IJITCE

ISSN 2347- 3657

International Journal of Information Technology & Computer Engineering

Email : ijitce.editor@gmail.com or editor@ijitce.com

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 1, 2024

SQL CHATBOT USING CONTEXT FREE GRAMMER

[1] Kondapalli Krishna , [2] CH. Sai Kumar, [3] C. Pranavi Reddy, [4] CH. Mahendra Sai, [5] A. Sai Kumar

[1]Assistant Professor in Department of CSE Sri Indu College Of Engineering And Technology

krishna.kondapalli25@gmail.com

[2,3,4,5] UG Scholars Department of CSE Sri Indu College Of Engineering And Technology

Abstract

we derive the semantics of a given English query and convert it into its equivalent SQL query. Instead of using neural networks for semantic analysis of English queries, we opt for a Context Free Grammar approach. Most neural network based systems can handle only one semantic at a time, whereas, because of the flexibility offered by our CFG approach, our system manages to handle simultaneous usage of conjunctive, disjunctive, and negative semantics. It also handles complex statements comprising of main as well as dependent clauses. In addition, the system also takes into account aggregate functions and constructs the required GROUP-BY and HAVING clauses. We describe how the system analyzes English queries by understanding the role that each part-of-speech has to play in constructing SQL queries. Numerous examples demonstrate the effectiveness of our approach where state-of-the-art techniques relying on deep learning algorithms fail to deliver.

I INTRODUCTION

Traditionally, relational databases have relied on a standard query language (SQL), a powerful tool for extracting data. But, often times this can be very challenging, specifically for non technical users. Even for power users, expertise is still required to understand the specific schemas, and entity relationships further add to the complexity of using SQL. However, in the real-world, many people use natural languages such as English to ask questions and find solutions to their problems. Programs, for example, Chatbots, have been developed to answer queries or deliver personalized responses to various requests. Chatbots, a word derived from “chat robots”, have dramatically revolutionized how people interact with online information and services, a trend that requires a rethinking of user needs during the development of chatbots. Chatbots are also known by other names, for example: virtual assistants, conversational agents, dialogue systems, personal assistants and conversational interfaces . In recent times, Chatbot adoption in many applications is on the rise due to automation of labor-intensive tasks at low costs, improved customer experiences, and other benefits. According to recent research, the global Chatbot

ISSN 2347–3657

Volume 12, Issue 1, 2024

market is predicted to grow at the rate of 31 percent by 2024 across different sectors such as healthcare, retail, travel and BFSI (Banking, financial services, and insurance services).Additionally, the retail sector is projected to spend $142 billion worldwide by leveraging chatbots.These trends are not driven only by technological advancements but largely by the desire to improve customer service and delivery experiences on ubiquitous devices, social media, customer relationship management and more . With the increase in demand by consumers, the Chatbot ecosystem is rapidly expanding, providing more opportunities to develop technological-focused platforms. The primary goal of this paper is to develop a Chatbot system that provides non-technical users with an interface to interact with RDBMS using English-like, human-readable statements. In this paper, we propose an SQL-Chatbot System (SCB) consisting of mainly three steps. In the first step, the given English query is decomposed into granular components (individual keywords). In the second step, the equivalent SQL query is built. In the final step, the query is executed to return the results back to the user. To the best of our knowledge, our system is the first to use Context Free Grammar (CFG) for an SQL-based Chatbot platform. Our system leverages

Context Free Grammar (CFG) to generate semantically equivalent SQL queries while contextualizing the role of parts-of-speech (POS).

II LITERATURE SURVEY

Chatbots: changing user needs and motivations

Chatbots have been around for decades. However, the real buzz about this technology did not start before the spring of 2016. Two reasons for the sudden and renewed interest in chatbots were (1) massive advances in artificial intelligence (AI) and (2) a major usage shift from online social networks to mobile messaging applications such as Facebook Messenger, Telegram, Slack, Kik, and Viber. The first of these reasons holds promise that intelligent chatbots may well be within reach.Then second concerns service providers' need to reach users in the context of mobile messaging. However, in spite of these drivers, current chatbot applications suggest that conversational user interfaces still entail substantial challenges, in general as well as for the field of human-computer interaction (HCI). Chatbots imply not only a change in the interface between users and technology; they imply changing user dynamics and patterns of use.

An Ontology-Based Dialogue Management System for Banking and Finance Dialogue Systems

Keeping the dialogue state in dialogue systems is a notoriously difficult task. We introduce an ontology-based dialogue manager (OntoDM), a dialogue manager that keeps the state of the conversation, provides a basis for anaphora

ISSN 2347–3657

Volume 12, Issue 1, 2024

resolution and drives the conversation via domain ontologies. The banking and finance area promises great potential for disambiguating the context via a rich set of products and specificity of proper nouns, named entities and verbs. We used ontologies both as a knowledge base and a basis for the dialogue manager; the knowledge base component and dialogue manager components coalesce in a sense. Domain knowledge is used to track Entities of Interest, i.e. nodes (classes) of the ontology which happen to be products and services. In this way we also introduced conversation memory and attention in a sense. We finely blended linguistic methods, domain-driven keyword ranking and domain ontologies to create ways of domain-driven conversation. Proposed framework is used in our in-house German language banking and finance chatbots. General challenges of German language processing and finance-banking domain chatbot language models and lexicons are also introduced. This work is still in progress, hence no success metrics have been introduced yet

A Review of Current Trends in the Development of Chatbot Systems

The main aim of any technological advancement is to lend a helping hand in making lives of humans easy. This is also true for the field of natural language processing. This is also the reason why conversational systems, also called as chatbot systems have gained popularity in the resent times. Chatbot system have been adapted and developed for many different domains. The paper conducts a detailed study of some of the recent chatbot systems/papers developed in different domains. These recent papers have been reviewed while keeping special attention to the type of knowledge given to these systems, the domain for which these systems have been developed, among other parameters in order to understand the recent trends in the development of chatbot systems

III EXISTING SYSTEM

User Interface: Develop a user interface where users can input English queries.

Lexical Analysis:Implement a lexical analyzer to break down the input query into tokens, identifying keywords, identifiers, operators, etc.

Lexical Analysis:Implement a lexical analyzer to break down the input query into tokens, identifying keywords,identifiers, operators, etc.

Semantic Analysis: Implement a semantic analyzer that assigns meaning to the parsed query. Understand the role of each part of speech and derive the corresponding SQL semantics.

SQL Query Generation:Design a module that generates SQL queries based on the derived semantics. Ensure it constructsvalid and meaningful SQL statements. Conjunctions, ***Disjunctions, and Negations***: Enhance your system to handle simultaneous usage of conjunctive, disjunctive, and negativesemantics, as mentioned in your abstract.

Clause Handling and Aggregate Functions:

ISSN 2347–3657

Volume 12, Issue 1, 2024

Extend your system to recognize and handle various SQL clauses (e.g., GROUP BY, HAVING) and aggregate functions.

Database Interaction: Develop a module that interacts with the database to execute the generated SQL queries and retrieve results.

Error Handling:Implement robust error handling to deal with incorrect queries or ambiguous input.

Testing and Validation:Test your system with a diverse set of English queries to ensure correctness and effectiveness. Validate the generated SQL queries against expected results.

Semantic Analysis: Implement a semantic analyzer that assigns meaning to the parsed query. Understand the role of each part of speech and derive the corresponding SQL semantics.

Documentation: Create comprehensive documentation explaining the system architecture, CFG rules, and usage instructions

Comparison with Neural Network Approaches: As per your abstract, compare the performance f your CFG-based approach with neural network Approaches, showcasing situations where your system excels.

Clause Handling and Aggregate Functions: Extend your system to recognize and handle various SQL clauses (e.g., GROUP BY, HAVING) and aggregate functions.

IV PROPOSED SYSTEM

The proposed system, titled "SQL ChatBot – using Context Free Grammar," is designed to derive the semantics of English queries and convert them into equivalent SQL queries. Departing from neural network-based semantic analysis, this system employs a Context Free Grammar (CFG) approach, offering a more flexible solution capable of handling conjunctive, disjunctive, and negative semantics simultaneously. Unlike many neural network systems that struggle with complex sentence structures and dependent clauses, this CFG-based system accommodates such intricacies.

Notably, it incorporates the ability to recognize and handle aggregate functions, constructing necessary GROUP-BY and HAVING clauses. The system's efficacy lies in its analysis of English queries by understanding the distinct roles played by each part-of-speech in constructing SQL queries. The proposed approach is demonstrated through numerous examples, showcasing its effectiveness in scenarios where state-of-the-art techniques reliant on deep learning algorithms may falter. This system aims to provide a robust and versatile solution for translating natural language Queries into SQL, offering advantages in handling diverse semantic structures and complex statements.

International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 1, 2024

V IMPLEMENTATION

User Interface Module:

This module serves as the entry point, providing a user-friendly interface for users to input English queries. It includes components for input validation and presentation of results, creating a seamless interaction between the user and the system.

Lexical Analysis Module:

Responsible for breaking down the input English query into tokens, the lexical analysis module identifies keywords, identifiers, operators, and other relevant elements. It plays a crucial role in preparing the query for subsequent syntactic and semantic analysis.

Syntactic Analysis (CFG Parsing) Module:

The syntactic analysis module utilizes Context Free Grammar (CFG) rules to parse the tokenized input and understand the grammatical structure of the English query. This module establishes the foundation for subsequent semantic analysis by identifying the syntactic elements of the query

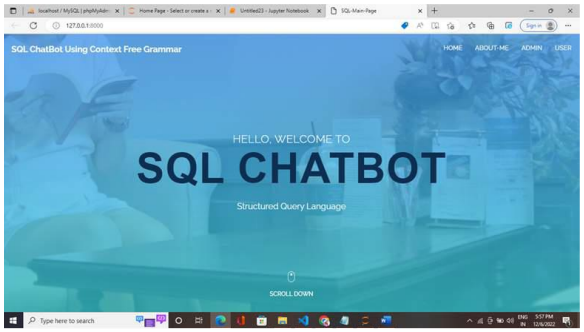
Semantic Analysis Module:

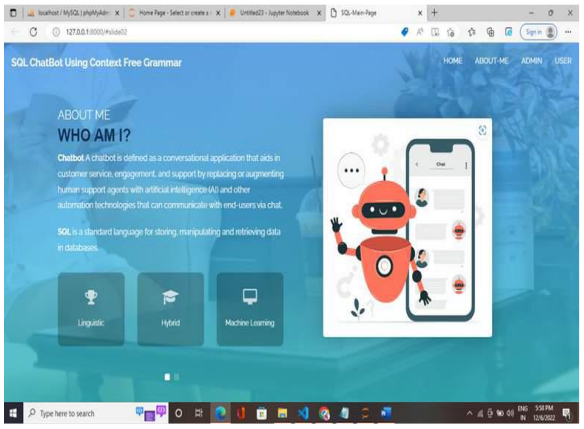
Building on the parsed structure provided by the syntactic analysis module, the semantic analysis Module assigns meaning to the English query. It interprets the roles of each part-of-speech and Derives the corresponding semantics, setting the stage for the generation of the equivalent SQL Query.

SQL Query Generation Module:

This module takes the derived semantics from the semantic analysis phase and constructs the equivalent SQL query. It maps the recognized elements of the English query to SQL keywords,operators, and clauses, ensuring the generated SQL statement is syntactically correct and semantically aligned with the user's intent

VI RESULTS

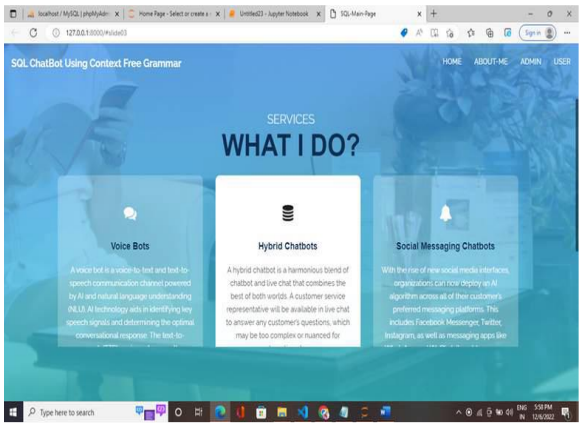


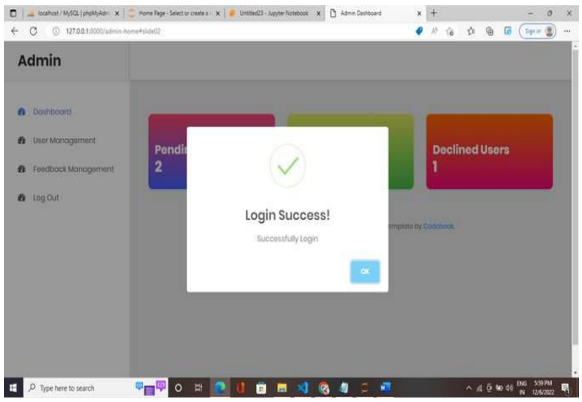


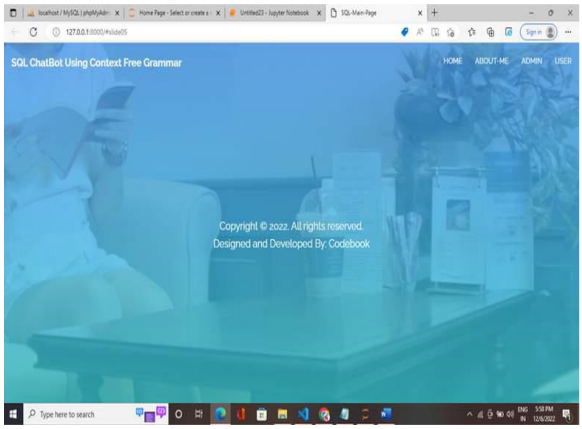
International Journal of Information Technology & Computer Engineering

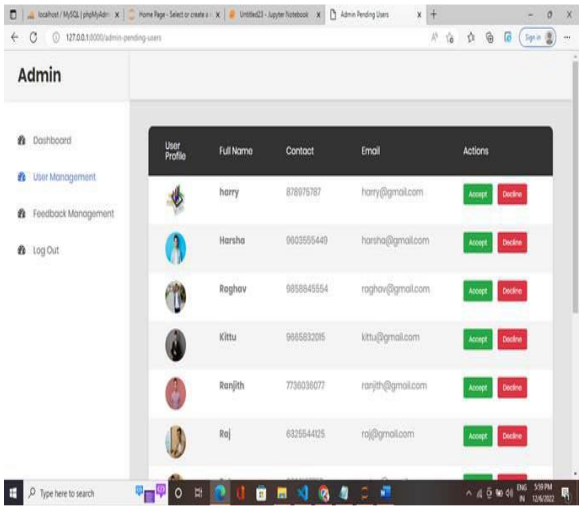
ISSN 2347–3657

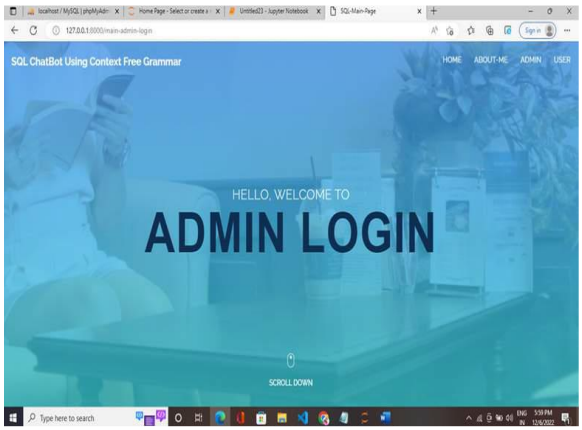
Volume 12, Issue 1, 2024

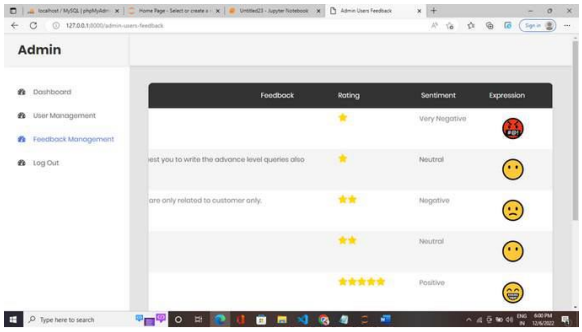








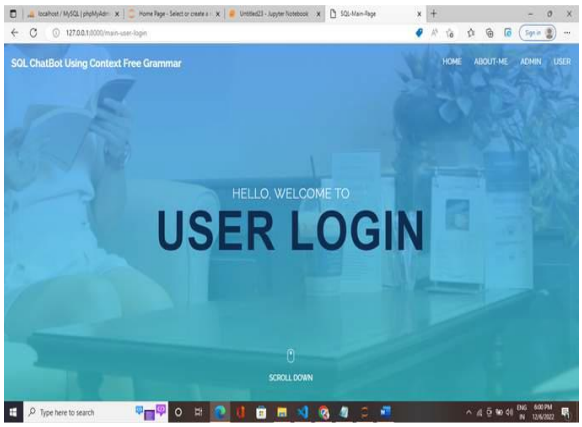


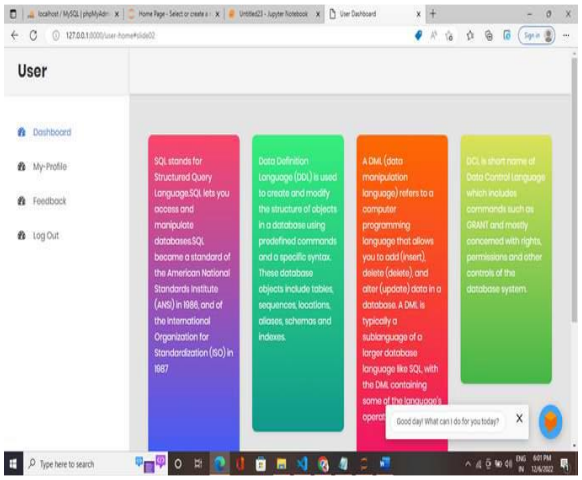


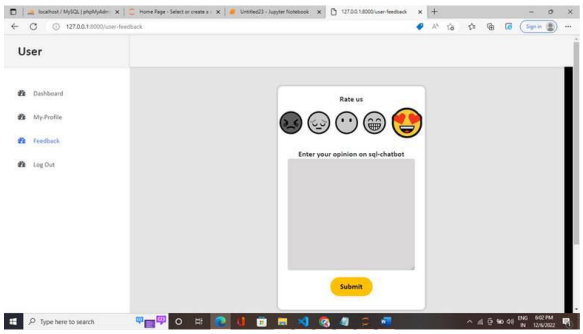
International Journal of Information Technology & Computer Engineering

ISSN 2347–3657

Volume 12, Issue 1, 2024







VII CONCLUSION

As discussed, the extended keyword search facility in our search engine initially deduces the structure of an English query with the help of CFG. Conditions within the original query are then

grouped/parenthesized in a way so as to preserve the precedence of the involved operators, and

hence the overall semantics of a query. The Framing Procedure goes a step ahead to deal with queries having more than one subordinate clause by informing the search engine of the way clauses of complex English queries can be associated. Roles played by parts of speech in composing English statements are understood and applied in the context or terminology of relational databases so as to construct their equivalent SQL query. The paper also illustrates the technique used to identify and further map the variety of attributes used in the English query statements to their SQL counterparts.The Clause-Construction algorithm is then introduced to detect the presence of any aggregate functions in the query, and if needed, to construct the required GROUP-BY and HAVING clauses. By default, the search system allows the simultaneous use of disjunctive (OR), conjunctive (AND), negative (NOT) semantics, along with the possible combinations of NAND and NOR. Thus, the semantics derived by GA are used to construct the WHERE clause. The construction mechanism of SELECT and FROM

clauses works in a way that is similar to traditional DBKWS. This paper also refines and exploits the list of delimiters and connectors used for decomposing the English queries. Overall, the paper demonstrates how the CFG approach enables extended keyword search and benefits DBKWS and NLIDB by capturing contextual information in English queries. In upcoming work, we plan to explore and compare the performance in construction of queries of deep-learning approaches using sentiment-analysis against our SCB system.

REFERENCES

[1] P. B. Brandtzaeg and A. Følstad, "Chatbots: changing user needs and motivations," interactions,
vol. 25, no. 5, pp. 38–43, Aug. 2018, doi: 10.1145/3236669.

[2] D. Altinok, "An Ontology-Based Dialogue Management System for Banking and Finance Dialogue Systems," p.9.

[3] T.P. Nagarhalli, V. Vaze, and N. K. Rana, "A Review of Current Trends in the Development of Chatbot Systems," in 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), Coimbatore, India, Mar. 2020, pp. 706-710.
doi:10.1109/ICACCS48705.2020.9074420.

[4] "Chatbot Market Size & Share | Growth Forecast Report 2024," Global Market Insights Inc. https://www.gminsights.com/industry-analysis/chatbotmarket (accessed Mar. 18, 2022).

[5] I. Intelligence, "Chatbot market in 2022: Stats, trends, and companies in the growing AI chatbot industry," Business Insider. https://www.businessinsider.com/chatbot-market-stats-trends (accessed
Mar. 18, 2022).

[6] Clickatell, "Chatbot Market to grow at 31 Percent CAGR from 2018 to 2024," Supply and Demand Chain Executive, Jul. 02, 2018. https://www.sdcexec.com/softwaretechnology/news/21011880/chatbotmarket-to-grow-at-31-percent-cagr-from-2018-to-2024 (accessed
Mar. 18, 2022).

[7] M. J. Minock, "A STEP Towards Realizing Codd's Vision of Rendezvous with the Casual User," p. 4, 1977.

[8] V. Zhon C. Xiong, and R. Socher, "Seq2SQL: Generating Structured Queries from Natural
Language using Reinforcement Learning," arXiv:1709.00103 [cs], Nov. 2017, Accessed: Mar. 18,
2022. [Online]. Available: http://arxiv.org/abs/1709.00103

[9] T. Yu et al., "Spider: A Large-Scale Human-Labeled Dataset for Complex and Cross-Domain Semantic Parsing and Text-to-SQL Task," arXiv:1809.08887 [cs], Feb. 2019, Accessed: Mar. 18, 2022. [Online]. Available: http://arxiv.org/abs/1809.08887