

Introduzione al testo filosofico mediante forme di annotazione semantica

Il campo delle applicazioni di *trattamento semantico* dei dati e dei documenti si è sviluppato già da molto tempo, ma non ha ancora raggiunto risultati, in un certo senso, conclusivi, com'è invece accaduto per la trattazione *sintattica* del linguaggio naturale, di documenti e informazioni espressi attraverso di esso. Le ragioni di questa incompletezza sono molte e complesse. A livello intuitivo è però facile comprendere come la struttura sintattica di una frase o di un discorso siano qualcosa di esplicitabile in modo relativamente semplice in virtù della *finitezza* delle forme sintattiche che strutturano il linguaggio umano (in numero non basso, ma non certo infinito) e nonostante la ricorsività di cui il linguaggio stesso “si serve” dal punto di vista sintattico per costruire strutture complesse in linea di principio senza limiti (di fatto, i vincoli sono quelli posti dalle risorse computazionali dell'entità che produce strutture sintattiche, umana o non umana che sia).

La *semantica* di un testo invece non deriva, come la sintassi, da un numero di forme finite attraverso cui si possono costruire strutture lunghe a piacere (per fare un'analogia con la matematica, si pensi all'insieme dei numeri naturali che, pur infinito, si può costruire componendo un numero finito di elementi). Essa è caratterizzata da una sorta di indeterminatezza costitutiva. Gli elementi base della semantica, i pezzetti su cui costruire l'impalcatura di significato di un discorso sono già in partenza infiniti, data la *natura produttiva delle categorizzazioni e delle concettualizzazioni* che gli esseri umani mettono in atto e che si riflette direttamente nelle forme linguistiche, cioè nei termini o associazioni di termini del discorso stesso (per fare un'analogia con la matematica, qui ci muoviamo fra ordini diversi d'infinito). La semantica, in breve, sfugge sempre, da un lato o dall'altro – leibnizianamente, da quello dell'analisi del significato dei termini a quello delle associazioni dei termini in termini o concetti più complessi – a un trattamento univoco, deterministico e limitato (almeno in linea di principio).

Una considerazione positiva di quest'aspetto – che può essere ritenuto un enorme ostacolo al progresso di svariate discipline, quali la linguistica computazionale, la filosofia, l'informatica e molte altre al confine fra queste – è il fatto che abbia spinto fin dagli esordi del suo presentarsi, in particolare con l'affermarsi del mondo della rete e con la moltiplicazione esponenziale dei documenti di testo disponibili per gli utenti, all'individuazione di tecniche per il *trattamento semantico* dei dati e delle informazioni contenuti in ogni tipo di “luogo” virtuale. Il campo da cui i nuovi strumenti

COMPLETEZZA SINTATTICA E
INCOMPLETEZZA SEMANTICA
DEL LINGUAGGIO NATURALE

LA NATURA PRODUTTIVA DELLE
CATEGORIZZAZIONI E
CONCETTUALIZZAZIONI

LE TECNICHE DI
“RAPPRESENTAZIONE” DELLE
CONOSCENZE

LE ONTOLOGIE INFORMATICHE:
PROGRAMMAZIONE A OGGETTI
E LINGUAGGI DI MARKUP

si sono sviluppati è certamente quello della *rappresentazione delle conoscenze* nell'ambito dell'intelligenza artificiale, anche in relazione alle ricerche di più esplicito tenore cognitivo, relative cioè alla comprensione della formazione e dell'utilizzo dei *concetti* come fenomeni mentali. Perciò, a partire dai *frame* e dalle *reti semantiche*, in un ambito più strettamente di intelligenza artificiale, ma anche dagli *script*, dalle mappe mentali e da quelle concettuali, in un ambito più specificamente cognitivo, si è arrivati ai moderni sistemi di strutturazione concettuale dei testi e, dunque, a enucleare il concetto di *ontologia informatica* e alla realizzazione di software per la costruzione di ontologie. In particolare, esse possono essere considerate il precipitato più prossimo di due filoni sviluppatisi all'interno dell'informatica: la programmazione a oggetti, oggi certamente la più diffusa, e i linguaggi di markup o "taggati" (XML, RDF, RDFS, OWL solo per fare alcuni esempi) che hanno portato a una standardizzazione nella rappresentazione dei documenti – innumerevoli e sempre crescenti – resi disponibili in rete.

LE ONTOLOGIE COME
"SEMANTICHE LOCALI"

In realtà, le ontologie informatiche devono non poco alla filosofia e alla logica del ventesimo secolo. Di fatto, ne sono una diretta applicazione pratica. Le ontologie sono, infatti, tecniche di rappresentazione della conoscenza che permettono di strutturare un dominio sulla base di *categorie specifiche*, delle relazioni fra loro intercorrenti e secondo il parametro di uno o più specifici obiettivi. Un'ontologia può essere considerata una *semantica locale*, appositamente creata per gestire informazioni riguardanti un ambito determinato, modellandone i punti di salienza attraverso l'utilizzo dei concetti fondamentali coinvolti e delle loro relazioni. Le ontologie utilizzate in ambito informatico permettono di attuare processi di negoziazione e ragionamento sul significato dei termini o delle categorie impiegate, nella misura in cui ognuno di essi assume il proprio *significato contestuale* all'interno di un determinato ambito o dominio, che può variare in modo sensibile con l'evoluzione e la moltiplicazione dei contenuti disponibili e degli utenti interagenti.

LE ONTOLOGIE COME MODELLI
DI STRUTTURAZIONE DELLA
CONOSCENZA

Le ontologie informatiche, rispetto a quelle filosofiche, hanno il vantaggio di essere disancorate da implicazioni "realistiche", nel senso che, almeno quelle più fondamentali, sono usate a un livello di astrazione che permette la loro applicazione a domini diversi, cambiando gli individui o gli elementi a cui l'ontologia si riferisce. Esse sono *modelli di strutturazione della conoscenza*. Contano principalmente per le relazioni che istituiscono fra gli elementi di categorie diverse. Permettono quindi tutto ciò che rientra nelle possibilità di una struttura gerarchica concettuale: classificare automaticamente una certa istanza sulla base delle relazioni gerar-

chiche espresse nell'ontologia, far ereditare le proprietà fra classi, sussumere classi sotto altre classi, e così via.

Le ontologie ricadono nell'ambito del *web semantico* o, come viene più spesso definito, del web 2.0. In termini molto generali, un'ontologia può essere assimilata a un *calcolo predicativo del primo ordine* basato su una teoria assiomatica, cioè su alcuni assiomi che definiscono i predicati interpretati in modo estensionale, i quali a loro volta definiscono sottoinsiemi del dominio degli individui. Tuttavia, le ontologie, che pure sono strutturazioni gerarchiche della conoscenza in cui i legami "is-a" e "part-of" costituiscono una parte fondamentale, non permettono operazioni di livello più complesso, come le *predicazioni di predicati*, cioè quelle tipiche delle *logiche del secondo ordine*. Questo fatto non costituisce un problema da un punto di vista strettamente computazionale, dato che ogni calcolo predicativo di un qualsiasi ordine può essere sempre ridotto a un calcolo predicativo del primo ordine, nel senso che ne esiste un *modello* al primo ordine. Ma dal punto di vista *pratico* la questione di questo tipo di riduzione pone serie limitazioni sul piano dell'*espressività*, che è ciò che maggiormente conta nel momento in cui lo scopo prefissato è quello di strutturare la conoscenza di un determinato ambito o dominio in modo pragmaticamente esaustivo. Infatti, è proprio la sua illimitata capacità espressiva che, dal punto di vista comunicativo e descrittivo, rende così potente il linguaggio naturale. Uno strumento che ne volesse catturare tutte le sfumature e le potenzialità in modo strutturato, dovrebbe essere altrettanto espressivo o, almeno, avvicinarsi a questo traguardo.

Dunque, pur essendo le ontologie annotazioni semantiche che vanno già in questa direzione, ne esistono di più potenti, pensate per ovviare a simili problemi e pervenire a una modellazione ancora più esaurente della conoscenza implicata in un determinato contesto: le *topic map*. Esse nascono da un'evoluzione di alcuni strumenti pensati per creare e trattare in modo automatico gli *indici analitici*. Un *indice analitico*, infatti, si pone a un livello semantico *superiore* rispetto al testo e i richiami fra le varie voci costituiscono un'associazione fra le categorie implicate nel discorso che s'istituisce a un meta-livello semantico. Una mappa topica ha tre elementi fondamentali: i *topic*, che costituiscono i *nodi* della mappa e corrispondono ai vari argomenti (*subject*) del discorso; le *association*, che costituiscono le relazioni fra i *topic*, specificando non solo che due *topic* sono in relazione, ma anche in che modo e perché sono in relazione; le *occurrence* che indicano le risorse o i riferimenti, anche esterni alla *topic map*, che possono essere utili ad essa (nel caso di un testo sono le occorrenze all'interno del testo di particolari istanze del *topic*). I *topic* svolgono un particolare

ONTOLOGIE, WEB SEMANTICO
E CALCOLO DEL PRIMO ORDINE

LE LOGICHE DEL SECONDO
ORDINE E IL PROBLEMA
DELL'ESPRESSIVITA'

LE MAPPE TOPICHE COME
EVOLUZIONE DEGLI INDICI
ANALITICI

TOPIC, SUBJECT, ASSOCIATION,
OCCURRENCE

ruolo (*association role*) ogni volta che sono all'interno di una *association* e dunque tale loro funzione si definisce *topic characteristic*.

Anche senza entrare troppo nei dettagli tecnici, non è difficile scorgere la rilevanza che le mappe topiche possono assumere per la strutturazione di testi complessi, come ad esempio quelli della tradizione filosofica. Spesso in questo genere di testi un medesimo termine riveste diversi significati a seconda del contesto (*scope*) in cui ricade, e quindi delle relazioni che stabilisce con altri *topic*. È facile comprendere che, individuando le varie associazioni fra *topic*, è possibile in maniera *metatestuale* scoprire se un determinato concetto fondamentale assume diversi significati, magari in modo progressivo, all'interno del testo. Così, mentre la negoziazione fra ontologie diverse, ad esempio nell'operato degli agenti software, permette la rinegoziazione del significato di un termine fino a una "visione condivisa", nelle *topic map* sono già presenti gli eventuali sviluppi del concetto e i suoi spostamenti di significato, inglobando esse in tal modo, *dinamicamente*, le funzioni che svolgono, *staticamente e informalmente*, le *mappe concettuali* come strutturazioni non formali e non trattabili del tutto automaticamente del significato dei termini. Inoltre, le *topic map*, similmente alle ontologie, possono essere unificate attraverso operazioni di fusione (*merging*) se si riferiscono allo stesso concetto anche se espresso con una diversa terminologia. Questo tipo di flessibilità le rende strumenti estremamente potenti dal punto di vista dell'analisi semantica testuale.

Dal punto di vista didattico, e con riferimento alla filosofia, le *topic map* permettono, oltre all'analisi del testo secondo meta-parametri concettuali (una sorta di filigrana semantica dei *subject* coinvolti nel discorso), anche un esercizio autonomo di concettualizzazione testuale. Infatti, le *topic map*, molto più che le ontologie, spingono a chiedersi non solo quali sono i concetti implicati, ma anche quali sono le *relazioni rilevanti* che vanno create per poter strutturare il testo. Il problema è squisitamente filosofico e costruire una *topic map* di un testo secondo una certa visione concettuale può portare a nuove associazioni fra *topic*, così come la scelta di altre associazioni iniziali può portare ad ancora nuove associazioni. Le *topic map* possono dunque essere sia uno strumento di autoapprendimento, sia un potente "motore di ricerca" cognitivo all'interno dei testi della tradizione filosofica, così come, se si vuole, di insiemi di testi di letteratura secondaria dedicati a uno specifico argomento che può presentare molteplici sfumature semantiche relative ai concetti implicati nella discussione (fenomeno non raro, ad esempio, in filosofia della mente).

LE MAPPE TOPICHE E IL TESTO
FILOSOFICO

MAPPE TOPICHE (DINAMICHE)
E MAPPE CONCETTUALI
(STATICO-FIGURATIVE)

L'ASPETTO CREATIVO DELLE
MAPPE TOPICHE IN QUANTO
"MOTORI DI RICERCA"
COGNITIVI