**1 .INRODUCTION**

During Google I/O 2018, Google revealed “Google Duplex” to the world. The technology, which is both impressive and a bit on the creepy side, featured a human-sounding robot having a conversation with a person who couldn’t even tell that they were talking to a robot. Google Duplex impressive is that it sounds like a real person with natural pauses and speech disfluencies.

The demonstration freaked some out but impressed others, including our Mobile Editor Julian Chokkattu , who [got a chance to demo Duplex recently](https://www.digitaltrends.com/mobile/google-duplex-hands-on/). Google Duplex is a pretty big leap the evolution of artificial intelligence. Google Duplex is a huge step in terms of A.I.’s ability to more naturally converse with humans.

Google revealed Duplex, an extension of the Google Assistant that allows it to carry out natural conversations by mimicking human voice. The assistant can autonomously complete tasks such as calling a hair salon to book an appointment, scheduling a restaurant reservation, or calling businesses to verify holiday store hours.While Duplex can complete most of its tasks fully autonomously, it is able to recognize situations that it is unable to complete and can signal a human operator to finish the task. Duplex was created to speak in a more natural voice and language by incorporating [speech disfluencies](https://en.wikipedia.org/wiki/Speech_disfluency) such as [filler words](https://en.wikipedia.org/wiki/Filler_(linguistics)) like "hmm" and "uh" and using common phrases such as "mhm" and "gotcha", along with more human-like [intonation](https://en.wikipedia.org/wiki/Intonation_(linguistics)) and response latency. Duplex is currently in development and will receive further testing during summer 2018.

Announced at this year’s I/O developer conference, Google Duplex is a new Assistant feature that can carry out specific tasks for you over the phone. It can make a reservation at your favorite restaurant, schedule an appointment at a hair salon, or call businesses to check their opening hours.

**2. GOOGLE ASSISSTANT**

The **Google Assistant** is a [virtual assistant](https://en.wikipedia.org/wiki/Virtual_assistant) powered by [artificial intelligence](https://en.wikipedia.org/wiki/Artificial_intelligence) and developed by [Google](https://en.wikipedia.org/wiki/Google) that is primarily available on [mobile](https://en.wikipedia.org/wiki/Mobile_device) and [smart home](https://en.wikipedia.org/wiki/Home_automation) devices. Unlike [Google Now](https://en.wikipedia.org/wiki/Google_Now), the Google Assistant can engage in two-way conversations. Assistant initially debuted in May 2016 as part of Google's messaging app [Allo](https://en.wikipedia.org/wiki/Google_Allo" \o "Google Allo), and its voice-activated speaker [Google Home](https://en.wikipedia.org/wiki/Google_Home).

Users primarily interact with the Google Assistant through [natural voice](https://en.wikipedia.org/wiki/Natural_language_processing), though keyboard input is also supported. In the same nature and manner as Google Now, the Assistant is able to search the Internet, schedule events and alarms, adjust hardware settings on the user's device, and show information from the user's Google account. Google has also announced that the Assistant will be able to identify objects and gather visual information through the device's camera, and support purchasing products and sending money, as well as identifying songs.

The Google Assistant, in the nature and manner of [Google Now](https://en.wikipedia.org/wiki/Google_Now), can search the Internet, schedule events and alarms, adjust hardware settings on the user's device, and show information from the user's Google account. Unlike Google Now, however, the Assistant can engage in a two-way conversation, using Google's [natural language processing](https://en.wikipedia.org/wiki/Natural_language_processing) algorithm. n May 2017, Google announced that the Assistant would support a keyboard for typed input and visual responses, support identifying objects and gather visual information through the device's camera, and support purchasing products and sending money.

**2.1. GOOGLE ASSISTANT vs SIRI**

Siri is comparatively same as Google assistant. Google assistant used in Android whereas, siri in iPhone. Siri is already exists in iPhone. Siri is difficult to use than Google assistant. Because, **Google assistant** is doing what we say. When English is no clear, it will handle. Butthe **Siri,** it needs a clear pronunciation in English, unless it can’t accept.

* **Recognizing your voice**

On the [HomePod](https://www.cnet.com/reviews/apple-homepod-review/), unlike the [iPhone](https://www.cnet.com/reviews/apple-iphone-x-review/), Siri can't differentiate between voices. That means anyone who says "Hey, Siri" will elicit a response from the Apple smart speaker. In contrast, Google smart speakers have voice training capabilities. With voice training, the whole family can use the speakers, but each of you can also get customize responses (and perform custom actions, like making purchases) based on who the speaker is talking to.

## Making calls

If I say "Hey Siri, call [Tyler Lizenby](https://www.cnet.com/profiles/tylerlizenby/)" (CNET multimedia producer) on my iPhone, she responds, "Calling Tyler Lizenby, mobile." If I give HomePod the exact same command, she'll reply: "I wish I could, but I can't help you make calls on HomePod." Bummer. Yep, Siri (on HomePod) doesn't let you dial in or out. Instead, the HomePod acts as a speakerphone. Either dial or receive a call on your iPhone and use the Audio button on the call screen to transfer the audio to your Apple speaker (see screenshot).

Google-Assistant-enabled speakers work differently. Both can currently dial out with simple voice commands. Google Assistant can't receive calls. Unlike calling, you can send messages from HomePod, as well as Google Assistant.

### Access

### One of the most important parts of any AI assistant is how easy it is to access.

### ****Siri**:** On Apple’s hardware, Siri still has a massive home court advantage when it comes to actually starting the digital assistant. Namely, Siri is baked into iOS on a core level — its accessible from anywhere, whether you’re in an app or at your lock-screen by simply holding down the home button. Siri is also always listening, too, and can be activated anytime through a “Hey, Siri” command on newer iPhones (although the feature, like much of Siri these days, feels largely half-baked.

### ****Assistant****: Google is making a big play here with Assistant, but it’s fighting an uphill battle here with the inherent limitations Apple has on iOS. That means that Assistant lives in the Assistant app. But Google is doing its best to put Assistant in as many places as possible: there’s an Assistant widget to quick launch the app, which is accessible from both the notification shade, the lock screen, and the home screen. Plus, there’s a 3D Touch action to also start Assistant

### . https://cdn.vox-cdn.com/thumbor/O5wDNXEsUiRWy1W3G8j1e4hv6wQ=/0x0:1500x1334/1200x0/filters:focal(0x0:1500x1334):no_upscale()/cdn.vox-cdn.com/uploads/chorus_asset/file/8546919/Untitled.png

### Hardware integration

Unlike when Siri first launched back in 2011, AI assistants can’t just work with a single device anymore. Working on an iPhone is one thing — but working across an entire ecosystem is a whole other ballgame.

**Siri:** The best thing that can be said about Apple’s assistant is that it exists on a variety of hardware platforms. If you’ve got an Apple Watch, a Mac, or an Apple TV, you can use Siri on it, but that’s about the extent of it. Siri on different platforms largely exist as separate things — you can’t use the native Siri on your iPhone to play a song on your Apple TV, for example, or use Siri on your Apple TV to make a call Controlling smart home utilities with Siri only works if manufacturers support Apple’s Home Kit, and you’ll need either an Apple TV or iPad to serve as a hub. Plus, you’re still locked strictly to Apple’s ecosystem — if Cupertino doesn’t make hardware for it (like, say, an Amazon Echo-like speaker), then you can’t use Siri on it.

**Assistant:** Obviously, just like Siri works with Apple stuff, Assistant works with Google stuff. That means Chrome cast, Android phones, and of course, Google Home. But one of the announcements to Assistant at I/O was that third-party hardware manufacturers would be able to integrate Assistant into their own devices, which means that there could be a much larger ecosystem for Assistant in the near future.

### Functionality:

Of course, an assistant is only useful if it can actually assist you with things. And at this point, there's basic table stakes that both platforms are able to accomplish — basic queries of the weather and sports scores, calendar reminders, alarms, etc. That said, there are clear differences between the two in terms of the extent of their abilities.

**Siri**: Siri is able to do more advanced queries like tell you the time difference between two cities or what the price of gasoline is in San Francisco, typically drawing on resources like Wolfram Alpha and Wikipedia to answer, but in practice I’ve generally find it easier and more reliable to just Google the question.

**Assistant**: Even with the limitations of iOS, Assistant gets a lot done. You can play videos in YouTube, add reminders and events through Google Calendar, and send emails through Gmail. Plus, there’s the usual AI assistant queries of stocks, the weather, and maps that’s common to both Siri and Assistant. Google has also tried its best to offer messaging and call support on iOS, but it can only get so far — both will get you 99 percent of the way there before passing you back to Apple’s apps for actually sending the message or placing a call. Meanwhile, things like setting an alarm just aren’t possible with Assistant on iOS, and others — like playing music, which shunts you over to YouTube — work, but are a bit of a hack.

## Music services

Google support a wide range of third-party music services, so your Google Assistant speaker can play a Spotify song with a simple voice command.

Siri on the HomePod isn't able to do this. You can only use voice commands to play Apple-approved music services. So, if you want to use Spotify, YouTube or another third-party service, you can, but you have to send the song manually from your phone or other [iOS](https://www.cnet.com/reviews/apple-ios-11-preview/) device.

## General questions

Siri actually did a great job responding to my basic, everyday questions, roughly the same as Google Assistant. If I wanted to know the weather forecast, the at-a-glance driving time to a movie theater across town, the name of some local restaurants that deliver, the latest news updates or information about a song playing on HomePod, the Apple voice assistant worked very well.

Siri also did a pretty good job with follow-up questions and natural language. Google Assistant also successfully responded to a variety of natural language questions.

## Smart home

I'm conflicted about using "Hey, Siri" on the HomePod to control smart home devices. During the HomePod set up, all of the info about your local Wi-Fi network, contacts and preferences are automatically sent over to the HomePod. That includes any existing [HomeKit](https://www.cnet.com/reviews/apple-homekit-preview/) products you have set up on your iOS device. It's ridiculously simple.

And every smart-home integration is so closely regulated by Apple that using voice control to open smart shades, adjust a thermostat or turn on lights is easy to achieve with a basic "Hey, Siri" command, because Apple supports all of those device types directly. In contrast, Google Assistant give you "invocation" words to interact with device types they don't support out of the box.

**3. GOOGLE DUPLEX**

Google Duplex is a new Assistant feature that can carry out specific tasks for you over the phone. In May 2018, Google revealed Duplex, an extension of the Google Assistant that allows it to carry out natural conversations by mimicking human voice.

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. Duplex is able to understand "complex sentences, fast speech and long remarks" according to Google.

During Google I/O 2018, Google revealed “Google Duplex” to the world. The technology, which is both impressive and a bit on the creepy side, featured a human-sounding robot having a conversation with a person who couldn’t even tell that they were talking to a robot. Google Duplex is a pretty big leap the evolution of artificial intelligence.

It is announced at this year’s I/O developer conference. It's a new tool from Google that aims to use Artificial intelligence (AI) to "accomplish real-world tasks over the phone" according to Google's AI researchers and developers. Google Duplex impressive is that it sounds like a real person with natural pauses and speech disfluencies like “hmm” and “umm,” which people generally use to gather their thoughts — Assistant doesn’t sound like a robot in those calls. Assistant also understands the context of conversations. So when a call doesn’t go as expected, as in the second example where the restaurant employee was not a native English speaker, Assistant still got the job done.

Duplex is designed to change the way a computer "talks" on the phone. The goal for duplex is to make things sound natural and for Assistant to think on the fly to find an appointment time that works. For years, businesses have been trying to create a way for people to have conversations with computers. Almost every time we call a business, we encounter an automated phone system. We have virtual assistants on our phones and virtual assistant-powered speakers in our homes. But although these computer systems can be helpful, they have their shortcomings. Google Duplex helps with problems by allowing the computer to have a natural conversation with a human. The A.I. system adjusts to the person, instead of the person adjusting to the system. Therefore, the person can speak normally, just as they would if they were speaking to another person. Google Duplex also makes it so the computer system sounds like a human. It uses a natural tone and uses words and phrases like “um” and “uh” just like a human person would. During a conversation, the A.I. system can also handle interruptions and elaborate. At the center of Google duplex is a Recurrent Neural Network that was built using a machine learning platform called [Tensor Flow Extend (TFX)](https://www.tensorflow.org/tfx/). When the system makes a phone call, it is pretty much indistinguishable from a live human being.

Google Duplex could be a lifesaver for travelers who don't speak the local language well.

Google Duplex can seriously boost productivity. It’s a great tool for those too busy to deal with less important tasks. It’s also great for people who struggle with hearing problems or phone anxiety (yes, that’s a real thing). Even better, the feature could be a lifesaver for travelers or people who don’t speak the local language well.

Google Duplex is amazing, creepy, and too good to go to waste

Although Google Duplex seems like a fantastic feature with the potential to make our lives easier, it’s also a bit creepy. Talking to a robot that sounds like a real person is not something most people are excited about.

**3.1. WHAT CAN IT DO?**

Initially Google Duplex will focus on three kinds of task:

1. Making restaurant reservations
2. Scheduling hair appointments
3. Finding out businesses’ holiday opening hours.

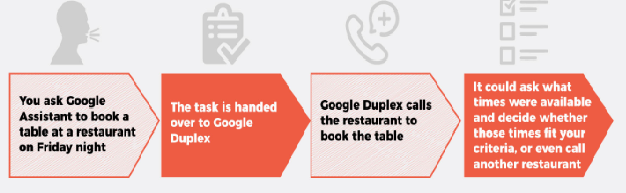
If the business you've asked Google Assistant to book an appointment with accepts online reservations, Assistant will use that to complete the task. Otherwise, Assistant/Duplex will call the business on your behalf. The idea of [not knowing if you're talking to a computer or a real person](https://www.cnet.com/news/google-duplex-assistant-bot-deception-scary-ethics-question/) is incredibly creepy. For its part, Google has said it will build in the requirement to inform those on the line that [the call is with Google Assistant](https://www.cnet.com/news/google-says-its-designing-duplex-with-disclosure-built-in/) or Google Duplex for transparency. Google said it will start testing Duplex this summer, focusing first on making restaurant reservations, scheduling hair salon appointments, and getting holiday hours over the phone. The feature will get smarter over time and become capable of taking on other tasks as well.

**3.2. HOW DOES IT WORK**

Google Duplex is the missing link between the Google Assistant and any business, because it enables the Assistant to get information that isn’t available digitally. It is a new technology for conducting natural conversations to carry out easy tasks that involve calling. Duplex is using recurrent neural network. It's built using Google's Tensor Flow Extended technology. Google trained the network on all those anonymized voicemails and Google Voice conversations agreed to let it listen to if opted in with a mix of speech recognition software and the ability to consider the history of the conversation and details like the time of day and location of both parties.

It works for scheduling certain types of appointments and also it cannot carry out general conversations. An example, you might want to know a bussiness’s holiday opening hours but they haven’t listed it on their website or you might want to know if a shop has a particular item in stock and it doesn’t have online stock availability.

Example:



To train the system in a new domain, it uses **real-time supervised training**. In the Duplex system, experienced operators act as the instructors. By monitoring the system as it makes phone calls in a new domain, they can affect the behavior of the system in real-time as needed. The network uses the output of Google’s automatic speech recognition (ASR) technology, that converts spoken word in to written text. . Google trained the understanding model separately for each task but leveraged the shared corpus across tasks.Finally, they used hyper-parameter optimization from TFX to further improve the model. Incoming sound is processed through an ASR system. This produces text that is analyzed with context data and other inputs to produce a response text that is read aloud through the text-to-speech (TTS) system.

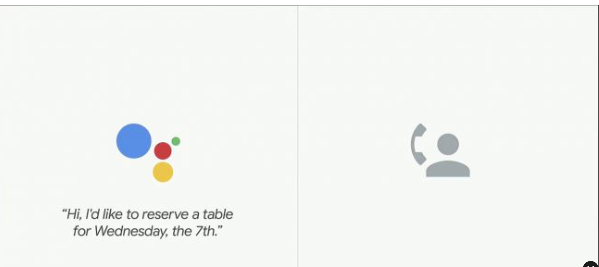
* 1. **WHAT’S SO CLEVER ABOUT IT?**

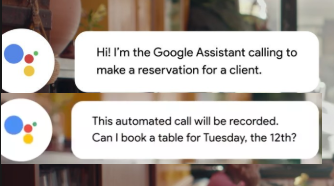
[Google Assistant](http://www.innov8tivedd.com/how-google-assistant-can-help-you-with-your-work/)  allows to communicate in a human-like way with a person over the phone to complete basic real-world tasks.Google CEO Sundar Pichai demonstrated the new AI system at Google I/O 2018. He presented two demos to the audience, one where a female voice called a hair salon to schedule a hair appointment; and the second where a male voice booked a reservation in a restaurant. Duplex talks like a normal person, and that makes it a natural – and natural-sounding – extension to the OK Google functionality already know.

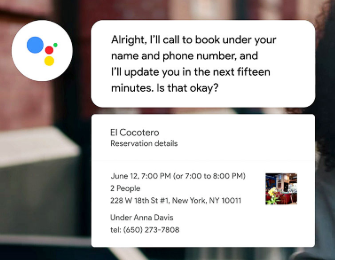
With a restaurant example:

With Duplex, we could say “OK Google, find me a table for Friday night” and the Google app would then call restaurants on your behalf. Not only that, but it would have conversations – so if you wanted a table for around 7:30 but there wasn’t one, it could ask what times were available and decide whether those times fit your criteria. If not, the Google app would call another restaurant. Similarly if you wanted to arrange a meeting with Sarah, the Google app could call Sarah (or Sarah’s AI) to talk through the available time slots and agree which one would be best.

The key here is that this is all happening in the background. You tell Google to do something and it goes and does it, only reporting back after the task is complete.

1) 





A long-standing goal of human-computer interaction has been to enable people to have a natural conversation with computers, as they would with each other. In recent years, we have witnessed a revolution in the ability of computers to understand and to generate natural speech, especially with the application of deep neural networks .Still, even state of the art systems, it is often frustrating having to talk to stilted computerized voices that don't understand natural language. Now announce Google Duplex, 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.

The Google Duplex technology is built to sound natural, to make the conversation experience comfortable. It’s important to us that users and businesses have a good experience with this service, and transparency is a key part.

**Conducting natural conversation**

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. 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.

**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 [Tensor Flow 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 and more. To understand the model separately for each task, but leveraged the shared corpus across tasks. Finally, we used hyper parameter optimization from TFX to further improve the model.

**Sounding natural**

Use a combination of a concatenative text to speech (TTS) engine and a synthesis TTS engine (using [Tacotron](https://research.googleblog.com/2017/12/tacotron-2-generating-human-like-speech.html) and [WaveNet](https://deepmind.com/blog/wavenet-generative-model-raw-audio/)) to control intonation depending on the circumstance. The system also sounds more natural thanks to the incorporation of speech disfluencies like “hmm”s and “uh”. These are added when combining widely differing sound units in the concatenative TTS or adding synthetic waits, which allows the system to signal in a natural way that it is still processing. It’s important for latency to match people’s expectations.

**System operation**

The Google Duplex system is capable of carrying out sophisticated conversations and it completes the majority of its tasks fullyautonomously, without human involvement. The system has a self-monitoring capability, which allows it to recognize the tasks it cannot complete autonomously (e.g., scheduling an unusually complex appointment). In these cases, it signals to a human operator, who can complete the task.  To train the system in a new domain, we use real-time supervised training. In the Duplex system, experienced operators act as the instructors. By monitoring the system as it makes phone calls in a new domain, they can affect the behavior of the system in real time as needed. This continues until the system performs at the desired quality level, at which point the supervision stops and the system can make calls autonomously.

**Benefits for bussinesses and users**

Businesses that rely on appointment bookings supported by Duplex, and are not yet powered by online systems, can benefit from Duplex by allowing customers to book through the Google Assistant without having to change any day-to-day practices or train employees. Using Duplex could also reduce no-shows to appointments by reminding customers about their upcoming appointments in a way that allows easy cancellation or rescheduling.

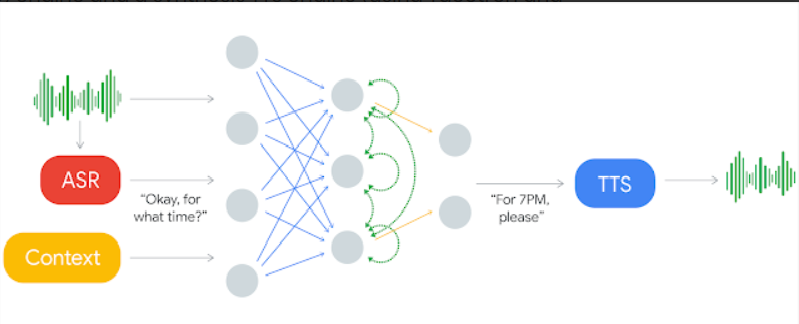
  Duplex can call the business to inquire about open hours and make the information available online with Google, reducing the number of such calls businesses receive, while at the same time, making the information more accessible to everyone. Businesses can operate as they always have, there’s no learning curve or changes to make to benefit from this technology.

For users, Google Duplex is making supported tasks easier. Instead of making a phone call, the user simply interacts with the Google Assistant, and the call happens completely in the background without any user involvement. Another benefit for users is that Duplex enables delegated communication with service providers in an asynchronous way, e.g., requesting reservations during off-hours, or with limited connectivity. It can also help address accessibility and language barriers, e.g., allowing hearing-impaired users, or users who don’t speak the local language, to carry out tasks over the phone.

Allowing people to interact with technology as naturally as they interact with each other has been a long standing promise. Google Duplex takes a step in this direction, making interaction with technology via natural conversation a reality in specific scenarios. We hope that these technology advances will ultimately contribute to a meaningful improvement in people’s experience in day-to-day interactions with computers.

**4. TECH BEHIND GOOGLE DUPLEX**

At the heart of Duplex is something Google calls a Recurrent Neural Network (RNN), which works out not only the words spoken but the context in which they are being used. At the core of Duplex is RNN, that has been built using Tensor Flow Extended. It makes the voice behind Duplex sound human-like. Google’s developers used a combination of Text-to- Speech (TTS) engine and a synthesis TTS engine (using Tacotron and WaveNet) to vary the tone of the machine.

****

* *Incoming sound is processed through an ASR system. This produces text that is analyzed with context data and other inputs to produce a response text that is read aloud through the TTS system.*

RNN uses the output of Google’s automatic speech recognition technology, as well as features from the audio, the history of the conversation and the parameters of the conversation and more.

At the core of it, Duplex is a recurrent neural network that can be trained for highly specialized tasks and it uses Google’s automatic speech recognition technology, so it can interface with the user. WaveNet and an AI-based generative program that’s part of Google’s DeepMind division, Duplex can sound a lot closer to a human than ever before, complete with “umms" and “aahs”.Speech disfluencies have been added to the system to make it sound even like a human.

Google Duplex tackles all issues and makes conversations sound natural with advances in understanding, interacting, timing, and speaking. It involves a [recurrent neural network](https://en.wikipedia.org/wiki/Recurrent_neural_network) (RNN) that is built with [TensorFlow Extended](https://www.tensorflow.org/tfx) (TFX). 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.

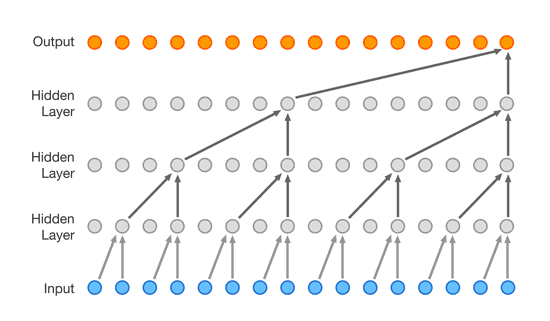
The system is capable of carrying out sophisticated conversations and it completes the majority of its tasks fully autonomously, without human involvement. It has a self-monitoring capability, which allows it to recognize the tasks it cannot complete autonomously scheduling an unusually complex appointment. In these cases, it signals to a human operator, who can complete the task.

**Voice commands and ASR** have been around for a while with Google products and most people are used to that kind of stuff at this point (Alexa, Siri, Cortana… actually, does anybody ever talk to Cortana), so I won’t go into it that much.

**Contextual information** probably starts with a strong prior on *what the conversation is about*. Duplex currently only works in very specific scenarios and it’s been emphasized that it’s *not* a general-conversation agent. That really helps. The search space for what we might talk about if I’m calling about a haircut is a lot smaller than the kinds of things I’d cover in conversation in a normal week. It’s also pretty clear that Duplex keeps track of where in the conversation it is - like greeting the person, making a request, providing clarification, confirming information, etc. That also means tracking the states of the conversation partner - for example, how confused are they? The restaurant reservation demo and associated commentary shows tracking the human state allows Duplex to nudge the conversation back into familiar territory (second half of the video).

[**Recurrent neural nets**](https://en.wikipedia.org/wiki/Recurrent_neural_network) are a type of neural net that is popular for dealing with temporal information, because it feeds some of the output back into previous parts of the network. This structure lets it exhibit “memory” of previous events and adjust its output based on the sequence of inputs rather than just a one-time input. This network was trained on a large dataset of calls that were relevant to the specific task, though it’s not clear where this corpus came from (*\*stares at Android tablet sitting on my desk\**). This part seems like it’s also responsible for generating the text of Duplex’s response to the human.

**Text-to-speech**has also been around for a while, but this one seems so very different from what we’re used to! It’s also probably the most creepy element of the whole product, because the speech is so natural that it’d be pretty hard to tell the difference between this output and a real human. Actually, I wasn’t all that surprised by this bit of the demo, because DeepMind (owned by Google) released a [post](https://deepmind.com/blog/wavenet-generative-model-raw-audio/) a while back about WaveNet, a [convolutional neural net](https://en.wikipedia.org/wiki/Convolutional_neural_network) for audio generation that was trained on raw audio data (very rough structure shown below). I’d recommend looking at the post. The audio samples are amazing. I can’t tell the difference between the WaveNet English samples and a human (I don’t have enough experience with Mandarin to judge those samples, but I imagine it’s similar).



To prevent it from sounding too robotic, the system was also taught a number of so-called “speech disfluencies and other noises people make in casual conversation. Like humans, the AI makes those sounds to convey that it’s still gathering its thoughts.”

**5. HOW NATURAL DOES IT SOUND?**

Google at the I/O 2018 , introduced one of the most controversial Assistant feature; Google Duplex which mimics itself as human and makes appointments. Google Duplex is a new Assistant feature that can carry out specific tasks for you over the phone that involve calling. Duplex is designed to respond in a natural manner and adapt to responses in real time, while understanding the context of the conversation.



* *A user ask the Google assistant for an appointment, which the assistant then schedules by having Duplex call the business.*

Generally the technology has four main blocks.

1. User: Ask what he needs through commands.
2. Google Assistant: Understands what the user is trying to say and converts it into text.
3. Google duplex: Calls the respective business and helps Google Assistant understand what the person on the other end is trying to say using Neural Networks.
4. Business: The person who handles business calls.

There are several challenges that Google has faced in conducting a natural conversation like natural language is hard to understand, a natural behaviour is tricky to model, latency expectations require fast processing and generating natural sounding speech with the appropriate intonations. The Google Duplex technology is built to sound natural, to make the conversation experience comfortable,” the company wrote in [a blog post](https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html) tied to the announcement. “It’s important to us that users and businesses have a good experience with this service, and transparency is a key part of that. We want to be clear about the intent of the call so businesses understand the context.

Duplex is the AI system designed to make human-sounding voice calls on your behalf so as to automate things like booking restaurant tables and hair appointments and in fact, each of us in turn took a call from Duplex as it tried to book a reservation.  The behavior of Duplex arises out of Google’s recently published [core AI principles](https://www.blog.google/technology/ai/ai-principles/). “We’re going to be very slow, very careful, and very thoughtful as we go here”.

An AI system to achieve real-world tasks over the phone. Technology plays an important part in altering our lifestyle. Likewise, Human-Computer Interaction has expanded rapidly. In recent years, we have witnessed a revolution in the ability of computers to understand and to generate natural speech. Google Duplex, a new technology for conducting natural conversations to carry out “real world” tasks over the phone. It is an automated voice assistant that can book restaurant reservations, check opening hours and accomplish other tasks over the phone. For such tasks, the system makes the conversational experience as natural as possible, allowing people to speak normally.

There are several challenges in conducting natural conversations. For example, natural spontaneous [speech](https://www.techexplorist.com/new-speech-recognition-chip-saves-99-power/4887/) people talk faster and less clearly or they use more complex sentences than when talking to computers. The problem is aggravated by phone calls, which often have loud background noises and sound quality issues.

For users, Google Duplex is making supported tasks easier. Instead of making a phone call, the user simply interacts with the Google Assistant, and the call happens completely in the background without any user involvement. Another benefit for users is that Duplex enables delegated communication with service providers in an asynchronous way, e.g., requesting reservations during off-hours, or with limited connectivity. It can also help address accessibility and language barriers, e.g., allowing hearing-impaired users, or users who don’t speak the local language, to carry out tasks over the phone.

Google uses a combination of a concatenative text to speech (TTS) engine and a synthesis TTS engine (using Tacotron and WaveNet) to control intonation depending on the circumstance.

One of the key important factor to manage people’s expectation is latency. For example, after people say something simple, e.g., “hello?” they expect an instant response and are more sensitive to latency. When low latency is required, Google uses faster, low-confidence models .

In extreme cases, it doesn’t even wait for RNN, instead use faster approximations .In some situations, it was actually helpful to introduce more latency to make the conversation feel natural — example, when replying to a really complex sentence.

6 points that raise a doubt :

 Talks about Google duplex still under development

 No business greets by saying “How may I help you”

 No background noises during the call

 for booking an appointment, phone numbers are asked usually

 Enough time was given for Google assistant to answer every question

 Google not giving answers to the questions raised after the duplex presentation

**6.ADVANTAGES**

* **Makes life easier**

You won’t have to deal with the menial task of booking a restaurant reservation or hair appointment.

Just activate Google Assistant and it’ll do it for you!

This way, you can focus more on important tasks that need your attention more.

* **Help accessibility and language barriers**

Since the virtual assistant is voice-activated, people with disability like impaired hearing, and immigrants or tourists having trouble speaking the local language, can use their phones to talk to someone on the phone to book a reservation.

* **Benefits traditional businesses**

Businesses that haven’t had a digital transformation and have no online booking system can still be accessed by Google Assistant through a call.

It could also [make local online data more accurate](https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html). For example, stores have different operation hour during holidays, and with Duplex, you can call a local store and ask its business hours.

* Google Duplex could be a lifesaver for travelers who don't speak the local language well.
* It’s a great tool for those too busy to deal with less important tasks.
* It’s also great for people who struggle with hearing problems or phone anxiety.

**7.DISADVANTAGES**

* **No control over the call**

Google Duplex is still under testing but from what has been shown in I/O, you can’t interfere with a call.

You have no way of knowing if the call is going off the rails. You only get to listen to what happened to the call after it’s been done.

* **Easier to abuse small businesses**

It’s just as easy to ask the AI to constantly reschedule or miss your appointment since you don’t feel an urgency to actually make your appointment because there’s no connection.

* Another issue is whether Google should get permission to record calls using Duplex.

**8. FUTURE SCOPE**

At Google I/O 2018, the internet company just demoed Duplex, a powerful upgrade to the Google Assistant, which holds the potential to revolutionise both our personal productivity as well as the expectations we have of artificial intelligence (AI) in our lives.With the latest development promising leaps in AI use in daily lives and time-saving. Google Duplex could be a lifesaver for travelers who don't speak the local language well.

Google Duplex can seriously boost productivity. It’s a great tool for those too busy to deal with less important tasks. It’s also great for people who struggle with hearing problems or phone anxiety (yes, that’s a real thing). Even better, the feature could be a lifesaver for travelers or people who don’t speak the local language well.

Google Duplex is amazing, creepy, and too good to go to waste.Although Google Duplex seems like a fantastic feature with the potential to make our lives easier, it’s also a bit creepy. Talking to a robot that sounds like a real person is not something most people are excited about.

**9.CONCLUTION**

Google Duplex is still very much an early experiment so its deployment depends entirely on how successful its trial runs are and how people respond to it. Google Duplex is a new Assistant feature that can carry out specific tasks for you over the phone.

The technology, which is both impressive and a bit on the creepy side, featured a human-sounding robot having a conversation with a person who couldn’t even tell that they were talking to a robot.

**10. BIBILOGRAPHY**

* [**www.what-is-google-duplex.com**](http://www.what-is-google-duplex.com)
* [**www.cnet.com/how-to/what-is-google-duplex**](http://www.cnet.com/how-to/what-is-google-duplex)
* **www.googles-duplex-takes-one-step-closer-to-help-humans-may- run-call-centres**
* [**www.androidcentral.com/google-duplex**](http://www.androidcentral.com/google-duplex)