Connecting People Through Virtual Assistant on Google Assistant

Anurag Batra, Ankit Yadav, and Sudhir Kumar Sharma(✉)

Institute of Information Technology and Management, Delhi, India
anuragbatra1999@gmail.com, ankit4607@gmail.com,
sudhir_sharma99@yahoo.com

Abstract. Technology is a major driving force in our lives today. Social media platforms like Facebook, Twitter, Whatsapp, Hike etc. have taken the world by storm and have become a nonseparable part of life. Google assistant (GA) is an Artificial Intelligence (AI) based virtual assistant developed by Google. Initially, the GA lacked the feature of chatting with random people. This research paper is about developing such a feature for GA by creating an "Action on Google". Action on Google is a custom built application that adds features to Google Assistant. This application (named as Chat Hub) connects two random people available online. People are free to chat as per their area of interest. Since GA is inbuilt into all android devices, the subscriber count is increasing significantly each day. The chat facility was introduced in GA on October 3, 2018. The objective of developing such a feature is to make GA more user friendly and benefit people by helping to find others with similar interests.

Keywords: Artificial Intelligence · Google Assistant · NLP · Python · Firebase

1 Introduction

Artificial Intelligence (AI) has become an inseparable element of computer science today. AI lies at the centre of intelligent computation whether one asks to play his/her favourite song or to guess a buyer's likes based on his/her shopping history. AI is the attempt to make machines intelligent. The goal could be anything. But the approach is to make it think, evaluate and react like a human. In technical terms, an AI based machine is capable of interpreting data from its environment and react accordingly in order to achieve its goal.

An AI machine must have the following traits: Knowledge, Problem Solving, Learning, Planning, Perception etc. Knowledge is the centre point of all AI enhancements. A machine simply cannot act without a knowledgebase which helps to evaluate options and make decisions. Like any other machine, AI machine must be capable of solving problems but on a larger dimension. A standard computer program is capable of taking inputs and providing specific outputs. An AI program not only solves its intended problem but is also capable of learning from past input (machine learning), adapting its behaviour and solve generalized class of problems. AI utilizes sensor inputs to evaluate the world around. This includes various methodologies like facial recognition, object

P. K. Singh et al. (Eds.): *Proceedings of ICETIT 2019*, LNEE 605, pp. 407–417, 2020.
https://doi.org/10.1007/978-3-030-30577-2_35

detection, voice recognition, gesture recognition etc. The field of robotics is totally dependent on AI as a robot needs to think in order to perform its tasks.

According to Williams [1], the goal of AI is to develop automation. The automation is focussed on tasks that require human intelligence. As humans are capable of numerous kinds of intelligence, researchers have tried to identify different ways of thinking which has led to subdivision of the AI field. William defines few central issues related to AI. Firstly, effective AI is possible only with the availability of knowledge in the form that can be readily applied to whatever problems that may occur. Secondly, AI requires the ability to draw conclusions from available knowledge.

Liu et al. [2] describe the growth of AI in the current 21st century. The authors mention the famous "Turing Test" to define the intelligence of a machine. In order to study the growth rate of AI, metrics have been defined to measure the number of authors, growth rate, authors per paper and papers per author. It is observed that between 2002 and 2015, there has been a significant explosion of publication in AI giving birth to numerous sub-fields within it. There was also an increase in number of published authors during same period. The increase in publications resulted in the increased number of authors. There was also an even more dramatic boost in the number of citations in the same period. This shows how authors were constantly exploring and learning from other people's work. To conclude, it is realized that with increase in publications, there was an increase in collaboration with other AI researchers and a decrease in average productivity of individual researcher. AI gradually became an open playing field.

Google's virtual assistant (GA) lacked the feature of chatting with random people. This research paper is about developing such an application for GA (names as Chat Hub). The application has been developed successfully and is integrated to GA since October 3, 2018.

The upcoming sections of this paper are structured as follows. Section 2 describes the related works. Section 3 discusses the experimental setup for developing the application. Section 4 is about the architecture used, algorithms implemented and python code used to build the application. Section 5 discussed the results and future scope of work presented in this paper. Conclusion has been presented in Sect. 6.

2 Related Work

A virtual assistant (VA) is a software agent with one goal, to assist. Just like any real life assistant, VA is meant to take verbal commands and perform operations. Some VAs are capable of replying in pre-recorded voices. A user may require a variety of tasks to be carried out, such as asking questions (online search), controlling home automation devices (switch on the air conditioner remotely), media playback, emails, to-do lists, appointments etc.

The popularity and usage of VA has been expanding significantly in recent years. Famous companies like Google, Amazon, Appleetc have all developed VA projects to enhance their products. Google has developed Google Assistant. Microsoft has launched Cortana. Apple has got Siri. Samsung handheld device feature Bixby. Besides the usual capability of call/text, managing emails, calendars, these VA are capable of engaging in making informed suggestions, cracking funny one liner jokes and much more. These VA use artificial intelligence principles to run and have become important part of our digital lives. The table below shows feature comparison between Google Assistant, Siri and Cortana. Y stands for yes meaning the given feature is available for the given virtual assistant. N stands for No.

Music service	Alexa	Google assistant	Siri
Chat Hub application	N	Y	N
Weather updates	Y	Y	Y
Calling	N	Y	Y
Maps	N	Y	Y
Schedule appointments	N	Y	Y
Amazon music	Y	N	N
Apple music	Y	iOS only	Y
Youtube music	N	Y	N
Google Play music	N	Y	N
Continuous talk	Y	Y	N

Popenici et al. [3] explore the impact of AI in teaching and higher education. The development of AI is bound to impact the learning in higher education. It is essential that the universities develop a mechanism to integrate "Intelligent Learning" into the education system and combine efforts with AI service providers to make learning experience fulfilling for the student.

Wang et al. [4] propose building a virtual assistant based on eight basic emotions and drives to reflect a person's intent. Using facial expression and voice recognition, different psychologies have been integrated into the VA. It is shown that we can predict the underlying psychology using the facial and voice inputs.

In [5], Rodriguez et al. introduce T-BOT and Q-BOT, two virtual assistants with the purpose of guiding students and teachers on e-learning platform. The T-BOT is intended to help students access the learning platform resources. The Q-BOT helps teachers to monitor and evaluate student progress.

Casal et al. [6] merge the world of intelligent PA with Internet of Things (IoT). The idea is to have intelligent virtual assistants interact with both users and smart devices. Interacting with smart devices is an important source of information and learning that improves the knowledgebase of the virtual assistant and thereby enhance its capability to serve users.

Okamoto et al. [7] evaluate the impact of various personal assistant abilities on the behaviour and performance of an organization. The organization performance is measured on parameters like success rate, reward and speed (difference between deadline and delivery date).

Campagna et al. [8] address the challenges of privacy, generality, usability and interoperability in virtual assistant technologies by proposing a programmable virtual assistant that uses Internet of Things to provide online services.

Pereira et al. [9] use anthropomorphic virtual assistant to improve adherence to oral anti-diabetics for senior citizens. Campagna et al. [10] propose personal virtual assistants responsible for sharing consumer data through ThinkTalk commands from their web accounts.

Sokol et al. [11] explore the area of automated decision making using voice-enabled Glass Box which users can interact regarding automated decisions. Oliveira et al. [12] survey existing Ambient Assisted Living technologies to propose intelligent assistants capable of helping visually impaired people for locating objects and avoid collision with objects placed at wrong locations.

3 Experimental Setup

This section presents the tools and open libraries used to setup and manage databases. Hardware configuration for development of this project includes an Intel Core i3 processor clocked at 3.36 GHz and 8 GB DDR4 RAM. This section is further divided into two subsections:

3.1 Languages Used

Python. Python is an interpreted, high-level, general-purpose programming language. That was created by Guido van Rossum and first released in 1991. Python has a design philosophy that emphasizes code readability, notably using significant whitespace and tabs. It provides constructs that enable clear programming on both small and large scales.

PHP. It stands for Hypertext Pre-processor and it is a server-side scripting language designed for Web development. It was originally created by RasmusLerdorf in 1994 and now its reference implementation is now produced by The PHP Group.

Node.js. Node.js is an open-source, cross-platform JavaScript run-time environment that executes JavaScript code outside of a browser i.e., at the backend servers. It was built using chrome v8 engine.

3.2 Libraries and Framework Used

Firebase Realtime Database. The Firebase Realtime Database is a cloud-hosted NoSQL database owned by Google that lets you store and synchronize data between your users at real-time with minimum latency. That makes it the most suitable database for using it for chatting database.

PyreBase. Firebase is a backend as a service used to build Android Apps, iOS Apps and web development too without creating a custom backend. It includes various libraries for authentication, storage, database, notifications etc. Pyrebase is an unofficial Python wrapper for all those firebase SDKs.

Time Module. Python has predefined a module, “time” which allows us to handle various operations regarding time, its conversions and representations, which find its use in various applications in life. The beginning time is set to 12:00 am, 1 January, 1970 and this time is termed as “epoch” in Python.

Operating System. The OS module in Python provides a way of using operating system dependent functionality. The functions that the OS module provides allows you to interface with the underlying operating system that Python is running on – be that Windows, Mac or Linux.

JSON. In computing, JavaScript Object Notation (JSON) is an open-standard file format that uses human-readable text to transmit data objects from one location to another they consist of attribute-value pairs and array data types.

4 Development of Chat Hub Application Based on Three Tier Architecture

The chat bot works on the basis of three tier architecture comprising of the client, server and database. Such architecture helps in decentralizing the bot structure and ensures the security and authenticity of database by blocking direct access (Fig. 1).

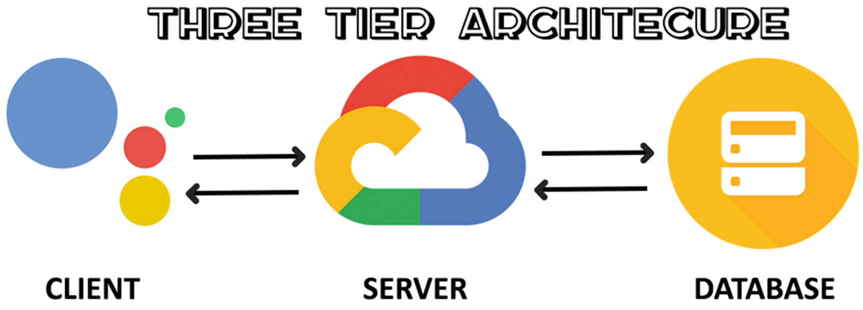


Fig. 1. Three tier architecture

Authors developed a web application in July 2018 which executed for two months with a monthly user interaction of eight thousand. The purpose of this application was to connect two random users for chatting. Later it was realized that such functionality was better suited as a feature for Google Assistant. Therefore, the web application was dropped and the new approach was adopted.

In the development phase of Chat Hub application, a selective group of users were taken and their cell phones were connected via Google Assistant. This setup was used to select two random users for chatting. Here, random implies “first come first serve” i.e., first user is connected to second, third to fourth, and so on. The above task was achieved using APIs and white listed Google IDs used to access logged in devices.

The section below describes the Action on Google functionality based on the three tiers described above.

4.1 Client Side

User calls the Action in Google Assistant on his/her device using voice command. The voice command is converted to text response using native API in the user device automatically. This text is then sent to the server where Google servers interact with the request and pass the request to dedicated servers for further processing which finally, provide a useful response for user queries.

4.2 Server Side

Once the client request reaches the server side, the Google servers process the text response using Natural Language Processing. The purpose is to categorize each text into intent. Finally, the processed message is sent as a post request to dedicated servers. The forwarded message includes its intent name.

4.3 Working with Database

The post request is received in JSON format. This JSON data is further processed for authentication. This is to ensure that data is coming from Google itself. After authentication, it is parsed by PHP to extract useful information. The code below describes the extraction process.

Pseudocode: index.php

```
1. // receive JSON response
   $requestBody = file_get_contents('php://input');

2. // Decode
   $json = json_decode($requestBody);

3. Assign useful data into the following variables: session id, response id, message and request intent.

4. Useful data fetched from JSON file is sent to *script.py* using a curl request.

5. After processing the received data *script.py* replies back which will be used to generate new JSON similar to request JSON.
```

The python program script.py includes a logic that checks whether the user was connected to anyone, what messages/responses exist for ongoing conversations and store new messages in Firebase Realtime database.

4.4 Flowchart

This subsection provides the flowchat of the algorithm developed for implementing Chat Hub application (Fig. 2).

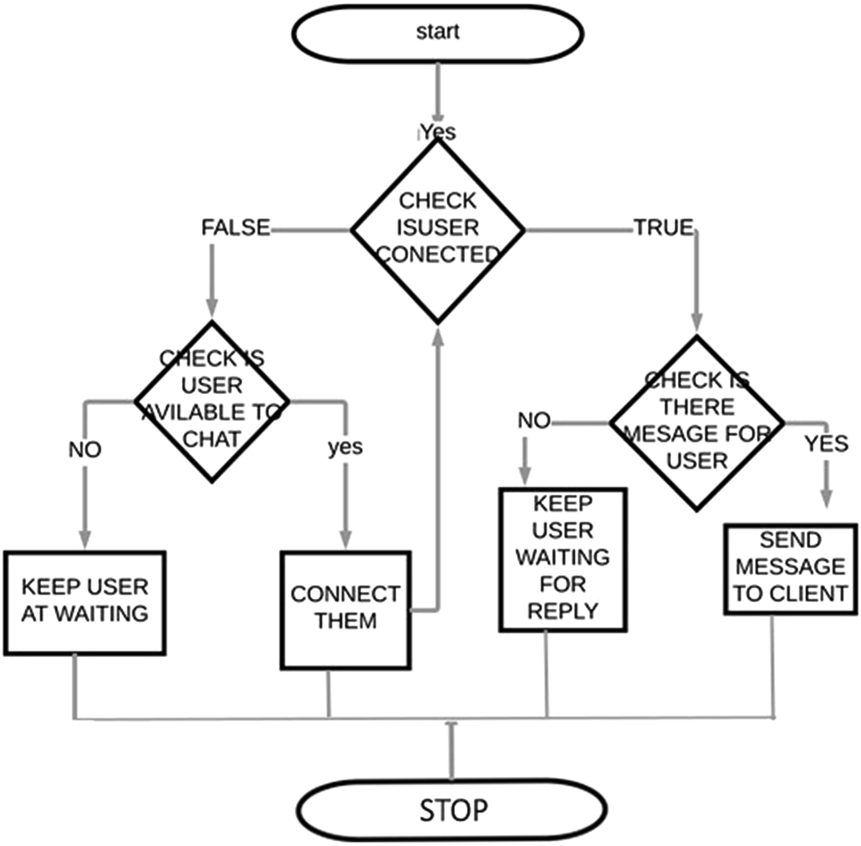


Fig. 2. Flowchart for script.py

4.5 Pseudocode

This section provides the pseudocode for Chat Hub python program. The detailed steps used in this program have been written in pseudocode mentioned below.

Pseudocode: chathub.py

```
1. Initialization: Database connection established.
2.  If (First User not connected to anyone):
         if (No User is in waiting state):
                 Keep First User in waiting state and display message
                 “Kindly wait for 30 seconds and then try again. Working on
                 finding a new friend.”

                  if (No User available):
                          Display message “Finding your perfect friend. Why
                          don’t you tell your friends about Chat Hub. It will
                          ensure no one has to wait this long.”

3.       else:
           Connection established between two users (First User and Second
         User).
                Display to Second User "Initiating Chat. Please Be Patient."
         Remove waiting state for First User.
                 Display to First User, "Be polite and respectful in greeting
                 your random friend.”
                 If (Any User sends multiple messages)
                          Display to that user “You are only allowed to send
                          one message at a time. Kindly wait for a reply from
                          the other end."
4. else:
         Check database for user messages.
         If (no message is available):
                 Display message "Please wait for 10 sec and then try again.”
         Else:
           Send message to other user.

5. Messages are sent back to index.php in form of variables.

 6. Store all messages in database that were exchanged between each pair of
      users to prevent spam and for checking abuse reports.
```

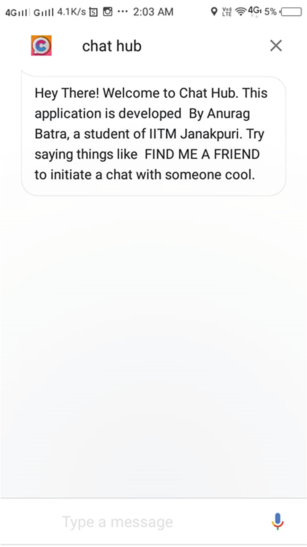
So far, user has been validated and his/her response has been stored. Next, a chat response is received from the other user. The response is then sent back to PHP in the form of variables where it is converted into a new JSON response.

This JSON response is automatically grabbed by the Google servers and sent to the client side. Hence, the first user is finally able to receive reply from second user. This complete process is repeated multiple times for the entire duration of the chat.

5 Results

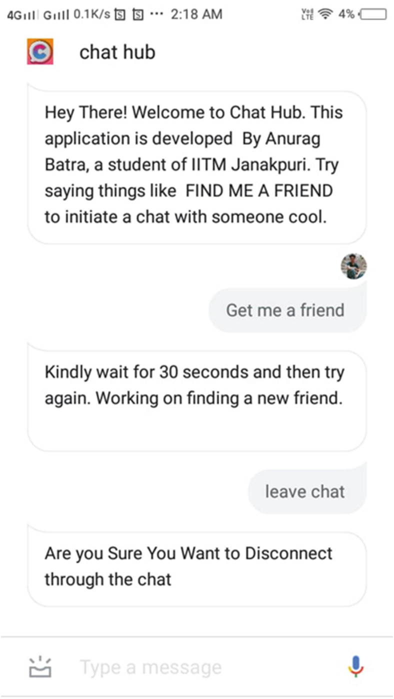
The first draft of this application was submitted on September 1, 2018. The alpha version was released on October 3, 2018. After numerous revisions, the application is currently running on its eighth version which was released on October 8, 2018.

Figure 3 shows the screenshot of welcome message when user opens Chat Hub via Google Assistant. Figure 4 describes the flow of interaction between user and the application. The user asks to find a chat partner by saying expressions like "Find me a Friend". While the application searches for potential chat partners the user is asked to wait for 30 s before he/she can try again. Once two users are connected, they can initiate chat based on their interests.



Hey There! Welcome to Chat Hub. This application is developed By Anurag Batra, a student of IITM Janakpuri. Try saying things like FIND ME A FRIEND to initiate a chat with someone cool.

Fig. 3. Welcome message of Chat Hub application



Hey There! Welcome to Chat Hub. This application is developed By Anurag Batra, a student of IITM Janakpuri. Try saying things like FIND ME A FRIEND to initiate a chat with someone cool.

Get me a friend

Kindly wait for 30 seconds and then try again. Working on finding a new friend.

leave chat

Are you Sure You Want to Disconnect through the chat

Fig. 4. Functioning of Chat Hub application

6 Conclusion and Future Work

The aim of this paper was to describe the implementation of "Chat Hub", a random Chat Application on Google Assistant. This was successfully implemented using Python programming language, FireBaseRealTime Database, PHP and JSON technologies. This project can be enhanced in a number of ways such as categorizing messages based on their intent, applying data analytics to gain valuable insights. Work can also be done on data to track user interests.

For future work, the application can be augmented to analyse stored data from chat history using data analytics and machine learning. Such analysis will be crucial to classify messages based on user interests, purpose of conversations and spam prevention. User interests can further be used to find relevant chat partners and thereby improve user experience. Spam prevention can be implemented using data analytics to block users who use unpleasant language and have conversations that produce negative results of sentimental analysis.

References

1. Williams, C.: A brief introduction to artificial intelligence. In: OCEANS'83, pp. 94–99. IEEE (2019)
2. Liu, J., Kong, X., Xia, F., Bai, X., Wang, L., Qing, Q., Lee, I.: Artificial intelligence in the 21st century. IEEE Access **6**, 34403–34421 (2018)
3. Popenici, S.A., Kerr, S.: Exploring the impact of artificial intelligence on teaching and learning in higher education. Res. Pract. Technol. Enhanced Learn. **12**(1), 22 (2017)
4. Wang, Z., Cheng, N., Fan, Y., Liu, J., Zhu, C.: Construction of virtual assistant based on basic emotions theory. In: International Conference on Affective Computing and Intelligent Interaction, pp. 574–581. Springer, Heidelberg (2005)
5. Rodriguez, E., Burguillo, J.C., Rodriguez, D.A., Mikic, F.A., Gonzalez-Moreno, J.C., Novegil, V.: Developing virtual teaching assistants for open e-learning platforms. In: 19th EAEEIE Annual Conference, pp. 193–198. IEEE (2008)
6. Santos, J., Rodrigues, J.J., Casal, J., Saleem, K., Denisov, V.: Intelligent personal assistants based on internet of things approaches. IEEE Syst. J. **12**(2), 1793–1802 (2018)
7. Okamoto, S., Sycara, K., Scerri, P.: Personal assistants for human organizations. In: Handbook of Research on Multi-Agent Systems: Semantics and Dynamics of Organizational Models, pp. 514–540. IGI Global (2009)
8. Campagna, G., Ramesh, R., Xu, S., Fischer, M., Lam, M.S.: Almond: the architecture of an open, crowdsourced, privacy-preserving, programmable virtual assistant. In: Proceedings of the 26th International Conference on World Wide Web, International World Wide Web Conferences Steering Committee, pp. 341–350 (2017)
9. Pereira Guerreiro, M., Brito Félix, I., Cavaco, A., Cláudio, A.P., Mendes, A., Balsa, J., Henriques, A.: Development of a complex intervention to improve adherence to antidiabetic medication in older people using an anthropomorphic virtual assistant software. Front. Pharmacol. **10**, 680 (2019)
10. Campagna, G., Xu, S., Ramesh, R., Fischer, M., Lam, M.S.: Controlling fine-grain sharing in natural language with a virtual assistant. Proc. ACM Interact. Mob. Wearable Ubiquit. Technol. **2**(3), 95 (2018)
11. Sokol, K., Flach, P.A.: Glass-box: explaining AI decisions with counterfactual statements through conversation with a voice-enabled virtual assistant. In: IJCAI, pp. 5868–5870 (2018)
12. Oliveira, J.D., Bordini, R.H.: A survey on the needs of visually impaired users and requirements for a virtual assistant in ambient assisted living. In: 16th International Conference on Information Technology-New Generations (ITNG 2019), pp. 449–453. Springer, Cham (2019)