

SISTEMI DI REASONING AUTOMATICO PER KNOWLEDGE GRAPHS:

Sempre più aziende e “Internet giants” (ad es., Facebook, Amazon, Google) desiderano modellare la conoscenza sotto forma di *corporate knowledge graphs* e gestirla attraverso *knowledge graph management systems* (KGMS), sistemi in grado di rispondere a interrogazioni e svolgere ragionamenti complessi in modo automatico. Allo stesso tempo, un KGMS deve garantire alte prestazioni e scalabilità in ambito Big Data, interfacciandosi con una molteplicità di sistemi: RDBMS e NoSQL stores aziendali, package di machine learning e analytics, sorgenti web.

La Divisione Ricerca sulle Tecnologie Avanzate del Dipartimento IT della Banca d'Italia è attiva in questo ambito e contribuisce allo sviluppo di un KGMS, il sistema Vadalog.

In questo ambito, si propongono le seguenti tesi di laurea magistrale, rivolte a laureandi in Ingegneria Informatica. Il lavoro di tesi sarà collocato nell'ambito di un progetto aziendale strutturato, il lavoro sarà organizzato secondo pratiche agili e sotto la costante guida di ricercatori-tutor della Banca d'Italia.

1. OPERATORI DI JOIN RICORSIVO

La possibilità di trattare efficientemente interrogazioni ricorsive è un elemento essenziale nel ragionamento automatico su knowledge graphs, poiché consente la navigazione secondo percorsi specificati in maniera compatta e non completamente noti a priori.

Nel sistema Vadalog, gli algoritmi di ragionamento fanno ricorso ad estensioni e varianti degli operatori dell'algebra relazionale e dell'SQL, con implementazioni che si ispirano alle tecniche utilizzate nei RDBMS.

La tesi ha come obiettivo lo studio e lo sviluppo di algoritmi di join ottimizzati per supportare interrogazioni con un uso pervasivo della ricorsione.

2. REASONING PROBABILISTICO

La possibilità di descrivere e interrogare knowledge graphs probabilistici è una caratteristica sempre più richiesta nei KGMS, poiché si desidera combinare la capacità dei knowledge graphs di rappresentare domini di interesse complessi, ricchi di entità in relazione tra loro, con il trattamento dell'incertezza.

La tesi ha come obiettivo lo studio e lo sviluppo di algoritmi di reasoning probabilistico per Vadalog. In particolare, il lavoro avrà come oggetto l'implementazione di un modello di ragionamento probabilistico, basato su varianti di *Probabilistic Graphical Models* ampiamente utilizzate nei moderni approcci di machine learning (ad es. reti Bayesiane, Markov Random Fields, Markov Networks).

3. PROGETTAZIONE CONCETTUALE DI KNOWLEDGE GRAPHS

L'obiettivo della tesi è lo studio e la predisposizione di una metodologia di progettazione concettuale per knowledge graphs. In particolare, il tesista dovrà proporre e sviluppare formalismi e tecniche per specificare un dominio di business sotto forma di knowledge graph.

Il formalismo ideato dovrà essere facilmente comunicabile ad un'utenza non tecnica, rappresentabile graficamente in maniera intuitiva e traducibile in maniera semplice in un knowledge graph trattabile dal sistema Vadalog.

Parte integrante del lavoro di tesi consisterà nella definizione di metodologie per la traduzione di schemi concettuali di knowledge graphs in specifiche trattabili dal sistema.