UNIVERSITÀ DEGLI STUDI DI MILANO

Facoltà di Scienze e Tecnologie

*Corso di Laurea Magistrale in Informatica*

Sistema visuale per la progettazione di flussi di dialogo per assistenti virtuali

Relatore: Prof. Stefano VALTOLINA

Controrelatore: Prof. Marco MESITI

Correlatrice: Dr.ssa Sara BONFITTO

Tesi di:

Lorenzo NERI

Matricola: 901009

Anno Accademico 2018-2019

Prefazione

Organizzazione della tesi

La tesi è organizzata come segue:

* Nel Capitolo 1, viene introdotto ed esposto l’obiettivo di questa tesi di laurea e le problematiche che propone di risolvere.
* Nel Capitolo 2, viene esposta la panoramica delle soluzioni messe a disposizione sul mercato, confrontandole tra di esse ed esponendo i punti a favore e a sfavore.
* Nel Capitolo 3, viene presentata la soluzione realizzata per raggiungere l’obiettivo su cui si basa questa tesi di laurea, includendo l’analisi effettuata a seguito del confronto con le soluzioni presenti attualmente sul mercato.
* Nel Capitolo 4, si approfondisce l’architettura del sistema implementato per la soluzione proposta.
* Nel Capitolo 5, si espone l’analisi effettuata a seguito dei test sottoposti a potenziali utilizzatori finali e i risultati ottenuti dalla soluzione proposta.

Ringraziamenti

Indice

* Capitolo 1: Pagina 6
* Capitolo 2:
* Capitolo 3:
* Capitolo 4:
* Capitolo 5:
* Bibliografia:

Capitolo 1

Introduzione

Un chatbot è un software che dialoga con l’utente all’interno di un contesto definito. Si potrebbe quindi definire come un assistente virtuale capace di rispondere ad una serie di richieste, ma esse devono essere ben studiate a priori in fase di progettazione del chatbot.

Per farlo in maniera ottimale è quindi necessario conoscere molto bene il contesto di utilizzo, in particolar modo il linguaggio e gli obiettivi.

Dal nome chatbot si intuisce quindi che questo assistente virtuale può essere interpellato dall’utente attraverso la chat di una piattaforma, ma sarebbe riduttivo definirlo soltanto in questo modo.

Infatti, esistono due tipologie di chatbot: la prima interagendo tramite chat e la seconda interagendo verbalmente. Esiste infatti la possibilità di interpellare il software attraverso l’uso della voce.

Da questo si capisce che con un chatbot è possibile avere una conversazione in cui il nostro interlocutore non è un essere umano, ma un robot virtuale al quale possiamo porre una serie di domande.

Esso è quindi un interlocutore sempre disponibile in qualsiasi momento della giornata, che ci può guidare e aiutare nel compiere alcune attività nella vita di tutti i giorni. Ovviamente, essendo un robot è preparato solo su alcuni argomenti sui quali può darti una serie di risposte o sui quali può dialogare.

Un chatbot ben programmato riesce comunque a condurre una conversazione attraverso un flusso di dialogo sensato e coerente, inoltre, con il machine learning riesce anche ad apprendere informazioni aggiuntive e utilizzarle in un secondo momento.

Questi robot dispongono di un’intelligenza artificiale che diventerà sempre più importante soprattutto utilizzato in certi ambiti, come ad esempio quello medico.

Ad esempio, potrebbe riconoscere in base a dei sintomi specifici che tipo di medicinale un paziente dovrebbe assumere o quali potrebbero essere le possibili cause dei sintomi.

Ovviamente per realizzare questo è sempre necessario un esperto del settore oltre che a svariati test. Potrebbe anche essere utile per velocizzare alcune attività come l’acquisto di prodotti attraverso un sito web. Parliamo quindi di robot virtuali che possono darci una mano in diversi momenti della giornata e in contesti differenti.

La realizzazione di un chatbot generalmente, impiega due figure professionali: la prima si occupa di disegnare il flusso di dialogo tra chatbot stesso e utente finale, implementando le frasi legate al contesto conversazionale, le risposte, le parole chiave di rilievo.

La seconda invece, si occupa di realizzare il software vero e proprio finalizzato all’implementazione del chatbot vera e propria.

Questo processo, richiede parecchio tempo e un continuo confronto tra le due figure prese in esame: generalmente, chi si occupa di disegnare i flussi di dialogo non ha esperienza o abilità legate alla programmazione, ha comunque conoscenze informatiche limitate che non gli permettono di realizzare software autonomamente.

Al contempo, chi ha notevoli competenze informatiche non è generalmente una figura professionale che possa realizzare flussi di dialogo in autonomia.

In diverse situazioni, siano esse legate a contesti aziendali, sanitari, di servizio pubblico o private, vi è non solo la necessità di offrire un chatbot, d’ora in avanti nominato “agente”, che permetta un’interazione nuda e cruda con un utente finale fatta di pure conversazioni testuali: è necessario permettere la creazione di promemoria, invio di messaggi, visualizzazione di contenuti multimediali e ottenimento di dati provenienti da terze parti (API, piattaforme di messaggistica, siti web, informazioni personali, codici, ecc.).

In aggiunta al contesto di dialogo, la necessità di dati e informazioni provenienti da terze parti, subentra il bisogno di offrire le utilità del proprio agente non solo in formato visivo, ma anche sonoro.

Si pensi a tutti gli utenti non vedenti o che necessitano dei servizi offerti dall’agente in situazioni in cui non possono fare un uso diretto di uno schermo: durante la guida, nel mezzo di una camminata, oppure tramite una telefonata tradizionale.

Un altro aspetto di rilievo, è legato agli errori effettuati dall’utente. In differenti situazioni, l’agente si trova impotente di fronte a frasi incomplete, con errori grammaticali, sconnesse.

Esso di fatto, per migliorare quanto più possibile l’esperienza offerta all’utente finale, deve avere a disposizione un sistema di apprendimento affinché possa imparare lui stesso dagli errori commessi dall’utente finale.

Il nostro obiettivo, è di fatto fornire un sistema di semplice utilizzo per i disegnatori di flussi di dialogo, affinché possano in totale autonomia realizzare un agente completo senza saper programmare.

Capitolo 2

Stato dell’arte

Il mercato offre numerose soluzioni di questo genere, tuttavia hanno diverse lacune.

Nella totalità dei casi, non prevedono la possibilità di istruire il bot. In molti casi gli utenti finali commettono errori di sintassi, questo implica che l’agente non sia in grado di carpire ciò che l’utente vuole comunicare, trovandoci a un bivio.

La prima strada, consiste nel realizzare un numero non indifferente di condizioni affinché il bot riconosca gli errori di battitura, di sintassi o frasi sconnesse dal contesto conversazionale: è un lavoro non indifferente che richiede un notevole quantitativo di tempo.

La seconda soluzione invece, è quella di far rientrare tutti questi problemi in un unico stato dove il bot invita l’utente a ripetere ciò che desidera comunicare (rispondendo ad esempio con “Non ho capito, potrebbe ripetere per favore?”).

Un’altra limitazione di cui molti framework soffrono, è legata alla piattaforma d’uso. Spesso, si tratta di agenti da usare per le pagine Facebook, widget per siti web o altri social network che permettono conversazioni tra utenti.

Per tanto, l’utente si vede costretto a usare una specifica piattaforma senza troppa libertà di scelta.

In altri casi, si è limitati a una conversazione “botta e risposta”: non esiste un vero flusso di dialogo. L’utente ha la possibilità di cliccare su bottoni, oppure scrivere frasi ben specifiche dettate dall’agente stesso. Non si va a creare un flusso di dialogo vero e proprio, bensì un botta e risposta tra utente e agente.

In diversi casi invece, è necessario interrogare API esterne per offrire all’utente i dati di cui ha bisogno.

Nella totalità dei casi, è un servizio non previsto dai framework presenti a mercato, ad eccezione di pochi casi in cui non si ha troppo margine di manovra, impedendo al designer dell’ agente di impostare i parametri necessari a invocare le API e gestire le eventuali risposte con i parametri ad esse associate (ovvero, i dati da offrire all’utente finale).

Capitolo 3

Soluzione Proposta

Dialogflow, è un servizio di Google che permette la realizzazione di agenti offrendo non pochi strumenti per un’implementazione a 360 gradi:

* Machine learning: permette l’auto miglioramento del chatbot in funzione di ciò che gli utenti scrivono. Questo significa che possiamo comprendere frasi sgrammaticate, scritte male o in modo parziale.
* Svincolo dalle piattaforme: gli agenti realizzati con Dialogflow permettono di essere usati sia come soluzione standalone, sia come integrazione per servizi di messaggistica (Telegram, Slack, Facebook Messenger, ecc.), sia come servizio aggiuntivo per siti web o applicazioni di terze parti quali Skype, sia come assistente vocale.   
  Quest’ultimo punto è di grande rilievo poiché permette di essere utilizzato con la voce. Dialogflow dà la possibilità di essere abbinato ad assistenti vocali quali Google Home, Alexa e non solo: siamo in grado di integrarlo a sistemi “fai da te” quali Raspberry Pi o altre piattaforme embedded. Tutto ciò permette di fornire supporto a persone con problemi visivi.
* Interrogazione di API esterne: Dialogflow, previa attivazione di un abbonamento a consumo, permette di integrare in totale libertà l’uso di API esterne. Questo permette di integrare al nostro agente informazioni legate a servizi di terze parti, quali ad esempio servizi metereologici, di trasporto e dati legati all’ente stesso che intende usarlo.
* Flusso di dialogo libero: siamo in grado di realizzare un flusso di dialogo in totale libertà. Partendo da frasi ipotetiche che l’utente finale può scrivere, possiamo carpirne informazioni di rilievo (parole, frasi, numeri, codici, date, ecc.) in totale libertà. In funzione del flusso di dialogo che stiamo realizzando, abbiamo piena autonomia e totale controllo sulle frasi scritte dall’utente, termini chiave e risposte che l’agente deve fornire di conseguenza.

Questo framework tuttavia, ha due limitazioni.

Pur offrendo un’interfaccia grafica abbastanza lineare, un utente con scarse conoscenze di programmazione può trovare difficoltà nell’implementazione dell’agente.

Nel caso in cui si desideri implementare chiamate ad API esterne, ha la necessità di saper programmare in Javascript: un agente realizzato in Dialogflow ha ad esso associato un file JS che prevede eventuali invocazioni ad API esterne in cui va gestita la risposta e gli eventuali dati forniti.

La soluzione per far fronte a ciò, è dare la possibilità di disegnare l’agente stesso tramite un sistema puramente visuale, impiegando DrawIo, software libero che permette la realizzazione di grafi. L’utente, tramite un linguaggio visuale ad hoc, ha la possibilità di realizzare un grafo seguendo delle linee guida rendendogli totalmente trasparente il processo di costruzione dell’agente, eliminando i passi in cui la programmazione era precedentemente necessaria.

Questo processo è possibile come segue: una volta realizzato il grafo che rappresenta il flusso di dialogo dell’agente con le eventuali chiamate ad API esterne, ne si realizza un file XML che verrà interpretato dal software di mia realizzazione.

Una volta terminato il processo di interpretazione del file, estraendo tutte le informazioni di rilievo per la realizzazione dell’agente, vengono impiegate le API di Google per poter dare vita all’agente.

Questo processo di interpretazione e realizzazione dell’agente è totalmente trasparente all’utilizzatore finale: potrà visionare in automatico l’agente stesso accedendo al suo account di Dialogflow per poi poterlo impiegare in base alle sue necessità.

Dialogflow: come funziona

Dialogflow, come anticipato nella sezione precedente, è un framework acquisito da Google (precedentemente nominato come “api.ai”) che permette la creazione di agenti intelligenti con tutte le peculiarità già esposte.

Alla base del suo funzionamento, c’è l’agente.

L’agente all’interno di Dialogflow

L’agente in Dialogflow, non è nient’altro che l’entità automatizzata che provvede a sostenere il dialogo con l’utilizzatore umano, carpire il contesto del dialogo, eventuali parole chiave e fornire risposte in funzione delle esigenze espresse dalla sua controparte reale.

Ogni agente, si compone di diversi elementi.

Gli intent

Gli intent rappresentano l’unità elementare del dialogo tra agente ed utente. A fronte delle frasi scritte e/o pronunciate dall’utente, l’agente reagirà secondo quanto da noi impostato.

Vale a dire: rispondere con frasi, immagini, dati o contenuti di natura multimediale.

L’utente ovviamente, all’interno delle frasi espresse può fornire dati di rilievo per l’agente affinché quest’ultimo possa rispondere nella maniera più opportuna al contesto del dialogo e alle esigenze dell’utente stesso.

Motivo per cui possiamo definire i parametri. I parametri sono nient’altro che variabili il cui contenuto varia a seconda di ciò che l’utente dice.

Nella quasi totalità dei casi, un dialogo ben formato non si limita mai a un semplice botta e risposta: avviene su più fasi. Di fatto, per ogni punto del dialogo, è necessario costruire un intent, ma non basta.

Un dialogo per sua natura, si basa su un flusso differenti informazioni scambiate tra gli attori coinvolti: l’agente per sostenerlo, ha necessità di memorizzare tale flusso e le informazioni di rilievo.

Facendo un esempio concreto, l’agente può fornire l’orario del primo treno che partirà dalla stazione A alla stazione B.

Possiamo notare immediatamente che l’agente dovrà tenere conto di tre informazioni cruciali: l’orario di partenza, la stazione in cui l’utente desidera arrivare e ovviamente, quella da cui parte.

In questo caso, l’agente potrà chiedere in prima battuta da dove desidera partire l’utente, in un secondo momento la stazione dove vuole arrivare, l’orario in cui vuole partire e infine, chiedere conferma qual ora l’utente si sia espresso male.

Già al secondo punto, l’agente deve tenere traccia della stazione di partenza espressa dall’utente per mantenere il flusso di dialogo coerente per due motivi: fornire una risposta finale concreta e valida all’utente e rendere quanto più naturale il dialogo.

Questo esempio esprime ciò che è il contesto. Gli intent infatti, possono avere dei contesti di input e dei contesti di output per far fronte a ciò.

Questo meccanismo, permette di mantenere in memoria i parametri di rilievo carpiti dalle frasi espresse da parte dell’utente finale per tutta la durata del dialogo.

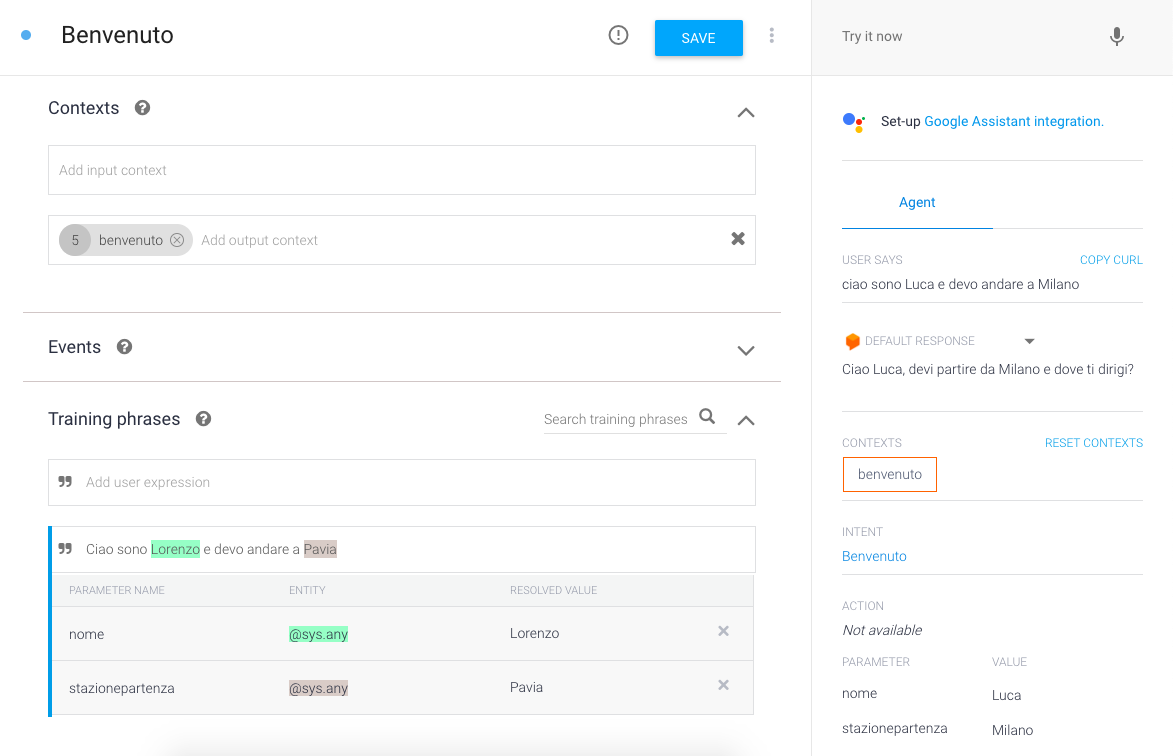


Figura 1: un esempio di intent

Come si evince dall’immagine, ci troviamo di fronte a un Intent il cui nome è “Benvenuto”.

L’intent, ha come contesto di output se stesso: questo perché ci sono due parametri che serviranno per il punto del dialogo successivo, il nome dell’utente e la stazione da cui desidera partire.

Abbiamo infatti la sezione “Training Phrases”: queste, sono le frasi potenziali che l’utente può pronunciare e ad esse, corrisponderà la risposta associata a questo intent.

Nelle training phrases abbiamo modo di specificare quali sono le informazioni di rilievo che l’utente sta offrendo all’agente: nell’esempio le parole “Lorenzo” e “Pavia” sono evidenziate con un rispettivo colore e riportate nella tabella sottostante.

Questa tabella rappresenta ciò che sono i parametri. Essi possono essere nominati con una parola chiave (in questo caso rispettivamente “nome” e “stazionepartenza”) per poi essere impiegati per tre operazioni differenti:

1. Nella risposta diretta all’utente
2. Come dati per chiamate ad API esterne
3. In altri intent purché abbiano nei propri contesti di input quello dell’intent in questione

Di fatto, l’intent ha un’ultima parte:

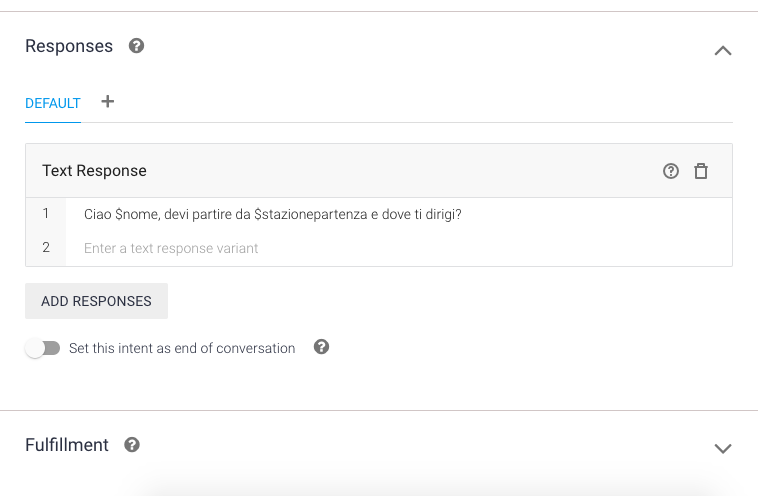


Figura 2: le risposte dell'intent

Abbiamo modo di impostare le risposte dell’agente: esse possono essere testuali, multimediali oppure risultato di una chiamata ad API esterne. Nel caso in esame, abbiamo impostato una semplice frase di testo dove sono indicati i due parametri esplicati in precedenza.

Come infatti si può notare dalla Figura 1, alla frase dell’utente “Ciao sono Luca e devo andare a Milano”, l’agente carpisce le due parole chiave e risponde “Ciao Luca, devi partire da Milano e dove ti dirigi?”.

­­Per comprendere a pieno i contesti, guardiamo la successiva immagine:

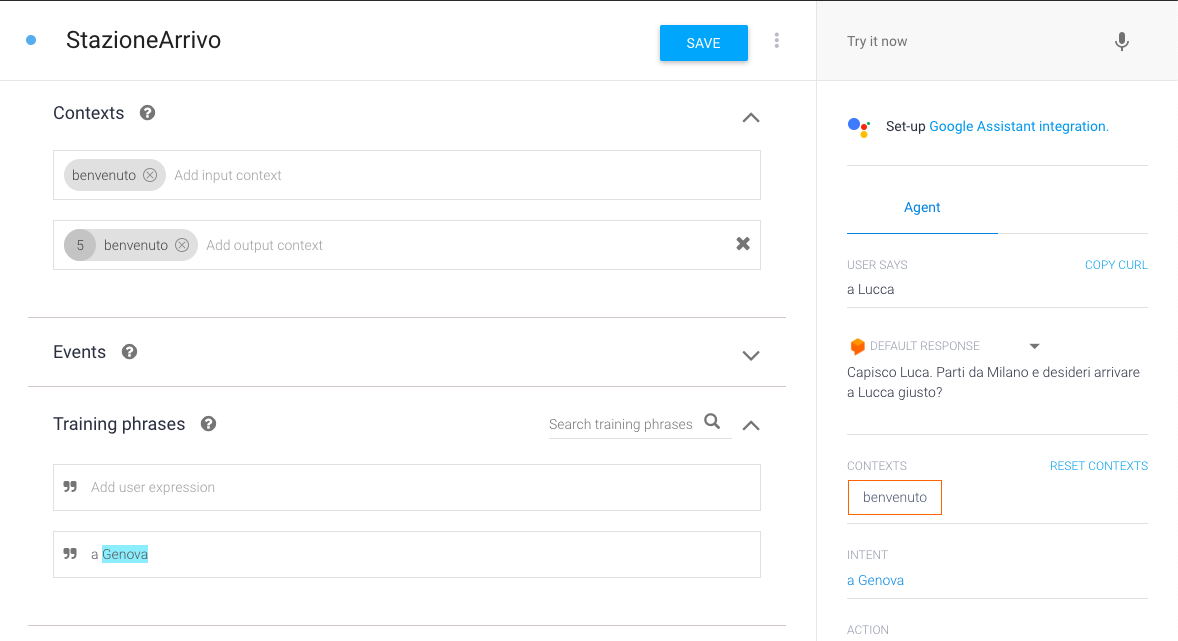


Figura 3: l'intent successivo dell'agente

Questo intent rappresenta il secondo punto del dialogo: ha infatti come contesto di input quello dell’intent esposto in precedenza.

Possiamo sempre notare dall’immagine, che oltre a carpire un altro parametro (ovvero la stazione di arrivo), l’agente risponde all’utente tenendo traccia del suo nome e della sua stazione di partenza: “Capisco Luca. Parti da Milano e desideri arrivare a Lucca giusto?”.

In precedenza, abbiamo detto che, oltre ai contenuti multimediali, l’agente può fornire risposte a seguito dei risultati ottenuti da chiamate ad API esterne: ciò è possibile grazie ai “Fulfillment”

Fulfillment

Un agente creato su Dialogflow permette di invocare API esterne. Queste, tornano utili in diversi scenari, in particolar modo per ottenere informazioni utili all’utente finale.

Ciò che avviene, è la generazione di un file JavaScript in cui si specifica la mappatura degli intent e le chiamate alle API: quest’ultimo punto, avviene tramite protocollo HTTP e l’uso di una callback.

Poiché in diversi casi le chiamate vanno parametrizzate in funzione di ciò che ha espresso l’utente, è possibile ottenere i parametri degli intent o dei contesti coinvolti per le chiamate stesse.

Riprendendo l’esempio espresso in precedenza, una concreta applicazione di questa utilità è invocare un’API di Trenitalia. Una volta ottenute le stazioni del tragitto che intende effettuare l’utente e l’orario in cui desidera partire, possiamo prendere questi tre parametri e invocare il servizio.

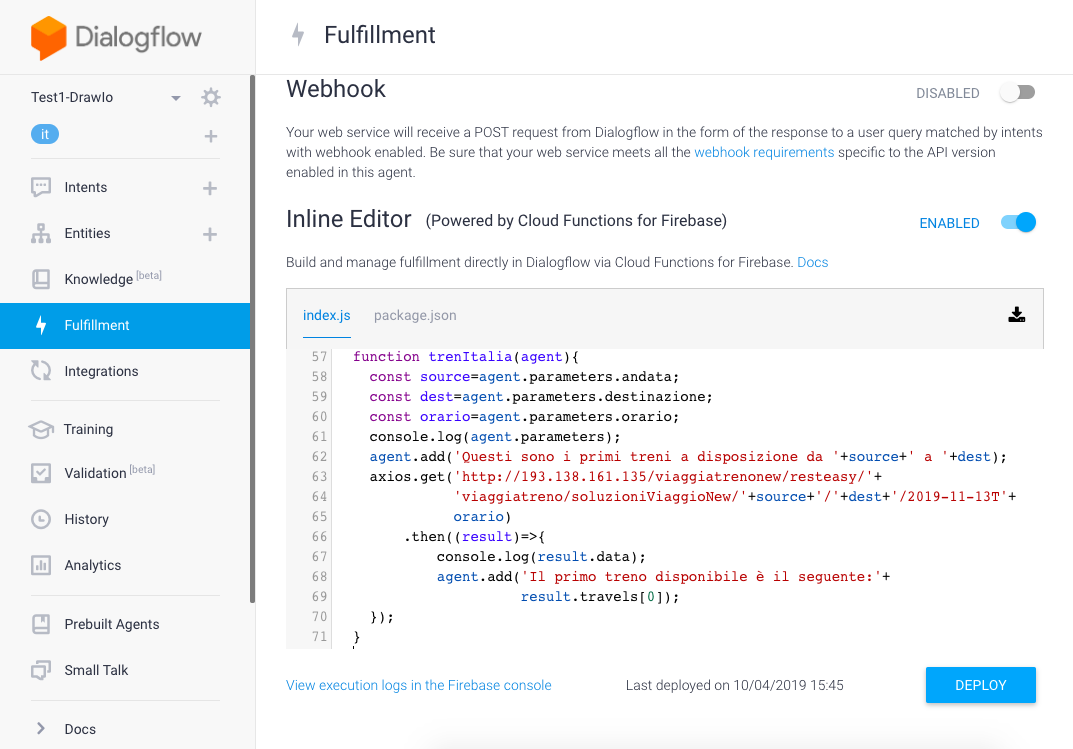


Figura 4: richiesta ad API Trenitalia

Come possiamo notare, nella sezione “Fulfillment” di Dialogflow e dalla funzione “trenItalia”, otteniamo i tre parametri dal contesto conversazionale. Tramite la libreria “axios”, che permette di eseguire richieste HTTP, inseriamo l’URL necessario passando in GET i tre parametri.

Nella callback della chiamata stessa, tramite la funzione “agent.add”, forniamo una risposta all’utente finale tramite l’intent su cui abbiamo effettuato la mappatura:



Figura 5: la mappatura della chiamata API sull'intent

La funzione “intentMap.set”, dato il nome dell’intent e la funzione, permette di mappare la funzione contenente la chiamata HTTP e l’intent stesso: in questo modo possiamo generare risposte provenienti da API esterne all’utente.

Naturalmente questo processo richiede un’esperienza consistente sull’uso di JavaScript e della programmazione, nonché dei protocolli di comunicazione, motivo per cui sarà totalmente trasparente al creatore finale tramite il sistema creato.

Capitolo 4

Architettura della Soluzione Proposta

Stando a quanto esposto nel capitolo precedente, dobbiamo rendere quanto più trasparente possibile il processo di generazione dell’agente ed eliminare la necessità di saper programmare all’utente finale.

Da un punto vista macroscopico, l’architettura della soluzione proposta si compone come segue:



Il secondo, terzo e quarto passo sono trasparenti all’utente finale: una volta “disegnato” l’agente con Draw.io sarà direttamente in grado di utilizzare l’agente nella sua interezza.

Per rendere quanto più minima l’esperienza di programmazione necessaria all’uso di questa architettura, è stato implementato un linguaggio visuale da impiegare per realizzare il grafo rappresentante l’agente.

Affinché l’architettura funzioni, è stato necessario realizzare un interprete in grado di comprendere il contenuto del grafo, estrapolarne le informazioni necessarie per la creazione dell’agente e infine, inoltrare i dati a Dialogflow attraverso le API di Google.

L’interprete, che solo interprete non è, è stato realizzato in Javascript/Typescript con l’ausilio di NodeJS e una libreria sviluppata direttamente da Google.

Infatti non si occupa solo di interpretare il grafo, ma una volta estrapolati i dati utili per costituire l’agente, si interfaccia con le API di Google affinché si possa effettivamente rendere vivo l’agente stesso.

La libreria in questione, è “Dialogflow: Node.js Client”. Essa, permette di interfacciarsi ai servizi di Dialogflow, includendo quindi il nostro agente, attraverso differenti metodi messi a disposizione direttamente mappati alle chiamate HTTP da effettuare a Google.

Di tale libreria, sono stati usati i metodi che permettono la creazione e l’aggiornamento degli intent, cuore pulsante di ogni dialogo.

Limitazioni

Le librerie impiegate non concedono totale margine di libertà e, poiché sono ancora in fase di sviluppo, hanno presentato alcuni bug che, pur non limitando il raggiungimento del nostro obiettivo, ci hanno costretto ad aggirare i problemi.   
In aggiunta a ciò, in questi mesi Google sta effettuando l’upgrade delle API di Dialogflow e migrando tutti i servizi nonché endpoint messi a disposizione.

Il primo limite, è stato quello della creazione dell’agente: esso, non può essere attualmente creato “ex novo”. È necessario crearne uno vuoto accedendo al proprio account sul sito ufficiale di Dialogflow.

Solo a questo punto, è possibile utilizzare le API di Google e quindi il sistema realizzato.

Un'altra limitazione, è stata il riconoscimento dei parametri. Pur forzando tramite le chiamate API gli attributi dei parametri, cosa necessaria dal momento che vanno creati, ce n’è uno in particolare che specifica se il nostro parametro è una lista oppure no.

Nonostante in fase di creazione e aggiornamento si specifica che tutti i parametri dell’agente non sono liste, le API li vedono come tali.

Google permette sì di ricorrere all’uso di API esterne e il sistema realizzato altrettanto, permettendo agli utenti di includere chiamate ad API esterne senza saper programmare in JavaScript.

Google tuttavia, pur permettendo di ricorrere a questa funzionalità, richiede un account a cui sia associato un metodo di pagamento. Il motivo è molto semplice: trattandosi di chiamate HTTP a servizi esterni, si va a consumare traffico dati, di conseguenza è necessario attivare un piano dati a consumo direttamente tramite Google.

Affinchè l’utente finale possa fare un uso completo del sistema, è necessario integrare le sue credenziali Google tramite un file JSON e un ID di sessione ottenibile direttamente dalla console di Dialogflow. Entrambe le procedure vanno eseguite manualmente poiché senza questi dati sarebbe impossibile interfacciarsi con le API di Google e quindi creare l’agente.

Come già esposto in precedenza, per includere API di terze parti nell’agente, è necessario attivare un sistema di pagamenti associato al proprio account Google. Questo servizio, poiché implica consumo dati, si basa su una politca “pay per consume”: in base al numero di richieste effettuate alle API esterne, Google ci addebiterà il costo.

Le risposte di eventuali API esterne, devono essere fornite in JSON ben formato: Dialogflow non accetta altri formati per ottenere risposte da esse.

Linguaggio Visuale - Introduzione

Il sistema realizzato ha come obiettivo principale di permettere la realizzazione di agenti a coloro che non sanno sviluppare codice.

Per tanto, ho realizzato un linguaggio visuale attraverso l’uso di Draw.io: un software gratuito online che permette la realizzazione di grafi.

Questo, permetterà poi la generazione di un file XML: verrà impiegato dall’interprete per estrapolare tutti i componenti che costituiranno l’agente finale, per poi crearlo effettivamente.

Il linguaggio si basa su quattro elementi principali che costituiscono qualsiasi agente: gli intent, i contesti, le credenziali di accesso e i fulfillment.

Linguaggio Visuale - Intent

L’intent viene costruito con il seguente elemento:

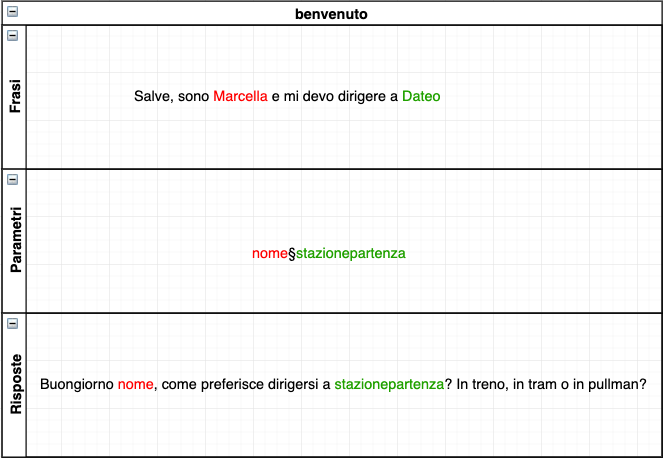


Figura 6: un intent realizzato con il linguaggio visuale

Esso segue la logica di Dialogflow, ha di fatto tre sezioni e un nome (in questo caso “benvenuto”). Le sezioni rappresentano rispettivamente:

1. Frasi: sono le frasi che l’utente può pronunciare durante il dialogo. Nel caso in esempio, troviamo “Salve, sono Marcella e mi devo dirigere a Dateo”.
2. Parametri: sono i parametri a cui associare le parole chiave da carpire all’utente.
3. Risposte: contiene le risposte testuali impiegate dall’agente per rispondere all’utente.
4. API: utilizzata in sostituzione alla sezione “Risposte”. Essa conterrà il nome dell’API che si intende invocare

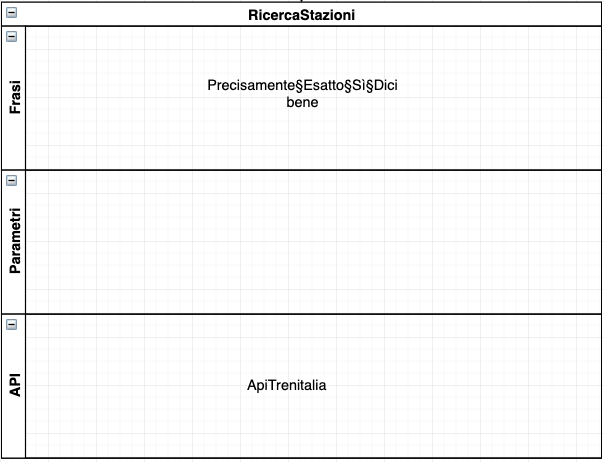


Figura 7: un intent in cui avviene una chiamata ad un API

Linguaggio Visuale – I parametri

Al fine di rendere completa la possibilità di creare gli intent, quindi associare i parametri alle parole chiave delle frasi espresse dall’utente e usarli poi in seguito nelle risposte o nelle API dell’agente, viene fatto l’uso dei colori.

Riprendendo l’esempio presente in figura 6, il nome dell’utente (ovvero “Marcella”), lo si vuole associare al parametro “nome”: quest’ultimo dev’essere colorato con un colore scelto liberamente dall’utilizzatore del sistema.

Affinché l’associazione venga poi rispettata in fase di creazione dell’intent, la parola chiave presente nella trainingphrase (quindi “Marcella”) deve essere colorata con lo stesso colore usato in precedenza.

In questo modo, permettiamo di integrare nell’agente i parametri carpiti dalle frasi pronunciate dall’utente: dobbiamo permettere di utilizzarli nelle risposte. Per far fronte a ciò, è sufficiente impiegare lo stesso colore usato per il parametro in quesitone e utilizzare il nome impostato per il parametro stesso all’interno della risposta.

Nel caso presente in figura 6 infatti, noteremo che la risposta dell’intent “benvenuto” sarà “Buongiorno nome…”.

In diversi casi, i parametri di un intent possono essere utilizzati per fornire risposte in altri intent e, naturalmente, come parametri per invocare API esterne.



Figura : uso di parametri provenienti da intent differenti

Per far fronte a ciò, è sufficiente procedere come esposto in figura 8, rispettando la seguente sintassi:

**nomeIntent.nomeParametro**

Nell’intent “viaggiotreno” vogliamo impiegare all’interno della risposta i due parametri “nome” e “stazionepartenza” impostati nell’intent “benvenuto”. Tenendo presenti le colorazioni in figura 6 e impiegandole nell’intent “viaggiotreno” dovremo appunto scrivere “benvenuto.stazionepartenza” e “benvenuto.nome”.

Questo però non basta: sono parametri provenienti da un altro intent, di conseguenza dobbiamo renderli accessibili tramite il contesto di dialogo.

Linguaggio Visuale – Contesti di dialogo

I contesti di dialogo, come già visto in precedenza, permettono di creare non solo un flusso, ma di tener traccia di tutto quello che l’utente ha espresso all’agente, incluse le parole chiave (i parametri).

Poiché si tratta di un flusso, per connettere due intent è sufficiente usare una freccia.



Figura creazione di un contesto

In questo modo, l’intent “viaggiotreno” avrà come contesto di input “benvenuto”, a loro volta, gli intent connessi con la freccia a “viaggiotreno”, avranno come contesto di input sia “benvenuto”, sia “viaggiotreno”.

Il meccanismo appena descritto, permette di utilizzare i parametri creati in altri intent e la realizzazione di un flusso di dialogo.

In Dialogflow abbiamo visto che in contesti di input e output possono essere creati a discrezione dell’utente. Per semplicità d’uso questo passaggio sarà trasparente all’utente finale.

Di fatto, qualsiasi intent che ha almeno una freccia uscente, avrà automaticamente impostato un contesto di output con il suo stesso nome (quindi l’intent “benvenuto” avrà come contesto di output “benvenuto”).

Allo stesso modo, un intent che ha una freccia entrante, avrà come contesti di input tutti i contesti di output dell’intent a cui è connesso (“viaggiotreno” ha come contesto di input “benvenuto”).

Seguendo questa logica e l’esempio guida mostrato fino ad ora, se si collegasse un terzo intent a “viaggiotreno”, accadrebbe ciò:

* “viaggiotreno” imposterebbe come contesti di output lui stesso più quelli provenienti dalla freccia entrante (ovvero “benvenuto”).
* Il terzo intent avrebbe come contesti di input sia “viaggiatreno”, sia “benvenuto”.

Questo meccanismo, è necessario per creare un automatismo totalmente trasparente nella creazione dei contest idi dialogo.

Linguaggio Visuale – Gestione parametri e frasi multiple

L’utente finale naturalmente, per completezza d’uso del sistema, può inserire quanti parametri, frasi e risposte desidera all’interno di un singolo intent.

Frasi e risposte tuttavia, contengono elementi di punteggiatura, di conseguenza è stato necessario individuare un simbolo di scarso utilizzo che potesse permettere all’interprete di capire come separarle.

In figura 6, noteremo che nelle frasi dell’intent è riportata la seguente stringa:

**“Precisamente§Sì§Esatto§Dici bene"**

Per inserire più di una trainingphrase, parametro o frase di risposta, è sufficiente separarle con il carattere “§”. In questo modo sarà possibile gestire più di un singolo elemento all’interno dell’intent.

Linguaggio Visuale – Gestione credenziali

Come specificato nelle limitazioni delle API di Dialogflow, è necessario creare un file JSON e ottenere dei parametri di sessione affinché il sistema funzioni.

I due parametri sono rispettivamente il nome del progetto per far capire a Dialogflow su quale agente vogliamo intervenire e un sessionID che permette di invocare i servizi.



Figura ottenimento dati sessione

Nella sezione a destra della console di Dialogflow, possiamo testare il nostro agente. Scrivendo una frase qualsiasi, a seguito della risposta dell’agente, apparirà la voce “COPY CURL”.

Cliccando, il contenuto è impostato come segue:

**curl -H "Content-Type: application/json; charset=utf-8" -H "Authorization: Bearer ya29.c.Kl6xBw8cTQHKE1igj99wI92\_rR3grgDgsOVi-7S47RNQM7VjJ-056U4-d5dmkVf1ilL2XHiPtdt2A3IOAQon7OnceJ9HyZf9GpHGzlr7GUsE7cEVbgGtaeOdclkULeNg" -d "{\"queryInput\":{\"text\":{\"text\":\"Ciao\",\"languageCode\":\"it\"}},\"queryParams\":{\"timeZone\":\"Europe/Rome\"}}" "https://dialogflow.googleapis.com/v2beta1/projects/test1-drawio-fsilir/agent/sessions/6bbd5860-a8d6-0e3d-675a-7b232b06d613:detectIntent"**

L’ultima sezione è un URL: da questo dobiamo estrapolare le due informazioni di rilievo.

Nella fattispecie: [https://dialogflow.googleapis.com/v2beta1/projects/**test1-drawio-fsilir**/](https://dialogflow.googleapis.com/v2beta1/projects/test1-drawio-fsilir/)

La parte in grassetto, è il nome del progetto.

Proseguendo: /agent/sessions/**6bbd5860-a8d6-0e3d-675a-7b232b06d613**:

È il sessionID.

Questi due parametri, verranno inseriti nel grafo dell’agente come segue:



Figura le credenziali del chatbot

Tramite la dicitura “ID: idAgente Chiave: chiaveDiSessione”, evitiamo all’utente finale di inserirli manualmente all’interno dell’interprete ogni qual volta è intenzionato a farne uso.

Linguaggio Visuale – API

Per poter invocare un’API, è necessario realizzare il seguente elemento grafico:

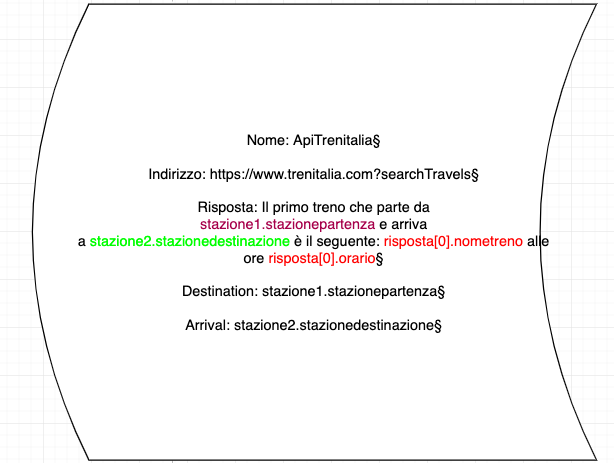


Figura 12: realizzazione di una chiamata ad API esterna

Esso si compone come segue:

1. Nome: è il nome di riferimento per riconoscerla
2. Indirizzo: è l’URL che si intende invocare
3. Risposta: è la risposta che l’agente deve fornire all’utente finale a seguito della chiamata effettuata all’API esterna.
4. Parametri: si tratta dei parametri da impiegare per effettuare la chiamata stessa.

Ognuna delle quattro sezioni termina con il carattere speciale “§” per far capire all’interprete dove terminano.

Linguaggio Visuale – I parametri delle API

Per invocare un’API esterna, è necessario spesso e volentieri impostare dei parametri: in molte situazioni questi derivano direttamente dal contesto di dialogo dell’agente.

Seguendo le direttive d’uso dell’API in questione, è sufficiente impostarle come segue:

**NomeParametroAPI: valore§ | parametroAgente§**

Stando alla sintassi e all’esempio in figura 12, per invocare l’API di Trenitalia, è necessario passare “Destination” come parametro GET ed è nostra intenzione valorizzarlo con il contenuto del parametro “stazionepartenza” presente nell’intent “stazione1”.

Naturalmente, per offrire una libertà d’uso completa all’utente finale, si possono inserire quanti parametri si desidera o quanti sono necessari per invocare l’API desiderata.

Linguaggio Visuale – Le risposte delle API

Un’API viene associata a un intent, di conseguenza sarà necessario offrire una risposta all’utente finale a seguito della chiamata della stessa.

L’utente ha libera discrezione: può usare testo libero, può usare parametri derivanti dal contesto di dialogo (purché colorati, come si vede in figura 12 con l’uso del parametro “stazione1.stazionepartenza”), ma soprattutto può usare i contenuti della risposta data dall’API.

Per quest’ultimo punto, è sufficiente usare la parola chiave “risposta” e colorarla: questa, non è nient’altro che la variabile ottenuta dalla callback a seguito della chiamata HTTP effettuata nel fulfillment dell’agente.

In funzione del contenuto (nel caso di figura 12, è un array e a noi interessa la sua prima cella) e della struttura di risposta, lasciamo libero arbitrio all’utilizzatore finale.

Il tutto è predisposto, come da limitazioni esposte, per oggetti JSON ben formati. Nell’esmepio di figura 12 infatti, usiamo la zeresima cella di un array i cui attributi sono rispettivamente “nometreno” e “orario”: segue infatti la stessa identica sintassi impiegata per leggere gli oggetti JSON.

Interprete – Introduzione

Draw.io come esposto in precedenza, permette di esportare i grafi creati in formato XML. Questi, affinché possano essere letti, estrapolate le informazioni di rilievo e inviati a Google, necessitano di essere interpretati.

Attraverso due librerie per NodeJS, è stato possibile realizzare un sistema di lettura del grafo: “libxmljs” e “xml2js”.

Seguendo la logica di costruzione di un agente, è necessario estrapolare tutte le informazioni legate agli intent, quindi le trainingphrases, i parametri e le risposte, i contesti di dialogo, le chiamate ad API esterne, nonché le credenziali dell’agente.

Bibliografia

* D. Fichter e J. Wisniewski, «Chatbots introduce conversational user interfaces,» *Online Searcher,* vol. 41(1), pp. 56-58, 2017.
* Valtolina, Stefano; Barricelli, Barbara Rita; Di Gaetano, Serena (2019): Communicability of traditional interfaces VS chatbots in healthcare and smart home domains. In: *Behaviour & Information Technology,* 195, pp. 1–25. DOI: 10.1080/0144929X.2019.1637025
* Følstad e P. Brandtzæg , «Chatbots and the new world of HCI,» *Interactions,* vol. 24, n. 4, pp. 38-42, 2017.
* Google, «What is conversation design?,» [Online]. Available: https://designguidelines.withgoogle.com/conversation/conversation-design/what-is-conversation-design.html#. [Consultato il giorno 01 Ottobre 2019].
* Manychat, «Flow Builder,» [Online]. Available: http://flowbuilder.manychat.com/. [Consultato il giorno 01 Ottobre 2019].
* Dialogflow: Node.js Client [Online]. Available: <https://github.com/googleapis/nodejs-dialogflow/>