



**INDEX**

|  |  |  |
| --- | --- | --- |
| **NO** | **TOPIC** | **PAGE NO** |
| **1** | **INTRODUCTION** | **5** |
| **2** | **ABSTRACT** | **6** |
| **3** | **CONTEXT ABOUT GOOGLE DUPLEX** | **8** |
| **4** | **ARCHITECTURE** | **9** |
| **5** | **ADVANTAGES** | **10** |
| **6** | **DISADVANTAGES** | **11** |
| **7** | **DNNS AND RNNS** | **12** |
| **8** | **CONDUCTING NATURAL CONVERSATIONS** | **13** |
| **9** | **DUPLEX IS GETTING SMARTER AND MAKING LIFE A LITTLE EASIER** | **15** |
| **10** | **CLOSED DOMAINS AND VANISHING GRADIENT PROBLEM** | **16** |
| **11** | **ADVANCE AI FOR EVERYONE** | **18** |
| **12** | **GOOGLE DUPLEX AND ASSISTANT** | **19** |
| **13** | **PROCESS FLOW AND CONCLUSION** | **20** |

**1.INTRODUCTION**

A digital assistant is a program capable of responding to textual commands or voice orders given by an individual. It is an AI powered software-based agent which can assist humans in performing day to day activities such as scheduling appointments, making calls, typing messages and operating other smart devices. Digital assistants are advanced form of chatbots and are capable of responding to complex questions requiring multiples source of information. They use Artificial intelligence, machine learning, natural language understanding and processing to assist the human user. We use digital assistants like Siri and Google voice assistant on our smartphones, Cortana on windows computers and Alexa, Google home on smart speakers. Google duplex is an advanced digital assistant which debuted on May 2018. It is a voice assistant system that works on artificial intelligence to accomplish real world tasks over the phone. It is an Artificial intelligence system which can replicate human intelligence by understanding and responding to voice conversations over the phone. Unlike other voice assistants which sound robotic, google duplex uses a natural-sounding human voice. On top of that the system can understand complex sentences, fast questions and long remarks. The system makes use of both machine learning and AI to conduct voice calls in the most human-like nature. Google showcased the assistant at the annual developer Conference-Google I/O 2018. The CEO of Google Mr. Sundar Pichai demonstrated the voice assistant system by calling a restaurant and a salon for reservation. Duplex went into action by calling the restaurant on behalf of the user, reserved the tables and notified the user of the reservation. This demonstration was interesting and spooky at the same time as the other parties receiving the call were not aware that they were actually responding to a robot. Google started rolling out the service to a select few US cities from November of 2018. As of 2019, It is available in 48 US states. Google is taking a steady approach towards expanding this system that is why it is still in the developmental stage. This system was recently expanded to New Zealand as a pilot project. Google Duplex would call local businesses in the country and would confirm working hours and holidays to help update them in the google maps for the ease of the consumers. It would not only help businesses update their operating hours but would also bring in more consumers. Here we have an Artificial Intelligence to talk about and it is a **Google Duplex**. In today’s world, every technology is getting smart with the help of smart ideas. You must have heard Google Assistant because it is our part of our life. Google Assistant does some small work for us like telling time, giving information about weather, setting an alarm, after asking some questions, it shows some accurate results, and much more. Google Assistant works for the user. We command them to complete some work by just saying **“Hey Google” or “Ok Google”.**

## **2. ABSTRACT**

Google Duplex is an Artificial Intelligence(AI) chat agent that can carry out specific verbal tasks, such as making a reservation or appointment, over the phone. It works to conduct natural conversations to accomplish certain types of tasks. It is a new technology for conducting natural conversations to carry out “real world” tasks over the phone. The technology is directed towards completing specific tasks, such as scheduling certain types of appointments. For such tasks, the system makes the conversational experience as natural as possible, allowing people to speak normally, like they would to another person, without having to adapt to a machine. Google Duplex is not able to carry out random casual conversation. Google Duplex is the technology behind a new Google Assistant feature. Duplex is a completely automated system that places calls on your behalf, complete with a natural-sounding human voice instead of a robotic one. Google Duplex is a new project from Google that is currently live in the majority of the US. It allows certain users to make a restaurant reservation by phone. However, instead of the user speaking directly to the restaurant employee, Google Duplex, with the help of Google Assistant, speaks for the user. This was no ordinary booking. It came through Google Duplex, a free service that uses artificial intelligence to call restaurants and mimicking a human voice speak on our behalf to book a table. The feature, which had a limited release about a year ago, recently became available to a large number of Android devices and iphones

## 3 CONTEXT ABOUT GOOGLE DUPLEX

• Defining a natural conversation

A natural conversation can be described with the following characteristics:

* Speaker is exhibiting goal-directed, cooperative, rationalbehavior.
* Speaker is using the appropriate tone.Speaker can understand and control the conversational flow and use the right timing.
* What is Google Duplex?

Google Duplex is the latest Al-voice based project from Google that is currently existing in the U.S (majorly in every area). It can provide power to individual users that they can restaurant reservations using smart-phones; in it, the user speaks indirectly to the restaurant employee, via Google Assistant. Google Assistant speaks on behalf of the user with an Al-based human-sounding voice software.

## • Google Duplex Applications

Google Duplex helps with these issues by granting the computer to have natural communication with a human. The A.I. system fine tunes to the person, instead of the person fine-tuning to the machine. Therefore, the person can generally utter, just as they would if they were communicating with another human. Google Duplex also generates it, so the computer systems sound like a real person. It uses a natural accent, as well as words and phrases like "um" and "uh," just like a human would. During a conversation. the A.I. system can also handle interruptions and elaborate.

How does Google Duplex model natural conversations?

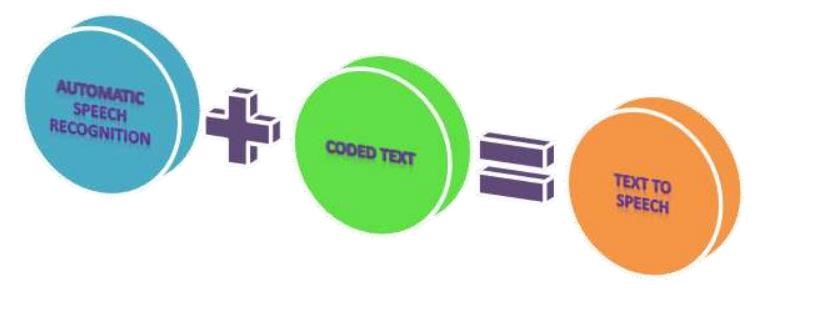
* Duplex uses a deep neural network (DNN); in more complex cases, it makes use of a recurrent neural network (RNN) which is more expensive, but better at modeling language.
* At the core of Duplex is a recurrent neural network (RNN)designed to cope
* with these challenges, built usingTensorFlow Extended (TFX).

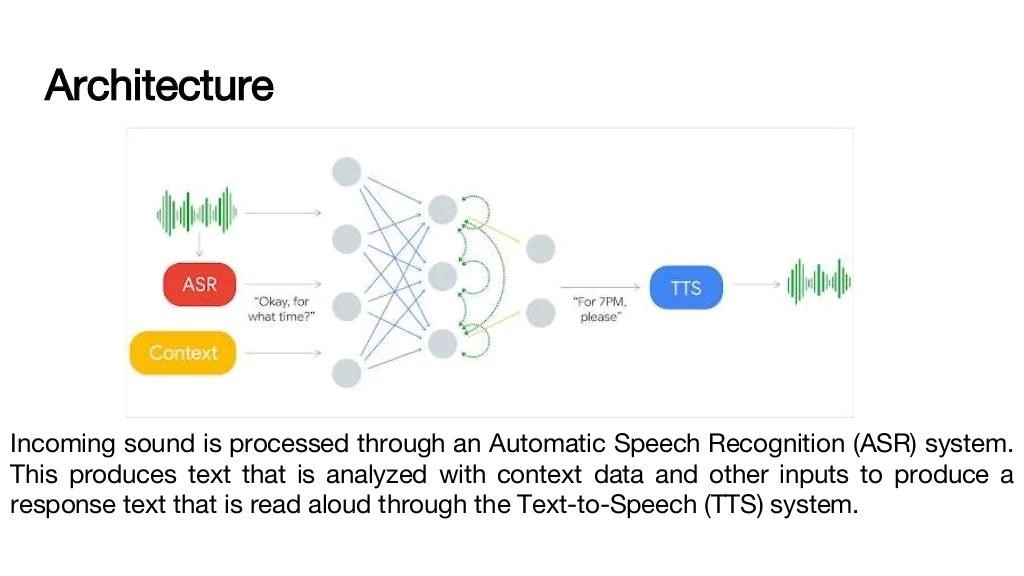
### TECHNIQUES BEHIND GOOGLE DUPLEX

A. TensorFlow Google’s TensorFlow — that ease the process of acquiring data, training models, serving predictions, and refining future results. The most famous deep learning library in the world is Google's TensorFlow. TensorFlow package of machine learning and deep learning models and algorithms and makes them helpful by way of a common figure of speech. Python is the front-end API for building applications in this framework.

All the Google product uses machine learning for the improvisation of search engine, translation, image capturing. Google Duplex’s conversations sound natural gratefulness to move ahead for accepting, interacting, timing, and verbalizing. The center of Duplex is a RNN (Intermittent neural system), which was manufactured utilizing TensorFlow Expanded (TFX).TensorFlow is a platform to run machine learning models.

To reach the high quality, we trained Duplex RNN on a couple of unknown phone discussion data. The system uses the Google’s automatic speech recognition technology (ASR) as well as the features of the audio, history of dialogue, the limitation of conversation etc. Continuously the system is trained for the fine understanding of the end user. At last Google used Hyper parameter optimization from TFX to the next improvisation. The oral language part uses text to speech (TTS) engine and a synthesis TTS engine





## **5.Benefits Of Google Assistant Duplex**

* One major benefit of Google Duplex is that it can save users time and effort by handling tasks that would otherwise require phone calls and appointments. For example, Duplex can be used to book a table at a restaurant, schedule a hair appointment, or make a reservation at a hotel, all without the user having to pick up the phone and make the call themselves.
* This is particularly helpful for busy individuals who don't have the time or patience to make phone calls, or for people with disabilities or impairments that make it difficult for them to use the phone. By automating these tasks, Google Duplex can improve productivity and accessibility for users, making their lives easier and more efficient.

## **6.Disadvantages Of Google Assistant Duplex**

While Google Duplex has several advantages, it also has some potential disadvantages:

1. Privacy Concerns: Google Duplex operates by recording and analyzing the user's voice data, which can raise privacy concerns. Some users may not feel comfortable having their voice recorded and stored by Google.

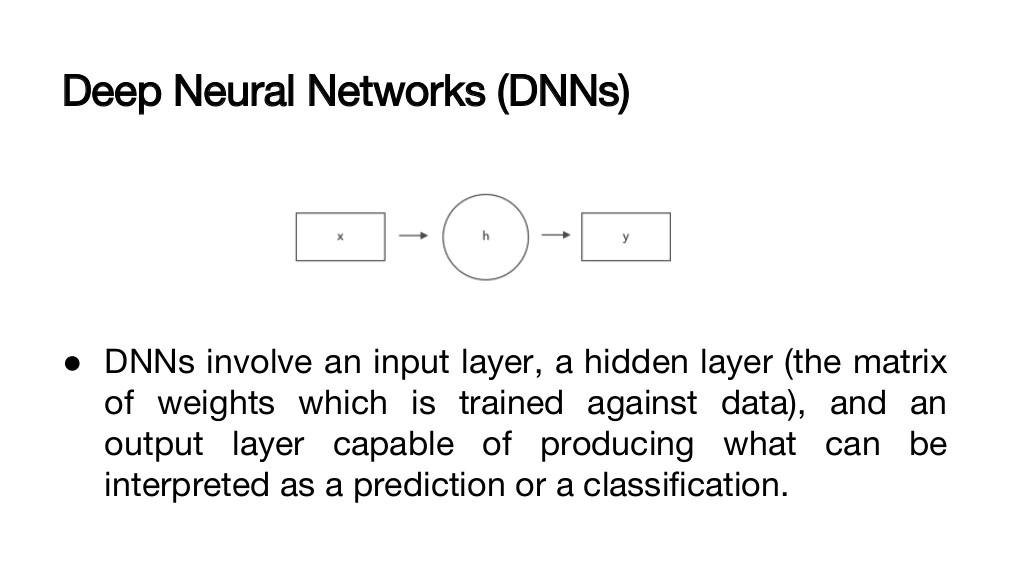
2. Dependence on Technology: Since Google Duplex relies on artificial intelligence and machine learning algorithms, it may not always work as intended. Users may become overly reliant on Duplex, which can be problematic if the technology fails to perform as expected.

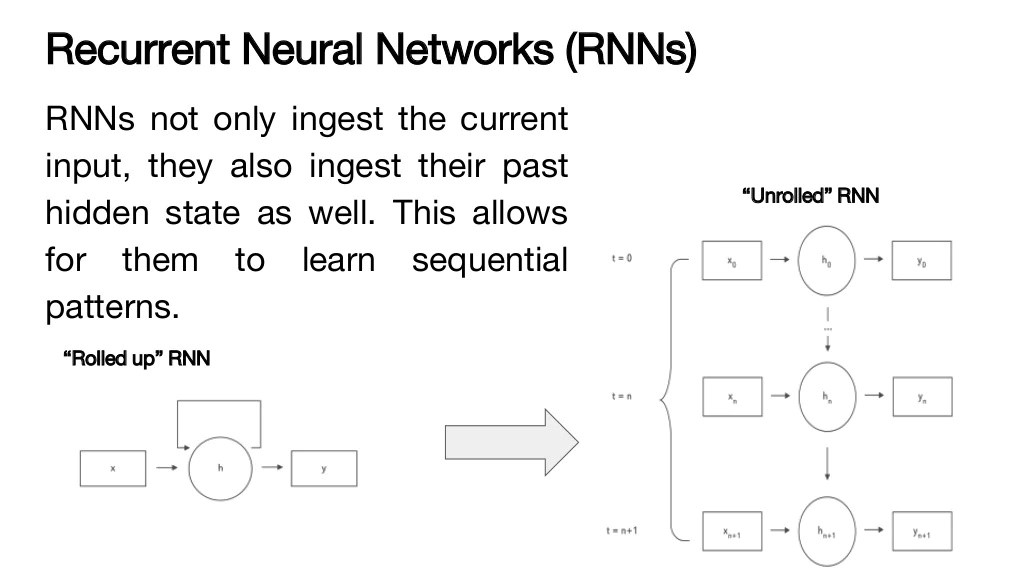
3. Limited Availability: As of now, Google Duplex is only available in select regions and for certain tasks. This means that users in other regions or who need to perform different tasks may not be able to benefit from the technology.

4. Job Losses: The automation of tasks that Google Duplex enables could potentially lead to job losses in industries such as customer service and hospitality, where human interaction is currently required.

Overall, while Google Duplex has the potential to make our lives easier, it is important to consider the potential drawbacks before using it.

## **5 DNNS AND RNNS**





Recurrent Neural Network is based on neural networks where link relation nodes create a directed graph behind a chronological sequence. It allows to show temporal active behavior. RNN is generally used to analyze the text, image capturing, emotion analysis and machine translation. How it work means the output of the previous step are attached with the input of current step not like traditional neural network, inputs and output are identical of each other. Sometime the previous word is helping to predict the next word of the sentence but it may be a problem. RNN used the Hidden layer overcome these problems. Hidden state are is plays prominent role in RNN which has the capable to remember few information about the next.RNN have a memory which recollect the information what has been identified. Hidden layer is processing point of all the inputs and outputs to produce the output

1. CLOSED DOMAINS AND VANISHING GRADIENT PROBLEM

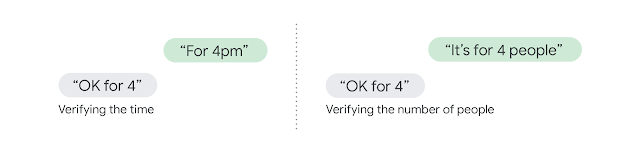
Closed domain operation

Google Duplex is not able to carry out random casual conversation. Rather, it was trained to autonomously handle three specific types of tasks:

* Scheduling a hair salon appointment
* Making a restaurant reservation, and • Asking about the business hours of a store.

**Conducting natural conversations**

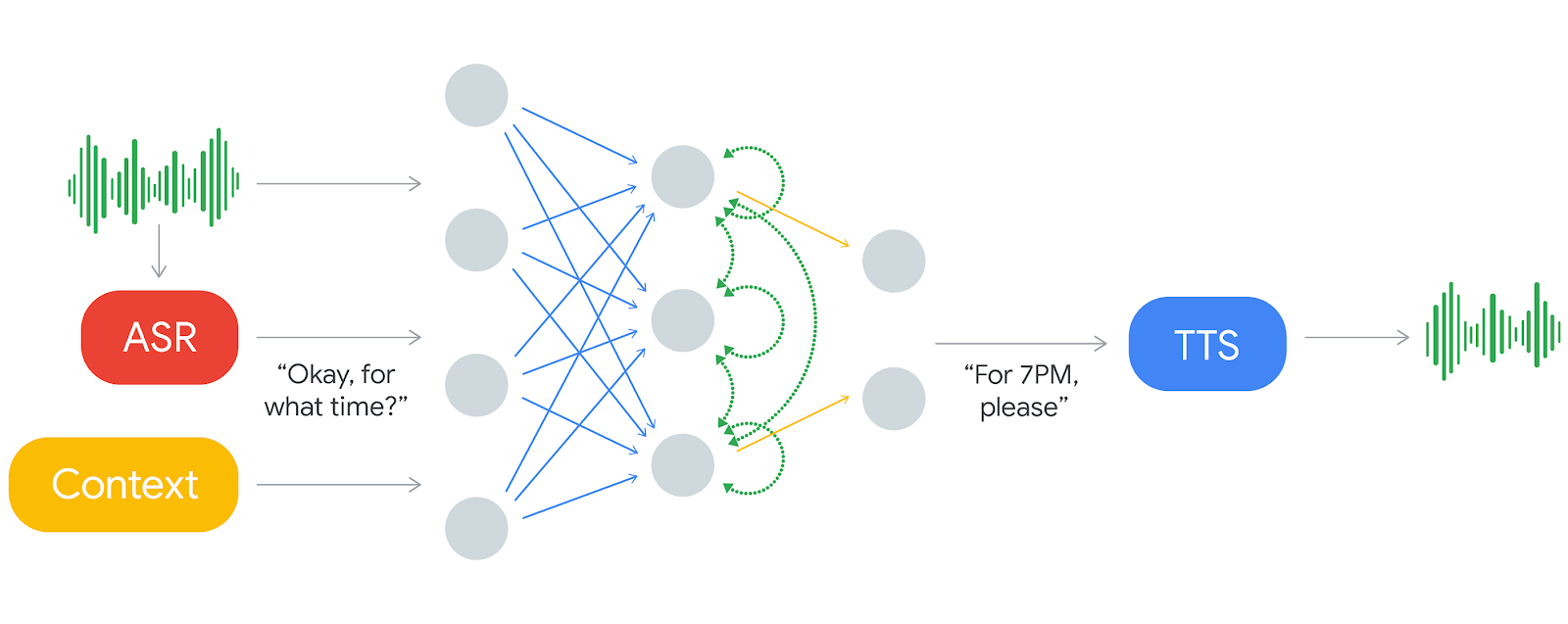
* There are several challenges in conducting natural conversations: natural language is hard to understand, natural behavior is tricky to model, latency expectations require fast processing, and generating natural sounding speech, with the appropriate intonations, is difficult.  
    
  When people talk to each other, they use more complex sentences than when talking to computers. They often correct themselves mid-sentence, are more verbose than necessary, or omit words and rely on context instead; they also express a wide range of intents, sometimes in the same sentence, e.g., *“So umm Tuesday through Thursday we are open 11 to 2, and then reopen 4 to 9, and then Friday, Saturday, Sunday we... or Friday, Saturday we're open 11 to 9 and then Sunday we're open 1 to 9.”*
* In natural spontaneous speech people talk faster and less clearly than they do when they speak to a machine, so speech recognition is harder and we see higher word error rates. The problem is aggravated during phone calls, which often have loud background noises and sound quality issues.  
    
  In longer conversations, the same sentence can have very different meanings depending on context. For example, when booking reservations “Ok for 4” can mean the time of the reservation or the number of people. Often the relevant context might be several sentences back, a problem that gets compounded by the increased word error rate in phone calls.



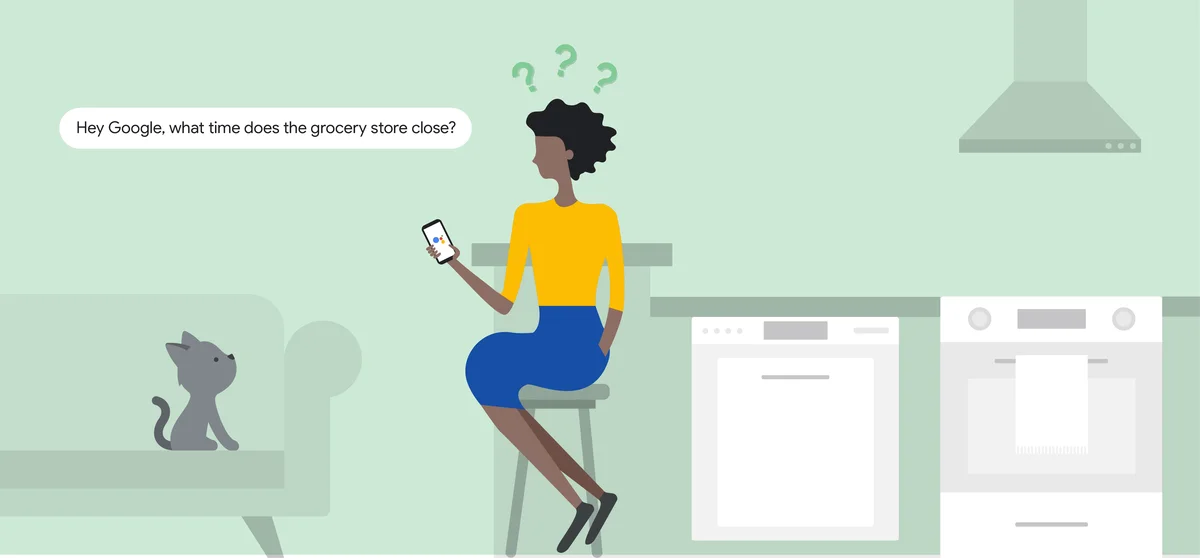
* Deciding what to say is a function of both the task and the state of the conversation. In addition, there are some common practices in natural conversations — implicit protocols that include *elaborations*(“for next Friday” “for when?” “for Friday next week, the 18th.”), *syncs* (“can you hear me?”), *interruptions* (“the number is 212-” “sorry can you start over?”), and *pauses* (“can you hold? [pause] thank you!” different meaning for a pause of 1 second vs 2 minutes).

**Enter Duplex**

* Google Duplex’s conversations sound natural thanks to advances in *understanding*, *interacting*, *timing*, and *speaking*.  
    
  At the core of Duplex is a [recurrent neural network](https://en.wikipedia.org/wiki/Recurrent_neural_network) (RNN) designed to cope with these challenges, built using [TensorFlow Extended](https://www.tensorflow.org/tfx) (TFX). To obtain its high precision, we trained Duplex’s RNN on a corpus of anonymized phone conversation data. The network uses the output of Google’s automatic speech recognition (ASR) technology, as well as features from the audio, the history of the conversation, the parameters of the conversation (e.g. the desired service for an appointment, or the current time of day) and more. We trained our understanding model separately for each task, but leveraged the shared corpus across tasks. Finally, we used hyperparameter optimization from TFX to further improve the model.



# Duplex is getting smarter and making life a little easier

****

**Why closed domain operation is important**

* Closed domains are loosely defined as any setting that has a limited number of conceivable interactions.
* Any closed domain has a sort of closed (and well-worn) number of conversational paths and options.
* When a domain is closed conversations are pigeonholed – the same sorts of conversations occur over and over building up a stronger dataset for harder-to-reach features such as natural timing knowing industry/trade slang and so on.

**Advantages of closed domain operation**

* It has a number of advantages but a major one is that it helps duplex avoid the “vanishing gradient problem” which is an issue for many DNNS and RNNS alike.
* It increases the sample size for particular conversational paths in duplex’s training data.

**Vanishing gradient problem**

* When many hidden layers are stacked such as in a multi layer DNN or between time steps in an RNN,

The network begins to “forget” the past.

* As the network goes through multiple layers of words,the original context gets lost so it fails to capture the relationship between the words that stand for apart in a conversation.
* This happens due to the underlying mechanics of backpropagation.

