

Seminario di Cultura Digitale

Relazione Finale

Il cuore dei bot

Tecnologie invisibili per un accesso
colloquiale alla conoscenza

Raffaele Masotti

455234

La relazione tratterà questioni legate all'elaborazione del linguaggio naturale e all'intelligenza artificiale, concentrandosi sulle applicazioni pratiche volte a sviluppare interfacce conversazionali. Verrà inoltre tracciata una panoramica dei principali SaaS (software as a service) utili per sviluppare prodotti in questo ambito e delle aziende innovative attive nel settore.

Indice

Introduzione	4
Metodi e tecnologie	6
Intelligenza Artificiale	6
Natural Language Processing	7
Matching learning	8
Interfacce conversazionali	9
La conversazione	9
Cosa sono i bot	11
I bot per la gestione delle relazioni con i clienti	13
Aziende italiane attive nel settore	14
Awhy	14
Userbot	15
SaaS per costruire esperienze conversazionali	17
Wit.ai	17
Api.ai	18
LUIS	20
Conclusioni	21
Bibliografia	22
Sitografia	23

Introduzione

Il mondo della fantascienza è sempre stato affascinato dalla possibilità di interagire con le macchine simulando una classica conversazione tra persone. L'idea che le macchine possano interagire con l'uomo al pari degli altri esseri viventi si fa largo nelle menti dei pensatori fin dalla prima metà del ventesimo secolo, per poi concretizzarsi in riflessioni più approfondite come l'articolo di Alan Mathison Turing, pubblicato nel 1950 sulla rivista *Mind* e intitolato *Computing Machinery and Intelligence*¹. Nell'articolo, il matematico propose una procedura pratica per stabilire quando una macchina possa definirsi intelligente, una sorta di “gioco dell'imitazione” in cui un uomo interagisce parallelamente con un calcolatore programmato per fingersi umano e con un terzo soggetto. Il test, etichettato successivamente come “il test di Turing”, ha esito positivo quando l'esaminatore non è in grado di distinguere l'uomo dalla macchina. Questa, che potrebbe sembrare una prestazione ormai superata dalla recente tecnologia, è in realtà una sfida che tutt'oggi affronta chi ad esempio propone tipologie di servizi per automatizzare il supporto ai clienti. Le teorie di Turing, spesso criticato per l'eccessivo ottimismo delle sue ipotesi, hanno costituito il punto di partenza di molte speculazioni teoriche e filosofiche, intrecciandosi successivamente con le più recenti scienze cognitive.

Qualche anno prima, nel libro *Controllo e Comunicazione nell'animale e nella macchina*, il matematico e filosofo statunitense Norbert Wiener conia il termine cibernetica e afferma:

I molti automi dell'epoca presente sono collegati al mondo esterno sia per quanto riguarda la ricezione di impressioni, che per l'esecuzione di azioni. Essi contengono organi sensoriali, effettori, e l'equivalente di un sistema nervoso per il trasferimento dell'informazione dagli uni agli altri ... Non è affatto miracoloso che possano essere trattati in un'unica teoria assieme con i meccanismi della fisiologia²

Equiparare il funzionamento degli organismi a quello delle macchine è il primo remoto passo verso l'idea che delle interfacce possano sostituirsi all'uomo. Un'idea che richiede non solo avanzamento dal punto di vista tecnologico, ma anche un

¹ A. M. Turing, *Computing Machinery and Intelligence*, *Mind* n. 49, 1950, pp. 433-460. Consultabile online nell'archivio della rivista *Mind*: <http://mind.oxfordjournals.org/content/LIX/236/433.full.pdf+html>

² N. Wiener, *Cybernetics, or control and communication in the animal and the machine*, prima edizione: The MIT Press, Cambridge (MA), 1948

cambio di paradigma sul piano culturale. Concetti ripresi di recente anche dallo studioso ungherese Tibor Vamos, che agli inizi degli anni ottanta, nell'articolo collaborativo della IIASA³, fa notare come il paradigma delle formiche di Hofstadter⁴ -secondo cui le società di formiche sono in grado di elaborare risposte collettive di molto superiori a quelle consentite dall'isolato sistema nervoso di ciascuno dei suoi membri, grazie ad una modalità cooperativa priva di gerarchia concettuale basata sulla propagazione dell'informazione- si possa applicare a sistemi hardware e software altamente interconnessi.

La creazione di un'interfaccia conversazionale non può prescindere dallo studio dei destinatari, dall'analisi dei dispositivi su cui verrà fruita e dalla tipologia di interazione richiesta. Negli ultimi anni, i grandi nomi dell'informatica hanno progressivamente aperto sempre più le porte allo sviluppo di software in grado di interagire con gli utenti via chat, andando a rispolverare un trend che già nei primi anni novanta muoveva i primi timidi passi, con soluzioni come l'assistente virtuale integrato nella suite Office proposta da Microsoft⁵. Al fine di migliorare il servizio offerto ai clienti, oggi sempre più aziende scelgono di integrare le proprie applicazioni di business a queste nuove tipologie di servizi. Un po' come nel test di Turing, la sfida che oggi affrontano è riuscire a ingannare (in senso buono) gli utenti, che al fine di varcare l'aspettativa qualitativa e godere di un'esperienza ottimale non devono essere in grado di riconoscere la differenza fra l'operatore virtuale e l'operatore reale.

³ *International Institute for Applied Systems Analysis*

⁴ Tibor Vamos, *Cooperative Systems - An evolution perspective*, p.2, 1983

⁵ L'Assistente di Office è stato incluso in Microsoft Office 97 e successivamente nelle versioni seguenti, fino a Office 2007. La caratteristica era una modalità di accesso tramite conversazione alla guida e presentava varie funzioni di aiuto, offrendo consigli basandosi su algoritmi bayesiani nelle versioni 97-2003 per Windows e 98-2004 per Macintosh

Metodi e tecnologie

Prima di parlare in dettaglio di cosa siano le interfacce conversazionali e di come possano applicarsi con successo in ambito business, è utile approfondire i metodi e le tecnologie alla base di questi sistemi. Sono molti i settori disciplinari che si sovrappongono quando si parla di soluzioni in grado di interagire con l'uomo. Qui vengono volutamente toccati solo i temi necessari a fornire un background necessario per comprendere nell'insieme le interfacce descritte.

Intelligenza Artificiale

L'intelligenza artificiale è presente nella letteratura ancor prima che il progresso informatico iniziasse il suo corso. La mitologia greca riporta storie di uomini meccanici⁶ progettati per imitare il nostro comportamento. Se inizialmente in Europa il tentativo di costruire cervelli meccanici è partito dall'idea di potenziamento progressivo della memoria di base e delle possibilità aritmetiche dei computer, poiché la comprensione di come funziona la nostra mente è progredita, il concetto di ciò che costituisce l'intelligenza artificiale è cambiato. Piuttosto che calcoli sempre più complessi, l'obiettivo si è spostato sull'imitazione dei processi decisionali e lo svolgimento di compiti in modo sempre più umani. Le intelligenze artificiali, entità progettate per agire in modo intelligente, sono spesso categorizzate in due gruppi fondamentali: applicati e generali. L'intelligenza artificiale applicata è quella più comune e si riferisce a sistemi progettati per effettuare operazioni in modo autonomo, come svolgere azioni specifiche o manovrare veicoli. Invece, l'intelligenza artificiale generalizzata (anche detta forte) si riferisce a sistemi o dispositivi che in linea teorica possono gestire qualunque compito e apprendere in maniera indipendente dall'intervento dell'uomo.

L'intelligenza artificiale così come la conosciamo oggi e legata all'ambito informatico, è nata negli anni sessanta del ventesimo secolo e ingloba in essa diversi

⁶ Nella mitologia greca, Talos è un gigante di bronzo il cui scopo era proteggere Creta dai pirati. Secondo una versione, Talos -il quale aveva una sola vena, attraverso la quale scorreva il ichor, il sangue divino degli olimpi- fu forgiato dal dio Hephaestus per regalarlo al re Minos di Creta.

sottoinsiemi sviluppatisi successivamente, come l'elaborazione del linguaggio e l'apprendimento automatico, che risultano fondamentali per lo sviluppo di sistemi conversazionali complessi.

Natural Language Processing

Il Natural Language Processing⁷ è un sottoinsieme della branca dell'intelligenza artificiale. La concezione di questo termine rientra nell'ambito del linguaggio naturale, riferendosi nel dettaglio al trattamento informatico del linguaggio senza uno scopo preciso. Il linguaggio naturale, inteso come la lingua parlata nel quotidiano, si contrappone al concetto di linguaggio formale matematico informatico. A differenza di quest'ultimo, il linguaggio naturale è molto più complesso da elaborare, a causa delle concretizzazioni creative del linguaggio e delle frequenti ambiguità di natura referenziale che caratterizzano le normali comunicazioni umane. Il Natural Language Processing può essere inoltre nei seguenti campi:

- signal processing, l'elaborazione numerica dei segnali;
- sentiment analysis, il significato dato dall'interlocutore alla frase;
- syntactic analysis, lo studio sintattico della struttura della frase.

Al fine di interagire con gli esseri umani, l'elaborazione del linguaggio naturale costituisce un tassello fondamentale, anche se indipendentemente non sufficiente a costruire interfacce conversazionali in grado di supportare adeguatamente la complessità dell'interazione umana.

⁷ NLP, nella traduzione italiana: elaborazione del linguaggio naturale

Matching learning

Il termine machine learning, tradotto letteralmente “apprendimento macchina”, comunemente reso in italiano come “apprendimento automatico”, è un ambito interno all’intelligenza artificiale finalizzato a realizzare sistemi in grado di imparare autonomamente, senza necessità di intervento umano per apprendere una determinata struttura di dati. Il machine learning si differenzia dalla programmazione di applicazioni in quanto l'intento non è programmare un'azione specifica, ma apprendere conoscenza grazie alla quale si eseguiranno le azioni. Si tratta dunque di automatizzare la creazione di modelli analitici, grazie ad algoritmi che apprendono iterativamente partendo da una base di dati esistente. L'aspetto iterativo è fondamentale nella costruzione di un sistema machine learning: una macchina programmata per l'apprendimento automatico deve essere in grado di adattarsi in maniera indipendente a nuovi dati, in modo tale da riuscire sempre a ricavare informazioni utili a risolvere problemi di natura differente.

Interfacce conversazionali

L'interrogativo che ci si può porre è se il linguaggio naturale potrà costituire multilateralmente l'interfaccia utente del futuro. L'obiettivo delle interfacce conversazionali sarà sempre più quello di offrire agli utenti un'interazione naturale nella vita di tutti i giorni, tramite lo strumento più spontaneamente utilizzato dall'uomo: la lingua in atto. Da un lato questo significa affrontare le sfide della comprensione del linguaggio naturale, dall'altro di convertire in valore il prodotto dall'interazione, applicandolo al dominio di interesse. Uno degli ambiti in cui questi sistemi possono dare prova della propria efficacia è certamente quello dell'assistenza clienti, in cui basi di conoscenza possono essere sfruttate per produrre in maniera intelligente e adattiva risposte alle richieste degli utenti.

La conversazione

Il fulcro di questa nuova tipologia di interfacce è costituito da un'interazione verbale molto comune per l'uomo: la conversazione. Con questo termine si intende uno strumento di conoscenza atto a scambiare idee e punti di vista differenti, oltre che a far emergere collegamenti fra le proprie conoscenze e quelle comunicate dagli altri. Per il cibernetico e psicologo inglese Gordon Pask, che negli anni settanta elabora un modello per spiegare la "costruzione di conoscenza", la conversazione è un atto mentale che si serve del linguaggio per aprirsi all'altro, esplicitando le proprie idee su una questione di interesse comune e facendo in modo che anche l'altro le manifesti. La conversazione serve inoltre ad attuare i propri pensieri, a confrontarli con gli altri e ad aprirsi a nuove conoscenze.

La conversazione può essere intesa anche nei presupposti del principio collaborativo delle massime conversazionali di Paul Grice. Il principio di cooperazione di Grice, ispirato alle categorie del filosofo tedesco Immanuel Kant esposte nell'opera *Fondamenta della metafisica dei costumi* di fine settecento, si articola su quattro massime⁸ che possono essere applicate anche alle interfacce

⁸ I dettagli della teoria di Grice sono consultabili nelle slide a questo indirizzo: <http://www.lettere.uniroma1.it/sites/default/files/407/Linguaggio%20e%20comunicazione%20II.pdf>

conversazionali:

- massima della quantità, “Dai un contributo tanto informativo quanto è richiesto. Non dare un contributo più informativo di quanto è richiesto”. Il volume dell'informazione deve quindi essere adeguato a quanto richiesto dallo scopo dello scambio linguistico. Più informazioni del dovuto, non rappresenta una violazione del principio di cooperazione, ma prolunga la conversazione rendendola più difficile da seguire. Di conseguenza, le interazioni con un bot dovrebbero essere brevi ed evitare generalizzazioni;

- massima della qualità, “Tenta di dare un contributo che sia vero. Non dire ciò che credi falso. Non dire ciò di cui non hai prove adeguate”. Un buon punto di partenza consiste nell'evitare presupposti poco verosimili sulle capacità di conversazione del bot con cui gli utenti iniziano una conversazione. L'impiego di tecniche di Natural Language Processing non sempre riesce a soddisfare le aspettative degli utenti. È necessario che le funzionalità e soprattutto i limiti di comprensione del linguaggio naturale del bot vengano assimilate gradualmente;

- massima della relazione, “Sii pertinente”. Nel realizzare un'interfaccia conversazionale, è necessario tenere conto necessariamente delle modalità con cui verrà fruita, stabilendo i limiti che aiutino gli utenti a ottenere le informazioni contestuali alle proprie richieste. Una conversazione può essere imprevedibile, quindi è utile disegnare dei percorsi per evitare che il flusso comunicativo esca fuori traccia;

- massima del modo, “Sii perspicuo. Evita l'oscurità. Evita l'ambiguità. Sii breve. Sii ordinato nell'esposizione”. L'utilizzo di un linguaggio colloquiale e la costruzione di un carattere riconoscibile può aiutare a mantenere la conversazione su un piano più naturale. Il modo non si riferisce a quanto detto, quanto al modo con cui si espone. È sempre utile incitare l'utente a compiere azioni specifiche, evitando vicoli ciechi che potrebbero costituire un fattore terminale per la comunicazione.

La possibilità di relazionarsi con le macchine con lo stesso strumento che da sempre usiamo per costruire e assimilare conoscenza, va nella direzione di rendere la tecnologia sempre più trasparente nel quotidiano, lasciando le persone libere di concentrarsi sui contenuti piuttosto che sugli strumenti utilizzati per recuperarli.

Cosa sono i bot

I bot, in passato concepiti come semplici sistemi di automatizzazione per task ripetitivi, sono oggi comunemente intesi come software in grado di conversare con l'uomo in modo naturale. Durante l'evento di Microsoft Build 2016, il CEO Satya Nadella ha dichiarato che i bot sono le nuove app (“bots are the new apps”⁹), a testimonianza del riaccessso interesse che negli ultimi anni stanno riscuotendo queste applicazioni.

I bot sono in grado di automatizzare dei compiti e tradurre lingua non strutturata in dati organizzati, promettendo di portare un nuovo livello di interazione tra uomo e computer. L'intelligenza artificiale non è sempre stata una componente presente nella realizzazione di bot, ma costituisce la chiave che consente ai bot di analizzare in maniera profonda quello che l'utente comunica, comprenderne gli intenti e agire di conseguenza in maniera adattiva e non semplicemente selezionare risposte pre-confezionate da un archivio statico.

A seconda delle tecnologie hardware che vengono utilizzate, i bot sono in grado di interagire con l'uomo in diversi modi, ad esempio: tramite interfacce di instant messaging, via e-mail e direttamente mediante la voce. La possibilità di interagire tramite la voce, consente ai bot di connettersi con gli utenti in qualsiasi situazione, anche senza l'utilizzo diretto delle mani e dei dispositivi che li rendono accessibili.

Per molti anni la declinazione più comune dei bot sono stati i chatbot. Il blog SEMrush formula questa definizione dei chatbot¹⁰:

Un chatbot è un servizio, caratterizzato da regole, schemi e (a volte) intelligenza artificiale, che permette agli utenti di interagire via chat, senza l'intervento umano. Può riguardare una serie infinita di argomenti, impieghi e applicazioni, da quelli più funzionali a quelli per mero divertimento ed è possibile trovarli quasi in ogni servizio di chat (Facebook Messenger, Slack,

⁹ <http://www.wired.co.uk/article/microsoft-build-bots-ai-cortana-keynote-conference>

¹⁰ <https://it.semrush.com/blog/chatbot-cosa-sono/>

Telegram, Sms, etc.).

Di fatto oggi molte soluzioni di chatbot, soprattutto in ambito business, prevedono una doppia presenza: da un lato i bot per offrire le risposte più semplici e rapide, dall'altro lato degli operatori reali, pronti a intervenire per colmare le richieste più complesse.

I bot per la gestione delle relazioni con i clienti

In un'analisi del 2011¹¹, la società di ricerca Gartner stimava che entro il 2020 i consumatori gestiranno l'85% delle loro relazioni con aziende e servizi senza nessuna interazione umana. Viene previsto inoltre che entro il 2020 le transazioni effettuate da dispositivi mobili saranno di circa 50 miliardi di euro, ovvero un terzo del mercato online globale. Un dato confermato anche relativamente al contesto italiano dall'Osservatorio Innovazione Digitale nel Turismo¹², che prevede una progressiva crescita della spesa digitale via mobile. In un contesto come questo, rispondere rapidamente alle domande dei clienti, spesso significa: non perdere clienti, offrire delle esperienze più soddisfacenti e sfruttare una base di informazioni esistente per risolvere problemi in maniera dinamica. I bot possono costituire una soluzione per assistere i consumatori con operazioni difficili, per offrire raccomandazioni e raccogliere dati utili, garantendo un supporto real-time e proattivo. Ad esempio, un bot integrato in un sito web di una compagnia di vendita al dettaglio potrebbe rispondere alle domande relative alle spedizioni, resi e suggerimenti sulla base delle domande poste. Anche se il bot non fosse in grado di assistere il cliente in tutte le sue richieste, potrebbe in ogni caso raccogliere informazioni preliminari (nome del cliente, richiesta, contatti) e inoltrare la richiesta al servizio clienti, risparmiando tempo considerevole per il call center dell'azienda.

Quindi in parte c'è la speranza che queste entità possano effettivamente lavorare, riflesso dell'esplosione recente di intelligenza artificiale, dall'altra offrono alle aziende un nuovo canale per raggiungere gli utenti, andando a occupare una fetta di mercato in crescita. La domanda diventa dunque se soluzioni di questo tipo abbiano bisogno dell'intelligenza artificiale generalizzata o se sia possibile restringere il dominio in modo da consentire alla statistica di prevedere con buona percentuale tutti gli scenari necessari a soddisfare l'esigenza di una determinata applicazione (ad esempio il supporto clienti per la vendita di una specifica categoria di prodotti). Tuttavia, creatività e libero arbitrio nella costruzione delle conversazioni possono sempre stimolare gli utenti ad allontanarsi da un certo dominio, e le aspettative degli utenti spesso crollano appena viene intravista la limitazione del mezzo, lasciando alla ricerca di risposte una nuova domanda: cosa possono chiedere? Questa e altre problematiche fanno emergere consapevolezza sul lungo cammino che le interfacce

¹¹ https://www.gartner.com/imagesrv/summits/docs/na/customer-360/C360_2011_brochure_FINAL.pdf

¹² https://www.osservatori.net/it_it/osservatori/osservatori/innovazione-digitale-nel-turismo

conversazionali devono percorrere per raggiungere reali possibilità di “conversazione”, ma ciò non toglie che un bot non possa funzionare in maniera ottimale su qualcosa di molto specifico, in cui l'utente è a conoscenza di cosa sia possibile chiedere.

Questo nuovo canale, inoltre, potrebbe costituire una via parallela alle tipiche modalità di fruizione, come siti web e app per il telefono. Il crescente utilizzo delle piattaforme di messaggistica istantanea e le API esposte dai principali fornitori di questi servizi, permette di distribuire i bot attraverso nuove piattaforme, dando alle aziende la possibilità di comunicare con i clienti in modo più diretto e naturale, mettendo a disposizione agenti virtuali che possano essere pronti a comprendere le richieste dei clienti in qualsiasi ora, inviando magari proposte e offerte, combinando sia contesto, sia dati personali e interessi.

Aziende italiane attive nel settore

Non mancano nel panorama italiano delle aziende che cercano di sfruttare il nuovo potenziale connesso a soluzioni conversazionali, il cui obiettivo è quello di proporre servizi innovativi che permettano di migliorare il rapporto fra le aziende e i clienti.

Awhy

Awhy nasce da studente di economia durante il suo periodo di stage in un call center per un'azienda che vende libri online e si sviluppa successivamente nell'ambito del PhD+, il programma dell'Università di Pisa finalizzato a promuovere e incoraggiare lo spirito imprenditoriale e di innovazione tra studenti di laurea magistrale, dottorandi, dottori di ricerca e docenti. L'azienda si è poi stabilita a Bologna, anche grazie a Telecom Italia che, attraverso il programma Working Capital, ha finanziato la startup nell'acceleratore Tim WCap.

Awby offre come servizio principale un'intelligenza artificiale per le aziende, costruendo un'assistenza automatizzata per migliorare il loro customer care. Mediante la semplice installazione di un widget sul sito dell'azienda, questo è in grado di interagire in linguaggio naturale, rispondendo alle domande ed esigenze dei clienti. Esistono due livelli di questo sistema, uno attivo e uno passivo. Il livello passivo è relativo alle domande che l'utente rivolge direttamente, mentre il livello attivo riguarda percorsi generati appositamente sulla base delle operazioni che l'utente svolge sul sito.

L'architettura che permette ad Awby di fornire questo servizio è una pipeline piuttosto articolata, composta dai seguenti moduli: sentence detector, tokenizer, typo correction, check greetings, statistical part of speech tagger, dictionary part of speech tagger, lemmatizer, synonyms e query vector.

Il sentence detector riconosce se l'utente ha scritto più di una frase. Il tokenizer si occupa di dividere ogni query in singole parole, eliminando i caratteri non necessari all'analisi. Il typo correction correggere possibili errori di battitura dell'utente, in modo da riconoscere in maniera precisa cosa l'utente vuole sapere. Il check greetings verifica che si tratti di semplici convenevoli e frasi di apertura, identificando in caso anche eventuali insulti da parte dell'utente. Lo statistical part of speech tagger è il primo step di analisi del linguaggio, in cui vengono associate le parole al ruolo che hanno nel discorso. Per effettuare questa operazione in maniera appropriata, viene usato un hidden markov model (un classificatore supervisionato). A conclusione del passaggio precedente, gli elementi esclusi vengono analizzati da un secondo part of speech tagger, controllando la presenza in un dizionario di lemmi ed eventualmente re-inseriti nella pipeline. I termini selezionati dal processo precedente vengono lemmatizzati dal lemmatizer per risolvere potenziali ambiguità. Il synonyms espande la comprensione della frase di partenza, aggiungendo ad ogni termine una lista di possibili sinonimi. Infine, la query iniziale e tutte quelle parallelamente generate nei passaggi precedenti vengono vettorizzate e confrontate con l'archivio di FAQ costruite per lo specifico dominio del cliente.

Userbot

Userbot è nata con un obiettivo molto simile a quello di Awby, con cui

condivide l'obiettivo generale e approccio. Si differenzia per l'esteso supporto che va oltre l'integrazione sui siti web, sconfinando su Messenger, Telegram, iOS, Android, su altri CRM.

Come altre aziende attive nel settore, anche Userbot ha incontrato i tipici problemi di questo tipo di applicazioni. Ad esempio: l'utilizzo frequente di abbreviazioni nelle chat¹³, la presenza di frasi grammaticalmente scorrette, le frasi spezzate in più intervalli, le domande fuori dominio e l'incomprensione della risposta ricevuta. Questi problemi sono stati affrontati con l'introduzione di un vocabolario delle abbreviazioni, di un correttore ortografico e delle possibilità di definire abbreviazioni personalizzate nel sistema. Se la frase ricevuta dal sistema non contiene nessuna domanda, questa viene automaticamente accorpata alla frase seguente nel tentativo di ricostruire il contesto e riuscire a dare una risposta realistica. L'argomento delle domande precedenti viene sempre memorizzato, nel caso in cui l'utente aggiunga informazioni nelle frasi successive. Inoltre, uno specifico livello dell'analisi si occupa di riconoscere le domande non tecniche, utilizzando un diverso dizionario di conversazione generica e cercando di riportare la conversazione sul dominio di interesse. Si tratta quindi di un sistema a due facce in grado di scegliere le frasi più pertinenti sulla base dell'addestramento sul contesto di riferimento. Infine, anche Userbot prevede la casistica in cui l'agente virtuale non sia in grado di rispondere alla domanda, re-indirizzando quindi il quesito ad un agente umano.

¹³ Ad esempio: c6, qnd?, ecc.

SaaS per costruire esperienze conversazionali

Molte delle tecnologie fin qui descritte vengono oggi inglobate in scatole nere pronte all'uso (i bot engine), che consentono di realizzare esperienze conversazionali senza conoscere approfonditamente tutti i livelli dello stack e le tecnologie alla base di questi strumenti. Alcune interfacce visuali di bot engine consentono addirittura di implementare bot rudimentali senza scrivere nemmeno una riga di codice. Se da un lato questo consente di ottenere dei risultati in modo molto semplice, dall'altro spesso vengono a mancare elementi di specificità del contesto necessari ad applicare questi prodotti a un dominio ristretto. Inoltre, spesso non permettono di interagire in modo generale con servizi terzi. Ad esempio, non sempre è possibile accedere a database esterni. Proprio per tali limiti, è necessario selezionare con cura i servizi che possano costituire singoli tasselli da incastrare a posteriori, piuttosto che soluzioni inclusive dalle possibilità spesso limitate.

L'attenzione dei grandi nomi dell'informatica a questo ambito -come ad esempio Google, Microsoft e Facebook- ha portato negli ultimi anni a una serie di acquisizioni di startup innovative, pensate per fornire il layer più alto di questo stack tecnologico, costruendo un'interfaccia facile da interrogare e adattare ai propri use case.

Wit.ai

Nel 2015 l'azienda Wit.ai, specializzata nel riconoscimento vocale e fondata appena 18 mesi prima, ha ufficialmente riportato sul proprio sito l'annuncio della fusione con Facebook. Wit.ai ha sviluppato una piattaforma flessibile per il riconoscimento vocale, che consiste in una serie di API che ogni sviluppatore, registrando un'apposita chiave, può includere nella propria applicazione. Le API si occupano di convertire parole e frasi in dati strutturati, interpretabili da una macchina. Una volta ottenuti questi dati, è compito poi del programmatore decidere come utilizzarli. Ad esempio, con l'utilizzo di queste API possono essere realizzate applicazioni di domotica e interfacce di controllo vocale.

Per aiutare gli sviluppatori a costruire interfacce conversazionali basate su testo, nel 2016 l'azienda ha rilasciato una soluzione bot engine, promuovendo la realizzazione di bot tramite l'ecosistema degli strumenti di Natural Language Processing. Il motore di bot è stato pensato in modo da fornire supporto allo sviluppo di interfacce anche se senza un'apposita interfaccia grafica. La piattaforma permette quindi di creare interfacce conversazionali high-end, in grado di tradurre il linguaggio naturale in dati strutturati e di apprendere nel corso delle interazioni, migliorando nel tempo grazie a tecniche di apprendimento automatico. Dal momento in cui il bot engine è stato rilasciato, la comunità è cresciuta da ventimila utenti a più di centomila sviluppatori in un tempo molto breve. La maggior parte degli sviluppatori ha creato bot per Messenger, Slack, Telegram e altre piattaforme. Tuttavia, gran parte delle chiamate API di Wit provenivano dall'API di Natural Language Processing. Questi dati hanno portato Facebook, nel corso di quest'anno, alla scelta di lanciare una piattaforma di Natural Language Processing integrata per Facebook Messenger, con l'intento di rendere i bot più affidabili, precisi e scalabili per ogni sviluppatore. Come confermato sul blog¹⁴ di Wit.ai, i servizi originariamente lanciati dall'azienda verranno chiusi nel giro di un anno per concentrare i propri sforzi sulla piattaforma Built-in NLP di Facebook.

Api.ai

Il sito di Api.ai riporta come frase d'apertura “Build delightful and natural conversational experiences”. L'azienda è stata fondata nel 2014 ed è stata acquisita da Google nel 2016. Api.ai fornisce un servizio per elaborare il linguaggio naturale, fornendo agli sviluppatori strumenti avanzati per costruire interfacce conversazionali. Come Wit.ai, permette di riconoscere ciò che l'utente comunica tramite messaggi testuali o vocali, trasformando questo input in dati utilizzabili per prendere decisioni. Si tratta di una soluzione all-inclusive che ingloba all'interno della sua piattaforma gli elementi necessari a progettare scenari di conversazione a cui corrisponde una specifica ontologia di azioni. La piattaforma include dei pacchetti di conoscenza pre-costruiti, relativi a diversi domini, come ad esempio: notizie, orari, meteo, riproduzione musicale e altri, permettendo di interpretare gli

¹⁴ <https://wit.ai/blog/2017/07/27/sunsetting-stories>

intenti delle conversazioni più comuni in modo immediato. È incluso inoltre un sistema di machine learning che può essere abilitato opzionalmente, il quale apprende dagli esempi forniti in fase di training e durante le interazioni con gli utenti finali al fine di individuare gli intenti in modo sempre più accurato. Uno dei punti di forza è certamente l'alto livello di integrazione sia con gli altri servizi cloud di Google, sia con servizi terzi come: Alexa, Skype, Slack, Cortana e altri sistemi di messaggistica. Il sistema supporta oltre 15 lingue e fornisce delle apposite librerie per consentire l'accesso alle API tramite diverse piattaforme e linguaggi di programmazione. Nonostante vengano elaborate milioni di richieste utente al giorno, il servizio riesce a garantire un uptime del 99%.

I concetti chiave di Api.ai sono: agenti, entità, intenti, azioni e contesti. Gli agenti corrispondono a una determinata applicazione sulla quale si vuole costruire un'esperienza conversazionale, il cui scopo generale è quello di trasformare linguaggio naturale in dati strutturati. La strutturazione di questi dati avviene tramite l'identificazione di diverse parti, fra cui le entità, che rappresentano i concetti specifici di un dominio che vengono associati alle frasi del linguaggio spontaneo. Ad esempio in una richiesta di informazioni meteo, le entità estratte potrebbero riguardare “@luogo” e “@giorno” e dipendono dai valori richiesti all'agente per ogni specifico use case. Per discriminare in prima battuta la strada da seguire per soddisfare una determinata interazione, sono particolarmente importanti gli intenti, che associano ciò che l'utente comunica all'azione necessaria per adempiere alla richiesta. L'intento può essere espanso in:

- quello che l'utente dice;
- l'azione che dovrà seguire all'atto comunicativo;
- la risposta che l'utente riceverà;
- il contesto d'applicazione.

L'azione può inoltre avere parametri variabili per distinguere cosa debba essere estratto da quello che l'utente comunica. Infine, i contesti sono la chiave referenziale legata al contesto di enunciazione dell'utente, aspetto necessario per discriminare frasi dal significato ambiguo prive di contesto da quelle pronunciate all'interno di una determinata sequenza di frasi.

Se l'utilizzo più comune del servizio di Api.ai è quello di interpretazione diretta dell'input dell'utente, nulla vieta applicazioni più creative, lasciando gli sviluppatori liberi di sfruttare la potenza interpretativa per costruire non solo interfacce conversazionali classiche, ma anche semplificare dei flussi utente sulla

base delle informazioni di contesto.

Nel corso dell'ultimo keynote pubblico¹⁵, Google ha annunciato la possibilità di integrare i servizi offerti da Api.ai all'interno del suo assistente virtuale. Questo significherà poter pubblicare bot intelligenti direttamente fruibili tramite punti d'accesso esistenti (Android, Google Now, Google Home, ecc.).

LUIS

LUIS è un acronimo di Language Understanding Intelligent Service. È uno dei tanti servizi cognitivi forniti nel pacchetto Azure Cognitive Services, ed è la risposta di Microsoft a soluzioni come Wit.ai e Api.ai. In maniera simile ad Api.ai, utilizzando questo servizio è possibile utilizzare delle API per ottenere intenti ed entità a partire da una frase in linguaggio naturale data come input. È possibile utilizzarlo per costruire applicazioni intelligenti che possano conversare e capire che cosa un utente stia cercando di trasmettere tramite una frase o una sequenza di frasi. Per LUIS ogni frase è un'espressione all'interno di cui si cela un determinato intento su cosa l'oratore intenda fare. Esattamente come Wit.ai, integra dei sistemi di apprendimento automatico per migliorare la propria capacità di riconoscere gli intenti con il tempo.

Nonostante sia da poco uscito dalla fase di beta, il sistema risulta essere il più performante di quelli analizzati¹⁶, anche se a differenza di Wit.ai e Api.ai, prevede un utilizzo gratuito limitato, con diversi piani a pagamento a seconda dei volumi di utilizzo.

¹⁵ <https://events.google.com/io/recap>

¹⁶ <https://stanfy.com/blog/advanced-natural-language-processing-tools-for-bot-makers>

Conclusioni

La conoscenza è da sempre considerata una risorsa importante. In particolare modo oggi lo è per le aziende che si trovano ad affrontare un ambiente competitivo sempre più dinamico in cui è necessario risolvere problemi in modo rapido per non perdere la fiducia dei clienti. Gli utenti finali utilizzano servizi e tecnologie in modo trasparente, ponendosi domande solo quando il loro flusso di interazione viene interrotto da fattori non previsti nella progettazione dell'interfaccia che stanno utilizzando. Anche se lontani dal paradigma di conversazione, i bot sono presenti da molti anni nei flussi di interazione di molti software, silenziosamente integrati per semplificare i processi e le decisioni degli utenti. Le interfacce conversazionali costituiscono una risorsa preziosa per migliorare l'usabilità e l'accesso a servizi e informazioni. Tuttavia, progettare bot in grado di garantire un'esperienza ottimale non è soltanto un'attività di tipo informatico, bensì l'incontro di diversi settori disciplinari al quale servirà ancora del tempo per mimetizzarsi nel quotidiano in modo davvero naturale.

Bibliografia

F. Varanini, *Macchine per pensare, L'informatica come prosecuzione della filosofia con altri mezzi*, Guerini e associati, Milano, 2015

M. McTear, Z. Callejas, D. Griol, *The Conversational Interface: Talking to Smart Devices*, Springer , 2016

C. Penco, *Introduzione alla filosofia del linguaggio*, Laterza, 2008

G. Graffi, S. Scalise, *Le lingue e il linguaggio*, Il Mulino, 2002

C. Pearl, *Designing Voice User Interfaces: Principles of Conversational Experiences*, O'Reilly Media, 2016

B.R. Marshall, *Le parole sono finestre (oppure muri). Introduzione alla comunicazione nonviolenta*, Esserci, Reggio Emilia, 2015

J. Pavlus, *The next phase of UX: designing chatbot personalities*, Fastcodesign, 2016

R. D. Lankes, J. Silverstein, S. Nicholson, *Le reti partecipative, la biblioteca come conversazione*, Information Institute of Syracuse University's School of Information Studies

R. Carli, R.M. Paniccchia, *L'analisi emozionale del testo. Uno strumento psicologico per leggere testi e discorsi*, FrancoAngeli, 2004

T. Vamos, *Cooperative Systems - An Evolution Perspective*, IIASA - International Institute for Applied Systems Analysis, 1983

T. Vamos, *Computer Epistemology, A Treatise on the Feasibility of the Unfeasible or Old Ideas Brewed New*, Hungarian Academy of Sciences, 1991

Sitografia

Cos'è il modello SaaS? Software come un servizio,
<https://azure.microsoft.com/it-it/overview/what-is-saas> (consultato il 05/08/2017)

Api.ai Documentation, <https://api.ai/docs/getting-started/basics>
(consultato il 05/08/2017)

Making Conversational Interfaces Easier to Build,
<https://developers.googleblog.com/2016/09/making-conversational-interfaces-easier-to-build.html> (consultato il 12/08/2017)

Api.ai vs Wit.ai (or is it Google vs Facebook?),
<https://www.themarketingtechnologist.co/api-ai-vs-wit-ai> (consultato il 13/08/2017)

An introduction to conversational commerce and bots,
<https://martechtoday.com/introduction-conversational-commerce-bots-199920> (consultato il 13/08/2017)

2016 will be the year of conversational commerce,
<https://medium.com/chris-messina/2016-will-be-the-year-of-conversational-commerce-1586e85e3991> (consultato il 14/08/2017)

Advanced Natural Language Processing Tools for Bot Makers – LUIS, Wit.ai, Api.ai and others,
<https://stanfy.com/blog/advanced-natural-language-processing-tools-for-bot-makers/> (consultato il 16/08/2017)

Launching Built-in NLP for Messenger and Sunsetting Bot Engine (beta), <https://wit.ai/blog/2017/07/27/sunsetting-stories> (consultato il 30/08/2017)

Built-in NLP,
<https://developers.facebook.com/docs/messenger-platform/built-in-nlp>
(consultato il 30/08/2017)