Documentazione Progetto CovidKB

Esame Ingegneria della Conoscenza 21/22

**Progetto a cura di:**

**-**Pasquale Monniello, mat. 716926, mail: p.monniello@studenti.uniba.it

-Emanuele Pasculli, mat. 718583, mail: e.pasculli4@studenti.uniba.it

-Giuseppe Volpe, mat. 716682, mail: g.volpe40@studenti.uniba.it

Sommario

[Introduzione 2](#_Toc94886635)

[Competenze Acquisite 2](#_Toc94886636)

[Guida all’utilizzo 2](#_Toc94886637)

[Assiomi 3](#_Toc94886638)

[Ontologia 4](#_Toc94886639)

# Introduzione

CovidKB-UNIBA22 è una KB proposizionale a scopo didattico, avente come dominio quello di una diagnostica semplice del virus covid-19.

La semplice assiomatizzazione è stata fatta sulla base di informazioni tratte dal sito della World Health Organization: <https://www.who.int/health-topics/coronavirus#tab=tab_3>.

Il sistema, in realtà, può effettuare ragionamenti su una qualsiasi assiomatizzazione di KB aciclica memorizzata in un formato compatibile; la scelta del dominio è dovuta all’attualità dell’argomento.

L’utente può porre domande circa il suo stato di salute al sistema consultando l’ontologia, e il sistema, a sua volta, si può interfacciare con l’utente chiedendogli il valore di verità di determinati atomi che sono stati dichiarati ‘askable’ (cioè che ha senso chiedere) nell’assiomatizzazione.

La procedura di risoluzione implementata è stata quella Top-Down, che risponde a una query posta nel formato corretto.

Per quanto riguarda la strategia di selezione delle clausole, è stata implementata una ricerca BFS (Breadth first search) non informata in un grafo (di cui abbiamo fornito una semplice implementazione), in cui gli stati goal sono quelli che non contengono atomi.

# Competenze Acquisite

Durante lo sviluppo di questo progetto, abbiamo preso familiarità con il linguaggio proposizionale, nello specifico abbiamo lavorato con clausole definite.

Abbiamo compreso il funzionamento delle KB assiomatiche in maniera più profonda, avendone esplorato la componente pratica.

Per quanto riguarda la strategia risolutiva, abbiamo approfondito il funzionamento della procedura Top-Down, imparando come impiegare gli algoritmi di ricerca nei grafi nella strategia di selezione.

# Guida all’utilizzo

All’interno del file CovidAPP-1.0.tar.gz, presente nella cartella target, che va estratto in qualsiasi cartella, vi sono cinque elementi:

* Una cartella ‘doc’ contenente la documentazione, ergo questo file.
* Una cartella ‘lib’ contenente i .jar.
* Una cartella ‘resources’ contenente le risorse utilizzate nell’APP.
* ‘BuildKB.bat’, eseguibile che costruisce il database da cui l’APP preleva informazioni
* ‘CovidKB.bat’, semplice eseguibile dell’applicazione

**Se si avvia l’applicazione per la prima volta su una macchina** avviare ‘BuildKb.bat’.

Immagine che contiene testo, monitor, interni, elettronico

Descrizione generata automaticamente

*Schermata principale dell’applicazione.*

Se si avvia l’applicazione su Windows il database sarà creato all’interno della cartella utente in una cartella appositamente creata chiamata “KBdatabases”.

Una volta caricata la schermata principale, sarà possibile interagire con l’applicazione tramite inserimento dei comandi nell’apposita barra.

* **Digitando il comando ‘help’**, sarà possibile vedere una lista dei comandi supportarti.
* **Digitando il comando ‘axioms’,** sarà possibile avere una lista degli assiomi caricati nell’applicazione
* **Digitando il comando ‘ontology’,** sarà possibile avere l’ontologia della KB caricata.
* **Digitando il comando ‘ask’,** seguito dall’atomo interessato, interrogherà l’applicazione.

Il sistema stampa una notifica di errore nel caso in cui l’utente inserisca comandi che non rispettano la sintassi stabilita.

Il sistema, inoltre, porrà delle domande all’utente attraverso finestre di dialogo che dispongono di bottoni per fornire la risposta (booleana), qualora sia necessario l’inserimento delle informazioni dall’utente.

I passi della visita del grafo vengono stampati su schermo durante la procedura, e una volta che questa è terminata, ne viene mostrato l’esito sull’interfaccia insieme alla clausola di risposta risultante.

# Assiomi

* common\_symptoms <- fever
* common\_symptoms <- cough
* common\_symptoms <- asthenia
* serious\_symptoms <- taste\_or\_smell\_loss
* serious \_symptoms <- chest\_pain
* common\_symptoms <- breathing\_difficulty
* is\_protected <- already\_had\_covid
* is\_protected <- vaccinated
* is\_at\_risk <- hanged\_out\_no\_protection
* is\_at\_risk <- contact\_with\_positive
* flu <- common\_symptoms ^ is\_protected
* covid\_mild <- common\_symptoms ^ is\_at\_risk
* covid\_serious <- serious\_symptoms ^ is\_at\_risk
* covid <- covid\_mild
* covid <- covid\_serious
* covid <- tested\_positive

# Ontologia

* common\_symptoms: il soggetto presenta sintomi comuni del covid
* fever: il soggetto ha la febbre
* cough: il soggetto ha la tosse
* asthenia: il soggetto ha spossatezza
* breathing\_difficulty: il soggetto presenta difficoltà respiratorie
* serious\_symptoms: il soggetto presenta sintomi gravi del covid
* taste\_or\_smell\_loss: il soggetto presenta la perdita di gusto o odore
* chest\_pain: il soggetto ha il dolore al petto
* is\_protected: il soggetto è protetto
* already\_had\_covid: il soggetto ha già avuto il covid
* vaccinated: il soggetto è vaccinato
* is\_at\_risk: il soggetto è a rischio
* hanged\_out\_no\_protection: il soggetto non ha usato protezione quando è all’esterno
* contact\_with\_positive: il soggetto ha avuto contatto con un positivo
* no\_symptoms: il soggetto non ha sintomi
* flu: il soggetto ha l’influenza
* covid\_mild: il soggetto ha il covid in maniera leggera
* covid\_serious: il soggetto ha il covid in maniera grave
* tested\_positive: il soggetto è risultato positivo al covid

# Valutazione

Poiché la correttezza di una KB dipende dalla sua semantica ed è indipendente dall’algoritmo eseguito, e data la semplicità del progetto, che non prevede nessun apprendimento o casualità, ergo la valutazione può avvenire offline, la valutazione quantitativa può avvenire su due principali assi:

* **Semantica**. Si può verificare la KB è corretta semanticamente. Una KB corretta semanticamente permette di derivare conseguenze logiche correttamente.
* **Procedura**. Si può verificare che la procedura di dimostrazione degli atomi sia corretta (sound) e completa.

Se entrambi questi elementi dovessero risultare corretti qualsiasi sia la loro istanza, allora tutti i risultati forniti dal programma (e dalla KB) risulterebbero corretti.

È importante aggiungere che prima di effettuare la valutazione, è stato eseguito il debugging della conoscenza, portando al rilevamento di alcune falle nella semantica

## Procedura

La procedura che abbiamo implementato è quella Top-Down. Poiché abbiamo seguito l’algoritmo ben noto (SLD), ogni KB aciclica permette alla nostra procedura di produrre risultati e di essere sia corretta sia completa.

## Semantica

Per verificare la correttezza della Semantica basta provare la sua corretta assiomatizzazione. A posteriori, la semantica della KB che abbiamo creato è *estremamente semplicistica*. Ad esempio: l’atomo covid\_mild (Covid leggero) segue common\_symptoms (Sintomi comuni) e is\_at\_risk (Presenza del rischio), ma se si è a rischio e si ha di sintomi comuni non è detto che si sia malati di covid. Poiché la KB creata non è in grado di processare la causalità, nell’implementazione intesa se il soggetto è a rischio è necessariamente covid, se il soggetto è protetto, o non è a rischio, non ha il covid. Questo problema avrebbe potuto essere parzialmente risolto dall’introduzione delle clausole di Horn e dalla presenza delle negazioni. Un’altra incongruenza tra assiomatizzazione e mondo reale è che la protezione non è dicotomica e non si traduce necessariamente in immunità.

Ergo, per valutare la semantica corretta dobbiamo fare delle assunzioni sul mondo, riducendo la sua rappresentazione e semplificandola. Solo in tal caso l’assiomatizzazione allora può essere definita corretta.

Il modo in cui questo può essere verificato (non strettamente necessario) sarebbe con l’esecuzione del programma ripetuta per ognuna delle configurazioni (modelli) della KB possibili e comparando i valori ottenuti con i valori reali. Considerando 11 atomi Askable, vi sono 2^11 possibili modelli di KB. Se la semantica è stata implementata correttamente, la KB avrà il 100% di True Positive rate.