UNIVERSITA' DEGLI STUDI DI ROMA

TOR VERGATA

FACOLTA' DI LETTERE E FILOSOFIA

CORSO DI LAUREA IN FILOSOFIA

TESI IN INFORMATICA E RAPPRESENTAZIONE DELLA CONOSCENZA

APPRENDIMENTO AUTOMATICO DI MODELLI PER LA SEMANTICA DISTRIBUZIONALE COMPOSIZIONALE: UNO STUDIO ESPLORATIVO

Relatore:
Chiar.mo Prof.
FABIO MASSIMO ZANZOTTO

Laureanda ANNALISA COLUZZI
Matr. 0109549

Anno Accademico 2009-2010

Cosa c'è in un nome?

" Che può mai significar la parola "Montecchi"? Non è una mano ,non un piede,non un braccio ,né un volto né alcuna altra parte che s'appartenga a un uomo. Oh sii qualche altro nome! Che cosa c'è in un nome? Quel che noi chiamiamo col nome di rosa , anche se lo chiamassimo d'un altro nome serberebbe pur sempre lo stesso dolce profumo. E così Romeo , pur se non fosse chiamato più Romeo, serberebbe pur sempre quella cara perfezione ch'egli possiede tuttavia senza quel nome..." (Shakespeare)

Introduzione

Per la nostra "amante avventurata nata sotto maligna stella" dal destino d'amore segnato è indescrivibilmente semplice riconoscere il suo amato anche senza l'ausilio del nome e lo stesso vale per tutti noi. Quando noi impariamo a riconoscere e identificare un oggetto, animato o inanimato che sia, quello che ci rende possibile il discernimento di esso dal resto del reale non è solo ed esclusivamente il "nome" (segno convenzionalmente affibbiatogli) bensì l'insieme delle peculiarità fisiche e non che costituiscono l'essenza e l'apparenza di esso. In fondo Giulietta chiede alla sua dolce metà di cambiare nome perché solo quest'ultimo gli è nemico. Se anche noi, per un'assurda ipotesi cominciassimo a cambiare i nomi agli oggetti o alle persone, non credo modificheremmo anche la loro conoscenza. Sicuramente avremmo molta difficoltà a comunicare tra di noi, infatti, l'assegnazione convenzionalmente riconosciuta di appellativi permette la comprensione comune e immediata all'interno di un determinato gruppo in cui gli individui condividono provenienza geografica, lingua e competenze, ma a livello di conoscenza personale non provocherebbe una rivoluzione destabilizzante. Per esempio se un giorno un ipotetico Sig. Michele Rambaldi dovesse decidere di indicare quello che noi comunemente chiamiamo tavolo con la parola "bicchiere", anche se alle nostre orecchie e al nostro intelletto apparirebbe bizzarro, il significato di quell'oggetto rimarrebbe comunque lo stesso (ovvero piano di appoggio di materiale forma e dimensione variabile con tre quattro o più gambe, atto a svariati utilizzi). Certo è che se in seguito a questa sua decisione chiedesse ad Anna Rossi di indicargli un bicchiere, sicuramente non potrebbe aspettarsi di vedere un ripiano, ma se rendesse partecipe la sua amica del cambiamento, non avrebbe problemi. Se invece Michele indicasse un tavolo e lo chiamasse bicchiere, sicuramente Anna sarebbe comunque in grado di capire che si tratta di un semplice cambiamento e riconoscerebbe il famigliare ripiano. Reindirizzando la nostra attenzione sull'oggetto vediamo che l'essenza del tavolo rimarrebbe invariata. Noi abbiamo deciso che si chiami tavolo ma il suo nome non è una peculiarità intrinseca dell'oggetto. Questi processi mentali per il nostro intelletto sono abbastanza elementari e di semplice comprensione ma nel momento in cui parliamo di una macchina, possiamo affermare che sarebbe tutto ugualmente scontato? Sicuramente no. Come può una macchina essere in grado di comprendere? Come può scindere il nome

dall'oggetto giacché ogni dato che immagazzina è sempre legato indissolubilmente alla sua denominazione, previa modifica voluta da uno input esterno (l'uomo). Se a una macchina indichiamo (non ostensivamente) un tavolo e lo chiamiamo bicchiere molto probabilmente, in base alle sue conoscenze, segnalerebbe un errore. Quindi il significato e il significante di una parola per una macchina sono inscindibili ? E per noi? A quale dei due conferiamo maggiore importanza? Una macchina può intelligentemente scegliere i caratteri essenziali di un oggetto e capire che quelli concorrono a rendere mostro il significato di tal elemento? Può scegliere autonomamente i dati da immagazzinare per ampliare la sua conoscenza? Se sì come può comprendere quali delle miriadi di parole presenti nel nostro linguaggio naturale hanno un significato simile? Può avere una conoscenza di base tale da poterle permettere di capire che, per esempio nelle frasi " Alice è proprio una bella ragazza" e " Mia dolce fanciulla non andar via" le parole "ragazza" e " fanciulla" sono apparentemente diverse ma hanno lo stesso significato?Questa è solo una frazione miserrima dei quesiti che sorgono nel momento in cui si approccia a qualcosa di così vasto e sconosciuto come l'Intelligenza Artificiale. Ovviamente non ho la presunzione di riuscire a dire qualcosa di nuovo o risolutivo con la mia tesi ma la voglia di studiare questi argomenti è tanta e mi pongo come unico obiettivo per questo mio lavoro introduttivo di riuscire a penetrare e sviscerare un po' più a fondo alcune di quelle domande cui forse non è ancora tempo di avere una risposta. In fondo senza curiosità il mondo si fermerebbe.

Capitolo 1

Chatbot : le macchine che parlano

Il sogno di realizzare una macchina, capace di intrattenere una conversazione tramite il Linguaggio Naturale è da sempre una delle spinte più potenti nella ricerca nel campo dell'Intelligenza Artificiale.

Come creare una macchina che sappia interloquire con un essere umano rasentando la perfezione discorsiva?

Eliza

Una possibile risposta è stata data nella seconda metà del ventesimo secolo (1966) da un informatico statunitense di nome Joseph Weizenbaum, nato a Berlino nel 1923 e cresciuto intellettualmente negli Stati Uniti d'America dal 1935, che riuscì, grazie al suo favoloso ingegno, a creare un programma per computer in grado di sostenere una parodia di una Seduta di psicoterapia di scuola Rogersiana. Questo programma prese il nome di ELIZA.

Weizenbaum scelse la psicoterapia "per evitare di dare al programma una vera

Origine del nome

Eliza Doolittle, la fioraia protagonista della commedia Pigmalione di George Bernard Shaw, che ripetendo le forme corrette di pronuncia riuscì a imparare il raffinato modo di esprimersi delle classi più colte trasformando il suo lessico incolto, ispirò l'appellativo da conferire a tale rivoluzionario programma.

conoscenza" poiché la seduta procede grazie a una concatenazione inesauribile tra domanda e risposta senza il bisogno di una conoscenza approfondita del paziente.

Funzionamento di ELIZA

ELIZA analizza e sostituisce semplici parole chiave in frasi preconfezionate. Non fa altro che estrapolare delle porzioni di frasi dalle risposte del paziente e le riformula sottoforma di domanda stimolando così la conversazione affinché vada avanti.

Negli anni successivi all'ideazione di questa prima chatbot sono state create numerosissime varianti, anche se ELIZA rimane comunque una pietra miliare per il semplice e allo stesso tempo sbalorditivo fatto, che per la prima volta un programmatore riuscì a rendere l'illusione di una conversazione tra uomo-uomo tramite un'interazione uomo –macchina.

A typical conversation is the following:

Men are all alike.
IN WHAT WAY?
They're always bugging us about something or other.
CAN YOU THINK OF A SPECIFIC EXAMPLE?
Well, my boyfriend made me come here.
YOUR BOYFRIEND MADE YOU COME HERE
He says I'm depressed much of the time.
I AM SORRY TO HEAR YOU ARE DEPRESSED
It's true. I am unhappy
DO YOU THINK COMING HERE WILL HELP YOU NOT TO BE UNHAPPY
I need some help, that much seems certain.
WHAT WOULD IT MEAN TO YOU IF YOU GOT SOME HELP
Perhaps I could learn to get along with my mother.
TELL ME MORE ABOUT YOUR FAMILY
My mother takes care of me.
WHO ELSE IN YOUR FAMILY TAKES CARE OF YOU
My father.
YOUR FATHER
You are like my father in some ways.
WHAT RESEMBLANCE DO YOU SEE
You are not very aggressive but I think you don't want me to notice that.
WHAT MAKES YOU THINK I AM NOT VERY AGGRESSIVE
You don't argue with me.
WHY DO YOU THINK I DON'T ARGUE WITH YOU
You are afraid of me.
DOES IT PLEASE YOU TO BELIEVE I AM AFRAID OF YOU
My father is afraid of everybody.
WHAT ELSE COMES TO YOUR MIND WHEN YOU THINK OF YOUR FATHER
Bullies.
DOES THAT HAVE ANYTHING TO DO WITH THE FACT THAT YOUR BOYFRIEND MADE YOU COME HERE

(Weizenbaum, 1966)

Come funziona una chatbot?

Il funzionamento di una chatbot si basa su una semplice funzione "MC" (Zanzotto, Macchine : Conoscenza e ragionamento, 2008) di STIMOLO- RISPOSTA.

MC:S→ R

S è l'insieme dei possibili stimoli e R delle possibili risposte e quindi a ogni possibile stimolo dovrebbe essere legata una possibile risposta. Si vengono a creare così delle coppie stimolo-risposta.

L'algoritmo di base ha questo tipo di funzionamento: con una lista ordinata di coppie stimolo-risposta la macchina sceglie di emettere la risposta ρ1 associata allo stimolo σ1 che è uguale e "unifica" con lo stimolo s che la macchina riceve dall'interlocutore. Una volta emessa la risposta torna ad attendere il prossimo stimolo

"Uguale" vuol dire che la sequenza di caratteri dello stimolo effettivo (quello ricevuto dalla macchina) deve essere necessariamente combaciante quella dello stimolo potenziale (quello che fa parte della coppia stimolo- risposta inserita dal programmatore).

Il problema è proprio questo!

Cosa succede se la sequenza di caratteri dello stimolo effettivo non ricalca alla perfezione quella che la macchina è in grado di riconoscere?

Linguaggio AIML

Il linguaggio usato per definire le coppie stimolo risposta è un'estensione del linguaggio di mark-up XML definito con l'acronimo AIML(Artificial Intelligence Markup Language). Questo tipo di linguaggio prevede che la sequenza di coppie stimolo- risposta sia contenuta all'interno di una zona di file di testo delimitata dai tag . Un esempio di coppia stimolo risposta è rappresentato nel modo seguente:

<AIML>

<CATEGORY>

<PATTERN>σ </PATTERN>

<TEMPLATE>ρ </TEMPLATE>

</CATEGORY>

</AIML>

Facciamo salutare la nostra chatbot

Nel momento in cui ci troviamo di fronte per esempio alla programmazione di un semplice saluto dobbiamo riuscire a coprire con le coppie ST l'insieme di tutte le possibili varianti; quindi dovremo cominciare a pensare e scrivere le CATEGORY di tutti gli eventuali saluti che la macchina potrebbe ricevere come stimolo.

- Buongiorno
- Salve
- Buonasera
- Ciao
- Hei
- Buona sera

…

Dopo di ché dovremo fare in modo che a ognuno di questi salut,i la macchina riconoscendoli, sappia rispondere. Di conseguenza dovremo scrivere le relative CATEGORY:

```
<CATEGORY>
        <PATTERN>Buongiorno</PATTERN>
        <TEMPLATE>Buongiorno</TEMPLATE>
</CATEGORY>
<CATEGORY>
        <PATTERN>Salve</PATTERN>
        <TEMPLATE>Buongiorno</TEMPLATE>
</CATEGORY>
<CATEGORY>
        <PATTERN>Buonasera</PATTERN>
        <TEMPLATE>Buongiorno</TEMPLATE>
</CATEGORY>
<CATEGORY>
        <PATTERN>Ciao</PATTERN>
        <TEMPLATE>Buongiorno</TEMPLATE>
</CATEGORY>
<CATEGORY>
        <PATTERN>Hei</PATTERN>
        <TEMPLATE>Buongiorno</TEMPLATE>
</CATEGORY>
```

…

Come si può notare questo metodo richiede moltissimo tempo e non garantisce la possibilità di coprire l'inesauribile ricchezza del linguaggio natural,e proprio perché lo stesso procedimento deve essere attuato con ogni singola parola che la macchina è chiamata a riconoscere.

Ovviamente si possono utilizzare delle variabili, come nel caso dei nomi propri (Buongiorno sono Annalisa – Ciao Annalisa cosa posso fare per te? / Buongiorno sono Fabio- Ciao Fabio cosa posso fare per te?)ma il processo rimane comunque infinito.

Quando una chatbot riceve uno stimolo effettivo che non sia l'equivalente grafico di quello potenziale, le regole di stimolo-risposta non si attivano e la macchina non è in grado emettere una risposta adeguata

Per esempio:

Se lo stimolo effettivo è "Salve sono Annalisa" e lo stimolo potenziale non prevede il "Salve" ma solamente" Buongiorno" ,"Ciao" o altro, la nostra macchina riconoscendo il saluto come diverso da quelli che essa ha in programma non attiva le regole adeguate e non emetterà risposta.

- Come è possibile inserire un numero infinito di parole con i loro innumerevoli sinonimi?

- Possiamo immettere tutte le varianti di struttura di ogni frase di probabile elaborazione?

Obiettivo

L'obiettivo di questo lavoro è trovare un modo che permetta alla macchina di trattare frasi con una composizione grafica diversa ma con un egual significato nello stesso modo per far sì che le regole di stimolo risposta siano comunque attivate.

Per proporre una possibile soluzione a questi problemi è necessario procedere per tappe tenendo presente che, il tema di cui ci stiamo occupando rientra in una branca di studio in continua evoluzione in cui ogni giorno personaggi di ben altro calibro apportano innovazioni.

Il mio schema di indagine consta di cinque punti:

1. Trattamento del Linguaggio Naturale come se fosse Formale
2. Semantica Distribuzionale rappresentante il significato delle parole
3. Semantica Distribuzionale Composizionale
4. Analisi e comprensione del Lessico Generativo
5. Induzione di modelli di Semantica Distribuzionale Composizionale da esempi

Capitolo2

Il linguaggio naturale

Alla base del comportamento intelligente dell'essere umano c'è certamente la capacità di elaborare simboli caricando gli oggetti di significati non del tutto riconoscibili nella loro natura intrinseca.

Il linguaggio verbale, che è un sistema di comunicazione tra individui ,è certamente l'esempio di elaborazione simbolica che maggiormente caratterizza l'essere umano; grazie ad esso si trasmettono informazioni veicolate da regole di grammatica.

Come si è sviluppato il Linguaggio Naturale?

La comparsa del Linguaggio Naturale avvenuta tra i cento e i duecentomila anni fa ha determinato quel salto evolutivo che ha differenziato l'homo sapiens dal resto degli ominidi .

A seguito di mutamenti strutturali della cavità orale, l'essere umano è stato in grado di sviluppare una capacità linguistica variegata e ricca.

Ci sono due scuole di pensiero che riguardano le origini : una che decreta la nascita di un unico linguaggio da cui si sono poi diramate le differenti lingue moderne, l'altra che prevede dei "ceppi primordiali" (Wikipedia L'encilopedia libera, 2011) già differenziati alle origini.

Sicuramente il Linguaggio naturale contemporaneo è il frutto di mutamenti, sviluppi ed elaborazioni avvenute nel corso dei millenni anche se non è ben chiaro se la nostra capacità linguistica si stata appresa man mano o se invece è qualcosa di originariamente innato nell'essere umano.

Come supporto teorico all'idea che il linguaggio non è altro che un condizionamento sociale, vi fu la scoperta del 1828 in Francia del "ragazzo selvaggio" ,che essendo vissuto fino all'età di dodici anni esclusivamente a contatto con gli animali non , anche a seguito di prove di insegnamento sostenute da esperti, a malapena ad articolare poche parole.

Da ciò fu dedotto che l'ambiente con cui si interagisce e la socializzazione dell'individuo sono fondamentali per lo sviluppo dell'intelligenza e quindi del linguaggio.

Linguaggio e intelligenza

Il linguaggio e l'intelligenza sono indissolubilmente legati tra loro giacché la capacità linguistica è allo stesso tempo spinta motrice per l'evoluzione, rafforzamento dell'intelletto e frutto del miglioramento dell'abilità celebrale avvenuto nel corso dei secoli.

Test di Turing

Per determinare se una macchina può essere definita intelligente gli studiosi ricorrono al cosiddetto test di Turing . Il suo ideatore Alan Turing (un grande matematico del '900 considerato uno dei padri dell'informatica) precisò la sua teoria nell'articolo *Computing machinery and intelligence*, apparso nel 1950 sulla rivista Mind.

Il test prende ispirazione dal "gioco dell'imitazione" in cui tre individui, un uomo A una donna B e una terza persona C sono tenuti separati; l'ultimo deve riuscire a capire con l'ausilio di domande, qual è l'uomo e quale la donna. Le informazioni fornite con le risposte devono essere dattilografate o trasmesse in un modo equivalente.

Nel Test di Turing l'unica differenza dal suddetto gioco è che invece di un uomo A vi è una macchina. Se l'individuo C non riesce a distinguere quale sia la macchina il test è superato con successo.

Per Turing quindi la capacità intellettiva si limita al solo saper elaborare espressioni cariche di significato. Nell'articolo è espressamente detto che in base alla complessità del software "emergeranno le funzioni intellettuali".

Tutto ciò sta alla base di quella che è definita come Intelligenza Artificiale il cui scopo è appunto, come abbiamo già detto, la costruzione di una macchina in grado riprodurre i processi cognitivi dell'essere umano.

Nonostante le aspettative del nostro matematico datassero nel 2000 il successo in questo campo, ancora nessuna macchina è stata in grado di superare il test.

Teoria linguistica di Noam Chomsky

Noam Chomsky fu linguista, filosofo statunitense nato il 7 dicembre 1928 a Filadelfia.
La sua teoria ammette che la capacità di ogni essere umano di produrre e capire delle frasi dipende da conoscenze preesistenti nella sua mente (frutto di esperienza continuamente arricchita dal mondo esterno) e dalle competenze linguistiche che sono:

- Inconsapevoli, perché la maggior parte degli esseri umani non sa quale sia l'origine di questa sua capacità né il modo in cui formuli e comprenda una frase
- Innate, in quanto senza questo carattere di preesistenza del linguaggio non si spiegherebbe come mai un bambino riesca, in brevissimo tempo e con un'elevata velocità, a parlare una lingua in modo corretto grazie ad un puro spirito imitativo.

Le competenze linguistiche sono di tre tipi:

- Fonologiche (capacità di produrre e capire i suoni della lingua parlata)
- Sintattiche (produrre e riconoscere frasi grammaticalmente corrette)
- Semantiche (capacità di assegnare significato a una frase)

I linguaggi artificiali

Ciò che principalmente differenzia un linguaggio naturale da uno artificiale è che il primo ha avuto uno sviluppo spontaneo nelle culture umane mentre il secondo è il frutto di una decisione consapevole e razionale di un linguista o di un team di studiosi che decidono la grammatica e il vocabolario di tale linguaggio.

Il fattore di pianificazione conferisce anche al linguaggio naturale una percentuale di "artificialità" proprio per il semplice fatto che ,nel momento stesso in cui si inizia a istituire una grammatica collettivamente riconosciuta e arbitrariamente codificata il linguaggio viene veicolato e disciplinato dalla razionalità umana.

La glossopoiesi è l'atto di creazione di una lingua artificiale e gli autori si possono definire glottoteti o glossopoeti.

L'Esperato (lingua nata più di cento anni con l'intento di essere un idioma internazionale grazie alla sua grammatica semplice e di facile apprendimento) e il linguaggio di

programmazione, sono finalizzati allo scambio di informazioni senza il rischio di incomprensioni e le loro regole ferree ne permettono l'uso universale ed eterno.

Confronto tra Linguaggio Naturale e Linguaggio artificiale

Aspetti positivi del linguaggio naturale:

- Adattabilità all'evoluzione
- Flessibilità
- Dinamicità
- Poesia e umorismo

Aspetti negativi del linguaggio naturale:

- Ambiguità
- Difficoltà normativa
- Localizzato non universale

Aspetti positivi del linguaggio artificiale

- Efficacia
- On the spot
- Disambiguità

Aspetti negativi del linguaggio artificiale.

- Staticità
- Inadatto a rappresentare il mondo del reale
- Settorialità

Esempi

LN

- Una frase anche se sintatticamente scorretta può essere in qualche modo comprensibile
- Una frase sintatticamente corretta può non avere alcun significato
- Uno stesso termine può avere significato diverso all'interno della stessa frase
- Una frase può essere ambigua se estrapolata dal suo contesto

LA

- Una comunicazione può fallire a causa di un minimo errore sintattico
- Un'istruzione non è necessariamente eseguibile solo per la correttezza sintattica
- Ogni simbolo ha un determinato significato specifico
- Le istruzioni non sono dipendenti dal contesto

Significato e significante

A questo punto abbiamo compreso che, il problema fondamentale è che la nostra macchina che parla non ha la capacità cognitiva necessaria per collegare due parole semanticamente equivalenti che però in apparenza a causa della loro raffigurazione grafica differente non sono rapportabili.

Come poter sciogliere quel legame apparentemente insolubile tra il significato di una parola e la parola stessa, per superare la difficoltà evidente dettata dall'infinita ricchezza del linguaggio naturale?

È necessario a questo stadio del discorso comprendere approfonditamente cosa si intende nel momento in cui definiamo il "significato " di una parola e di conseguenza di una frase.

Semiotica

La disciplina che tradizionalmente si occupa della significazione di un elemento è sicuramente la semiotica.

La definizione tradizionale circoscrive il suo campo di studio ai segni e il segno è tuttora l'oggetto delle sue ricerche.

Naturalmente le riflessioni riguardanti questo argomento non hanno avuto inizio con la nascita di questa nuova branca di studio bensì già ai temi di Platone se ne scorgeva l'importanza.

Nonostante l'interesse riscosso ai tempi dei padri della filosofia, la sua fondazione ufficiale è datata fine Ottocento inizi Novecento.

Oggetto di indagine

"Scopo di questo libro è esplorare le possibilità teoriche e le funzioni sociali di uno studio unificato di un fenomeno di significazione e/o comunicazione" (Eco, 1975)

All'inizio del suo *Trattato di semiotica generale*, Umberto Eco da questa definizione di qual è il punto focale della ricerca semiotica.

Sono due gli obiettivi di studio:

- Processi di comunicazione (Eco sostiene si debba studiare i processi culturali in quanto fondati su processi di comunicazione)

- Sistemi di significazione (quello che a noi più interessa)

Sistemi di significazione

Un sistema di significazione non è altro che un "dispositivo che collega entità presenti a entità assenti" (Traini) .

Ciò significa che viene a istaurarsi una relazione indissolubile tra un oggetto realmente materialmente presente e qualcosa che non lo è (un semaforo rosso può stare per un pericolo, un faro per un punto di riferimento ecc.)

La significazione è quel legame che nasce tra elementi percepiti con i sensi e elementi creati con l'intelletto.

I segni, data la loro intrinseca capacità di unire qualcosa che è visibile con qualcosa che non lo è, sono senza ombra di dubbio sistemi di significazione.

"Un sistema di significazione è un costrutto semitico autonomo che possiede modalità d'esistenza del tutto astratte, indipendenti da ogni possibile atto di comunicazione che le attualizzi. Al contrario […] ogni processo di comunicazione tra esseri umani -o tra ogni tipo di apparato intelligente sia meccanico che biologico- presuppone un sistema di significazione come propria condizione necessaria." (Eco, 1975)

La semiotica, considerando Ferdinand de Saussure e Charles Sanders Pierce come padri fondatori, ha oramai più di un secolo di vita e ha sempre visto contrapporsi al suo interno due anime guida :quella strutturalista e quella interpretativa.

Analizziamo un po' meglio le teorie riguardanti il significato elaborate dai capostipiti di questa disciplina.

Il “significato” da Saussure a Eco

Ferdinand de Saussure

Ferdinand de Saussure nato a Ginevra nel 1857 intraprende i suoi studi letterali all’età di diciannove anni e inizia la sua carriera scrittoria ad appena ventuno. Il suo testo più importante *Corso di linguistica generale* viene pubblicato postumo da due suoi allievi, creato dalla raccolta di appunti presi a lezione.

Tralasciando parte dei contributi di questo padre della semiotica analizziamo come definisce il significato.

Saussure fa una netta distinzione tra significato e significante.

La sua ipotesi è la seguente: nel momento in cui un parlante produce una fonia, compie quest’atto seguendo la guida psichica dettatagli da una sorta di schema mentale chiamata immagine acustica.

Quest’immagine acustica corrisponde al significante che è un modello, un’entità astratta costruita dal singolo individuo tramite l’educazione e la socializzazione avvenuta nella comunità d’origine.

Il significante è quindi un “modello superindividuale […] chi parla ha imparato, modificato, aggiustato questo modello nel tempo; si è esercitato a riprodurlo con la voce e a riconoscerlo con l’udito”. (Traini)

Per quanto riguarda il significato, il processo è lo stesso: da uno schema mentale astratto derivante da un collettivo si ottiene un senso comunicabile.

Il segno linguistico di Saussure ha due facce:

- concetto o significato
- immagine acustica o significante

Il legame che unisce questi due aspetti è:

- Arbitrario, poiché l’idea di fiore non è legata in alcun modo alla sequenza di suoni f-i-o-r-e quindi nella realtà, non vi sono dei collegamenti naturali.

- Lineare giacché "il significante essendo di natura auditiva si svolge soltanto nel tempo e ha i caratteri che trae dal tempo".

Louis Hjelmslev

Louis Hjelmslev nato a Copenaghen nel1899 all'età di diciotto anni si iscrive all'università della stessa città, dopo aver approfondito studi di linguistica nel '37 inizia a insegnare nella città natale.

Il suo testo principale è *I fondamenti della teoria del linguaggio* pubblicato nel 1943.

Secondo Hjelmslev i linguaggi sono strutturati in piani e per la precisione due:

- il piano del contenuto
- il piano dell'espressione

che corrispondono a quelli che Saussure definiva significato e significante.

Per quanto riguarda il primo possiamo pensare a una materia che come massa amorfa di suoni viene articolata dalla forma linguistica in segmentazioni che permettono il delineare delle sostanze.

Lo stesso vale per il piano del contenuto: una massa amorfa viene articolata dalla forma del contenuto secondo schemi lessicali specifici.

I due piani sono connessi tra di loro in modi differenti a seconda del tipo di linguaggio: se si tratta di un linguaggio monoplanare come per esempio l'algebra a ogni elemento del piano dell'espressione, ne corrisponde uno del contenuto quindi c'è un rapporto uno-a – uno.

Se invece si tratta di un linguaggio biplanare com'è il linguaggio naturale, allora non si può parlare di una corrispondenza biunivoca di elementi dei due piani.

Quindi non è necessariamente detto che un significato possa essere espresso da un solo significante anzi, grazie alla ricchezza lessicale abbiamo molto spesso la situazione opposta, che più significanti possono rappresentare lo stesso significato.

Charles Sanders Pierce

Nato a Cambridge nel 1839 indirizza la sua formazione in ambito scientifico e a soli sedici anni entra all'università di Harvard dove il padre insegna matematica.

Dopo essersi laureato a pieni voti, inizia una carriera universitaria con non poche difficoltà che lo ridurranno in miseria.

La semiotica di Pierce è definita interpretativa e alla base delle sue teorie pone il concetto di semiosi.

"Per semiosi intendo un'azione, un'influenza che sia, o coinvolga, una cooperazione di tre oggetti, come per esempio un segno, il suo oggetto e il suo interpretante."

I tre termini devono o essere necessariamente presenti e non si può avere una semiosi a due termini.

Il punto di partenza per Pierce è l'Oggetto, inteso come realtà esterna, chiamato anche oggetto dinamico. Per rendere conto dell'oggetto come realtà in sé noi abbiamo bisogno dei segni che mediano tra l'oggetto e l'interpretante in quanto è determinato dal primo e genera il secondo.

Pierce usa il termine representamen per indicare il significante e oggetto immediato per il contenuto. Quindi l'oggetto dinamico è l'oggetto in sé mentre quello immediato è il significato rappresentato dal segno. Pierce distingue i segni in icone simboli e indici.

Umberto Eco

Nato ad Alessandria nel 1932 si laurea ventidue anni dopo a Torino e nel '32 pubblica la sua prima opera *Opera aperta*.

Grandissimo linguista semiotico e filosofo del '900 continua la sua opera scrittoria tuttora.

Nel Trattato di semiotica generale Eco parla di due teorie fondamentali:

- Teoria dei codici
- Teoria della produzione segnica

Nella prima tutto ruota intorno alla funzione segnica, ovvero la relazione che intercorre tra espressione e contenuto.

Il concetto di codice stabilisce che le unità del sistema semantico e quella del sistema sintattico associate tra loro corrispondono a una data risposta.

“Quando un codice associa gli elementi di un sistema veicolante agli elementi del sistema veicolato, il primo diventa l’Espressione del secondo, il quale a sua volta diventa il Contenuto del primo. Si ha funzione segnica quando un’espressione è correlata a un contenuto, ed entrambi gli elementi correlati diventano funtivi della correlazione. “ (Eco, 1975)

Un segno per Eco è il luogo d’incontro di elementi indipendenti partecipi di una correlazione astratta.

In tutto ciò la decodifica da parte del destinatario può essere di tipo aberrante se sia lui che il mittente della comunicazione non condividono una conoscenza equivalente della struttura e del valore del codice.

Capitolo3

Semantica distribuzionale composizionale

Arrivati a questo punto abbiamo appurato che ,nel momento in cui dobbiamo programmare una macchina intelligente ci troviamo di fronte alla difficoltà di poter riconoscere il significato di una parola o di una frase espresso con diversi significanti. Tentando di rispondere ai suddetti quesiti proviamo ad analizzare una possibile soluzione rappresentata dalla semantica distribuzionale composizionale.

Semantica distribuzionale

La semantica distribuzionale si fonda sull'idea di poter trovare un sistema attendibile per determinare la similarità semantica tra le parole con un processo innovativo che prende in esame non solo la rappresentazione grafica del significato di un oggetto ma anche altro.

Uno dei suoi più importanti obiettivi è di comprendere che grado di sinonimia, semisinonimia e relazione intercorre tra due unità linguistiche.

Tutto questo per aggirare il limite oggettivo che la ricchezza, ricercatezza e cavillosità del linguaggio naturale pone nei nostri confronti nel momento in cui vogliamo far si che la macchina diventi intelligente.

La semantica distribuzionale attua una "rivoluzione copernicana" nel metodo di deduzione di significato delle parole.

Nei Modelli semantici distribuzionali (word space models) vi è "un accostamento analogico tra le proprietà del significato e le proprietà dello spazio" (Lenci, Spazi di parole: metafore e rappresentazioni semantiche)

Due principi costituiscono l'humus generativo della ricerca in semantica distribuzionale**:**

Firth's principle: "You shall know a word by the company it keeps"

Harris's Distributional Hypothesis: "Words that occur in the same contexts tend to have similar meanings"

John Rupert Firth

John R. Firth nato a Keighley nel 1890 fu un linguista inglese insegnò all'università di Punjab e in quella di Londra prima di recarsi alla Scuola di studi Oriantali e Africani dove divenne Professore di Linguistica generale.

La colonna portante della visione linguistica di Firth è l'idea che il significato di una parola sia strettamente dipendente dal contesto in cui essa si trova con la sua nozione di contesto di situazione. Il suo lavoro teorico fu talmente innovativo che diede vita all'aggettivo Firtiano.

Avendo insegnato per più di venti anni all'università londinese il suddetto professore riuscì a influenzare un'intera generazione di studiosi di linguistica riuscendo a indirizzare futuri personaggi di rilievo.

John Firth's principle

Il principio di John Firth, affermando che è possibile conoscere il significato di un termine dalle relazioni che intrattiene con gli altri, giustifica l'idea che il significato delle parole, o delle sequenze di esse può essere modellato usando le informazioni contestuali rappresentate in vettori spaziali.

Grazie a Firth vi è una valorizzazione del contorno di una parola e quindi il significato non è qualcosa di intrinsecamente presente nell'espressione grafica di un concetto bensì può essere dedotto dai legami che instaura con le altre espressioni linguistiche.

Harris's Distributional Hypothesis

Zellig Sabbettai Harris

Nato nel 1909 a Balta fu un grande linguista sintatista matematico e metodologista scientifico situandosi sulla scia della scuola bloomfieldiana, nel suo *Metodi in Linguistica Strutturale (1951)*formula i principi di analisi distribuzionale rifiutando l'utilizzo del concetto di senso e formulando l'ipotesi di una possibile somma degli ambienti delle unità linguistiche.

Dedusse che il problema dell'ambiguità sintattica (una frase due sensi) è causato dalla differente trasformazione che ha luogo dal nodo originario di costruzione.

L'Ipotesi distribuzionale di Harris afferma che le parole che occorrono nello stesso contesto tendono ad avere lo stesso significato e quindi giustifica il fatto che noi possiamo "comparare il significato delle parole mettendo in relazione la loro rappresentazione nello stesso spazio vettoriale rappresentante l'informazione contestuale" .(Zanzotto, 2010)

L'Ipotesi Distribuzionale è correlata alla "discovery procedures" ("metodi e operzioni impiegate nel campo della linguistica strutturale che cercano di rivelare per mezzo di segmentazioni e classificazioni le categorie fondamentali e la loro relazione di un determinato linguaggio sulla base di un numero finito di frasi" (BookRags))della tradizione strutturalista e il frutto ultimo di una "visione associazionista che prende come chiave

fondamentale di esplorazione paradigmatica del lessico la ricostruzione dei rapporti che intercorrono tra i suoi elementi nei contesti linguistici." (Lenci, Spazi di parole: metafore e rappresentazioni semantiche)

I significati delle parole quindi abbiamo visto che non prendono valore da una visone dizionariale bensì nascono dalla rappresentazione contestuale in cui sono immerse.

I più noti modelli di semantica distribuzionale

Vi sono vari modelli di Semantica distribuzionale i più noti sono:

- Latent Semantic Analysis :brevettato nel 1988 da Susan Dumais e Thomas Landauer E' una tecnica di elaborazione del linguaggio naturale con particolare attenzione volta alla semantica vettoriale per analizzare le relazioni tra un insieme di documenti e le condizioni contenute in essi. LSA può utilizzare una matrice termine-documento che descrive le occorrenze del termine nel documento;gli elementi della matrice sono proporzionali al numero delle volte che il termine appare nel documento. Dopo la costruzione della matrice di occorrenza del termine-documento LSA trova un'approssimazione per difetto dovuta a vari motivi. La polisemia(una parola più significati) rappresenta uno dei limiti di questo modello semantico
- Hyperspace Analogue to Language :ritiene che il contesto di una parola sia costituito solamente dalle parole che lo circondano immediatamente. In HAL viene calcolata una matrice dove il numero di parole del suo lessico è rappresentato da N usando un frame di 10 parole lette. La somiglianza semantica tra due parole è data dal coseno dell'angolo tra i vettori e quindi la correlazione semantica di due termini avviene nel momento in cui tendono a comparire con le stesse parole
- Random Indexing: alternativa alla LSA è un'operazione composta da due fasi:"la prima in cui ogni contesto parola o documento nei dati è assegnato a un'unica e casualmente generata rappresentazione chiamata indice vettoriale; successivamente i vettori contesto sono prodotti dalla scansione del testo e ogni volta che una parola occorre in un contesto , tale di-dimensonale indice vettoriale del contesto si aggiunge al vettore contesto per la parola in questione." (SAHLGREN) La rappresentazione delle parole è quindi data dai vettori contesto d-dimensionali che sono effettivamente la somma dei contesti delle parole. L'approccio di questo

modello avviene a ritroso rispetto a quello tradizionale che screpolava i vettori dal contesto .

Un ulteriore utilizzo dei modelli distribuzionali, oltre a quello principale di comprendere la similarità tra termini, è stato quello di applicarli per l'acquisizione lessicale da parte del bambino .

Lara e il suo il spazio semantico

Marco Baroni e Alessandro Lenci nel 2007 hanno presentato uno spazio semantico costruito a partire dagli input che Lara, una bambina di 2 -3 anni, riceveva dall'esterno. Si è notato che in tal spazio semantico i nomi di umani e i nomi di animali comparivano stranamente vicini, al punto di rischiare la sovrapposizione come se la categoria fosse una solamente e non due distinte.

Da questo si è evinto che il problema era sorto a causa del modo in cui i nomi venivano utilizzati dagli adulti che interagivano con Lara . Nel momento in cui la bambina riceveva l'input di un nome di animale, questo era sempre accostato o a un giocattolo o a un personaggio di una favola e quindi con caratteristiche umane.

Il carattere "figurato"del linguaggio infantile, si deduce, sia una conseguenza intrinseca della sensibilità del bambino al contesto delle rappresentazioni nello spazio semantico.

Quindi un uso simile di nomi nell'input determina una similarità dal punto di vista semantico nell'universo concettuale del bambino(Lenci)

Da tutto ciò ne emerge che il contenuto di un termine è estremamente radicato nel contesto,che ora acquista un ruolo costitutivo del significato, in cui è inserita

"La similarità distribuzionale diviene dunque la causa di correlazione semantiche che vengono poi a manifestarsi sottoforma di analogie."

Rappresentazione vettoriale

Vari modelli di Semantica distribuzionale sono stati elaborati ,ma il comune denominatore è l'idea che per determinare la similarità tra due parole sia necessario misurare il grado di sovrapposizione dei contesti linguistici in cui esse ricorrono.

Uno spazio semantico è strettamente legato a uno spazio geometrico e come in quest'ultimo ogni punto è definito da un vettore di n punti che rappresentano le coordinate, così, il significato di una parola è definito dalla posizione di questa in un sistema di coordinate determinato dai contesti linguisti in cui la parola può ricorrere"(Lenci)

Secondo Lapata uno spazio semantico di parole è definito da una quadrupla<T,B,M,S,>:

- T è l'insieme delle parole target che formano gli elementi dello spazio
- B è la base che definisce le dimensioni dello spazio e contiene i contesti linguistici
- M è una matrice di co- occorrenza che fornisce una rappresentazione vettoriale di ogni termine

Partendo dall'assunzione che, due parole che si pongono in relazione con elementi linguistici simili avranno una maggiore vicinanza, ogni parola verrà rappresentata da un vettore a n dimensioni ciascuna delle quali registra il numero delle volte che la parola occorre.

I modelli differiscono per la base che adottano ovvero che tipo di contesto scelgono.

	abbaia	coda	corre	domestico	zampe
cane	2	5	1	5	4
gatto	0	3	1	2	5
leone	0	3	3	0	2
luce	0	0	0	0	0
abbaia	0	1	1	1	1
macchina	0	1	4	0	0

Tabella 1 matrice di co- occorrenza tra parole

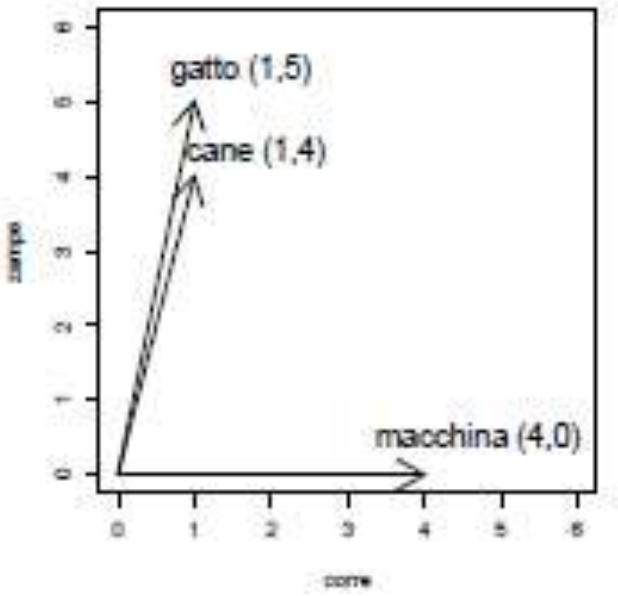
- S metrica è la misurazione della distanza tra i punti nello spazio. Se due parole hanno un numero elevato di corrispondenze di dimensioni con valori simili avranno una distanza minore e quindi per l'ipotesi distribuzionale una similarità semantica maggiore.

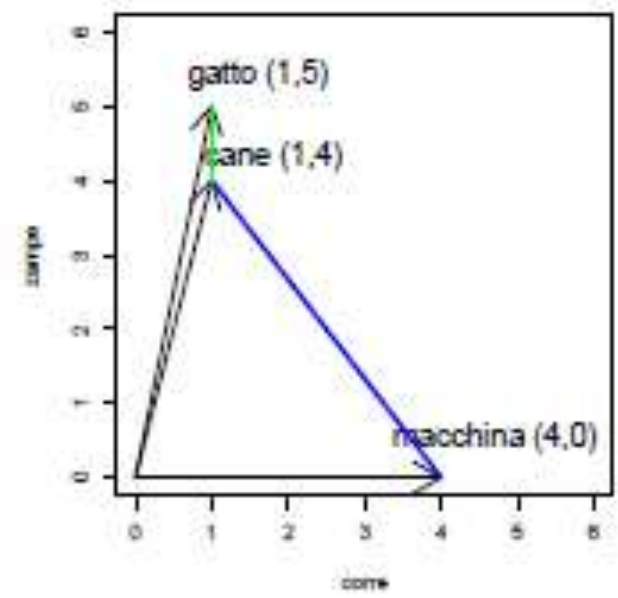
Se due vettori hanno un coseno pari a 1 e quindi un angolo pari a 0 le parole corrispondenti avranno una similarità massima; se invece il coseno è 0 e quindi l'angolo è di 90° allora la similarità sarà minima.

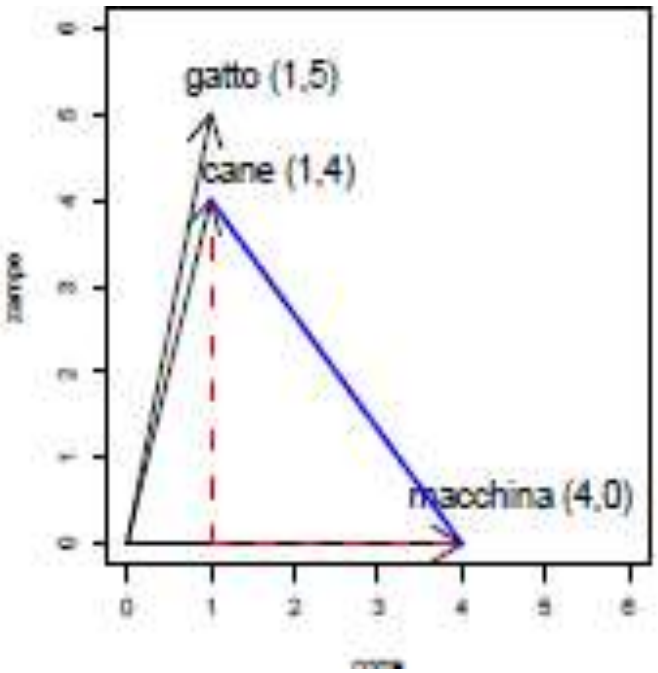
Il vettore assegnato a una parola non ha alcun significato intrinseco ma serve solamente a determinare la posizione nello spazio e il grado di similarità con le altre parole.

L'unica domanda degna di essere posta nel momento in cui ci troviamo di fronte alla deduzione di similarità semantica tramite la rappresentazione vettoriale è:

Quanto sono distanti due parole nello spazio vettoriale?






gatto (1,5)
cane (1,4)
macchina (4,0)
0
1
2
3
4
5
6

Principio di composizionalità

Tutto ciò che abbiamo detto fino ad ora sui modelli di semantica distribuzionale si riferisce esclusivamente alle parole.

Ricordiamo qual è il nostro obiettivo : rendere una macchina capace di riconoscere un significato, anche se la sua rappresentazione grafica non combacia perfettamente con quella che ha immagazzinato nel momento della programmazione.

Va da se che non spossiamo realizzare un modello adibito alla sola trattazione delle parole ma dobbiamo estenderlo a delle intere frasi.

Per far ciò dobbiamo ricorrere a quello che viene chiamato Principio di composizionalità.

Friedrich Ludwig Gottlob Frege

Nato a Wismar nel 1848 fu un grandissimo matematico logico e filosofo tedesco. Considerato dalla critica moderna come il più grande logico dopo Aristotele e padre del pensiero formale del ‘900.

Fu il primo fautore del logicismo, credendo di poter ridurre tutta l’aritmetica alla pura logica e tramite il suo progetto, avrebbe dimostrato che i giudizi dell’aritmetica essendo sintetici a priori si sarebbero potuti comprovare solo con l’ausilio della logica e delle regole razionali.

Tralasciando il suo immane lavoro, quello che ci interessa da vicino è solamente quello che lui stesso definisce principio di composizionalità.

Ludwig Wittgenstein riprende nel *Trattaus Logico-Philosophicus* del 1921, questo principio definendo le tavole di verità come una combinatoria di possibilità

Il principio di Frege appare per la prima volta in *Senso e significato* del 1892 e poi nei *Principi.*

"Il senso di ogni espressione linguistica è funzione del senso delle parti; in particolare il senso di un enunciato complesso è funzione del senso degli enunciati componenti."

Nel lungo dibattito sulla scelta di una maggiore adeguatezza di un approccio simbolico(che tratta le parole come se fossero unità elementari in analizzabili) o di un metodo empirico per l'elaborazione del Linguaggio naturale, si inserisce il perché di uno studio sui modelli di semantica distribuzionale COMPOSIZIONALE. Le semantiche distribuzionali si focalizzano sulla spiegazione del significato tramite i vettori ma non forniscono un modello per la composizione del significato.

Per gestire il significato di frasi o sentenze servirebbe comunque un approccio simbolico.

Alcuni modelli simbolici di semantica del linguaggio naturale propongono una via alternativa in cui gli elementi della frase non sono considerati delle unità assestanti e inanalizzabili.

Le operazioni di composizione usa le complesse rappresentazioni di parole e produce una rappresentazione omogenea per le sequenze di parole.

Il significato di queste sequenze è ottenuto dalla composizione del significato delle parole componenti.(input-output).

Uno dei modelli composizionali più rilevante è sicuramente il Lessico generativo di cui parleremo più aventi.

Abbiamo detto che la semantica distribuzionale estrapola il significato delle parole dallo studio e l'analisi del loro contesto.

Il "key point"(Zanzotto) è che ogni parola che appare significativamente in un corpus ha un significato distribuzionale.

La proprietà di una parola "target" è la co-orrenza che presenta in una determinata finestra di contesto di dialogo. Se le sue proprietà occorrono frequentemente allora probabilmente quelle saranno i suoi attributi salienti.

Anche se molto utilizzato questo modello presenta comunque dei limiti di ordine pratico.

Come nella contrapposizione tra l'approccio simbolico e quello distribuzionale nella semantica del linguaggio naturale, una significativa sfida è quella di ottenere modelli composizionali efficienti in grado di rappresentare strutture complesse utilizzando le rappresentazioni distribuite delle parti.

Questo auspicabile metodo deve avere due caratteristiche:

- immediata accessibilità alle informazioni immagazzinate nella rappresentazione distribuita
- rappresentazione analogica: gli oggetti simili dovrebbero avere rappresentazioni simili

Queste due proprietà differenziano la semantica distribuzionale dalla semantica distribuzionale composizionale.

Due fasi determinano la rappresentazione distribuzionale composizionale:

- composizione in cui vengono fusi i significati delle due parole
- estrazione in cui dalla mescolanza dei concetti se ne estrae un significato comune e differente.

La matrice di rappresentazione composizionale

Abbiamo detto che per le parole l'elemento grafico che viene utilizzato è il vettore .

Per quanto riguarda la semantica distribuzionale composizionale utilizzeremo le matrici.

Partendo da un modello generico addizionale che somma i due vettori x e y, rappresentanti rispettivamente la prima e la seconda parola e ne ottiene un terzo z,passiamo alla scrittura di una matrice singola:

$$\vec{z} = (A \quad B)\begin{pmatrix}\vec{x}\\ \vec{y}\end{pmatrix}$$

Riportando l'esempio del Professor Zanzotto:

$$\vec{contact} = (A \quad B)\begin{pmatrix}\vec{close}\\ \vec{interaction}\end{pmatrix}$$

La matrice precedente può essere riscritta in questo modo :

$$\begin{pmatrix}11\\0\\3\\0\\11\end{pmatrix} = (A_{5\times5} \quad B_{5\times5})\begin{pmatrix}27\\3\\2\\5\\24\\23\\0\\3\\8\\4\end{pmatrix}$$

Focalizzando sulla matrice (AB) si può trasporre così:

$$\begin{aligned}\vec{z}^T &= \left((A \quad B)\begin{pmatrix}\vec{x}\\ \vec{y}\end{pmatrix}\right)^T\\ &= (\vec{x}^T \quad \vec{y}^T)\begin{pmatrix}A^T\\ B^T\end{pmatrix}\end{aligned}$$

dove la prima matrice è conosciuta mentre la seconda viene stimata tramite algoritmi

Conseguentemente si ottiene:

$$\begin{pmatrix}\vec{z}_1^T\\ \vec{z}_2^T\\ \vdots\\ \vec{z}_n^T\end{pmatrix} = \begin{pmatrix}(\vec{x}_1^T \quad \vec{y}_1^T)\\ (\vec{x}_2^T \quad \vec{y}_2^T)\\ \vdots\\ (\vec{x}_n^T \quad \vec{y}_n^T)\end{pmatrix}\begin{pmatrix}A^T\\ B^T\end{pmatrix}$$

In cui dalle triple(z,x,y,)si ottengono i vettori che possono essere visti come due matrici di n righe.

L’equazione globale del sistema è :

$$Z = (X \quad Y)\begin{pmatrix} A^T \\ B^T \end{pmatrix}$$

Questo sistema di equazione rappresenta i vincoli che le matrici A e B devono soddisfar per essere un possibile modello di semantica distribuzionale composizionale.

A questo punto arriviamo proviamo ad affrontare il quarto punto del nostro schema di indagine.

Capitolo 4

Il Lessico Generativo

Il Lessico Generativo è un modello di Semantica Distribuzionale Composizionale.

Questo modello ci permette di trattare complessi fenomeni di linguaggio naturale e il suo "padre " è stato sicuramente James Pustejovsky, professore di scienza informatiche presso la Brandeis University.

Nel suo scritto *The Generative Lexicon* parte da due assunti fondamentali:

- senza un apprezzamento della sintattica di un linguaggio uno studio semantico è destinato a fallire

  Non c'è modo di separare il significato di una parola dal suo significante
- il significato delle parole dovrebbe riflettere le più profonde strutture di un sistema

Secondo P. la rappresentazione del contesto delle parole deve essere vista come lo sviluppo di fattori generativi che servono per coloro che utilizzano il linguaggio per creare e manipolare il contesto di vincoli, al fine di comunicare e comprendere le informazioni che vengono scambiate.

Ne segue che non esiste un' interpretazione universale ma dipende intrinsecamente dagli interpretanti.

Attraverso svariati test si possono distinguere le parole in base ai loro generi ,categorie e valori.

Tralasciando le varie teorie e ricerche condotte da Pustejovsky focalizzerei l'attenzione su quello che a mi avviso tocca da vicino la nostra trattazione

Teoria composizionale dei Qualia

Pustejovsky ritiene che la" rappresentazione semantica lessicale di una parola non è un'espressione isolata ma in realtà è collegata al resto del lessico, quindi l'integrazione globale di una semantica di un elemento lessicale si ottiene dall'ereditarietà strutturata attraverso i differenti qualia associati con una parola ." Io chiamo questa: struttura di ereditarietà lessicale." (Pustejovsky)

Il significato lessicale di una parola può essere meglio catturato assumendo i seguenti livelli di rappresentazione:

Argument Struttura: Il comportamento di una parola trattato come una funzione. Indica come è costruita l'espressione sintattica.

2. Evento Struttura: Identificazione di un particolare tipo di evento per una parola o una frase:stato, processo, o transizione.

3. Qualia Struttura: gli attributi essenziali di un oggetto come definito dall'elemento lessicale

4. Ereditarietà Struttura: come la parola è globalmente connessa ad altri concetti nel lessico

Ogni livello contribuisce a fornire informazioni sul significato delle parole.

$$\begin{bmatrix} \alpha \\ \text{TYPESTR} = \left[\text{ARG1} = \textbf{the type of } \alpha\right] \\ \text{ARGSTR} = \left[\text{D-ARG1} = \textbf{other arguments in the qualia}\right] \\ \text{EVENTSTR} = \left[\text{E1} = \textbf{events in the qualia}\right] \\ \text{QUALIA} = \begin{bmatrix} \text{FORMAL} = \textbf{isa-relation} \\ \text{CONSTITUTIVE} = \textbf{parts of } \alpha \\ \text{TELIC} = \textbf{purpose of } \alpha \\ \text{AGENTIVE} = \textbf{how } \alpha \textbf{ is brought about} \end{bmatrix} \end{bmatrix}$$

Riprendendo l'esempio preso in esame da Michael Johnston e Federica Busa:

$$\begin{bmatrix} \textbf{knife} \\ \text{TYPESTR} = \left[\text{ARG1} = \boxed{x}\,\textbf{artifact\_tool}\right] \\ \text{ARGSTR} = \begin{bmatrix} \text{D-ARG1} = \boxed{y}\,\textbf{physobj} \\ \text{D-ARG2} = \boxed{w}\,\textbf{human} \\ \text{D-ARG3} = \boxed{z}\,\textbf{human} \\ \text{D-E1} = \boxed{e_1}\,\textbf{transition} \\ \text{D-E2} = \boxed{e_2}\,\textbf{process} \end{bmatrix} \\ \text{QUALIA} = \begin{bmatrix} \text{FORMAL} = \boxed{x} \\ \text{CONSTITUTIVE} = \{\textbf{blade,handle,...}\} \\ \text{TELIC} = \textbf{cut\_act}(\boxed{e_2}, \boxed{w}, \boxed{x}, \boxed{y}) \\ \text{AGENTIVE} = \textbf{make\_act}(\boxed{e_1}, \boxed{z}, \boxed{x}) \end{bmatrix} \end{bmatrix}$$

La struttura dei qualia é formata da quattro regole di composizione:

- Regola di composizione del qualia formale : distingue l'oggetto da un più largo dominio
    - Orientamento
    - Magnitudine
    - Dimensionalità
    - Colore
    - posizione
- Regola di composizione del qualia costitutivo :valuta la relazione che intercorre tra un oggetto e i costituenti o le parti di sua proprietà
    - Materiale
    - Peso
    - Parti e elementi componenti
- Regola di composizione del qualia telico: prende in esame la funzione di un oggetto e il suo scopo
    - Proposito per cui un agente ha creato l'oggetto
    - Costruzione in funzione di qualcosa e le sue specifiche attività
- Regola di composizione del qualia agentivo: calcola i fattori coinvolti nella creazione dell'oggetto
    - Creatore
    - Artefatto
    - Specie naturale
    - Catena causale

Le regole frasali del Lessico Generativo catturano le differenze che si evidenziano nel momento in cui un nome" capo" entra in contatto con un nome modificante

Confronto italiano- inglese

Per fare degli esempi di che cosa si sta parlando utilizzerò quelli presi in esame da Busa – Johnston

a.	bread knife	b.	wine glass	c.	bullet hole
	coltello da pane		bicchiere da vino		foro di pallottola
d.	lemon juice	e.	glass door	f.	silicon breast
	succo di limone		porta a vetri		seni al silicone

:

In inglese l'ultimo termine è il termine capo e il primo il modificatore

In italiano raramete si utilizzano dei composti nominali.

Prendendo in esame il primo e il secondo esempi,o i nomi pane e vino indicano lo scopo per il quale l'oggetto è stato creato (il coltello creato per tagliare il pane e il bicchiere per contenere il vino) La preposizione che questo tipo di modificazione è *da*.

Per quanto riguarda il terzo e il quarto caso, i nomi pallottola e limone indicano l'origine dell'oggetto indicato da nome capo . Come è stato ottenuto?(il succo derivato dalla spremitura del limone) La preposizione relativa è *di*

Nel quinto e nel sesto caso, vetri e silicone indicano di che materiale sono costituiti gli oggetti porta e seni. Con che cosa sono fatti? (porta fatta con i vetri e i seni fatti con il silicone)

La differenza che si nota tra l'inglese e l'italiano è che nella lingua anglosassone la preposizione scompare.

Per quanto riguarda la modificazione apportata al nome capo nel momento in cui vige una regola di composizione del qualia telico, si può dire che ci si trova a determinare una modificazione dettata da un evento e la situazione è un po' più articolata.

In alcune forme in cui il modificatore definisce tale evento la preposizione italiana è *di* in altre *da.*

I casi in cui compare la preposizione di l'evento preso in esame è il risultato di un'attività mentre nel momento in cui rileviamo la preposizione da si ha solamente la descrizione di un'attività.

a. hunting rifle fucile da caccia	b. race car macchina da corsa	c. carving wood legno da intaglio
a. destruction weapons armi di distruzione	b. credit card carta di credito	c. rest home casa di riposo
d. concentration camp campo di concentramento	e. divorce procedure procedura di divorzio	

(Busa)

Nonostante il Lessico Generativo non riesca ad avere una copertura totale d'applicabilità, resta comunque un modello d'ispirazione per il corretto funzionamento di un modello semantico distribuzionale composizionale.

Capitolo 5

Induzione di modelli di Semantica Distribuzionale Composizionale da esempi

Per cercare di studiare come avvengono praticamente le modificazioni nel linguaggio e come si attiva il meccanismo di composizione del significato la cosa più utile da fare è prendere in esame svariati termini e metterli a confronto.

Per far ciò è stato necessario lavorare con un dizionario, che essendo archivio naturale di espressioni equivalenti ci offre la possibilità di utilizzare tali espressioni per estrarre esempi positivi per i CDS di formazione e di testare i modelli"(Zanzotto 2010)

L'idea di fondo è che nei dizionari vi è già una comparazione di significato tra la parola target e la sequenza di definizione che segue.

Per esempio:

target word (t)	definition sequence (s)
contact	close interaction
high life	excessive spending

(Zanzotto 2010)

Si può notare che nella definizione di contatto la combinazione del significato di "interazione" e "ravvicinata" da come significato complessivo proprio contatto; o meglio il significato del termine contatto è ottenuto dall'unione dei significati dei lessemi della definizione dizionariale.

Di conseguenza il vettore t è ciò che deve risultare dalla combinazione dei vettori s1(vettore rappresentante close) e s2(vettore rappresentante interaction).

Per valutare l'efficacia del metodo di Semantica Distribuzionale Composizionale sono stati presi in esame alcuni termini nella cui definizione compaiono o due nomi o un nome e un aggettivo.

WordNet ® è un grande database lessicale della lingua inglese, sviluppato sotto la direzione di George A. Miller (emerito). Sostantivi, verbi, aggettivi e avverbi sono raggruppati in insiemi di sinonimi cognitive (synsets), ognuno esprime un concetto distinto. Synset sono collegati per mezzo di relazioni concettuali-semantici e lessicali. La rete di parole e concetti relativi significato può essere navigato con il browser. WordNet è anche disponibile liberamente e pubblicamente per il download. struttura di WordNet lo rende uno strumento utile per la linguistica computazionale e l'elaborazione del linguaggio naturale.

Nel corso degli anni, molte persone hanno contribuito allo sviluppo di WordNet. Attualmente, il team WordNet comprende i seguenti membri, e il progetto WordNet è ospitato presso il Dipartimento di Informatica:

George A. Miller (emerito)
Christiane Fellbaum
Randee Tengi
Helen Langone
Adam Ernst (Wordnet a lexical database for English)
Lavanya Jose

Per lo svolgimento di questo lavoro è stato utilizzato Wordnet come dizionario di riferimento .

Esempi NN

1. 108895148 Aberdare mining town (S1 (NP (NP (DT a)(NN mining)(NN town))(PP (IN in)(NP (JJ southern)(NNP Wales)))))#
2. 101750027 Acanthophis elapid snakes (S1 (NP (JJ Australian)(NN elapid)(NNS snakes)))#
3. 109767700 actress female actor (S1 (NP (DT a)(NN female)(NN actor)))#
4. 114729737 actomyosin protein complex (S1 (NP (NP (DT a)(NN protein)(NN complex))(PP (IN in)(NP (NN muscle)(NNS fibers)))))#
5. 114108039 angina heart condition (S1 (NP (NP (DT a)(NN heart)(NN condition))(VP (VBN marked)(PP (IN by)(NP (NP (NNS paroxysms))(PP (IN of)(NP (NN chest)(NN pain)))(ADJP (JJ due)(PP (TO to)(NP (NP (VBN reduced)(NN oxygen))(PP (TO to)(NP (DT the)(NN heart)))))))))))#
6. 109882615 businesswoman female businessperson (S1 (NP (DT a)(NN female)(NN businessperson)))#
7. 112460308 chincherinchee South African (S1 (NP (NP (NNP South)(NNP African))(ADJP (JJ perennial)(PP (IN with)(NP (NP (JJ long-lasting)(NNS spikes))(PP (IN of)(NP (NP (JJ white)(NNS blossoms))(SBAR (WHNP (WDT that))(S (VP (AUX are)(VP (VBN shipped)(PRT (RP in))(PP (TO to)(NP (NP (NNP Europe))(
8. 106519369 pass complimentary ticket (S1 (NP (DT a)(NN complimentary)(NN ticket)))#
9. 102097298 Scottie old Scottish (S1 (NP (NP (NNP old)(NNP Scottish))(NP (NP (NN breed))(PP (IN of)(NP (NP (JJ small)(JJ long-haired)(ADJP (RB usually)(JJ black))(NN terrier))(PP (IN with)(NP (NP (VB erect)(NN tail))(CC and)(NP (NNS ears)))))))))#
10. 103071552 tintometer measuring instrument (S1 (NP (NP (DT a)(NN measuring)(NN instrument))(VP (VBN used)(PP (IN in)(NP (JJ colorimetric)(NN analysis)))(S (VP (TO to)(VP (VB determine)(NP (NP (DT

the)(NN quantity))(PP (IN of)(NP (NP (DT a)(NN substance))(PP (IN from)(NP (NP (DT the)(NN color))(SB

Esempi JN

1. 100002137 abstraction generalconcept (S1 (NP (NP (DT a)(JJ general)(NN concept))(VP (VBN formed)(PP (IN by)(S (VP (VBG extracting)(NP (JJ common)(NNS features))(PP (IN from)(NP (JJ specific)(NNS examples)))))))))#
2. 100021939 artefact man-made object (S1 (NP (NP (DT a)(JJ man-made)(NN object))(VP (VBN taken)(PP (IN as)(NP (DT a)(NN whole))))))#
3. 100023773 motive psychological feature (S1 (NP (NP (DT the)(JJ psychological)(NN feature))(SBAR (WHNP (WDT that))(S (VP (VBZ arouses)(NP (NP (DT an)(NN organism))(PP (TO to)(NP (NN action))))(PP (IN toward)(NP (DT a)(VBN desired)(NN goal)))))))#
4. 100127531 best supreme effort (S1 (NP (NP (DT the)(JJ supreme)(NN effort))(SBAR (S (NP (PRP one))(VP (MD can)(VP (VB make)))))))#
5. 100135637 haymaker hard punch (S1 (NP (NP (DT a)(JJ hard)(NN punch))(SBAR (WHNP (WDT that))(S (VP (VBZ renders)(S (NP (DT the)(NN opponent))(ADJP (JJ unable)(S (VP (TO to)(VP (VB continue)(S (VP (VBG boxing)))))))))))))#
6. 100144632 stroke light touch (S1 (NP (NP (DT a)(JJ light)(NN touch))(PP (IN with)(NP (DT the)(NNS hands)))))#
7. 100154433 support documentary validation (S1 (NP (JJ documentary)(NN validation)))#
8. 110477077gladiator professional boxer (S1 (NP (DT a)(JJ professional)(NN boxer)))#
9. 110532751 ritualist social anthropologist (S1 (NP (NP (DT a)(JJ social)(NN anthropologist))(SBAR (WHNP (WP who))(S (VP (AUX is)(NP (NP (NN expert))(PP (IN on)(NP (NNS rites)(CC and)(NNS ceremonies)))))))))#
10. 113910384 space empty area (S1 (NP (NP (DT an)(JJ empty)(NN area))(PRN (-LRB- -LRB-)(VP (ADVP (RB usually))(VBN bounded)(PP (IN in)(NP (NP (DT some)(NN way))(PP (IN between)(NP (NNS things))))))(-RRB- -RRB-))))#

Dopo aver estrapolato degli esempi e poste le adeguate domande l'ultimo passaggio è stato determinato dalla scelta di quale regola di composizione di qualia si trattasse.

Conclusioni

In questa tesi non essendomi riproposta di trovare una soluzione al problema inizialmente citato , l'unico obiettivo che spero di aver raggiunto è quello di aver compreso in quale direzione probabilmente si dovrà rivolgere l'attenzione di uomini illustri dal prodigioso intelletto, per realizzare quello che solamente un secolo fa sembrava solamente un'idea assurda partorita da menti di uomini folli: rendere le macchine intelligenti.

Ripercorrendo brevemente le tappe di questo studiosi è d'obbligo rielencare i punti fondamentali sotto la cui guida sono state scritte queste pagine.

- Definizione di una chatbot del suo funzionamento e scoperta dei limiti linguistici
- Comprensione della differenza tra il significato di una parola e il significante della stessa
- Definizione del problema dovuto all'impossibilità apparente di sciogliere il legame tra le due componenti di un termine
- Proposta di una soluzione
    - Trattamento del linguaggio naturale come se fosse formale
    - Spiegazione di modelli Semantica Distribuzionale
    - Modelli di Semantica Distribuzionale Composizionale
    - Analisi del Lessico Generativo
    - Induzione di modelli CDS da esempi positivi (Wordnet)

Negli ultimi anni nel campo dell'intelligenza artificiale sono stati fatti passi da gigante (ecco un tipico esempio di espressione che la nostra macchina che parla non riuscirebbe a comprendere) ma la strada è ancora molto molto lunga.

Chissà se veramente un giorno avremo la possibilità di intrattenere un dialogo con un computer o se i robots che vediamo nei film resteranno solo finzioni cinematografiche?

Forse la ricchezza e la fecondità della capacità intellettiva e la creatività sono destinate a restare una peculiarità umana? Certo è che ,indipendentemente dal risultato ultimo del lavoro di migliaia di scienziati impiegati in questa ricerca ,ne sarà comunque valsa la pena.

Ringraziamenti

Una delle domande più frequenti che mi sono sentita porre in questi ultimi mesi è stata:

“Perché fai la tesi con il professore di informatica ? Scusa non fai filosofia?”

Lo so che apparentemente può sembrare una scelta abbastanza insolita ma la risposta è veramente semplice.

Dando per scontato il fatto che pur essendo una studentessa di lettere e filosofia mi interessino incredibilmente gli studi scientifici, ci fu un episodio particolarmente significativo .

Nel mese di giugno 2010 precisamente il 30, ho avuto il piacere di sostenere l’esame di Filosofia della Scienza con il professor Arturo Carsetti.

A parte il fatto che parte del programma riguardava l’Intelligenza Artificiale, durante l’esame il professore mi disse :” Durante la mia lunga carriera di docente ho notato che a parità di possibilità c’è una profonda differenza tra voi che studiate lettere e filosofia e gli studenti di materie scientifiche. Loro sono molto più pronti di voi e anche se trovano qualche difficoltà non hanno alcun problema a sostenere un vostro esame mentre senza ombra di dubbio il 90 % di voi non riuscirebbe mai a terminare una prova di Analisi Matematica.”

Purtroppo gli ho dovuto dar ragione…

Riflettendo però mi sono chiesta come mai chi studia filosofia, dimenticando che i più grandi pensatori del nostro passato affiancavano sempre studi scientifici e studi umanistici, ha questa sorta di repulsione nei confronti di tutto ciò che comprenda numeri esperimenti e formule?

Nel mio piccolo mi piacerebbe portare aventi una formazione comprensiva dei due ambiti.

Per questo ringrazio moltissimo il professor Carsetti docente dalla bisbigliante forza devastante dell’esperienza .

Certo è che se inaspettatamente il Professor Zanzotto non mi avesse dato la possibilità di preparare la tesi con lui, non avrei avuto il piacere di approcciarmi a questo studio interessantissimo che spero di portare avanti. Quindi grazie Professore per l'infinita pazienza disponibilità e professionalità con cui mi ha seguito in questo lavoro. Non è una cosa che si vede molto spesso nella nostra società quindi grazie ancora.

Lo so che i ringraziamenti tendono a essere smielati e fini a se stessi però non posso non ringraziare quel genio ribelle del mio professore di filosofia del liceo : Professor Chierichini che ho avuto la fortuna di ascoltare per tre anni e che con la sua potenza intellettiva è riuscito a farci amare la Filosofia. Tutti dovrebbero far lezione con lui.

Grazie alle persone che mi sono state accanto in questo lunghissimo percorso : la mia amica …ops collega che sa bene quanto è stata decisiva, il mio fidanzato che mi ha sopportato con amore e pazienza ,non senza poche difficoltà e mia nonna senza la quale la mia vita non sarebbe la stessa, grazie nonna . In fine non posso non ringraziare il mio angelo custode senza il quale niente sarebbe accaduto, niente avrebbe senso… grazie mamma per esserci sempre stata soprattutto nei momenti bui e di sconforto. Senza di te non ce l'avrei fatta…

Bibliografia

(s.d.). Tratto il giorno 2011 da BookRags.

(s.d.). Tratto il giorno 2010 da Wordnet a lexical database for English.

(2011). Tratto da Wikipedia L'encilopedia libera.

Busa, M. J. (s.d.). *Qualia Structure and the Compositional Interpretetion of Compounds* .

Eco. (1975). *Trattato di semiotica generale.*

Firth, J. R. (1957).

Lenci, A. (s.d.). Spazi di parole: metafore e rappresentazioni semantiche.

Lenci, A. (s.d.). Spazi di parole: metafore e rappresentazioni semantiche.

Pustejovsky, J. (s.d.). *The Generative Lexicon* .

SAHLGREN, M. (s.d.). An Introduction to Random Indexing.

Shakespeare, W. *Romeo e Giulietta.*

Traini, S. *Le due vie della semiotica. Teorie strutturali e interpretative.* Strumenti Bompiani.

Weizenbaum, J. (1966). *ELIZA--A Computer Program For the Study of Natural Language Communication Between Man and Machine.*

Zanzotto, F. M. (2010). *Provisional: Estimating linear Compositional Distributional Semantics Models using Dictionaries* .

Zanzotto, F. M. (2008). Macchine : Conoscenza e ragionamento.