of user queries. Compared to demand-driven chatbots, generative ones are more difficult to create and keep up and running.

REFERENCES

1. Zadrozny, W., et al.: Natural language dialogue for personalized interaction. Commun. ACM 43(8), 116–120 (2000)
2. Hussain, S., Athula, G.: Extending a conventional chatbot knowledge base to external knowledge source and introducing user based sessions for diabetes education. In: 2018 32nd International Conference on Advanced Information Networking and Applications Work- shops (WAINA) 2018. IEEE
3. Levesque, H.J.: Common sense, the turing test, and the quest for real AI: Reflections on Natural and Artificial Intelligence. MIT Press, Cambridge (2017)
4. Wikipedia: Chatbot. 29 December 2018 (cited 30 December 2018). https://en.wikipedia.org/wiki/Chatbot
5. Weizenbaum, J.: ELIZA–a computer program for the study of natural language communication between man and machine. Commun. ACM 26(1), 23–28 (1983)
6. Wikipedia. Artificial Linguistic Internet Computer Entity. 19 November 2018 (cited 30 December2018). https://en.wikipedia.org/wiki/Artificial_Linguistic_Internet_Computer_Entity
7. Wallace, R.S.: The anatomy of ALICE. In: Parsing the Turing Test, pp. 181–210. Springer, Dordrecht (2009)

8. Ramesh, K., et al.: A Survey of Design Techniques for Conversational Agents. In: International Conference on Information, Communication and Computing Technology. Springer, Singapore (2017)
9. Budulan, S.: Chatbot Categories and Their Limitations (2018). (cited May 2018). https://dzone.com/articles/chatbots-categories-and-their-limitations-1
10. Chen, H., et al.: A Survey on Dialogue Systems: Recent Advances and New Frontiers (2017). arXiv preprint arXiv:1711.01731
11. Yan, Z., et al.: Building Task-Oriented Dialogue Systems for Online Shopping. In: AAAI (2017)
12. Masche, J., Le, N.-T.: A Review of Technologies for Conversational Systems. In: International Conference on Computer Science, Applied Mathematics and Applications. Springer, Cham (2017)
13. Abdul-Kader, S.A., Woods, J.: Survey on chatbot design techniques in speech conversation systems. Int. J. Adv. Comput. Sci. Appl. 6(7), 72–80 (2015)
14. Mathur, V., Singh, A.: The Rapidly Changing Landscape of Conversational Agents (2018). arXiv preprint arXiv:1803.08419
15. Bradeško, L., Mladenić, D.: A survey of chatbot systems through a loebner prize competition. In: Proceedings of Slovenian Language Technologies Society Eighth Confer- ence of Language Technologies (2012)
16. Weizenbaum, J.: ELIZA—a computer program for the study of natural language communication between man and machine. Commun. ACM 9(1), 36–45 (1966)
17. Clarke, D.: Three AI Technologies that could make chatbots intelligent. 24 March 2018 (cited 18 December 2018). https://hackernoon.com/three-ai-technologies-that-could-make-chatbots-intelligent-10f8c6e8b4b0
18. Shawar, B.A., Atwell, E.: Chatbots: are they really useful? Zeitschrift für Computer linguistik und Sprach technologie, p. 29 (2007)
19. Wikipedia contributors. ChatScript. Wikipedia, The Free Encyclopedia. 4 September 2018, 19:19 UTC. https://en.wikipedia.org/w/index.php?title=ChatScript&oldid=858055386. Accessed 22 Jan 2019
20. Robino, G.: How to build your first chatbot using Chatscript (2018). (cited 18 December2018). https://medium.freecodecamp.org/chatscript-for-beginners-chatbots-developers-c58b b591da8
21. Milward, D., Beveridge, M.: Ontology-based dialogue systems. In: Proceedings of 3rd Workshop on Knowledge and reasoning in practical dialogue systems (IJCAI 2003) 2003
22. Al-Zubaide, H., Issa, A.A.: Ontbot: Ontology based chatbot. In: 2011 Fourth International Symposium on Innovation in Information & Communication Technology (ISIICT). IEEE (2011)
23. Nuez Ezquerra, A.: Implementing ChatBots using Neural Machine Translation Techniques (2018). Universitat Politècnica de Catalunya
24. Csáky, R.: Deep Learning Based Chatbot Models (2017). https://doi.org/10.13140/rg.2.2.21857.40801
25. Pelk, H.: Machine learning, neural networks and algorithms (2016). (cited 18 January 2019). https://itnext.io/machine-learning-neural-networks-and-algorithms-4cd27a105b89
26. Wikipedia contributors. Artificial neural network. In: Wikipedia, The Free Encyclopedia.

https://en.wikipedia.org/w/index.php?title=Artificial_neural_network&oldid=878429976. Accessed 13:24, 25 Jan 2019
27. Young, T., et al.: Recent trends in deep learning based natural language processing. IEEE Comput. Intell. Mag. 13(3), 55–75 (2018)
28. Cahn, J.: CHATBOT: Architecture, design, and development. In: University of Pennsylvania School of Engineering and Applied Science Department of Computer and Information Science (2017)
29. Cho, K., et al.: Learning phrase representations using RNN encoder-decoder for statistical machine translation (2014). arXiv preprint arXiv:1406.1078
30. Sak, H., Senior, A., Beaufays, F.: Long short-term memory recurrent neural network architectures for large scale acoustic modeling. In: Fifteenth Annual Conference of the International Speech Communication Association (2014)
31. Ji, Z., et al.: An information retrieval approach to short text conversation (2014). arXiv preprint. arXiv:1408.6988.1
32. Duijst, D., Sandberg, J., Buzzo, D.: Can we Improve the User Experience of Chatbots with Personalization? (2017)
33. Liu, B., Xu, Z., Sun, C., Wang, B., Wang, X., Wong, D.F., Zhang, M.: Content-oriented user modeling for personalized response ranking in Chatbots. IEEE/ACM Trans. Audio Speech Lang. Process. (2018)
34. Pilato, G., Augello, A., Gaglio, S.: A modular architecture for adaptive ChatBots. In: Proceedings of 5th IEEE International Conference on Semantic Computing, ICSC 2011 (2011)
35. Vtyurina, A., Savenkov, D., Agichtein, E., Clarke, C.L.A.: Exploring conversational search with humans, assistants, and wizards. In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems – CHI EA 2017 (2017)
36. Niculescu, A.I., Banchs, R.E.: Strategies to cope with errors in human-machine spoken interactions: using chatbots as back-off mechanism for task-oriented dialogues. In: Errors by Humans Mach. multimedia, multimodal Multiling, Data Process (2015).