GRAPH-BASED SELF-ADAPTIVE CONVERSATIONAL AGENT WITH CONTEXT-AWARENESS BEHAVIOUR PREDICTIONS

A TECHNICAL REPORT SUBMITTED TO AUCKLAND UNIVERSITY OF TECHNOLOGY
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF COMPUTER AND INFORMATION SCIENCES

Supervisor

Dr. Weihua Li

Prof. Edmund Lai

A/Prof. Quan Bai

June 2021

By

Lan Zhang

School of Engineering, Computer and Mathematical Sciences

Abstract

Conversational agents have been widely adopted in dialogue systems for various business purposes. Many existing conversational agents are rule-based and require significant human intervention to adapt the knowledge and conversational flow. In this thesis, I propose a graph-based adaptive conversational agent model which is capable of learning knowledge from human beings and adapting the knowledge-base according to user-agent interactions. Extensive experiments have been conducted to evaluate the proposed model by comparing the responses from the proposed adaptive agent model and a conventional agent.

The user's personalised knowledge graph is generated through the user-agent interactions. To initiate conversations, it is important to develop a model for learning the users' preference and giving reasonable suggestions during the conversations. Therefore, based on the personalised knowledge collected by the conversational agent, I propose a novel graph-based context-aware user behaviour prediction and recommendation system. The user's preferences have been considered as an important factor in determining the recommendation outputs. On top of that, the model incorporates contextual information extracted from both users' historical behaviours and events relations, where the contexts have been modelled as knowledge graphs. By leveraging the advantages offered from the knowledge graph, events dependencies and their subtle relations can be established and have been introduced into the recommendation process. Experimental results explicitly indicate that the proposed approach can outperform state-of-the-art

algorithms and yield better recommendation outcomes.

Contents

List of Tables

List of Figures

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the qualification of any other degree or diploma of a university or other institution of higher learning.

Signature of student

Publications

Zhang, L., Li, W., Bai, Q., & Lai, E. (2021, May). Graph-based Self-Adaptive Conversational Agent. In Proceedings of the 20th International Conference on Autonomous Agents and MultiAgent Systems (pp. 1791-1793) (AAMAS)
(Video demonstration access link: https://youtu.be/DbcBdLXiHZg)

Zhang, L., Li, X., Li, W., Zhou, H., & Bai, Q. (2021, May). Context-Aware Recommendation System using Graph-based Behaviours Analysis. Journal of Systems Science and Systems Engineering. Doi: https://doi.org/10.1007/s11518-021-5499-z

Acknowledgements

My Master's thesis project turns out to be a challenge to me. It varies from previous teaching courses in that it has much tighter standards for critical thoughts and problem-solving skills. Furthermore, since research is an iterative process, patience and consistency are often needed when coping with persistent failures. It is difficult for me to complete my master project without sincere support from many people.

Firstly, I want to thank Dr Weihua Li, my primary supervisor. With a high level of enthusiasm, inspiration, and patience, he continues to assist me with my research. He encourages me to expand my mind and helps me pursue my research path through literature review. Also, he gave me the opportunity and encouraged me to explain my project accurately in front of people in the non-Computer-Science field. I remember that he often reminded me that it is always important to explain and demonstrate my research project. Secondly, I would like to thank my second supervisor, Prof. Edmund Lai, for his help in my research process. Also, my external supervisor, A/Prof. Quan Bai, provides me with constant inspirations. His professional knowledge in this area benefits my study to a large extent. He also offers valuable advice and encouragement while I am working on my thesis. During the conversations with him, he can always point out the deficiencies of my research and give suggestions for the amendments.

Finally, I would like to express my gratitude to all members of the Department of Computer Sciences at Auckland University of Technology for providing facilities and resources that have helped make this thesis a success.

Intellectual Property Rights

Chapter 1

Introduction

With the exponential growth of data and mobile devices over the last decade, Conversational Agents (CAs) have become an increasingly popular option to interact with the end-users. The popularity and adoption of CAs are rapidly spreading across many sectors, e.g., banking, insurance, and tourism. However, most CAs fail to satisfy the users' needs due to unclear purposes, nonsensical responses, or insufficient usability. Moreover, traditional CAs are not equipped with the capability of improving themselves via the interactions with end-users. They are supposed to be enhanced through the calibrations conducted by the domain experts. The traditional CAs turn out to be too inflexible due to the limited dialogue capabilities. To address this challenge, I proposed a novel model, called Graph-based Self-adaptive Conversational Agent (GSCA), which is capable of improving itself through the user-agent interactions. Meanwhile, the GSCA generates user's personalised Knowledge Graphs (KGs) according to user's input. By analysing the personalised KGs, the system can collect user preferences since it is recognised as one of features used by the Recommendation Systems (RSs). However, many RSs only consider the user's context, e.g., user's location, but ignore both user preferences and the context of historical behaviours. The results of such RSs will be relatively general, lacking personalised recommendations for different users and thus unable to satisfy the needs. To address this issue, I proposed the Graph-based Context-Aware Recommendation (GCAR) algorithm.

The rest of this chapter is organised as follows. Subsection 1.1.1 introduces the background of the GSCA, meanwhile Subsection 1.1.2 briefly introduces the background of GCAR. Research motivations are presented in Subsection 1.2. Subsection 1.3 illustrates the research questions raised in this thesis. In Subsection 1.4, the study design is described, which includes research methodology and evaluation method. The main contributions are outlined in Subsection 1.5. Lastly, the organisation of the entire thesis is structured in Subsection 1.6.

1.1 Background

1.1.1 The Graph-based Self-adaptive Conversational Agent

Nowadays, the CAs have been widely adopted by many businesses for a wide range of applications, including intelligent tutoring for improving learning and teaching, chatbots offering 24/7 supporting services and personalised conversation agents in health care (Kerlyl, Hall & Bull, 2006; Gnewuch, Morana & Maedche, 2017; Laranjo et al., 2018; Ranjbartabar, Richards, Bilgin, Kutay & Mascarenhas, 2020; Mendonça, Melo, Coheur & Sardinha, 2017; Battineni, Chintalapudi & Amenta, 2020).

Most existing intelligent CAs can be categorised as goal-driven and non-goal-driven agents (Logacheva, Malykh, Litinsky & Burtsev, 2020; Yan, 2018; Zhu et al., 2020). In the former, with a predefined service target, CAs participate in a Question-and-Answer (Q/A) system by providing domain-specific services to address the problems in a specific sector. In other words, CAs are equipped with a set of fixed rules involved in a particular field's services, such as customer support, ticket booking, etc. While, in the latter, the service target and conversational scope are not predefined. Non-goal-driven

CAs are capable of handling a variety of problems by leveraging various information and ontology in the universe. Such CAs are normally adopted in an open field, such as entertainment.

However, the advent limitations of these agents are presented. First, the knowledge curation of CAs relies on manual inputs from domain experts. CAs themselves have very limited capabilities of retaining and recalling knowledge obtained from the conversations (Chakrabarti & Luger, 2015). Second, the current state-of-the-art approaches for designing agents are based on sequence-to-sequence variants, enabling the automatically adaptive skill of the CAs. Whereas, such approaches generally suffer from an inability to bring memory and knowledge fail to consider the time-series context (Sutskever, Vinyals & Le, 2014; Vinyals & Le, 2015; Serban, Lowe, Charlin & Pineau, 2016). Third, CAs' responses may appear out of control when the models are black-box. Specifically, for those agents trained using neural networks, the model turns out to be untransparent, and it is a non-trivial task to revise the CAs' knowledge base without re-training. Furthermore, the GSCA framework will capture the personalised KG for each user, and obtain each user's personal preferences by analysing the KG.

Based on the user's personalised information captured by the GSCA, I further design a recommendation mechanism, allowing the GSCA to initiate conversations and give proper recommendations to the. The details are elaborated in the next subsection.

1.1.2 Context-Aware Recommendation System using Graph-based Behaviours Analysis

With the rapid development of modern technologies, massive volumes of information have been generated every day. Given such big data with numerous choices, it is difficult for users to identify useful information that they need and make decisions proactively. Therefore, on the basis of the GSCA framework, I further proposed the GCAR algorithm, which takes the user's personalised preferences into account. RSs play a significant role in assisting users in discovering interesting information, facilitating the process of decision making and leading to the development of many real-world applications (J. Lu, Wu, Mao, Wang & Zhang, 2015), such as music recommendation (Hu, Shi, Zhao & Yu, 2018), movie recommendation (C. Wang, Zhu, Zhu, Qin & Xiong, 2020; Zhuang et al., 2017) and online shopping (Han et al., 2019; Xu et al., 2018).

Traditional RSs usually require the specification of both users and items. Specifically, users' attributes, e.g., demographic information and historical behaviours, e.g., user-item ratings and event adoptions, become the key consideration for the recommendation. Collaborative filtering (Breese, Heckerman & Kadie, 2013) and content-based filtering (Lops, De Gemmis & Semeraro, 2011) have been recognised as two fundamental recommendation algorithms. The former makes recommendation according to choices made by other people with similar tastes and preferences derived from historical purchasing records (Sarwar, Karypis, Konstan & Riedl, 2001). The latter is developed based on the fact that users usually select similar items repeatedly (Pazzani, 1999). However, a user's contextual information, such as time, location, weather, and previous behaviours, plays an essential role in decision-making. For example, the facility should not be too far away from the user's location, and indoor activities should be recommended if it is raining. Such context is neglected in these traditional recommendation algorithms. Many researchers have studied Context-Aware recommendation Systems (CARS) to mitigate this limitation, generating more reasonable and intelligent recommendations by considering the user's contextual situation (Haruna et al., 2017). Whereas the context of an event (a.k.a. activity or item in the current setting) at the semantic level, reflecting the relations among the activities, influences the user's decision making to a large extent.

First, the event context implicitly extends a users' preference and update the user's profile promptly. For example, people who frequently attempted Chinese food might also be interested in Chinese culture, where the food and culture can be implicitly associated through the entity “Chinese”. People who worry about “food shortage” due to pandemics might also be concerned with “face-masks” since the relations between these two terms can be established via “pandemics”. Second, the event context can estimate the conditional probabilities for event sequences, even with insufficient user transactions. Through the contextual features, events and the corresponding relations implicitly get connected, forming a network. By utilising the network topology, the conditional probabilities of event occurrence can be estimated without user transactions. By adopting the previous example the probability of watching a Chinese show after the user has the Chinese food can be estimated since both are associated via “Chinese”. Therefore, the conditional probability estimation still works without user transactions. By considering the aforementioned factors, the KG leverages connected data to understand concepts and infer meanings, which is suitable for semantic representation and reasoning over numerous entities (Q. Wang, Mao, Wang & Guo, 2017).

1.2 Research Motivations

There are many research motivations to explore Knowledge Graph-based Conversational Agent and Context-aware Recommendation system by analysing user’s behaviours. In recent years, CAs are widely used in different industries to assist people in completing some specific tasks, such as mental health consultant, medical advice and customer services. Therefore, many researchers have been attempting to make a breakthrough in the emotional intelligence of CAs.

In the contemporary research fields, there are many studies describing the adaptive conversational agents and context-aware recommender systems. The primary research gaps are identified as follows:

- The performance of traditional CAs are proved to be more efficient. However, there are still some limitations to be addressed. Firstly, the expansion of the knowledge base relies on manual addition by domain experts. In other words, they are not capable of bringing memory to retain or recalling the knowledge. Many agents are unable to conduct a broader dialogue, understand the dialogue or check whether the conversation flow is developing in the desired direction (Chakrabarti & Luger, 2015). The current state-of-the-art methods, namely various sequence-to-sequence models, attempt to enable the automatically adaptive skills of the CAs, but generally suffer from an inability to bring memory and knowledge to bear (Sutskever et al., 2014; Vinyals & Le, 2015; Serban et al., 2016). Finally, most of the existing studies have not achieved the self-improvement of CAs during the interactions between users and agents. Therefore, it is difficult to analyse the personalised conversation flows and identify the user's preference.

- The user's contextual information, e.g., location, weather and past behaviours, plays a vital role in decision-making. However, for the most traditional RSs, these are usually ignored. Many researchers realise the importance of contextual information and use this point to launch many new RSs, but there are still some shortcomings. They ignore the contextual information of the event, which is also crucial in the user's decision making. On one hand, historical events play a significant role in reflecting users' preference. On the other hand, the relationship between two consecutive events also has a significant impact on the RSs.

1.3 Research Question

Based on the literature review and research motivations explained in previous subsections, the research questions are listed as follows.

Research Question 1: How to enable a CA to improve itself through the interactions with end users?

- **Sub-Research Question** How to develop a CA without a pre-trained knowledge base but is capable of adapting the knowledge obtained from the users?

Research Question 2: How to develop a model for CAs to predict user's personalised behaviours or provide event recommendations?

- **Sub-Research Question** How to enable the CAs to learn user's preference and model the relationship among the events?

1.4 Design of study

In this subsection, I will explain the research methodology adopted in this thesis. Next, the proposed approach, i.e., Graph-based Conversational Agent and User Behaviour Analysis, leveraged in this research is introduced. Finally, the evaluation methods of the proposed approach are elaborated.

1.4.1 Research Methodology

The research methodology employed in my thesis demonstrates an interactive process, which is illustrated in Figure 1.1. To explain this further, firstly, I conduct literature review, identifying the research gaps from the existing studies. Then, based on the identified research gaps, the research questions are proposed. The models are then developed to address these research issues. Next, I collect public datasets for model evaluation. On the other hand, the simulated dataset is also adopted. For example, in the proposed GCAR, Monte Carlo Simulations (Mooney, 1997) with the independent cascade model is adopted to simulate the relationship between user's behaviours and

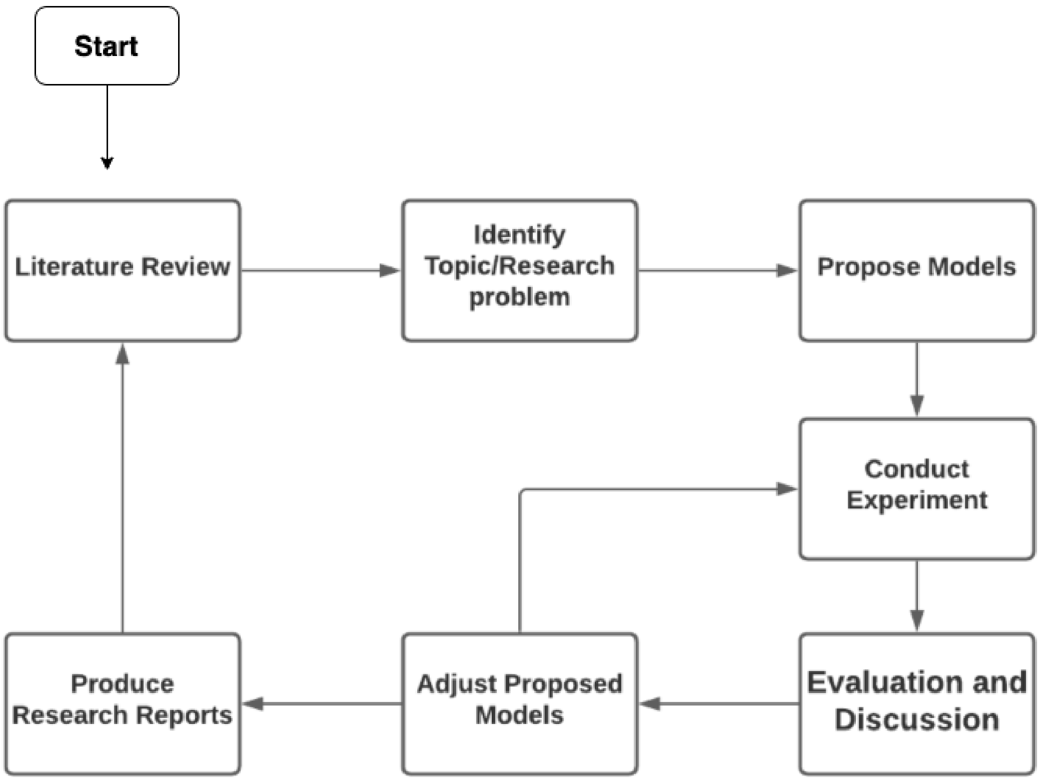

Figure 1.1: Research Methodology

the context information related. Then, I observe the experimental results, investigate the insights, and evaluate the effectiveness and efficiency of the model. If the performance is not promising, the model will be fine-tuned and the parameters will be revised. I will re-conduct the experiments with the updated parameters. After the model reaching a satisfiable performance, I will analyse the results and work on the report.

1.4.2 Evaluation Method

This thesis proposes two different models with novel algorithms. Therefore, I employ different evaluation methods.

Firstly, I propose a novel algorithm, named the Temporal-based Triple Retrieval (TTR), to effectively identify the most relevant triples with a confidence score, where both temporal features and user preferences are considered. The approach to measure

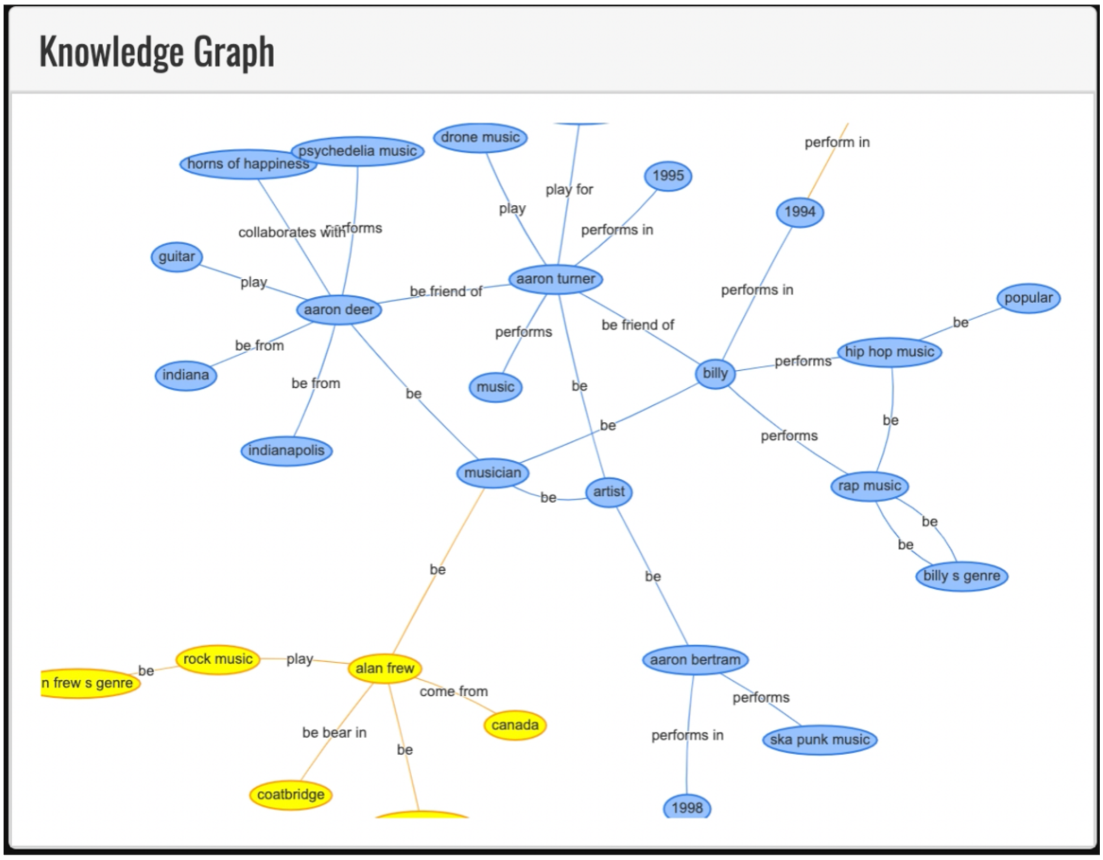

Figure 1.2: Knowledge Graph Visualization for TTR Algorithm

its performance is mainly through checking the rationality of the natural language outputs. Therefore, in Figure 1.2, the KG visualization has been employed to observe the performance of the proposed TTR algorithm. Namely, based on the pre-designed rules, the Triples with the relevant topic are supposed to be closer to each other.

Secondly, I propose the GCAR algorithm, where three traditional evaluation metrics serving the NER tasks are employed, i.e., Precision, Recall and F1-score (Goutte & Gaussier, 2005). The higher value proved the more efficient performance of the model. Four possible outcomes are presented below.

1) An event is recognised as “shopping" when the user truly chooses "shopping" as his/her next event (True Positive, TP)

2) An event is recognised as “shopping" when it is not the user’s next event (False

Positive, FP)

3) An event is recognised as not going shopping when the user truly chooses “shopping" as his/her next event (False Negative, FN)

4) An event is recognised as not going shopping and the user truly does not choose “shopping" as his/her next event (True Negative, TN)

Based on the above four possible results, Precision, Recall, and F1-score are defined as follows.

- **Recall** quantifies the number of positive class predictions made out of all the positive examples in the dataset:

$$Recall = \frac{TP}{TP + FN} \tag{1.1}$$

- **Precision** quantifies the number of positive class predictions that actually belong to the positive class:

$$Precision = \frac{TP}{TP + FP} \tag{1.2}$$

- **F1-score** provides a single score that balances both the concerns of precision and recall in one number:

$$F1 = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{1.3}$$

The evaluation metrics used in this thesis are not limited to the evaluation metrics discussed above. The needs of the expanded models and problems are used to determine new evaluation metrics. The details are introduced in each chapter’s experiment section.

1.5 Thesis Contribution

In this thesis, I have proposed two novel models based on KGs, i.e., GSCA and GCAR. One objective is to address the challenging issues of the traditional CAs. The second objective is to develop a novel framework allowing agents to make recommendations based on user's historical behaviours and the corresponding implicit associations. The thesis contributions are summarised as below.

1) I proposed GSCA framework to enable the CA to improve and adapt knowledge through the interactions with end users. The CAs are capable of automatically selecting the most relevant triples as the input for the response generation algorithm. On top of that, I proposed the Temporal-based Triple Retrieval (TTR) to effectively identify the most pertinent triples with a confidence score, where both temporal features and user preferences are considered.

   The models and relevant results are published in the 20th International Conference on Autonomous Agents and Multiagent Systems (AAMAS) (L. Zhang, Li, Bai & Lai, 2021).

2) I further improved the GSCA by proposing a novel recommendation algorithm, called GCAR, enabling the GSCA to learn the user preferences from the events and contexts. The KGs are utilised to model the relationships among the users, events and event contexts.

   The models and relevant results are published in the Journal of Systems Science and Systems Engineering (AISSM).

1.6 Thesis Organisation

The structure of the rest thesis is listed as follows:

- **Chapter 2** reviews relevant state-of-the-art research works regarding CAs and CRSs and the novel techniques adapted by those research works.
- **Chapter 3** introduces the proposed GSCA framework, which enables the CA to improve itself through the interactions with users. Besides selecting the most relevant triple as the input of the response sentence generation algorithm, I proposed the Temporal-based Triple Retrieval (TTR) algorithm to effectively identify the most relevant triples with a confidence score, where both temporal features and user preferences are taken into consideration. It addresses the Research Question 1.
- **Chapter 4** introduces the GCAR Algorithm, which considers the contextual information, and analyses the user's preferences to address user behaviour prediction problem. It tackles Research Question 2.
- **Chapter 5** concludes the thesis by summarising the advantages and limitations of the proposed method. Meanwhile, an overview of the future work is given.

Chapter 2

Literature Review

Since the thesis involves CAs and RSs with KGs, the the literature review of this thesis will cover these three major aspects. Different types of CAs, RS models and their variants have been reviewed and compared. Specifically, I explore the contemporary research works associated with CAs and the applications of self-adaptive CAs. Next, the studies associated with KG and triple-to-text generation have been reviewed because they form the technical foundation for my research work. Then, the existing research works of context-aware recommendations and user behaviour prediction have been reviewed. Finally, I summarise the findings and provide a brief overview of the research gaps.

2.1 Conversational Agents

2.1.1 Traditional Conversational Agents

In the last few years, many Intelligent Personal Assistants (IPAs) have been developed, such as Apple's Siri and Microsoft's XiaoIce, which have been applied to different fields. Specifically, Apple's Siri mainly provides users with personalised assistance on cell phones. XiaoIce, i.e., a social chatbot, also can be considered as a virtual assistant, and its function is limited to simple Q & A and images commenting (Shum, He & Li, 2018). Many researchers have dedicated great efforts in developing intelligent chatbots. For example, Zhou et al. suggest employing Deep Attention Matching Network to build the Multi-Turn Response Selection chatbots, where the idea is inspired by the recently proposed Transformer in machine translation (Zhou et al., 2018). Vaswani et al. utilise attention to solve sequence-to-sequence generation problems (Vaswani et al., 2017). The authors extend the key attention mechanism of Transformer in two ways, i.e., self-attention and cross-attention. The model is proved to be more efficient than that of state-of-the-art models. Kerlyl et al. bring the chatbot techniques to the education field and develop intelligent tutoring systems, aiming at exploring the feasibilities of using chatbots to support negotiations and transform negotiations (Kerlyl et al., 2006). Feine et al. illustrate the development of a proof-of-concept prototype, inventing an intelligent dialogue component developed using the chatbots technology (Feine, Morana & Maedche, 2019). To achieve accurate and intelligent Q & A, the authors propose to store different knowledge and technologies in various modules and use corpus callosum to start other modules by recognising specific features. Radziwill et al. evaluate the quality of existing chatbots and intelligent conversational agents (Radziwill & Benton, 2017).

Traditional chatbots system relies on a fixed set of rules. Whereas, modern chatbots are mainly developed based machine learning approaches, including supervised learning (Cunningham, Cord & Delany, 2008) , unsupervised learning (Barlow, 1989) and hybrid intelligence methods. These methods are costly but efficient in problem-solving. For example, Feine et al. propose a novel approach to designing socially adaptive chatbots where a two-step method is proposed to convert descriptive Ω-knowledge into a machine-executable representation, which makes the existing Ω-knowledge more suitable for Design Science Research (DSR) projects (Feine et al., 2019). The authors also provide theoretical guidance for the chatbots design process and support establishing a more comprehensive knowledge-based system, e.g., E-learning and customer services.

Chatbots require the a wide range of modern techniques, including built-in Artificial Intelligence (AI), Nature Language Processing (NLP), programming and conversation services. There are plenty of cloud-based chatbots services available for developers to use. However, the challenge of chatbots programming still exists. For example, Rahman et al. claim one of the challenges is handling the NLP issue, as well as the machine learning techniques. Furthermore, the training dataset turns out to be a big issue due to its difficulties in matching the intent by considering context (Rahman, Al Mamun & Islam, 2017). Abdul-Kader et al. conduct a survey based on the speech search engine chatbot design techniques. In this survey, authors compare the different technologies used in the chatbots in 9 research works with those used in the Loebner-Prize chatbots. The research results prove that chatbot design technology has not yet found a universal method. In addition, general chatbots need to be improved by designing a more comprehensive knowledge base (Abdul-Kader & Woods, 2015). Divya et al. discuss the Self-Diagnosis Medical Chatbot, where the technology is designed to reduce medical cost. The chatbots act as a medical reference book and help patients quickly and better understand their conditions. After interactions with the chatbots, the users will obtain information about specific diseases and previous chat histories stored in the database (Divya, Indumathi, Ishwarya, Priyasankari & Devi, 2018).

To improve customer satisfaction, many organisations plan to adopt CAs as a channel that offers rich 24/7 customer service (Gnewuch et al., 2017). Also, the FaceBook ParlAI turns out to be one of the chatbot frameworks which can be applied on messengers and the ultimate goal is to achieve a real dialogue with humans (Miller et al., 2017).

In previous studies, many researchers focus on improving the IQ of chatbots and evaluate chatbots based on the breadth of knowledge. In recent years, chatbots are widely used in different industries for specific tasks, such as mental health consultant, medical advice and customer services. Therefore, researchers attempt to make a breakthrough in the emotional intelligence of chatbots. The performance of traditional CAs is proved to be helpful and efficient in interacting with users. However, some general limitations of such CAs are presented. Firstly, the expansion of the chatbots' knowledge base relies on manual addition by domain experts. In other words, they are not capable of retaining and digesting the knowledge through the interactions. In other words, most CAs fail to "hold a longer conversation, understand the conversation, gauge whether the conversation is going in the desired direction, and act on it" (Chakrabarti & Luger, 2015). The current state-of-the-art methods, namely various sequence-to-sequence models, attempt to equip CAs with the capabilities of adapting knowledge, but generally suffer from an inability to bring memory and knowledge (Sutskever et al., 2014; Vinyals & Le, 2015; Serban et al., 2016).

2.1.2 Self-adaptive Conversational Agent

CAs are software programs that are designed to communicate with users through the natural language interaction interfaces (Shawar & Atwell, 2005; Zierau, Wambsganss, Janson, Schöbel & Leimeister, 2020). In recent years, researchers have been exploring adaptive CAs and applying them to various scenarios, e.g. online gaming, tutoring systems (Seering, Luria, Ye, Kaufman & Hammer, 2020; Latham, Crockett, McLean & Edmonds, 2011, 2012; Shawar & Atwell, 2007), and online collaboration support tasks (Tegos, Demetriadis & Karakostas, 2011). MentorChat is recognised as an online collaboration support chatbot (Tegos et al., 2011). It models a specific domain through a set of critical concepts, which allows the instructors to specify the type of agent actions to be used when a key concept is traced in the students' dialogue. This model is capable of automatically enhancing the cognitive model of a single student.

However, it still requires domain knowledge modelling, and re-training is indispensable. Wambsgans et al. propose a novel model based on argumentation mining and prototype a conversational interface on an open platform, i.e., ArgBot, which aims to tackle the challenges of adaptive argumentation support to individuals (Wambsganss, Guggisberg & Soellner, 2021).

2.2 Knowledge Graph

2.2.1 Knowledge Graph Expansion

The KG expansion refers to appending new nodes, relations or subgraphs to the existing KG. To satisfy the users' needs during the user-agent interactions, the CAs are supposed to continuously increase and adapt the information in the knowledge-based by expending the KGs. Yoo et al. propose an auto-growing KG called PolarisX (Polaris Expander). To enable self-improvement for PolarisX, the system crawls news sites and social media, extracts new relationships, generates knowledge subgraphs and appends them to the existing KGs (Yoo & Jeong, 2020).

Gawriljuk et al. propose a framework for constructing and incrementally scaling KGs, and the authors describe and evaluate techniques for efficient constructing KGs by generating candidate matches using the MinHash/LSH algorithm (Gawriljuk, Harth, Knoblock & Szekely, 2016). However, the domain-specific knowledge used in their studies are restricted to “artists". Furthermore, the performance of MinHash algorithm turns out to be inefficient when a similar name-value between the candidate and target entity is presented. Grainger et al. introduce a novel system - the Semantic Knowledge Graph (Grainger, AlJadda, Korayem & Smith, 2016). By using this novel model, the relationships between any entities in a document corpus can be dynamically specified and scored. However, the implementation of these functions is limited to processing individual documents as input data. Moreover, this system is not capable of processing new input data in real-time, since the important knowledge graph expansion in adaptive conversational agent fails to be implemented.

2.2.2 Relevance Search

Relevance search takes a query entity as input. It returns top-ranked entities in a KG that are most relevant to the query entity, which has been utilised Web search, RSs and CAs. Current graph querying techniques have developed various kinds of ranking functions. For instance, Pound et al. describe a ranking function that considers both the syntactic similarity and the semantic coherence between queries and matches (Pound, Ilyas & Weddell, 2010). In contrast, a Conditional Random Field (CRF) model is utilised by Structureless Graph Querying (SLQ) to learn and estimate the likelihood of a match applying to a question (Yang, Wu, Sun & Yan, 2014).

One of the key challenges in the field of adaptive CAs is identifying appropriate entities and relations from the large-scale KGs that can be used as responses. In previous studies, sequence-to-sequence generative models have become popular for response generation. However, the responses generated by these sequence-to-sequence models are not always coherent or contextually appropriate, turning out to be too general and lacking interesting content. On top of that, sequence-to-sequence methods require a large amount of training data to ensure reasonable responses. First of all, it is a non-trivial task to obtain sufficient data from the user side. For domain experts, it is also impractical to pre-define meta-paths for all types of information over a schema-rich KG. Researchers have been investigating how to conduct relevance searches in KGs. For example, Gu et al. propose a new approach, called RelSUE, to conduct relevance search in KGs (Gu et al., 2019). RelSUE facilities distance and degree-based heuristics to narrow down the search space since it searches all meaningful original paths, avoiding semantically duplicate entities. However, the model must operate with the predefined meta-path.

KG search strategies are different in various expressive and user-friendly tools. Standardised query languages, e.g., SPARQL, require users to understand the underlying data model fairly well (Pérez, Arenas & Gutierrez, 2009). Unstructured search strategies e.g., keyword search, which is easy to use but ignore the structural constraints in queries. It is challenging to formulate the graph query problem. Su et al. aim to address the graph pattern matching problem and propose a general Graph Relevance Feedback (GRF) framework (Su et al., 2015). The GRF employs the random forest algorithm to select the best candidate KGs and transfers the problem into a binary issue, where a ranking function is adopted to determine whether the all the KGs are supposed to be traversed. However, the GRF framework does not consider the time stamp of each entity in which it appears, making the retrieved data losing the timeliness. Timeliness is recognised as an essential factor of information validity.

2.2.3 Text Generation Task in Conversational Agent

With the continuous updating of chatbots technology, Human-Computer Interaction (HCI) is no longer limited to the present simple ways, but it shifts to natural language-based interfaces (Zadrozny et al., 2000). Specifically, text-based chatbots are triggered by users using natural language. Nowadays, with the increasing popularity of Artificial Intelligence, chatbots are able to answer relatively complex questions raised by users. However, it is difficult for traditional chatbots to adopt adaptive multi-turn and natural language to interact with human beings (Hussain, Sianaki & Ababneh, 2019).

The earliest chatbot program ELIZA was published by Joseph Weizenbaum in 1966. The early chatbot, e.g., ELIZA, is developed using a rule-based model that checks the keywords entered by the user and converts sentences according to the rules associated with the keywords in the pre-defined script (Weizenbaum, 1966, 1983).

A.L.I.C.E (Artificial Linguistic Internet Computer Entity) is a chatbot inspired by the ELIZA program (Wallace, 2018). A.L.I.C.E. is built using Artificial Intelligence Mark-up Language (AIML), which is based on eXtensible Markup Language (XML). AIML-based chatbots are the most popular because they are lightweight and easy to configure. Given a set of predefined responses to the questions, the appropriate context needs to be considered when generating these responses, which leads to constraints on the CAs. Gopalakrishnan et al introduce an open domain topical-chatbot, which covers eight broad topics, and the chatbot does not possess a clearly defined role (Gopalakrishnan et al., 2019). The authors train several latest encoder-decoder dialogue models on topical-chat and perform an automatic manual evaluation for benchmarking test. Also, the extensive research works have been dedicated to the neural models development, using graph encoder-decoder to produce text based on the extracted information (Devlin, Chang, Lee & Toutanova, 2018). Previous studies have proved that compared with the neural models, the template-based methods can achieve better performance on the text generation (Wiseman, Shieber & Rush, 2017). Konstas et al. model the graph structure by adopting linearisation and a sequence encoding (Konstas, Iyer, Yatskar, Choi & Zettlemoyer, 2017). The Graph Attention Network (GAN) and the graph Long Short-Term Memory (LSTM) model have been proved to improve the performance of information propagation (Veličković et al., 2017; Song, Zhang, Wang & Gildea, 2018).

The task model for generating coherent knowledge text needs to meet the following two requirements. Firstly, the model should consider global features of knowledge. Secondly, the local features of each entity should be taken into account. Hence, KGs turn out to be a suitable tool for representing knowledge.

2.2.4 Question Answering task over Knowledge Graph

Question Answering over Knowledge Graph (QAKG) attempts to use information in the KG to answer natural language questions (Huang, Zhang, Li & Li, 2019). It allows end-users to effectively and quickly access the valuable information in the KG. However, it is difficult for QAKG to capture the semantic meanings of natural language. In the meantime, several methods of information graph embedding have been suggested. The main idea is to represent each entity as a low-dimensional vector, so that it is possible to retain the relevant information in the KG.

Typically, unstructured text information will be stored in the form of text files, which turns out to be difficult to obtain the relations between different entities across the entire corpus. KG can intuitively express the relationship between entities. In the QAKG, the questions are expressed in natural language, and the key objective is to retrieve the correct answers presented as a triple or set of triples from the KG. The triples can be extracted by using existing tools and libraries, e.g., Open Information Extract (OpenIE), which can extract structured relational triples from plain text without a pre-specified relational pattern. Specifically, OpenIE first splits the sentences into groups and then shortens each group as much as possible to generate a set of short sentence fragments; these fragments are then divided into semantic triples as the system's output (Stanovsky, Dagan et al., 2015). These semantic triples will be treated as the input of the text generation systems to get a response sentence.

2.3 Recommendation System

Model CAs are capable of memorising users preferences through the interactions, and initiating conversations by providing reasonable recommendations. Therefore, in this section, the recommendation system as part of CAs has been reviewed. This subsection mainly covers two fields, i.e., the context-aware recommendation and KG-based recommendation.

2.3.1 The Context-aware Recommendation Systems

Traditionally, "context" in RSs refers to the contextual information, such as time and location, turning out to be an important feature of the environment. With the assistance of contextual information, RSs can be applied to a number of scenarios, including restaurant recommendation (Tewari, Youll & Maes, 2003; Tung & Soo, 2004), tour guide (Van Setten, Pokraev & Koolwaaij, 2004) and commercial recommendation (Yuan & Tsao, 2003). A series of context-aware recommendation systems have been designed to meet the increasing service demand in people's daily life. For example, Zhang et al. improve the collaborative filtering by taking neighbour users with high similarity into consideration (J. Zhang & Pu, 2007) and Xiao et al. utilise ontologies to make the recommendation promptly (H. Xiao, Zou, Ng & Nigul, 2010). In general, such recommendation systems usually depend on the spatial and temporal information obtained from the merchant, and they only provide universal recommendations to the customers according to the current position and time (Manotumruksa, Macdonald & Ounis, 2018). However, the events and activities' contextual information, reflecting the meaning of user's behaviours, are ignored.

2.3.2 Knowledge Graphs-based Recommendation System

To conduct personalised recommendations, KGs have been widely adopted to model users, items, and relations (Guo et al., 2020; Chaudhari, Azaria & Mitchell, 2017). Wang et al. identify relationships between items and entities, having the item-entity connections learnt to improve the recommendation accuracy (H. Wang et al., 2019). Similarly, Ma et al. propose a novel model to extract rules from item features using

KGs and incorporate these rules into the recommendation process (Ma et al., 2019). Wang et al. develop a KG-based approach, i.e., RippleNet, to make recommendations by considering users' previous behaviours, where the user preference is learnt from the historical data (H. Wang et al., 2018). Wang et al. investigate the utility of KG, breaking down the independent interaction assumption by connecting items with attributes (X. Wang, He, Cao, Liu & Chua, 2019).

A further improvement of KGs-based recommendation is to consider temporal features. Temporal knowledge graphs are utilised to conduct reasoning to infer the relations from users' behaviours (C. Xiao et al., 2019). Similarly, Zhang et al. incorporate spatial-temporal semantic information in KGs to extract knowledge from heterogeneous data (W. Zhang et al., 2018). To measure user preferences, the Session-based Temporal Graph (STG) model has been proposed to differentiate the long-term and short-term preference of users (Xiang et al., 2010). The interactions with the user also have been considered in temporal KGs to reflect the time-series impact on the recommendation accuracy (C. Xiao, Sun & Ji, 2020).

2.4 Summary of Literature Review

In this chapter, I have given a detailed review of the related works in the field of CAs, RSs and KGs. The gaps from the existing studies are summarised as follows.

- Very few studies of CAs focus on adaptive approaches.
- In the existing studies, very few researchers apply KG to model the CAs' knowledge base.
- In the existing studies, CAs usually require enormous data for training and domain experts are supposed to be involved.

The proposed novel GSCA significantly reduces the cost of model training, as well as the manual efforts from domain experts. It is capable of automatically adapting knowledge obtained from user-agent interactions. Furthermore, GSCA utilises KGs to personalised data, e.g., user preferences, and facts. Such information will be utilised in the recommendation process conducted by CAs.

Meanwhile, most existing research works map the users' behaviours to a KG without considering the context of events. In contrast, the GCAR conducts recommendation by considering user behaviours with the event context knowledge graph. There are two advantages of having event context involved. First, with evolution of user behaviours, the corresponding event mapped to the event context knowledge graph keeps revising. In this sense, the recommendation outcome is updated, and RS will carry out accurate and reasonable results. Second, as for the users without sufficient historical behavioural data, the event context knowledge graph can provide the basis for event prediction.

The literature review pointed out the research gaps in the adaptive CAs and context-aware RSs, revealing the need for in-depth research in this field. To address the limitations mentioned previously, in this thesis, I employ multi-layer KGs for information modelling and achieve the CAs adaptiveness. The details of the proposed approaches are introduced in Chapter 3 and Chapter 4.

Chapter 3

The Graph-based Self-adaptive Conversational Agent

3.1 Introduction

The challenges of developing CAs can be concluded as follows. First, the knowledge curation of agents relies on manual inputs from domain experts. The CAs are not capable of collecting knowledge during user-agent interactions. Second, the current state-of-art CAs cannot bring memory and knowledge and fail to consider the time-series context. Third, CAs' responses may appear out of control when the knowledge base of the model is unknown.

In this chapter, I propose a novel conversational agent model named Graph-based Self-adaptive Conversational Agent (GSCA) to address the challenging issues mentioned above. The proposed model enables the agents to learn from the human-agent interactions, continuously enriching the knowledge base. I represent agents' knowledge base as a dynamic and transparent knowledge graph, where the nodes denote key entities and links that describes the semantic relationship. On top of that, to obtain appropriate responses, I develop a temporal-based triple extraction algorithm for GSCA, where

Google T5 (Raffel et al., 2019) has been utilised for text generation.

The proposed model enables the CAs to achieve two innovations. First and foremost, it automatically generates responses according to the user's inputs, which is not wholly dependent on the existing corpus. Furthermore, the novel agent model is capable of enhancing the knowledge base through proactively obtaining the relevant information from external resources, e.g., Wikipedia, and interacting with users. By applying the proposed model, the CA can be customized and augmented in different sectors.

3.2 The GSCA Framework

3.2.1 Semantic triple Collection and Matching

The GSCA framework supports agents to learn and adapt the knowledge through conversations with end-users. Figure 3.1 demonstrates the overall picture of the framework.

A user can communicate with an agent by sending a text. Semantic triples will be extracted using information extraction techniques. Under normal circumstances, sentences will be expressed as triplets according to English grammar rules. Each triple includes three parts; each part can be a single vocabulary or a meaningful phrase. Among them, we named these three parts, head, relation and tail.

$$Triple : F = \{f_1, f_2, ..., f_n\}, n \in N, N \in (1, +\infty) \tag{3.1}$$

which can be represented as a collection of facts, and each fact comprises three entities ε, i.e., the integer N starts from 1, the range of N depends on how many triples generated from this conversation.

$$f_x = (h_x, r_x, t_x), f_x \in F, \tag{3.2}$$

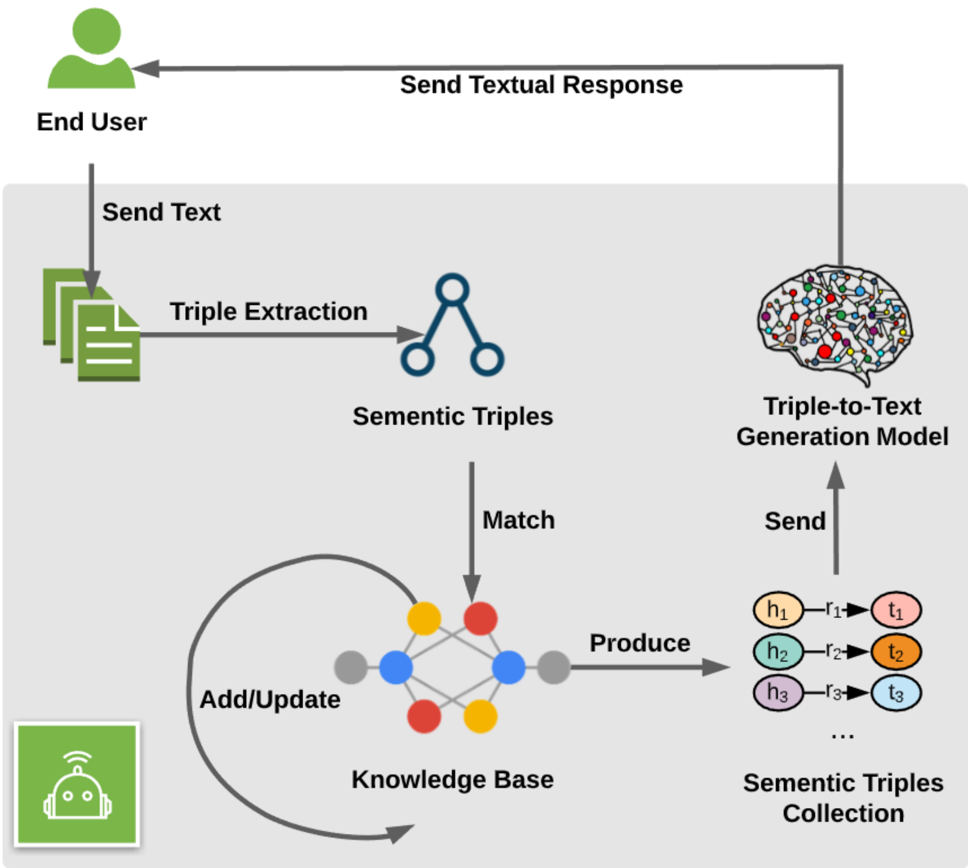

Figure 3.1: GSCA Framework

where h_x, r_x and t_x represent head, relation and tail of f_x, respectively.

In GSCA, each entity $\varepsilon \in \{h, r, t\}$, has been granted with enriched features, including the last accessed time τ_ε, frequency of being visited ω_ε and attention degree η_ε. The agent is capable of inferring the user's update-to-date preferences according to these features. Specifically, the attention degree of any entity η_ε can be derived by using a time decay function, which is represented in Equation 3.3.

$$\eta_\varepsilon = e^{-\alpha \Delta\tau}, \alpha > 0, \tag{3.3}$$

where $\Delta\tau = \tau_\varepsilon - \tau_{now}$, and α describes a constant controlling the degree of decay. Having the entities with enriched features, the agent utilises the extracted triples to match and enhance the contents of the existing knowledge base.

3.2.2 Temporal-based Triple Retrieval Algorithm

I propose a novel algorithm, named the Temporal-based Triple Retrieval (TTR), to effectively identify the most relevant triples with a confidence score, where both temporal features and user preferences are taken into consideration. Our algorithm involves three major steps.

First, the entities of each fact n are vectorised:

$$f'_x = (v(h_x), v(r_x), v(t_x)) \tag{3.4}$$

where $v(.)$ indicates a word-to-vector function, converting a token into a vector, in this thesis pre-trained word vector Glove.6B is utlised, the details are explained in Subsection 3.3.3. There are two reasons why the translate models for KG, e.g., TransE, are not adopted. (1) Such algorithms vectorise the entire triple, but GSCA intends to match entities for obtaining a sub-graph. (2) GSCA facilities a self-adaptive framework;

thus, it is a non-trivial task to keep training and updating the vectors in a dynamic environment.

Second, we estimate the distance between the hypothesis triple and the existing triples using Equation 3.5

$$dis(f_i, f_j) = \sum_{\varepsilon \in \{h,r,t\}} w_\varepsilon \cdot sim(v(\varepsilon_i), v(\varepsilon_j)) \tag{3.5}$$

where $sim(\varepsilon_i, \varepsilon_j)$ denotes the cosine similarity between ε_i and ε_j, having

$$sim(\varepsilon_i, \varepsilon_j) = \frac{v(\varepsilon_i) \cdot v(\varepsilon_j)}{\| v(\varepsilon_i) \| \times \| v(\varepsilon_j) \|} \tag{3.6}$$

While w_ε balances the trade-off among h, r and t with a restriction of $\sum_{\varepsilon \in \{h,r,t\}} w_\varepsilon = 1$.

Third, the agent identifies potential answer triples with extracted triples and the connected siblings. Given hypothesis triple f_i, the confidence score of an identified triple $f_j = (h_j, r_j, t_j)$ can be derived from distance $dis(f_i, f_j)$, normalized frequency of being accessed $\hat{\omega_\varepsilon}$ and attention degree of the entity η_ε, where $\varepsilon \in \{h_j, r_j, t_j\}$.

According to the confidence score, the triples are selected and become the inputs of the Triple-to-Text model for response generation. However, if the score falls below a certain threshold, the agent treats these triples as new knowledge and resolves them internally. On top of that, the agent can proactively initiate conversations with users. Without any inputs from the users, the agent can periodically trigger a semantic triples recommendation using the TTR algorithm, which dynamically produces up-to-date triples according to the recent conversations and preferences of the user. Another novelty of our model can be reflected as the capability of joining a variety of KGs produced by multiple users, forming a global KG for supporting the ongoing conversation.

3.3 Knowledge Graph Model

In this proposed model, we adapt both personalised KGs and global KG. This Subsection mainly introduces these two KGs and how they interact.

3.3.1 Knowledge Graphs Interactions

The proposed GSCA is capable of revising the agents' knowledge base without retraining, which means that the system must capture user input and store it for ease of use in future conversations. Figure 3.2 illustrates the interactions between personal KGs and the Global KG. The system will generate a separate knowledge graph for each user. Meanwhile, the Conversational Agent has a Global KG, which is shared by all users. When information overlaps, all nodes can be connected with corresponding relations, including two nodes from different KGs.

The response triple selection and sentence generation methods of the GSCA framework are listed as follows:

- The best match triples exist in the current KG; under these circumstances, the system will generate the most appropriate response by retrieving the KG.
- The knowledge base of the agent is insufficient. To address this issue, our system will retrieve relevant external information. For instance, the user's input text is related to the Covid-19. Then, the NewsAPI will be triggered to search for the related news by news article headlines; that information will be used as the raw materials for responses.

Notably, the model intends to generate new nodes from the external resources and extend the knowledge base. Thus, these triples will join to global KG. Apart from global KG, the CA will also have a personalised KG dedicated to each individual,

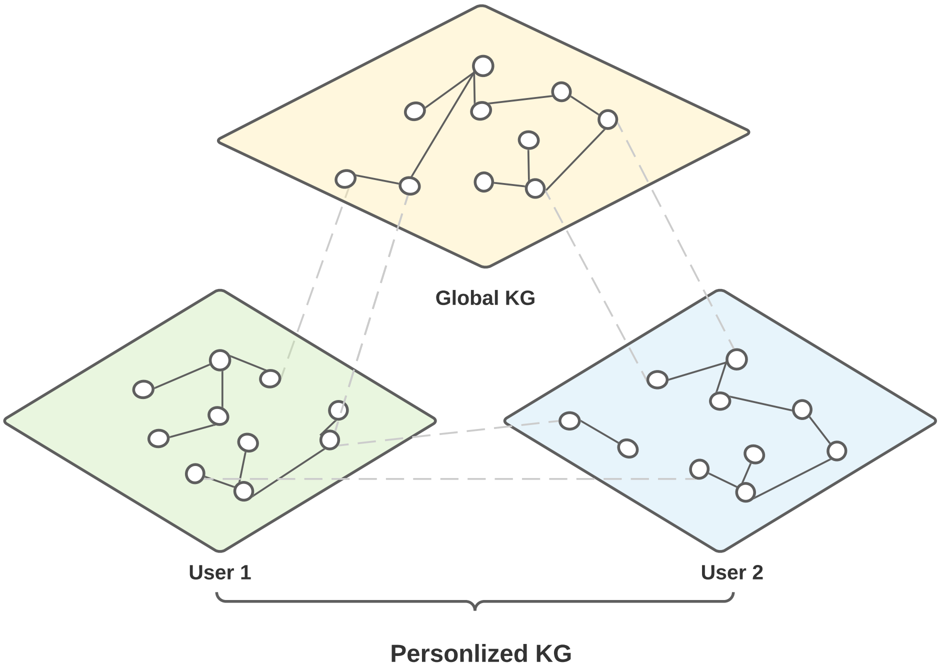

Figure 3.2: Knowledge Graphs Interaction Model

recording each user's conversational flows. In order to extract personal preference and other information, prepare for further research.

The personalised KG records all the information from user-agent interactions, including the triples that convert from chat content along with the timestamp and frequency. These contents can express user preferences to a certain extent.

Unlike existing conversational agents, our model is not trained with large datasets. Sometimes, the current user's knowledge base might appear insufficient. To address this issue, the KG augmentation is applied, i.e., the conversational agent automatically retrieves triples from other users' knowledge base and generate a response.

3.3.2 Triple-to-Text Model

In this proposed **GSCA** framework, the triple-to-text model was trained by using deep learning methods to achieve that the chatbot responds to users in natural language.

```xml
<entry category="Company" eid="Id21" shape="(X (X) (X) (X) (X))" shape_type="sibling" size="4">
    <modifiedtripleset>
        <mtriple>Trane | foundingDate | 1913-01-01</mtriple>
        <mtriple>Trane | location | Ireland</mtriple>
        <mtriple>Trane | foundationPlace | La_Crosse,_Wisconsin</mtriple>
        <mtriple>Trane | numberOfEmployees | 29000</mtriple>
    </modifiedtripleset>
</entry>
```


Figure 3.3: WebNLG Challenge 2020 Dataset Sample

Google pre-trained language model Transfer Text-to-Text Transformer (T5) (Raffel et al., 2019) has been utilised in text generation task.

The RDF-to-text task from WebNLG Challenge 2020 is used to train the T5 model. Figure 3.3, illustrates a sample of RDF triples in the WebNLG dataset. Given four RDF triples, the aim is to generate sentences, e.g., “Trane, which was founded on January 1st, 1913 in La Crosse, Wisconsin, is based in Ireland. It has 27000 employees.”

3.3.3 Word Embedding

In Nature Language Processing (NLP) tasks, one of the most common technique is word vectors representation, in the proposed model, this framework adopts pre-trained word vectors which are selected from the non-domain sources in Glove , in this project Glove.6B.100d, Glove.6B.200d and Glove.6B.300d are utlised (Pennington, Socher & Manning, 2014):

- Pre-trained word vectors, haveing 6B tokens and 400K vocabulary, come from Wikipedia 2014 and English Gigaword Fifth Edition.

Although our GSCA framework adopts the KG basis, the use of KG embedding will affect the randomness and flexibility of the triple combination. In addition, my GSCA framework has no pre-trained knowledge base, data collected from user-agent interactions seem not possible to meet the number of context requirements of training a KG embedding model.

3.4 Experiments and Analysis

Two experiments have been conducted to evaluate the performance of GSCA. I simulate two virtual end-users, i.e., Lindsay and Jason, who communicate with the GSCA based on two topics, i.e., food and artist.

Experiment 1 aims to evaluate the agent's self-adaptive capabilities, having the users' preference and attention degree considered. Table 3.1 illustrates only a few conversation examples, but it is obvious that the agent responds by extending the user's input.

The proposed TTR algorithm determines the selection of triples for response generation. For evaluation purpose, I intentionally increase the access frequency ω_ε and refresh the last visiting time stamp η_ε of node "hip hop music" via mentioning it a few times to the agent. I find that the agent is more inclined to use this node for generating responses.

I conduct 60 rounds of conversations covering the topics of both artist and food. After varying both ω_ε and η_ε through the interactive conversations, the experimental results show that 73% of the responses are adapted and appear reasonable. There are around 17% chances for a user to receive a response without knowledge adaptation, and merely 10% turn out to be unreasonable. This is due to the limited knowledge obtained by the agent.

Experiment 2 evaluates the knowledge augmentation capability of the agent. Though the knowledge base keeps increasing through agent-human interactions, the knowledge may still appear insufficient to handle all the user's inputs, especially in the early stage. The agent intends to borrow the knowledge obtained from other users for generating a reply.

I further initiate 30 rounds of conversation using user Lindsay. The agent properly handles 70% of the inputs through the TTR algorithm, matching the triples and

Table 3.1: Conversation examples

Topic	Input Sentence	Retrieved Triple	Reply Text
Artist	Aaron turner is a friend of Billy.	"aaron turner" "be friend of" "aaron deer"	Aaron turner was a friend of Aaron Deer.
Artist	Hip hop music is very popular.	"hip hop music" "be genre of" "rap music"	Hip hop music is a genre of rap music.
Artist	Billy performs rap music.	"Billy" "perform" "hip hop music."	Billly performs hip hop music.
Food	Almond is brown colour	"almond" "be" "flowering plant"	Almond is a flowering plant.
Food	Almond has colour	"Almonds" "be from" "rosaceae family"	Almonds are from the rosaceae family.
Food	Almonds is food.	"sandwich" "be" "food"	The sandwich is a food.

generating responses. 30% is missing from the knowledge base, but new triples are retained. It is worth noting that 10% of the responses utilize the knowledge of both users, Lindsay and Jason. This implies the agent is capable of effectively augmenting a user's knowledge by using others.

3.5 Summary

In this chapter, I proposed a GSAC framework to address the research gap of the Conversational Agents without re-training and adaptive information during user-agent interactions. Firstly, I demonstrate the data collection processes, including text-to-triplets generation, knowledge graph joining and expansion. Secondly, I explain the procedure of answer triple matching, and the text generation deep learning model T5, pre-trained word vectors representation Glov. Finally, I conduct two experiments on different topics: “Artists" and “food".

Based on the experimental results, I can conclude that our proposed GSAC framework enables agents to learn, adapt and augment knowledge through interactions. Since the adaptive conversational agent's knowledge is transparent, it is easy to modify the response by updating the knowledge graph. The GSAC framework can be applied to various scenarios, including education, user profiling, knowledge correction, knowledge augmentation, etc.

This chapter mainly answers the Research Question 1 mentioned in Chapter 1. The research work of this chapter has been published in (L. Zhang et al., 2021).

The GSAC framework proposed in this chapter serves as a basis for the next chapter. In chapter 4, I future proposed a model Context-aware Recommendation System using Graph-based Behaviour Analysis.

Chapter 4

Context-Aware Recommendation System using Graph-based Behaviours Analysis

4.1 Introduction

Based on the GSCA framework introduced in Chapter 3, the users' personalised KGs are utilised for user preference analysis, which is an important feature adopted in the Graph-based Context-Aware Recommendation (GCAR) algorithm. This chapter introduces a novel recommendation approach, i.e. GCAR, by considering user behaviours and event context. The former models the user-event interactions, and the latter formulates the event-event relations through the KG. The recommendation for events is based not only on both features but also on the temporal features. In other words, the impact from the recently happened events has been involved in the recommendation. Extensive experiments are conducted to evaluate the performance of GCAR. On top of that, I further explore the effectiveness of applying knowledge graph by conducting parameter analysis. The experimental results demonstrate that GCAR can perform better by comparing with other traditional algorithms, especially when there are complex relations among user behaviours and events.

The rest of this chapter is organised as follows. Subsection 4.2 introduces the preliminaries, formal definitions and problem formulation. Subsection 4.3 systematically elaborates on the proposed GCAR. In Subsection 4.4, experimental results are presented to evaluate the performance of the GCAR. The summary of this chapter is given in Subsection 4.5.

4.2 Preliminaries

This section introduces the formal definitions required by GCAR and formulates the context-aware recommendation problem for the current setting.

4.2.1 Formal Definition

Definition 4.2.1 ***A user*** *$u_i \in U$ refers to a person who can initiate a number of activities, where $U = \{u_1, u_2, \cdots, u_n\}$ represents a finite set of users. A user can be denoted by using a tuple, i.e., $u_i = (A_i, E_i')$, where A_i describes the demographic attribute set of the user, E_i' indicates the historical events that u_i has participated in (refer to Definition 4.2.2).*

Definition 4.2.2 ***An event*** *generally refers to a thing that happens or take place. In the current settings, event $e_n = (u_i, t, l), e_n \in E_i$ describes a specific activity that is participated by user u_i at time t in location l, where $E_i = \{e_1, e_2, \cdots, e_m\}$ denotes all the possible events can be actioned by u_i, and $E_i' \subset E_i$.*

Definition 4.2.3 ***A transaction*** *$txn_i^k \in T_i$ refers to an ordered sequence of events initiated by user u_i within a reasonable time-frame, e.g., one day. $T_i = \{txn_i^1, txn_i^2, \cdots, txn_i^n\}$ denotes user u_i's transaction set and $txn_i^k =< e_1, e_2, \cdots, e_m >$.*

Definition 4.2.4 ***User Behavioural Graph (UBG)*** *$G_{ub} = (U, E, L_{ub})$ projects users U and events E to the same space, having both connected through the user-event interactions. In other words, the link set of G_{ub}, i.e., L_{ub}, represents the users' behaviours.*

Definition 4.2.5 ***Event Knowledge Graph (EKG)*** *$G_k = (V_f, L_f)$ refers to a collection of interlinked semantic descriptions of entities that are associated with events. For example, user v_i initiates an event e_i, "watch a rugby game", the associated entities, e.g., "rugby", "sports" and "New Zealand", as well as their relations, are all included in the G_k. V_f describes the entities referring to as facts, and L_f indicates the semantic relationship among the elements of V_f.*

Definition 4.2.6 ***Event Context Graph (ECG)*** *$G_c = (E, G_k, L_e)$ models the events E and the associated entities in the same space. G_c also presents the implicit event-event relationship through the common entities in G_k.*

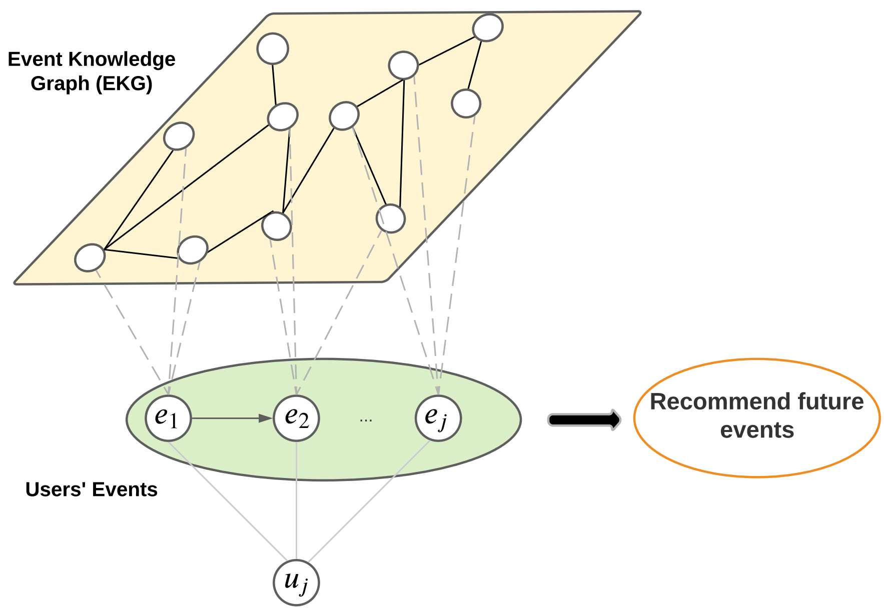

Figure 4.1: Context-aware Recommendation with EKG

4.2.2 Problem Formulation

In this subsection, I formally define the context-aware recommendation problem. Given user u_j and recent historical events $E'_j = < e_1, e_2, \cdots, e_m >, E'_j \subset E_j$ with the associated EKG G_k, the objective is set to recommend u_j with a future event, i.e., e_{m+1}. In other words, given the user, current state and event context, the system will estimate the occurrence probabilities of the events $\{e_x | e_x \in E_j\}$, where the probability of e_x is represented as $+p(e_x | e_1, e_2, \cdots, e_m)$.

Figure 4.1 demonstrates the general idea of the context-aware recommendation problem defined in this chapter. A user conducts a chain of actions, and these behavioural events are not directly associated with each other, but all of them connect with the nodes of EKG.

To address this problem, our proposed approach using Graph-based Context-Aware Recommendation (GCAR) Algorithm is elaborated in the next section.

4.3 Graph-based Context-Aware Recommendation

In this section, I give an overview of the GCAR and explain each module in detail. The framework of GCAR is presented in Figure 4.2, where three layers, i.e., users, events and EKG, are incorporated.

Users conduct a series of actions, and these user-event interactions can be modelled as UBG (refer to Definition 4.2.4) by merging both layers. Likewise, ECG models the relations between the events and EKG. Figure 4.2 shows that all the events are semantically linked with the EKG nodes. There is no direct relation among the events, but the implicit long-range connectivities can be established through the events' connections with EKG. In other words, EKG forms the context of events.

Figure 4.3 illustrates an example of the GCAR framework, which also demonstrates the events' long-distance connectivities through the UBG and EKG interactions. UBG shows the user-event relations, and the events' contextual information can be obtained via ECG.

4.3.1 User Behaviour Graph

The UBG module captures the users' behaviours and generates user-event interactions. Figure 4.4 demonstrates the graph extraction process, starting from the data acquisition. The sensors on users' IoT devices or intelligent mobile return update-to-date information

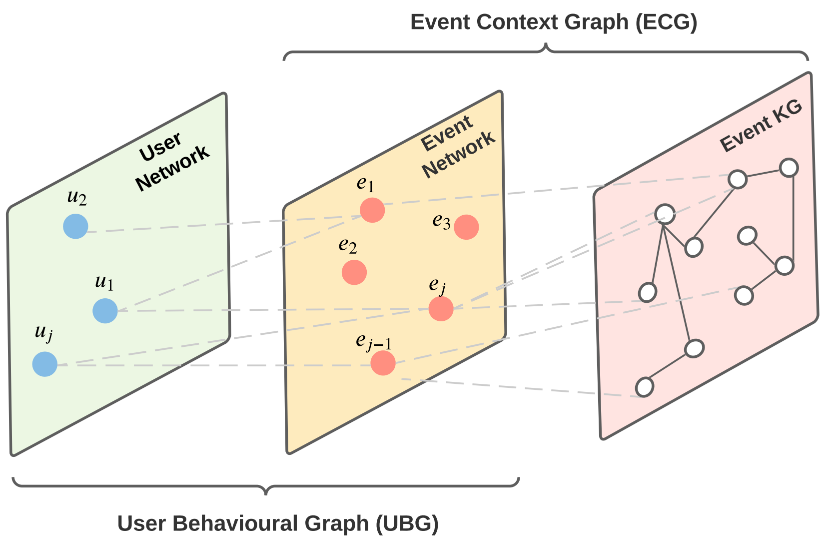

Figure 4.2: The GCAR Framework

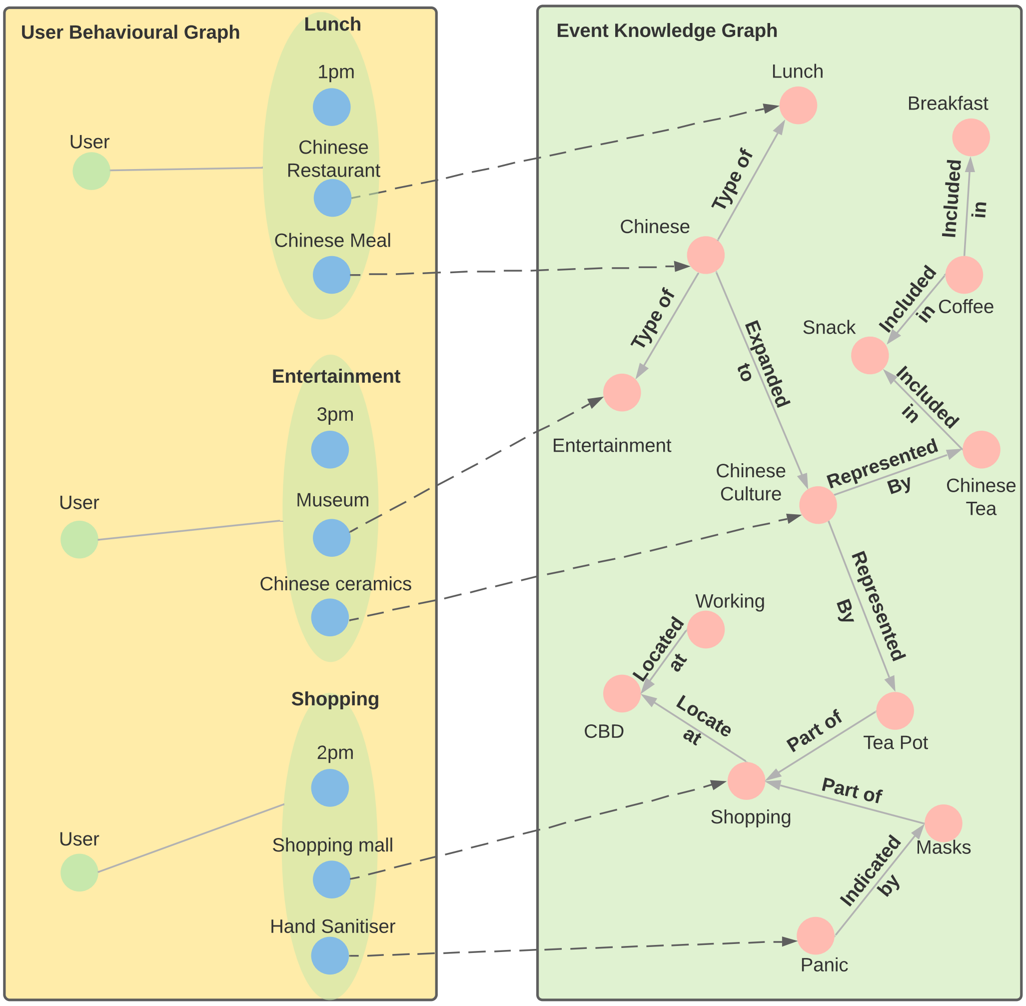

Figure 4.3: An Example of GCAR

about the user's activities, including the current time and location. The time and location have been acknowledged as two essential features of an event.

Based on people's daily routine, I define five intervals listed in Table 4.1. The occurrence of an event has one or more "preferred" time intervals. For example, people usually consider visiting morning cafe in the early morning of weekdays, rather than going shopping. The temporal feature plays an important role in determining the users' actions.

Location is classified into various address categories, e.g., school and gym. With

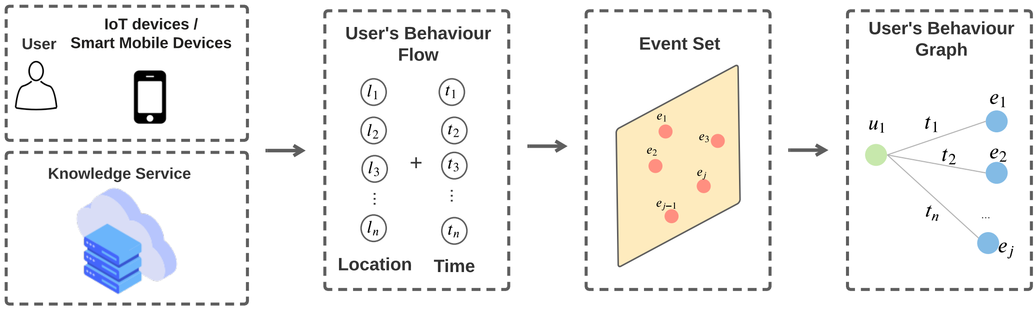

Figure 4.4: User's Behaviour Graph: User-Event Interactions

Table 4.1: Time intervals for Events

Time Stamp	Time Interval	Range of the time
EM	Early Morning	6 a.m. – 9 a.m.
AM	Morning	9 a.m. – 12 p.m.
Noon	Midday	12 p.m. – 15 p.m.
PM	Afternoon	15 p.m. – 18 p.m.
Evening	Evening	18 p.m. – 21 p.m.

the assistance of cloud-based services, e.g., Google Map, the activities can be inferred. The collected data will be transformed into a sequence of user behaviours with simple inferences, which model the user-event interactions. For example:

- in a restaurant at noon for 20 minutes ⇢ having lunch
- in the gym, for 20-40 minutes ⇢ doing workout
- out of the office from 5:30 pm ⇢ finished work

The advantage of adopting behaviour graph expression is to link multi-dimensional information such as time, location and user's events together and project into a two-dimensional model.

4.3.2 Event Context Graph

Most existing works only focus on modelling the context of users but ignore the event context. In the proposed model, event context models events and event-event relations through EKG, and both are projected to the same space.

One event can reach the other through linking with the corresponding entities in EKG. Some events may have a very close distance and are strongly associated with each other, e.g., both events are linked with the same set of entities. In contrast, some may possess long-range connectivity, which requires several hops from one event to reach the other.

To estimate the probabilities of one event impacting the other, I utilise one of the classic influence-diffusion models, i.e., the Independent Cascade (IC) model, which presents the influence as a hopping and infecting process (Kempe, Kleinberg & Tardos, 2003; Li, Bai & Zhang, 2018). The propagation is initiated from the active events node, i.e., recent events conducted by the users. There is a uniform probability to activate the adjacent nodes for each hop, and the neighbours will attempt to activate the connected nodes if being activated. I regard *Event 2* gets impacted by *Event 1* if *Event 2* becomes active when *Event 1* is conducted by the user.

Figure 4.5 demonstrates the main idea of event-event influence simulation in Event-Context Graph. Nodes in different colours, i.e., red, yellow and white, represent

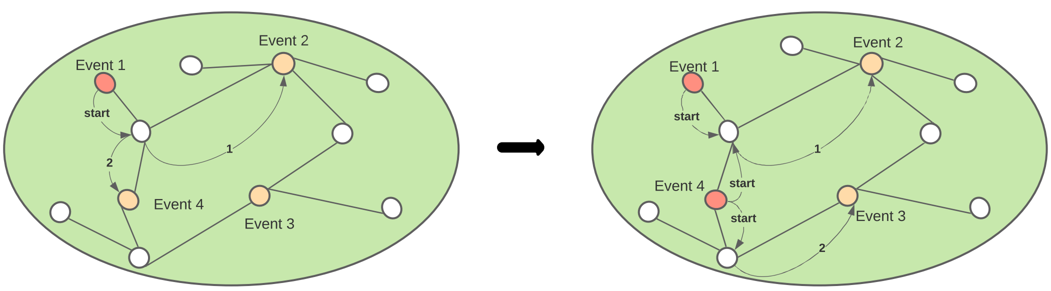

Figure 4.5: Event-Context Graph with IC Model

activated event nodes, inactive event nodes and context nodes. The arrow refers to the direction of influence propagation, and the number indicates possible directions.

Monte Carlo Simulations (Mooney, 1997) with IC model is adopted. For example, if *Event 1* is the initial activity that the user has conducted, *Event 1* will spread influence to the rest of the nodes in the ECG. Given 1000 times simulations, if *Event 2* appears activated 600 times, we can conclude there is a probability of 60% for *Event 1* to impact *Event 2*.

4.3.3 GCAR Algorithms

This subsection introduces the proposed algorithms to recommend events for users. The algorithm considers both user's behaviours and event context, modelled as UBG and ECG, respectively.

First, based on the UBG and current time step t, the user-event occurrence degree $\sigma(u_i, e_m, t+1)$ is estimated by considering the time and location of both user and event using Equation (4.1).

$$\begin{aligned}&\sigma(u_i, e_m, t+1) = \\ &\alpha \cdot \mathrm{dis}^{\tau}_{t+1}(u_i, e_m) + (1-\alpha) \cdot \mathrm{dis}^{l}_{t+1}(u_i, e_m),\end{aligned} \tag{4.1}$$

where $\mathrm{dis}^{\tau}_{t+1}(u_i, e_m)$ estimates the distance between u_i and e_m in terms of time, and $\mathrm{dis}^{l}_{t+1}(u_i, e_m)$ calculates the location distance. Constant $\alpha \in [0, 1]$ balances the trade-off between the weight of time and location.

Second, we obtain event-event strength according to the ECG. If no recent events presented, the probability for user u_i to adopt event e_m, i.e., $p(e_m, u_i)$, is calculated using Equation (4.2). In this case, there is no impact received from other events.

$$p(e_m, u_i) = \sigma(u_i, e_m, t + 1) \tag{4.2}$$

Given a set of recently occurred events of user u_i, i.e., $E'_i \neq \varnothing, E'_i \subset E_i$, without considering the user's historical behaviours, the probability for u_i to perform event e_m is formulated using Equation (4.3).

$$p(e_m, u_i|E'_i) = 1 - \prod_{e_j \in E'_i} (1 - p(e_m, u_i|e_j)) \tag{4.3}$$

In Equation (4.3), $p(e_m, u_i|e_j)$ refers to the probability of u_i adopting e_m after conducting event e_j, which is formulated in Equation (4.4).

$$p(e_m, u_i|e_j) = 1 - \prod_{r \in R(e_m, e_j)} (1 - \epsilon^{h(r)}), \tag{4.4}$$

where ϵ denotes the uniform activation probability, which is a constant in the current setting. $R(e_m, e_j)$ represents a set of paths, bridging e_m and e_j. While $h(.)$ refers to a function calculating the length of a particular path. For example, if r shows $e_m \to v_i \to v_j \to e_j$, we have $h(r) = 3$.

Finally, considering both UBG and ECG, the probability of user u_i to conduct event e_m at time step $t + 1$, i.e. $\eta(e_m, u_i, t + 1)$, is formulated in Equation (4.5), in which $\beta \in [0, 1]$ is applied to balance the UBG/ECG trade-off.

$$\begin{aligned} \eta(e_m, u_i, t + 1) &= \beta \cdot \sigma(u_i, e_m, t + 1) \\ &+ (1 - \beta) \cdot p(e_m, u_i|E'_i) \end{aligned} \tag{4.5}$$

Algorithm 1 describes the Graph-based Context-Aware Recommendation Algorithm. The inputs incorporate user u_i, current time t, the user's recent events E'_i, all the possible events for u_i, i.e., E_i, and the number of events to be recommended k. The output is a list of recommended events L with scores. Line 1 initialises the event candidate list. Line 2 iterates all the potential events that have not been actioned by u_i recently. Line 3 aims to calculate user-event occurrence degree. Lines 4-6 check if u_i has any recently conducted events. If not, the probability of u_i to adopt e_m is fully determined by $\eta(e_m, u_i, t+1)$. Lines 7-10 estimate the event adoption impacted by the event-event strength, and Line 11 finalises $\eta(e_m, u_i, t+1)$ for each candidate event. Lines 13-14 update the score and append the event to the candidate list L. Lines 16-17 sort L based on $e_m.\eta$ and return top k events for the recommendation.

4.4 Experiment and Analysis

Two experiments are conducted in this chapter to evaluate the proposed algorithm, i.e., the GCAR algorithm. Experiment 1 compares GCAR against the state-of-the-art

Algorithm 1: Graph-based Context-Aware Recommendation Algorithm

Input: u_i, t, E_i', E_i, k
Output: A list of recommended events with scores L
Initialise the event candidate list $L := []$
for $e_m \in (E_i - E_i')$ **do**
    Compute $\sigma(u_i, e_m, t+1)$ using Equation (4.1)
    if $E_i' = \emptyset$ **then**
        $\eta(e_m, u_i, t+1) := \sigma(u_i, e_m, t+1)$
    else
        for $e_n \in E_i'$ **do**
            Calculate $p(e_m, u_i | e_n)$ using Equation (4.4)
        end
        Calculate $p(e_m, u_i | E_i')$ using Equation (4.3)
        Calculate $\eta(e_m, u_i, t+1)$ using Equation (4.5)
    end
    Update $e_m\eta := \eta(e_m, u_i, t+1)$
    $L := L \cup \{e_m\}$
end
Sort list L based on $e_m\eta, e_m \in L$ in a descending order
return $L[k]$

algorithms, including Naive Bayes, SVM, decision tree and MLP. In Experiment 2, parameter analysis is conducted to investigate the contribution of event context in the process of recommendation.

4.4.1 Experiment Settings

Datasets and Evaluation Metrics

I simulate the dataset based on the Ego-Facebook (Panzarasa, Opsahl & Carley, 2009). A particular number of users and events with the corresponding attributes are generated. The network topological structure of Ego-Facebook is adopted to simulate EKG. The links between events and EKG are randomly generated. The interactions among user and events are created through event-driven simulations with temporal features. Based on this approach, three datasets have been created, having 10 users with 5, 10, 15 events, respectively. All the generated events are randomly linked with the nodes of the simulated EKG. Each user has 100 transactions, represented as a sequence of events.

According to the problem definition, the objective is to recommend a user with events based on a number of factors, including users' and events' attributes, as well as recently occurred events. The problem can be categorised as a classification problem. Traditional evaluation metrics, including precision, recall and F1 score, have been adopted to evaluate the performance.

Baselines

Four traditional algorithms have been involved as the counterparts of the GCAR Algorithm.

- **Naive Bayes** (Namahoot, Brückner & Panawong, 2015). It is one of the scalable and straightforward probabilistic classifiers, which is based on the Bayes' theorem.

- **Support Vector Machine (SVM)** (Oku, Nakajima, Miyazaki & Uemura, 2006). It is a supervised machine learning model, which can be used for classification and regression analysis.

- **Multilayer Perceptron (MLP)** (S. Lu et al., 2018). It is recognised as a typical example of feed-forward Artificial Neural Networks (ANN) which utilises supervised learning techniques.

- **Decision Tree** (Patel & Prajapati, 2018). It is a simple approach for classification problems, which can predict the value of a target variable based on several input variables.

4.4.2 Experiment 1

In Experiment 1, I evaluate the performance GCAR algorithm by comparing it against four state-of-the-art algorithms by using three datasets. Tables 4.2–4.4 demonstrate

Table 4.2: Performance evaluation using the dataset with 10 users and 10 events

	Recall	**Precision**	**F1-Score**
SVM	0.424	0.501	0.459
Naive Bayes	0.556	0.573	0.564
Decision Tree	0.660	0.552	0.601
MLP	0.706	0.678	0.692
GCAR	**0.760**	**0.754**	**0.757**

Table 4.3: Performance evaluation using the dataset with 10 users and 15 events

	Recall	**Precision**	**F1-Score**
SVM	0.565	0.736	0.639
Naive Bayes	0.587	0.555	0.570
Decision Tree	0.603	0.688	0.643
MLP	0.797	0.761	0.778
GCAR	**0.890**	**0.919**	**0.904**

Table 4.4: Performance evaluation using the dataset with 10 users and 5 events

	Recall	**Precision**	**F1-Score**
SVM	0.725	0.728	0.727
Naive Bayes	0.580	0.556	0.570
Decision Tree	**0.853**	0.749	0.798
MLP	0.793	0.783	0.790
GCAR	0.850	**0.817**	**0.833**

the results. As can be observed from these tables, our algorithms outperform others, especially when the number of events increases. This is because, given the same EKG, more connections with the EKG will be established when having more events. Subsequently, the relations among the events become complicated. GCAR considers the event context and can handle this situation better than the others.

4.4.3 Experiment 2

Experiment 2 conducts parameter analysis and explores the impact of EKG by adjusting the weights of both event-event and user-event relations. The same sets of datasets are adopted in this experiment.

Table 4.5: GCAR parameter analysis using the dataset with 10 users and 10 events

Weight		**Evaluation Metrics**		
Event-event	**User-event**	**Recall**	**Precision**	**F1-Score**
0	1.0	0.080	0.006	0.012
0.1	0.9	0.150	0.023	0.039
0.2	0.8	0.410	0.168	0.238
0.3	0.7	0.520	0.401	0.453
0.4	0.6	0.540	0.457	0.495
0.5	0.5	0.560	0.446	0.497
0.6	0.4	0.620	0.680	0.648
0.7	0.3	0.760	0.754	0.757
0.8	0.2	0.720	0.730	0.725
0.9	0.1	0.720	0.701	0.701
1.0	0	0.670	0.673	0.671

Tables 4.5–4.7 illustrate the results. In Table 4.5, GCAR yields the best performance when having the weights of 0.7 and 0.3 assigned to event-event and user-event, respectively. If event-event relation is removed from the recommendation process, i.e., allocating 100% to user-event, the GCAR will give a very Unsatisfactory performance. The other two datasets also demonstrate the same pattern. By comparing Table 4.6 and Table 4.7, we can find GCAR carries out better recommendation results with a higher number of events. This also implicitly proves GCAR's superior ability in handling complex event-event relations.

Table 4.6: GCAR parameter analysis using the dataset with 10 users and 5 events

Weight		**Evaluation Metrics**		
Event-event	**Event-user**	**Recall**	**Precision**	**F1-Score**
0	1.0	0.100	0.010	0.018
0.1	0.9	0.100	0.010	0.018
0.2	0.8	0.430	0.372	0.399
0.3	0.7	0.600	0.553	0.575
0.4	0.6	0.790	0.776	0.783
0.5	0.5	0.800	0.748	0.773
0.6	0.4	0.780	0.785	0.783
0.7	0.3	0.850	0.817	0.833
0.8	0.2	0.820	0.799	0.810
0.9	0.1	0.780	0.789	0.785
1.0	0	0.820	0.741	0.779

Based on the experimental results, we can conclude that the proposed GCAR algorithm can produce more superior results than traditional algorithms, especially when the event context appears complex. Even without considering user-event interactions, namely, not enough transactions, the GCAR still performs well. By leveraging the context of events, GCAR can also tackle the cold-start and data sparsity problem in RSs.

Table 4.7: GCAR parameter analysis using the dataset with 10 users and 15 events

Weight		Evaluation Metrics		
Event-event	**User-event**	**Recall**	**Precision**	**F1-Score**
0	1.0	0.0100	0.0001	0.0002
0.1	0.9	0.01000	0.00010	0.00020
0.2	0.8	0.130	0.286	0.179
0.3	0.7	0.210	0.588	0.309
0.4	0.6	0.580	0.898	0.705
0.5	0.5	0.800	0.914	0.853
0.6	0.4	0.880	0.903	0.891
0.7	0.3	0.890	0.919	0.904
0.8	0.2	0.930	0.932	0.931
0.9	0.1	0.900	0.919	0.909
1.0	0	0.940	0.932	0.936

4.5 Summary

This chapter proposed a novel algorithm called GCAR to produce user activities recommendations. In the proposed model, both user-event interactions and event-event relations are considered. GCAR algorithm facilitates knowledge graph to model the event-event relations. Specifically, I developed a three-layer architecture, i.e., user network, event network and event knowledge graph. Next, I combine the user network and event knowledge graph as the user behaviour graph, event network and event knowledge graph as the event context graph. Both user behaviour graph and event knowledge graph are utilised for the event recommendations. The classic information diffusion model has been adopted to estimate the conditional probability for event occurrence. Finally, the proposed GCAR algorithm has been evaluated by comparing against state-of-the-art algorithms on three simulated datasets. The experimental results demonstrate the superior of GCAR and excellent ability to handle complex relations among the events.

Therefore, the proposed GCAR can be applied to achieve intelligent behaviour prediction and recommendation. It is capable of serving various intelligent predictions as further developments in the specific domains such as personal reminders and elderly healthcare services.

This chapter mainly answers the Research Question 2 mentioned in Chapter 1. The research work of this chapter has been published in (L.Zhang et.al, 2021).

The conclusions of the thesis and future research directions will be given in the next chapter.

Chapter 5

Conclusions and Future Work

5.1 Introduction

This chapter summarise the findings based on the proposed novel approaches, e.g., Graph-based Self-adaptive Conversational Agent and Graph-based Context-Aware Recommendation system. The major contributions of my research work have been explained in Subsection 5.2. The limitations and possible directions for the future work are discussed in Subsection 5.3.

5.2 Summary of Major Contributions

My research work brings contributions to the adaptive conversational agents from the following two aspects.

5.2.1 Graph-based Self-Adaptive conversational Agent

- I proposed the GSCA framework, which is capable of improving the agent itself through the interactions with the users. The knowledge-based of CAs can be expanded and adapted without human intervention. Namely, no domain expert or re-training procedure is required.

- I proposed a novel approach for CAs to store, retrieve and generate information with KGs. The unstructured text data are transformed into a multi-dimensional representation form. The originally unrelated conversational flow is associated by the relations.
- A personalised KG is proposed, preparing for the recommendation system to analyse user preferences. The KG joining and KG expansion techniques are adopted to expand and fine-tune the knowledge base. Moreover, the KG augmentation method is applied when the personalised knowledge appears insufficient to handle the user's inquires.
- I further proposed the semantic triple extractions methods, where the user input is automatically converted into triples with corresponding timestamp. Also, the system automatically calculates the nodes' frequency, which can capture the user preferences.
- I developed a novel Temporal-based Triple Retrieval (TTR) algorithm, which can retrieve the best match triples as the input of the text generation method. TTR takes user preference and temporal features of nodes into consideration.

5.2.2 Graph-based Context-Aware Recommendation

- I formally defined the interactive process of multi-layer KGs, formulated the user-event interactions and event-event interactions by proposing a three-layer architecture. The architecture consists of the user network, event network and event knowledge graph. I modelled the user-event interactions with the user network and event network. In addition, I modelled the event-event interactions with the event network and event knowledge graph.

- I proposed the Graph-based Context-Aware Recommendation (GCAR) algorithm to achieve intelligent behaviour prediction and recommendation. In this algorithm, event-context, event-event influence and user preference are taken into consideration. The conditional probability is used to estimate the probability of the next event.

5.3 Limitations and Future Work

In my master studies, I developed a framework of Graph-based Self-adaptive Conversational Agent (GSCA) and proposed the novel Graph-based Context-aware Recommendation (GCAR) algorithm. The limitations of the GSCA framework and GCAR algorithm can be concluded as following aspects.

Firstly, in the GSCA framework, the response generation module requires relatively high computer hardware performance. As a result, the users have to wait for some time to get the response from the GSCA. Secondly, the truth discovery is not applied. As for information collected from user-agent interactions, GSCA is not capable of distinguishing between authenticity and falsehood. The users might get a response with the incorrect information. Thirdly, the GCAR algorithm shows lower accuracy when it makes recommendations to new users. Because the information of a specific user's personalised KG appears insufficient, impacting the recommendation results.

In the future, I plan to enhance the existing framework by addressing the conflict knowledge learnt from multiple users, where the truth discovery will be applied. With the GCAR algorithm, I will consider modelling each user to address insufficient personalised information. Moreover, I will also consider combining the ideas of the traditional recommendation system with the GCAR algorithm to build a hybrid recommendation system. The primary purpose is to provide accurate and intelligent personalised recommendations in dealing with the complex event.

References

Abdul-Kader, S. A. & Woods, J. (2015). c. *International Journal of Advanced Computer Science and Applications*, *6*(7).

Barlow, H. B. (1989). Unsupervised learning. *Neural computation*, *1*(3), 295–311.

Battineni, G., Chintalapudi, N. & Amenta, F. (2020). Ai chatbot design during an epidemic like the novel coronavirus. In *Healthcare* (Vol. 8, p. 154).

Breese, J. S., Heckerman, D. & Kadie, C. (2013). Empirical analysis of predictive algorithms for collaborative filtering. *arXiv preprint arXiv:1301.7363*.

Chakrabarti, C. & Luger, G. F. (2015). Artificial conversations for customer service chatter bots: Architecture, algorithms, and evaluation metrics. *Expert Systems with Applications*, *42*(20), 6878–6897.

Chaudhari, S., Azaria, A. & Mitchell, T. (2017). An entity graph based recommender system. *AI Communications*, *30*(2), 141–149.

Cunningham, P., Cord, M. & Delany, S. J. (2008). Supervised learning. In *Machine learning techniques for multimedia* (pp. 21–49). Springer.

Devlin, J., Chang, M.-W., Lee, K. & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Divya, S., Indumathi, V., Ishwarya, S., Priyasankari, M. & Devi, S. K. (2018). A self-diagnosis medical chatbot using artificial intelligence. *Journal of Web Development and Web Designing*, *3*(1), 1–7.

Feine, J., Morana, S. & Maedche, A. (2019). Leveraging machine-executable descriptive knowledge in design science research–the case of designing socially-adaptive chatbots. In *International conference on design science research in information systems and technology* (pp. 76–91).

Gawriljuk, G., Harth, A., Knoblock, C. A. & Szekely, P. (2016). A scalable approach to incrementally building knowledge graphs. In *International conference on theory and practice of digital libraries* (pp. 188–199).

Gnewuch, U., Morana, S. & Maedche, A. (2017). Towards designing cooperative and social conversational agents for customer service. In *Icis*.

Gopalakrishnan, K., Hedayatnia, B., Chen, Q., Gottardi, A., Kwatra, S., Venkatesh, A., … AI, A. A. (2019). Topical-chat: Towards knowledge-grounded open-domain conversations. In *Interspeech* (pp. 1891–1895).

Goutte, C. & Gaussier, E. (2005). A probabilistic interpretation of precision, recall and f-score, with implication for evaluation. In *European conference on information* *retrieval* (pp. 345–359).

Grainger, T., AlJadda, K., Korayem, M. & Smith, A. (2016). The semantic knowledge graph: A compact, auto-generated model for real-time traversal and ranking of any relationship within a domain. In *2016 ieee international conference on data science and advanced analytics (dsaa)* (pp. 420–429).

Gu, Y., Zhou, T., Cheng, G., Li, Z., Pan, J. Z. & Qu, Y. (2019). Relevance search over schema-rich knowledge graphs. In *Proceedings of the twelfth acm international conference on web search and data mining* (pp. 114–122).

Guo, Q., Zhuang, F., Qin, C., Zhu, H., Xie, X., Xiong, H. & He, Q. (2020). A survey on knowledge graph-based recommender systems. *IEEE Transactions on Knowledge and Data Engineering*.

Han, J., Zheng, L., Xu, Y., Zhang, B., Zhuang, F., Philip, S. Y. & Zuo, W. (2019). Adaptive deep modeling of users and items using side information for recommendation. *IEEE transactions on neural networks and learning systems*, *31*(3), 737–748.

Haruna, K., Akmar Ismail, M., Suhendroyono, S., Damiasih, D., Pierewan, A. C., Chiroma, H. & Herawan, T. (2017). Context-aware recommender system: A review of recent developmental process and future research direction. *Applied Sciences*, *7*(12), 1211.

Hu, B., Shi, C., Zhao, W. X. & Yu, P. S. (2018). Leveraging meta-path based context for top-n recommendation with a neural co-attention model. In *Proceedings of the 24th acm sigkdd international conference on knowledge discovery & data mining* (pp. 1531–1540).

Huang, X., Zhang, J., Li, D. & Li, P. (2019). Knowledge graph embedding based question answering. In *Proceedings of the twelfth acm international conference on web search and data mining* (pp. 105–113).

Hussain, S., Sianaki, O. A. & Ababneh, N. (2019). A survey on conversational agents/chatbots classification and design techniques. In *Workshops of the international conference on advanced information networking and applications* (pp. 946–956).

Kempe, D., Kleinberg, J. & Tardos, É. (2003). Maximizing the spread of influence through a social network. In *Proceedings of the ninth acm sigkdd international conference on knowledge discovery and data mining* (pp. 137–146).

Kerlyl, A., Hall, P. & Bull, S. (2006). Bringing chatbots into education: Towards natural language negotiation of open learner models. In *International conference on innovative techniques and applications of artificial intelligence* (pp. 179–192).

Konstas, I., Iyer, S., Yatskar, M., Choi, Y. & Zettlemoyer, L. (2017). Neural amr: Sequence-to-sequence models for parsing and generation. *arXiv preprint arXiv:1704.08381*.

Laranjo, L., Dunn, A. G., Tong, H. L., Kocaballi, A. B., Chen, J., Bashir, R., … others (2018). Conversational agents in healthcare: a systematic review. *Journal of the American Medical Informatics Association*, *25*(9), 1248–1258.

Latham, A., Crockett, K., McLean, D. & Edmonds, B. (2011). Oscar: an intelligent adaptive conversational agent tutoring system. In *Kes international symposium* *on agent and multi-agent systems: Technologies and applications* (pp. 563–572).

Latham, A., Crockett, K., McLean, D. & Edmonds, B. (2012). Adaptive tutoring in an intelligent conversational agent system. In *Transactions on computational collective intelligence viii* (pp. 148–167). Springer.

Li, W., Bai, Q. & Zhang, M. (2018). Siminer: a stigmergy-based model for mining influential nodes in dynamic social networks. *IEEE Transactions on Big Data*, *5*(2), 223–237.

Logacheva, V., Malykh, V., Litinsky, A. & Burtsev, M. (2020). Convai2 dataset of non-goal-oriented human-to-bot dialogues. In *The neurips'18 competition* (pp. 277–294). Springer.

Lops, P., De Gemmis, M. & Semeraro, G. (2011). Content-based recommender systems: State of the art and trends. *Recommender systems handbook*, 73–105.

Lu, J., Wu, D., Mao, M., Wang, W. & Zhang, G. (2015). Recommender system application developments: a survey. *Decision Support Systems*, *74*, 12–32.

Lu, S., Chen, H., Zhou, X., Wang, B., Wang, H. & Hong, Q. (2018). Graph-based collaborative filtering with mlp. *Mathematical Problems in Engineering*, *2018*.

Ma, W., Zhang, M., Cao, Y., Jin, W., Wang, C., Liu, Y., ... Ren, X. (2019). Jointly learning explainable rules for recommendation with knowledge graph. In *The world wide web conference* (pp. 1210–1221).

Manotumruksa, J., Macdonald, C. & Ounis, I. (2018). A contextual attention recurrent architecture for context-aware venue recommendation. In *The 41st international acm sigir conference on research & development in information retrieval* (pp. 555–564).

Mendonça, V., Melo, F. S., Coheur, L. & Sardinha, A. (2017). A conversational agent powered by online learning. In *Proceedings of the 16th conference on autonomous agents and multiagent systems* (pp. 1637–1639).

Miller, A. H., Feng, W., Fisch, A., Lu, J., Batra, D., Bordes, A., ... Weston, J. (2017). Parlai: A dialog research software platform. *arXiv preprint arXiv:1705.06476*.

Mooney, C. Z. (1997). *Monte carlo simulation* (No. 116). Sage.

Namahoot, C. S., Brückner, M. & Panawong, N. (2015). Context-aware tourism recommender system using temporal ontology and naïve bayes. In *Recent advances in information and communication technology 2015* (pp. 183–194). Springer.

Oku, K., Nakajima, S., Miyazaki, J. & Uemura, S. (2006). Context-aware svm for context-dependent information recommendation. In *7th international conference on mobile data management (mdm'06)* (pp. 109–109).

Panzarasa, P., Opsahl, T. & Carley, K. M. (2009). Patterns and dynamics of users' behavior and interaction: Network analysis of an online community. *Journal of the American Society for Information Science and Technology*, *60*(5), 911–932.

Patel, H. H. & Prajapati, P. (2018). Study and analysis of decision tree based classification algorithms. *International Journal of Computer Sciences and Engineering*, *6*(10), 74–78.

Pazzani, M. J. (1999). A framework for collaborative, content-based and demographic filtering. *Artificial intelligence review*, *13*(5), 393–408.

Pennington, J., Socher, R. & Manning, C. D. (2014). Glove: Global vectors for word representation. In *Empirical methods in natural language processing (emnlp)* (pp. 1532–1543). Retrieved from `http://www.aclweb.org/anthology/D14-1162`

Pérez, J., Arenas, M. & Gutierrez, C. (2009). Semantics and complexity of sparql. *ACM Transactions on Database Systems (TODS)*, *34*(3), 1–45.

Pound, J., Ilyas, I. F. & Weddell, G. (2010). Expressive and flexible access to web-extracted data: a keyword-based structured query language. In *Proceedings of the 2010 acm sigmod international conference on management of data* (pp. 423–434).

Radziwill, N. M. & Benton, M. C. (2017). Evaluating quality of chatbots and intelligent conversational agents. *arXiv preprint arXiv:1704.04579*.

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., … Liu, P. J. (2019). Exploring the limits of transfer learning with a unified text-to-text transformer. *arXiv preprint arXiv:1910.10683*.

Rahman, A., Al Mamun, A. & Islam, A. (2017). Programming challenges of chatbot: Current and future prospective. In *2017 ieee region 10 humanitarian technology conference (r10-htc)* (pp. 75–78).

Ranjbartabar, H., Richards, D., Bilgin, A. A., Kutay, C. & Mascarenhas, S. (2020). User-models to drive an adaptive virtual advisor. In *Proceedings of the 19th international conference on autonomous agents and multiagent systems* (pp. 2117–2119).

Sarwar, B., Karypis, G., Konstan, J. & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on world wide web* (pp. 285–295).

Seering, J., Luria, M., Ye, C., Kaufman, G. & Hammer, J. (2020). It takes a village: Integrating an adaptive chatbot into an online gaming community. In *Proceedings of the 2020 chi conference on human factors in computing systems* (pp. 1–13).

Serban, I. V., Lowe, R., Charlin, L. & Pineau, J. (2016). Generative deep neural networks for dialogue: A short review. *arXiv preprint arXiv:1611.06216*.

Shawar, B. A. & Atwell, E. (2007). Fostering language learner autonomy through adaptive conversation tutors. In *Proceedings of the the fourth corpus linguistics conference.*

Shawar, B. A. & Atwell, E. S. (2005). Using corpora in machine-learning chatbot systems. *International journal of corpus linguistics*, *10*(4), 489–516.

Shum, H.-Y., He, X.-d. & Li, D. (2018). From eliza to xiaoice: challenges and opportunities with social chatbots. *Frontiers of Information Technology & Electronic Engineering*, *19*(1), 10–26.

Song, L., Zhang, Y., Wang, Z. & Gildea, D. (2018). A graph-to-sequence model for amr-to-text generation. *arXiv preprint arXiv:1805.02473*.

Stanovsky, G., Dagan, I. et al. (2015). Open ie as an intermediate structure for semantic tasks. In *Proceedings of the 53rd annual meeting of the association for computational linguistics and the 7th international joint conference on natural language processing (volume 2: Short papers)* (pp. 303–308).

Su, Y., Yang, S., Sun, H., Srivatsa, M., Kase, S., Vanni, M. & Yan, X. (2015). Exploiting relevance feedback in knowledge graph search. In *Proceedings of the 21th acm sigkdd international conference on knowledge discovery and data mining* (pp. 1135–1144).

Sutskever, I., Vinyals, O. & Le, Q. V. (2014). Sequence to sequence learning with neural networks. In *Advances in neural information processing systems* (pp. 3104–3112).

Tegos, S., Demetriadis, S. & Karakostas, A. (2011). Mentorchat: Introducing a configurable conversational agent as a tool for adaptive online collaboration support. In *2011 15th panhellenic conference on informatics* (p. 13-17). doi: 10.1109/PCI.2011.24

Tewari, G., Youll, J. & Maes, P. (2003). Personalized location-based brokering using an agent-based intermediary architecture. *Decision Support Systems*, *34*(2), 127–137.

Tung, H.-W. & Soo, V.-W. (2004). A personalized restaurant recommender agent for mobile e-service. In *Ieee international conference on e-technology, e-commerce and e-service, 2004. eee'04. 2004* (pp. 259–262).

Van Setten, M., Pokraev, S. & Koolwaaij, J. (2004). Context-aware recommendations in the mobile tourist application compass. In *International conference on adaptive hypermedia and adaptive web-based systems* (pp. 235–244).

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., … Polosukhin, I. (2017). Attention is all you need. *arXiv preprint arXiv:1706.03762*.

Veličković, P., Cucurull, G., Casanova, A., Romero, A., Lio, P. & Bengio, Y. (2017). Graph attention networks. *arXiv preprint arXiv:1710.10903*.

Vinyals, O. & Le, Q. (2015). A neural conversational model. *arXiv preprint arXiv:1506.05869*.

Wallace, R. (2018). *Artificial linguistic internet computer entity (alice)(2001).*

Wambsganss, T., Guggisberg, S. & Soellner, M. (2021). Arguebot: A conversational agent for adaptive argumentation feedback.

Wang, C., Zhu, H., Zhu, C., Qin, C. & Xiong, H. (2020). Setrank: A setwise bayesian approach for collaborative ranking from implicit feedback. In *Proceedings of the aaai conference on artificial intelligence* (Vol. 34, pp. 6127–6136).

Wang, H., Zhang, F., Wang, J., Zhao, M., Li, W., Xie, X. & Guo, M. (2018). Ripplenet: Propagating user preferences on the knowledge graph for recommender systems. In *Proceedings of the 27th acm international conference on information and knowledge management* (pp. 417–426).

Wang, H., Zhang, F., Zhao, M., Li, W., Xie, X. & Guo, M. (2019). Multi-task feature learning for knowledge graph enhanced recommendation. In *The world wide web conference* (pp. 2000–2010).

Wang, Q., Mao, Z., Wang, B. & Guo, L. (2017). Knowledge graph embedding: A survey of approaches and applications. *IEEE Transactions on Knowledge and Data Engineering*, *29*(12), 2724–2743.

Wang, X., He, X., Cao, Y., Liu, M. & Chua, T.-S. (2019). Kgat: Knowledge graph attention network for recommendation. In *Proceedings of the 25th acm sigkdd* *international conference on knowledge discovery & data mining* (pp. 950–958).

Weizenbaum, J. (1966). Eliza—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *9*(1), 36–45.

Weizenbaum, J. (1983). Eliza—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, *26*(1), 23–28.

Wiseman, S., Shieber, S. M. & Rush, A. M. (2017). Challenges in data-to-document generation. *arXiv preprint arXiv:1707.08052*.

Xiang, L., Yuan, Q., Zhao, S., Chen, L., Zhang, X., Yang, Q. & Sun, J. (2010). Temporal recommendation on graphs via long-and short-term preference fusion. In *Proceedings of the 16th acm sigkdd international conference on knowledge discovery and data mining* (pp. 723–732).

Xiao, C., Sun, L. & Ji, W. (2020). Temporal knowledge graph incremental construction model for recommendation. In *Asia-pacific web (apweb) and web-age information management (waim) joint international conference on web and big data* (pp. 352–359).

Xiao, C., Xie, C., Cao, S., Zhang, Y., Fan, W. & Heng, H. (2019). A better understanding of the interaction between users and items by knowledge graph learning for temporal recommendation. In *Pacific rim international conference on artificial intelligence* (pp. 135–147).

Xiao, H., Zou, Y., Ng, J. & Nigul, L. (2010). An approach for context-aware service discovery and recommendation. In *2010 ieee international conference on web services* (pp. 163–170).

Xu, Y., Yang, Y., Han, J., Wang, E., Zhuang, F. & Xiong, H. (2018). Exploiting the sentimental bias between ratings and reviews for enhancing recommendation. In *2018 ieee international conference on data mining (icdm)* (pp. 1356–1361).

Yan, R. (2018). " chitty-chitty-chat bot": Deep learning for conversational ai. In *Ijcai* (Vol. 18, pp. 5520–5526).

Yang, S., Wu, Y., Sun, H. & Yan, X. (2014). Schemaless and structureless graph querying. *Proceedings of the VLDB Endowment*, *7*(7), 565–576.

Yoo, S. & Jeong, O. (2020). Automating the expansion of a knowledge graph. *Expert Systems with Applications*, *141*, 112965.

Yuan, S.-T. & Tsao, Y. W. (2003). A recommendation mechanism for contextualized mobile advertising. *Expert systems with applications*, *24*(4), 399–414.

Zadrozny, W., Budzikowska, M., Chai, J., Kambhatla, N., Levesque, S. & Nicolov, N. (2000). Natural language dialogue for personalized interaction. *Communications of the ACM*, *43*(8), 116–120.

Zhang, J. & Pu, P. (2007). A recursive prediction algorithm for collaborative filtering recommender systems. In *Proceedings of the 2007 acm conference on recommender systems* (pp. 57–64).

Zhang, L., Li, W., Bai, Q. & Lai, E. (2021). Graph-based self-adaptive conversational agent. In *Proceedings of the 20th international conference on autonomous agents and multiagent systems* (pp. 1791–1793).

Zhang, W., Gu, T., Sun, W., Phatpicha, Y., Chang, L. & Bin, C. (2018). Travel attractions recommendation with travel spatial-temporal knowledge graphs. In *International conference of pioneering computer scientists, engineers and educators* (pp. 213–226).

Zhou, X., Li, L., Dong, D., Liu, Y., Chen, Y., Zhao, W. X., … Wu, H. (2018). Multi-turn response selection for chatbots with deep attention matching network. In *Proceedings of the 56th annual meeting of the association for computational linguistics (volume 1: Long papers)* (pp. 1118–1127).

Zhu, Q., Zhang, Z., Fang, Y., Li, X., Takanobu, R., Li, J., … Huang, M. (2020). Convlab-2: An open-source toolkit for building, evaluating, and diagnosing dialogue systems. *arXiv preprint arXiv:2002.04793*.

Zhuang, F., Zhang, Z., Qian, M., Shi, C., Xie, X. & He, Q. (2017). Representation learning via dual-autoencoder for recommendation. *Neural Networks*, *90*, 83–89.

Zierau, N., Wambsganss, T., Janson, A., Schöbel, S. & Leimeister, J. M. (2020). The anatomy of user experience with conversational agents: A taxonomy and propositions of service clues.