

Accesso a banche dati remote ed eterogenee mediante interfacce di browsing su thesauri

O. Signore - A.M. Garibaldi - M. Greco*

(*) CNR - CNUCE (Pisa) e Università di Lecce
via S. Maria, 36 - 56126 Pisa

Sommario

La consultazione di banche dati produce spesso risultati insoddisfacenti dal punto di vista del richiamo e della precisione. Le cause possono essere individuate nella scarsa conoscenza della struttura dei documenti e del contenuto dei singoli campi. Una soluzione è quella di decentrare a livello di periferia intelligente le funzioni di interfaccia, sviluppando del software applicativo ad hoc, che consenta l'interazione con una rappresentazione concettuale del documento, e l'interrogazione di thesauri strutturati.

Viene presentato uno strumento che consente la visualizzazione in forma grafica della struttura di un thesaurus, permettendo la navigazione sulla struttura e la selezione dei termini di interesse, per sottomettere successivamente la query al sistema residente sull'host remoto.

1 - L'informazione on-line

All'inizio degli anni '80 erano state formulate ottimistiche previsioni per la crescita del mercato dell'informazione online, ma di fatto tale crescita è stata molto rallentata da vari fattori, alcuni di natura essenzialmente tecnologica (velocità e affidabilità delle reti di trasmissione dati, supporto solo di terminali TTY sulle reti geografiche, vincoli valutari e problemi di addebito, alti costi di trasmissione e di interrogazione, complessità delle procedure di logon, etc.), ma altri legati alle caratteristiche intrinseche dei sistemi di Information Retrieval.

In effetti, all'utente interessato ad accedere alle banche dati disponibili in rete viene fornita una risposta semplicistica ("Nulla di più semplice, procurati un modem e un identificatore di rete, e poi accedi alla banca dati che ti interessa") che privilegia gli aspetti meramente tecnici rispetto a quelli culturali.

La conseguenza è che spesso, dopo la non semplice trafila tecnica e amministrativa necessaria per poter accedere al mercato dell'informazione on line, l'utente, lungi dall'aver soddisfatto le proprie esigenze, si trova sovente nella situazione delineata in fig. 1.

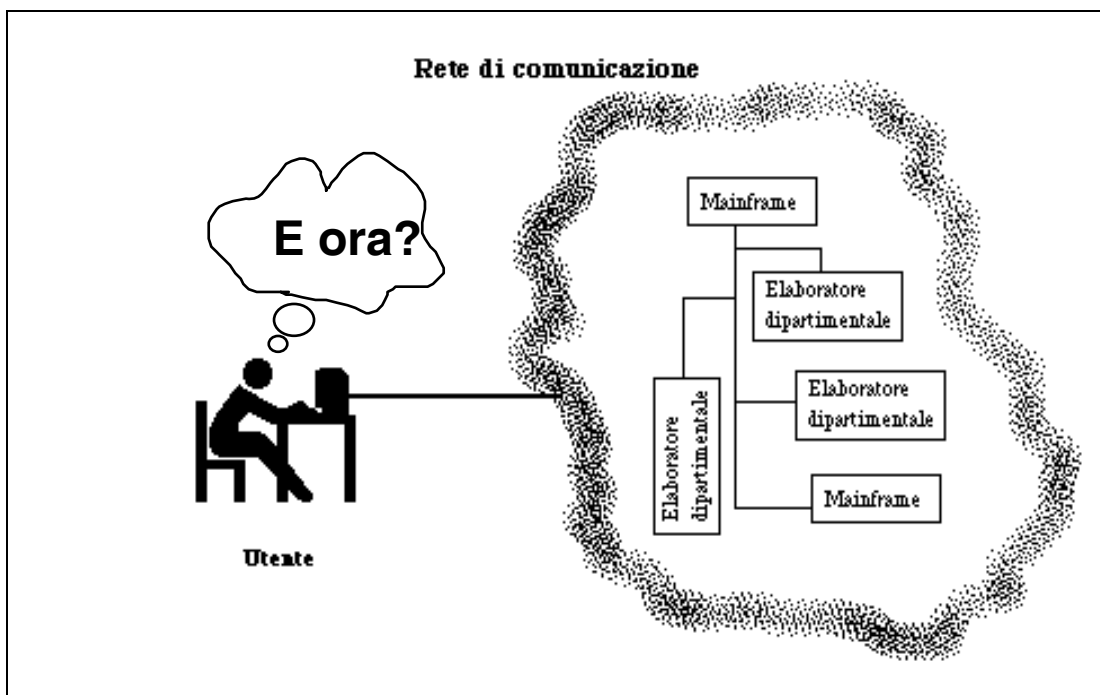


Fig. 1 - Accesso a banche dati mediante rete geografica

I sistemi esistenti, infatti, sono sostanzialmente equivalenti dal punto di vista funzionale, ma differiscono, a volte in modo netto, altre volte per piccole particolarità, soprattutto nel linguaggio di interrogazione. A questo particolare problema si era tentato di dare una soluzione in ambito CEE, mediante la definizione ([Negus77], [Negus80]) di un linguaggio standard di interrogazione (Common Command Language, o CCL) che costituiva un sovrainsieme dei linguaggi più diffusi. Questo linguaggio è stato implementato con successo da vari host, ([Bartoli82], [Schreiber80]), e ha costituito la base per una proposta di standard in ambito ISO.

2 - Le caratteristiche dei sistemi di Information Retrieval

Alla base di tutti i problemi resta il fatto che le banche dati residenti sugli host sono gestite da Sistemi di Information Retrieval.

Ogni sistema per il reperimento dell' informazione può essere descritto, da un punto di vista funzionale, come un insieme di informazioni (*Documenti*), un insieme di richieste utente (*Domande*) e un meccanismo (*Similitudine*) che provvede a identificare quali documenti sono in grado di soddisfare la richiesta dell' utente, cioè identifica quale frazione dei documenti disponibili ha buone probabilità di contenere

informazioni che corrispondono ai criteri di selezione specificati dall' utente. In linea di principio, e nell' ipotesi che sia possibile effettuare un confronto diretto tra l' insieme delle possibili domande e l' insieme dei contenuti dei documenti, il funzionamento di un sistema per il reperimento dell' informazione potrebbe essere rappresentato dalla fig. 2.

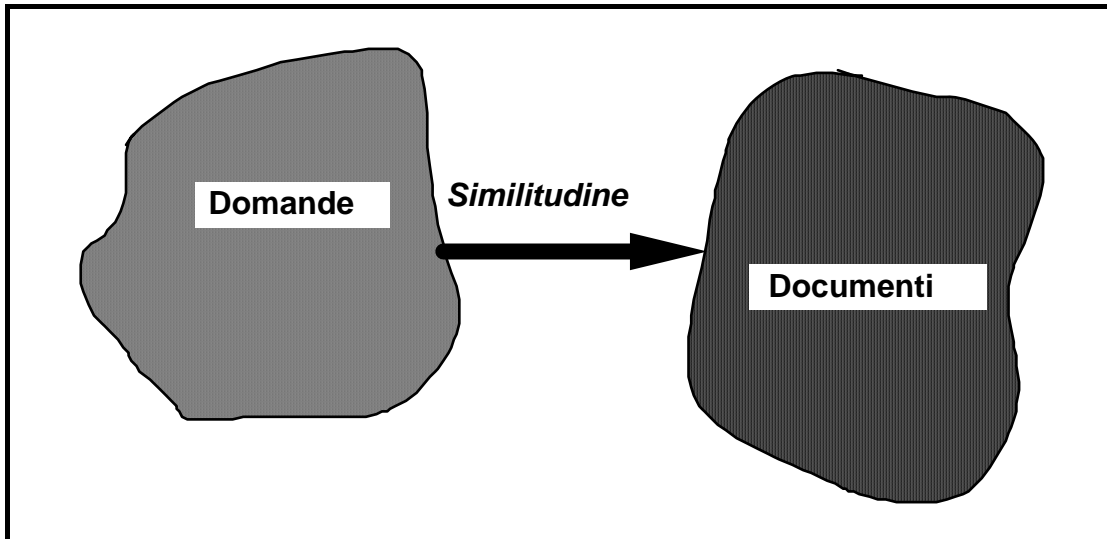


Fig. 2 - Schema funzionale di un Information Retrieval System

In pratica, il contenuto informativo del documento viene convertito in un formato particolare, facendo ricorso a un *linguaggio di indicizzazione* (o *linguaggio di classificazione*) grazie al quale è possibile effettuare l' analisi della corrispondenza tra l' insieme delle domande e l' insieme dei documenti (fig. 3).

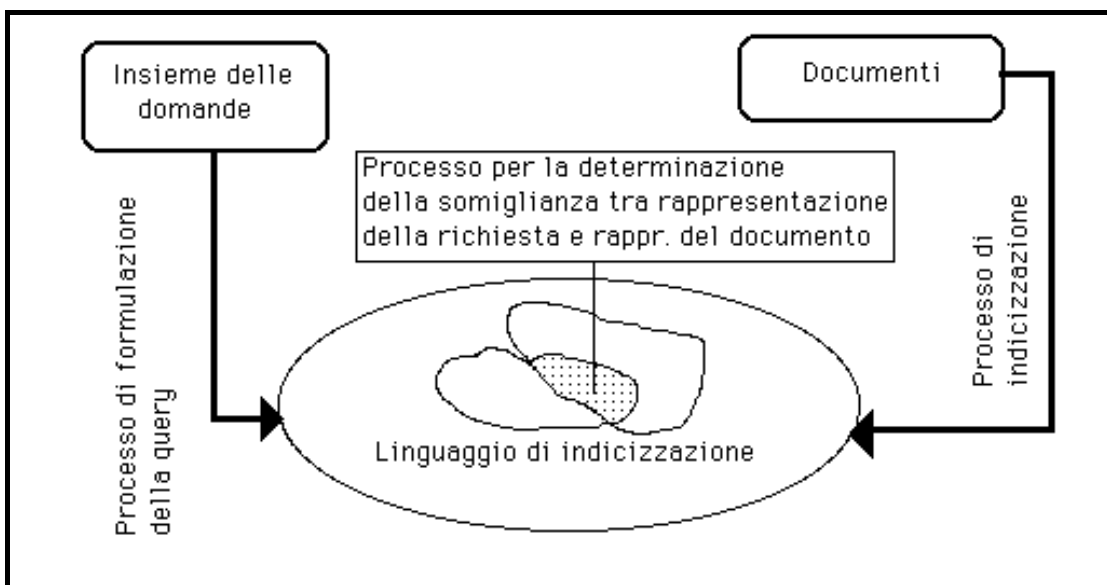


Fig. 3 - Ruolo del linguaggio di indicizzazione

La determinazione dell' equivalenza tra domande e documenti necessita pertanto di un processo di analisi della domanda e di un processo di indicizzazione del documento. E' sui risultati di questi due processi che viene effettuata l' analisi della corrispondenza tra i due insiemi di dati. Sia i meccanismi di formalizzazione delle domande che i problemi connessi con l' indicizzazione dei documenti sono stati oggetto di studio e sperimentazione.

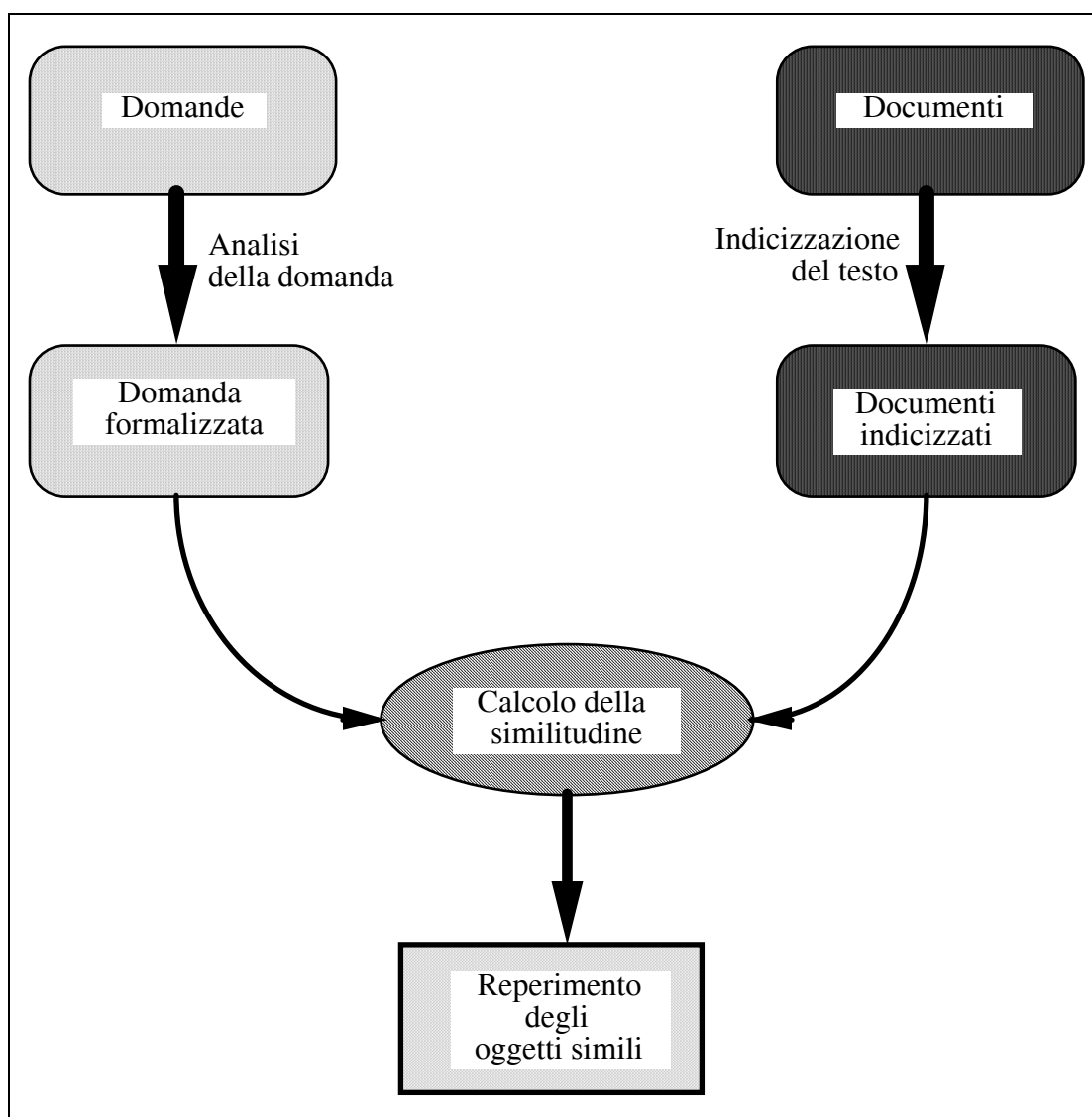


Fig. 4 - *Processo di reperimento dei documenti*

In definitiva, un Information Retrieval System è un sistema in grado di rappresentare documenti e recuperare quelli che sono probabilmente rilevanti per una richiesta

dell' utente. A tale scopo, esso dispone di metodi per la rappresentazione dei documenti e di metodi per il recupero dei documenti.

In un sistema ideale, a una query utente corrisponde la selezione di tutti e soli i documenti pertinenti alla domanda, cioè il sistema reperisce tutti i documenti rilevanti. In realtà, la situazione reale è quella schematizzata nella fig. 5.

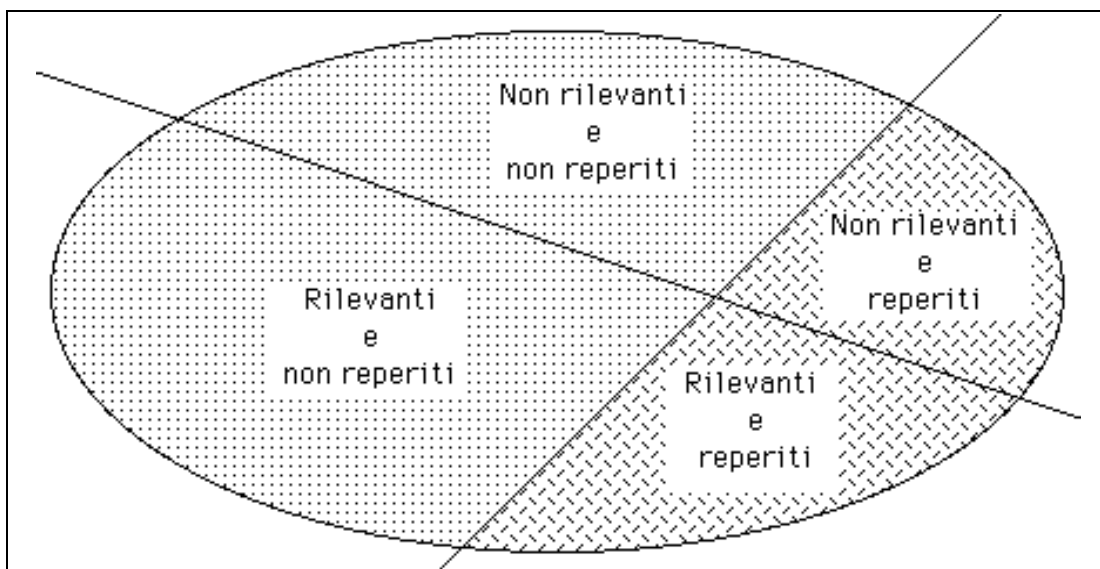


Fig. 5 - Misura dell' efficacia di un sistema di Information Retrieval

Di conseguenza, l' efficacia di un sistema per il reperimento dell' informazione viene misurata valutando *congiuntamente* due fattori:

- *Richiamo*, definito come il rapporto tra il Numero di Documenti Rilevanti Recuperati e il Numero di Documenti Rilevanti Esistenti;
- *Precisione*, definito come il rapporto tra il Numero di Documenti Rilevanti Recuperati e il Numero di Documenti Recuperati.

Appare evidente che i due fattori assumono valori compresi tra zero e uno.

A parte alcuni sistemi più recenti, tutti i sistemi commerciali di Information Retrieval gestiscono una unità di informazione (il *documento*) che viene strutturata in modo piatto, e suddivisa in paragrafi o campi. L' utente può ricercare tutti i documenti che contengono alcuni specifici concetti o nel documento inteso come insieme indistinto, o in sue particolari sezioni. Una richiesta ben mirata fa normalmente riferimento ad una specifica porzione di testo, in modo da ottenere un rumore abbastanza basso (ovvero una fattore precisione elevato). Ciò comporta che l' utente debba conoscere bene la struttura del documento, e quindi i nomi dei singoli campi o paragrafi, e il loro contenuto. Ovviamente i problemi aumentano quando,

per necessità implementative, il contenuto dei campi non è trasparente, in quanto le informazioni sono memorizzate in modo codificato. Normalmente, i sistemi dispongono di meccanismi di aiuto (*on line help*) che non sono tuttavia sempre soddisfacenti.

In conclusione, in un approccio Information Retrieval è necessario:

- identificare i *contenuti dei documenti* da archiviare
- identificare i *contenuti della query* dell'utente
- avviare una ricerca per *argomento*.

3 - Il ruolo dei thesauri nell' interfaccia utente

Gli utenti dei sistemi di reperimento delle informazioni hanno spesso bisogno di supporto per definire o raffinare gli argomenti di loro interesse, in altri casi devono migliorare le loro conoscenze sull' argomento per decidere esattamente quali sono gli elementi di interesse, infine esistono utenti che hanno una idea chiara dei concetti da ricercare, ma hanno delle difficoltà ad esprimerli. Esistono anche casi di utenti che partono da argomenti ben definiti e chiari, ma scoprono che il concetto di partenza è o troppo generico (e quindi produce in risposta un numero troppo elevato di documenti), o troppo specifico (determinando il ritrovamento di un numero troppo esiguo di documenti). E' perciò evidente che l' utente ha bisogno di un supporto che gli permetta di operare un affinamento del concetto ricercato.

Un modo per agevolare le ricerche degli utenti è quello di rendere disponibili *thesauri strutturati*, cioè *liste di termini tra cui esistono delle relazioni semantiche di gerarchia, equivalenza e preferenza* [ISO86]. I thesauri costituiscono perciò uno strumento fondamentale per permettere all' utente di identificare nella banca dati i concetti di interesse, e quindi di reperire i documenti rilevanti, oltre a costituire, in alcuni casi, una vera rappresentazione della conoscenza specifica in un dato dominio. Un sistema di questo tipo richiede all' utente di muoversi nella gerarchia dei concetti, e soprattutto di comprendere le interrelazioni esistenti tra di essi.

A livello internazionale ([ISO86]) sono stati standardizzati i tipi di relatori tra termini (NT, BT, LT, USE, etc.), e molti sistemi di Information Retrieval permettono all' utente di vedere, dato un termine, tutti i termini ad esso collegati, per poter espandere o rendere più specifica la domanda posta al sistema. I meccanismi adottati da tali sistemi, però, non appaiono adeguati a permettere la navigazione dell' utente in una realtà complessa come quella costituita da thesauri comprendenti migliaia di termini. In questi casi, viene implicitamente assunta una competenza

elevata da parte dell' utente, il quale può invece sentirsi disorientato, e aver bisogno di ritornare al punto iniziale per poter seguire un diverso cammino tra i concetti rappresentati nel thesaurus.

L' importanza di una appropriata interfaccia utente è stata ampiamente evidenziata nel corso degli ultimi anni. Una buona interfaccia utente deve mettere l' utente in grado di formulare ricerche significative mediante descrittori appropriati, quindi deve disporre di un meccanismo di comunicazione per i concetti chiave.

A livello implementativo, si tenta spesso di dare soluzione al problema mediante la realizzazione di interfacce a menù, che si presentano tipicamente come in fig. 6.

Immetti i valori richiesti

Autore = _____ **OR** _____ **OR** _____

Anno = _____

Argomento = _____ **OR** _____ **OR** _____

Fig. 6 - Una tipica interfaccia a menù

Tuttavia, le interfacce di questo tipo risultano piuttosto rigide, e di fatto si limitano a evitare all' utente l' utilizzo della logica booleana e dei livelli di parentesi, privilegiando quindi un aspetto tecnico elementare rispetto al problema fondamentale di *rendere accessibile la stessa base di conoscenza sia agli utenti che agli indicizzatori*.

Questo particolare problema, assai più rilevante ai fini della formulazione di una query efficace, può invece venir risolto da una *interfaccia grafica* orientata al *browsing*.

Una interfaccia di tale tipo può riuscire di particolare utilità per gli utenti estemporanei (ma non vanno trascurate le difficoltà a cui può andare incontro anche un utente esperto), che possono non conoscere perfettamente l' organizzazione e il contenuto della banca dati, e consente il raggiungimento di due risultati: il raffinamento dei concetti ricercati e il miglioramento della conoscenza della banca dati.

Una interfaccia grafica permette all' utente di orientarsi bene, anche meglio di quanto accade con le interfacce in linguaggio naturale, che hanno il difetto di non essere in grado di comunicare con chiarezza la struttura e gli obiettivi della banca dati.

4 - Il sistema Proteus

L'idea di realizzare il sistema Proteus è basata sul fatto che, come già notato in precedenza, e sottolineato in [Thompson89], mentre molto lavoro teorico e pratico è stato svolto per quanto concerne una migliore rappresentazione dei documenti e lo sviluppo di metodi più efficaci per l'organizzazione e la ricerca dei documenti (indicizzazione automatica, modelli probabilistici), non altrettante energie sono state profuse nello sviluppo di strumenti che fossero di ausilio all'utente per formulare la query. In altri termini, il problema da risolvere è quello di mettere l'utente nella condizione di esprimere la query nel modo migliore possibile (massimo Richiamo, massima Precisione).

Al raggiungimento di questo obiettivo, osta sostanzialmente la difficoltà per l'utente di definire in modo preciso il contenuto della query. Inoltre, per quanto concerne l'utilizzo di sistemi tradizionali, disponibili sulle reti pubbliche, non vanno dimenticati i già citati problemi della difficoltà di connessione e dell'eterogeneità del linguaggio.

Da queste considerazioni scaturisce l'opportunità di consentire all'utente di interagire con il sistema utilizzando una descrizione del documento indipendente dalla particolare implementazione fisica, analogamente a quanto avviene nella progettazione della basi di dati, in cui si definisce uno schema concettuale dei dati e una serie di operazioni su tale schema.

Tali considerazioni ci hanno portato a definire alcune delle caratteristiche fondamentali dell'organizzazione del thesaurus e dell'interfaccia grafica per i thesauri.

Una soluzione praticabile è sembrata la separazione delle funzioni di interfaccia e supporto alla formulazione della query, da quelle di interrogazione della banca dati.

Le due funzioni vengono svolte come due task separati (Fig. 7), di cui il primo si occupa della formulazione della query, e il secondo della sua esecuzione, mediante connessione in rete in emulazione di terminale.

Per il momento, non si prevede di intercettare direttamente i dati provenienti dall'host, ma semplicemente di permettere all'utente di passare da un task all'altro, mantenendo quindi il protocollo di comunicazione e la formattazione dei dati previsti dall'host.

L'interfaccia realizzata è adattabile a qualunque sistema di Information Retrieval (da qui il nome Proteus), ed è dotata delle seguenti caratteristiche:

- disponibilità su un personal computer di larga diffusione;

- realizzazione dell' interazione utente-sistema mediante una rappresentazione astratta del documento;
- possibilità di browsing grafico sui concetti.

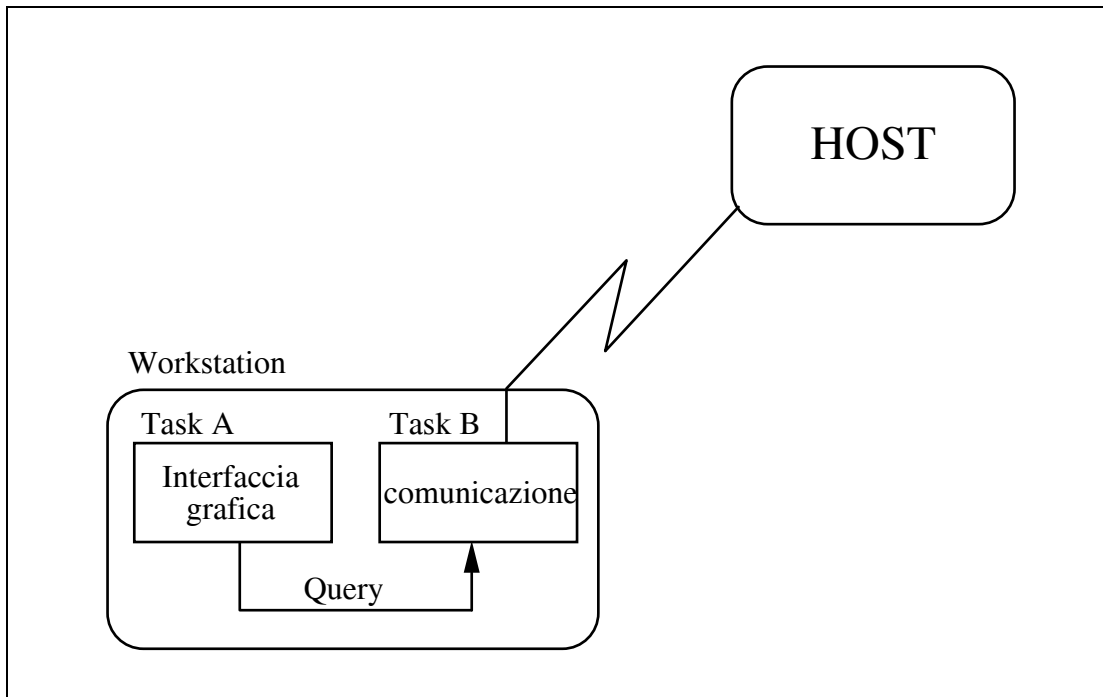


Fig. 7 - *L'architettura generale di Proteus*

La scelta implementativa di utilizzare una architettura hardware di grande diffusione non ha comportato particolari vincoli sulla portabilità dell' applicazione, anche se una sua riscrittura in ambienti industry standard non è automatica.

La possibilità di interazione mediante una rappresentazione astratta del documento offre alcuni vantaggi, quali:

- indipendenza dai vincoli sui nomi imposti dai singoli sistemi di Information Retrieval;
- possibilità per l' utente di interagire con i campi utilizzando una "sintassi esterna" nel caso di campi di struttura particolare ([Signore90]);
- preparazione della query in un linguaggio uniforme, indipendente dal sistema di gestione, e quindi disponibilità di un paradigma di interazione semplice e consistente.

Per quanto concerne il browsing, si è ritenuto opportuno privilegiare la soluzione in cui il browsing viene utilizzato come supporto alla formulazione della query, piuttosto che come meccanismo per la ricerca diretta del documento. Tale scelta, dettata anche dall' ipotesi di poter disporre di sistemi di classificazione e di thesauri

di elevata qualità, è molto simile a quella operata da CANSEARCH ([Pollitt84]), CALIBAN ([Frei83]), e CoalSORT ([Monarch87]), mentre differisce dall' approccio dei sistemi di tipo ipertestuale.

Proteus si avvale di un database di supporto costituito da un insieme di tabelle suddiviso in tre gruppi:

- l' insieme di tabelle utili ai fini della scelta del database e dell' host remoto, che contengono una descrizione della struttura dei vari database e dei diritti di accesso. Per ogni campo, viene indicato se esso è un campo a testo libero, o con vocabolario controllato, o con thesaurus, o con codifica interna particolare.
- L' insieme di tabelle necessarie per formulare la query, che, sui campi con dizionario o con thesaurus, consentono all' utente non esperto di scegliere in modo adeguato i valori da inserire nella query sul campo, consultando la lista dei valori che esso può assumere, ordinata in modo alfabetico e ricavata dinamicamente, su cui poi effettuare delle selezioni.
- L' insieme di tabelle per realizzare la navigazione sul thesaurus, che contengono le informazioni necessarie a rappresentare un termine come nodo di un grafo. Essa è utilizzata dalla parte di programma che si occupa del disegno del grafo che rappresenta il thesaurus.

La query, preparata in una sintassi standard, viene memorizzata in una tabella, e verrà successivamente tradotta nella sintassi adeguata per poterla presentare all' host remoto per l' esecuzione.

5 - Il funzionamento dell' interfaccia

L'insieme di funzionalità messe a disposizione può essere sostanzialmente suddiviso in tre gruppi, ognuno corrispondente a una ben precisa fase dell' interazione.

Nel **primo gruppo** troviamo tutte le funzionalità necessarie per permettere all' utente la scelta dell' host remoto cui dovrebbero essere inviate le query preparate, e la scelta del data base da interrogare.

Il **secondo gruppo** di funzionalità è indispensabile alla preparazione della query. In tale fase possono essere d'aiuto il Dict_Mgr ed il Thesaurus Manager; quest' ultimo racchiude l' insieme di funzionalità del **terzo gruppo**.

Mentre le funzionalità del primo e secondo gruppo sono di utilizzo assolutamente generale, la gestione del thesaurus presenta aspetti implementativi dipendenti dall' organizzazione del thesaurus stesso.

In figura 8 è illustrato il modo in cui i gruppi di funzionalità si succedono nel dialogo prima di giungere alla richiesta di esecuzione.

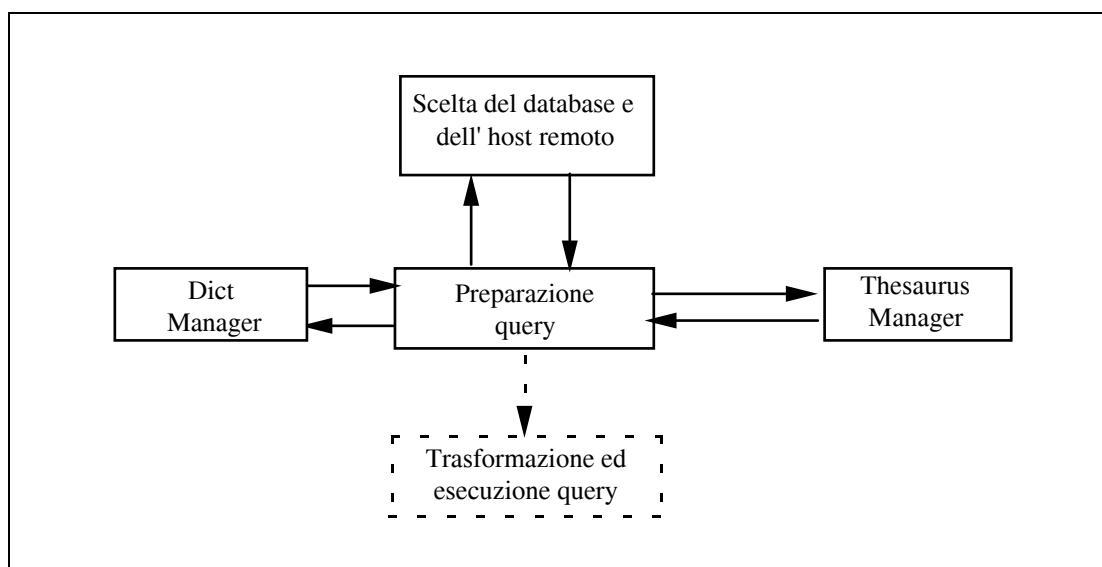


Fig. 8 - I gruppi di funzionalità messi a disposizione dall' interfaccia

Il dialogo si svolgerà quindi in tre fasi ben distinte, l' ultima delle quali può essere "aggirata" nel caso in cui l' utente sappia già con precisione quali termini del thesaurus inserire nella query. Dato che sono state sfruttate le possibilità di realizzare applicazioni multitasking, offerte dall' ambiente prescelto per l' implementazione, l' interazione con il Thesaurus Manager *può essere eseguita in parallelo* con la fase del dialogo riguardante la preparazione della query sui campi del documento.

5.1 - La scelta del database e dell' host remoto

In questa fase si presenteranno all' utente, in sequenza, tre finestre di dialogo con cui si richiede l' immissione di valori o la scelta fra più alternative proposte.

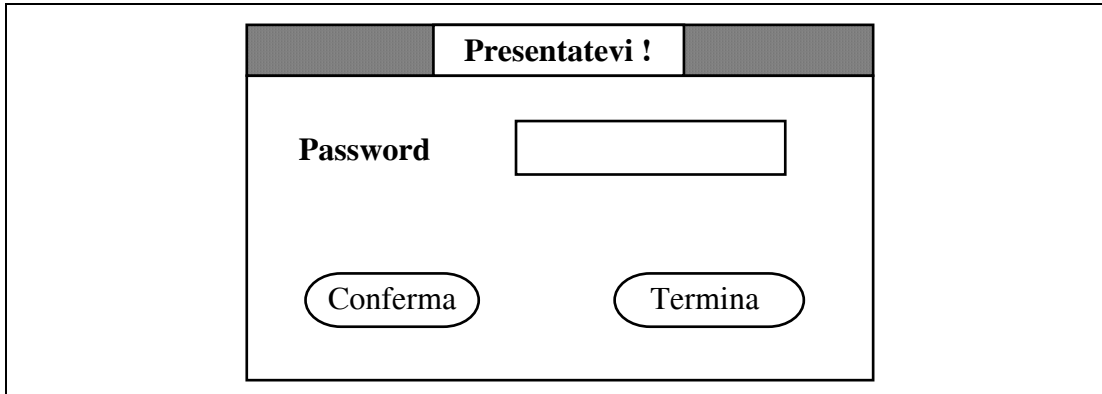


Fig. 9 - *La finestra di presentazione: l'utente digita la propria password*

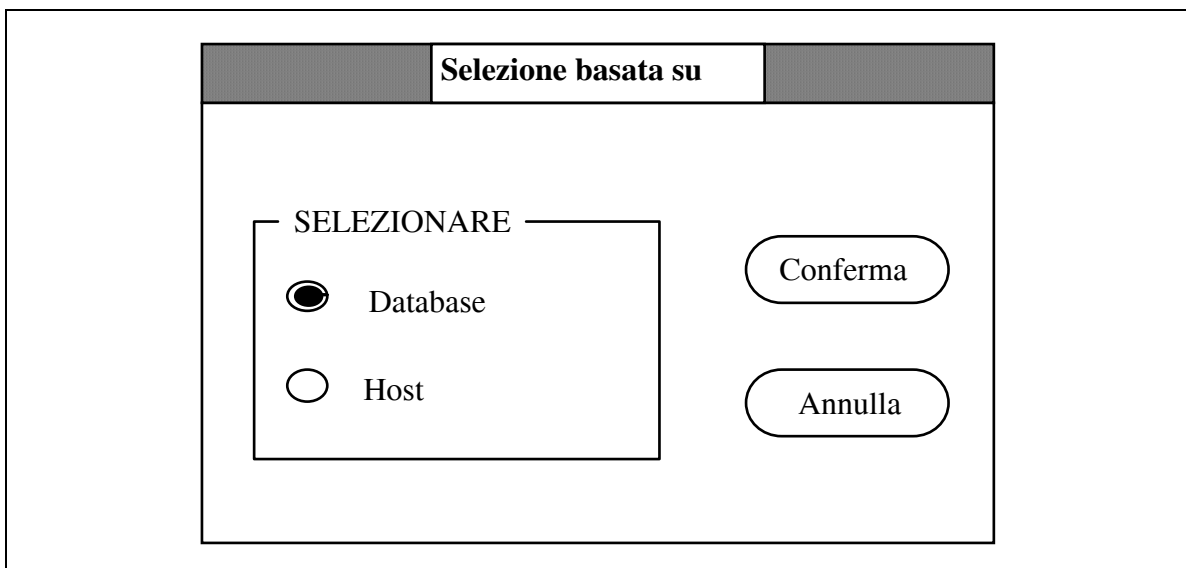


Fig. 10 - *La finestra per la scelta della priorità: o il particolare database, o il particolare host*

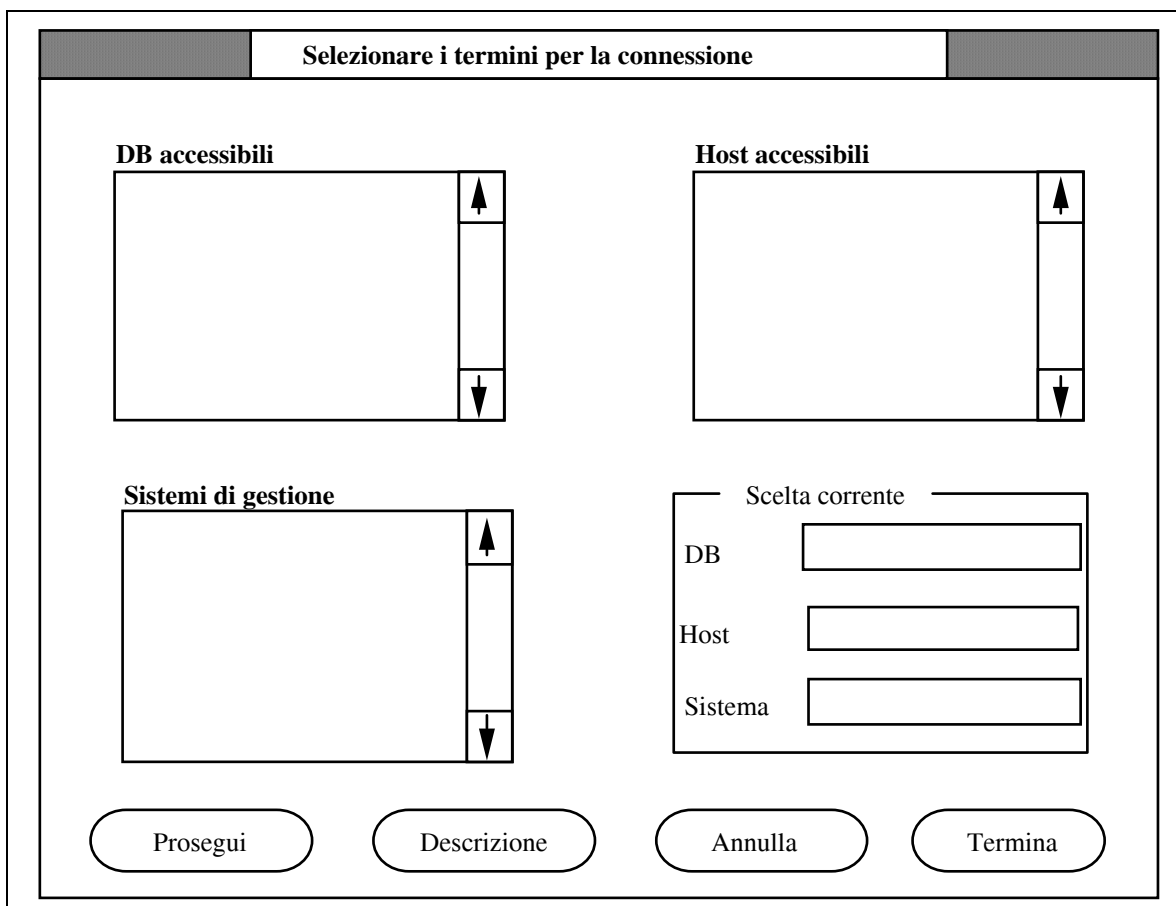


Fig. 11 - *La finestra per la scelta del database: essa varia a seconda del tipo di priorità indicato (qui è mostrato il caso in cui sia stata data priorità al database).*

5.2 - La formulazione della query

Nella fase di formulazione della query l'utente interagisce con una rappresentazione "simbolica" o "concettuale" del documento (Fig. 12), e non con la sua struttura come definita e utilizzata dal sistema di gestione.

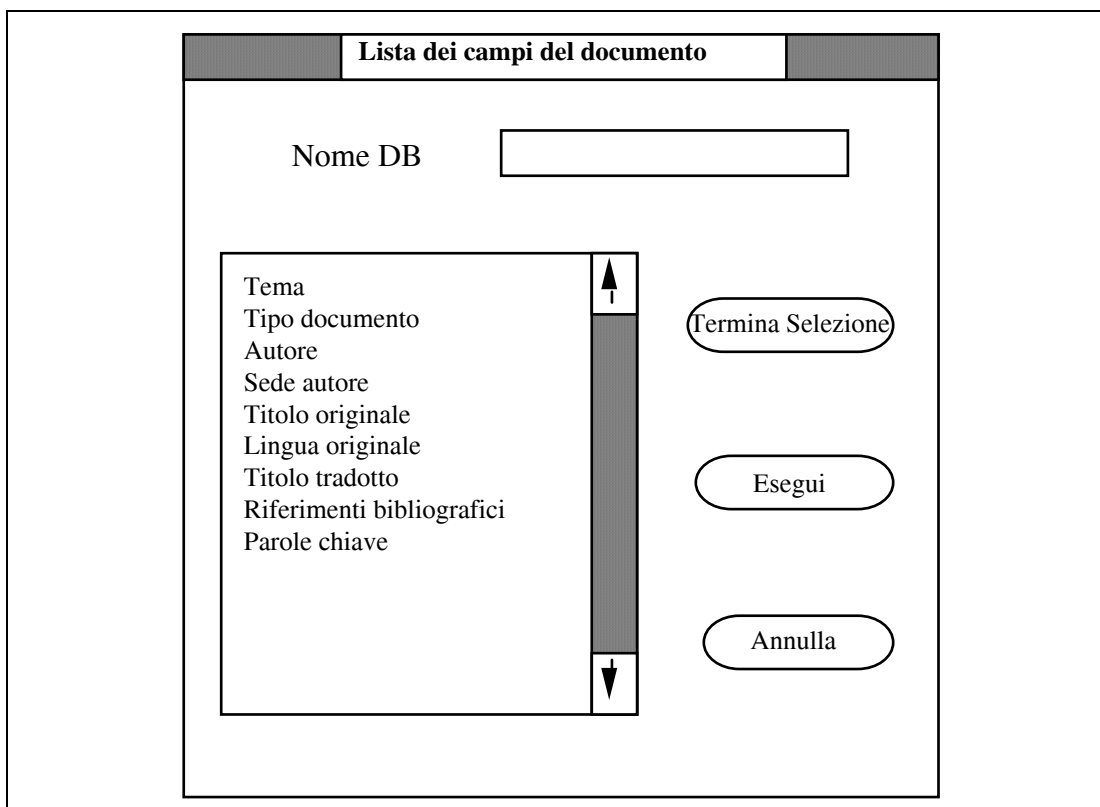


Fig. 12 - La finestra con il "documento virtuale"

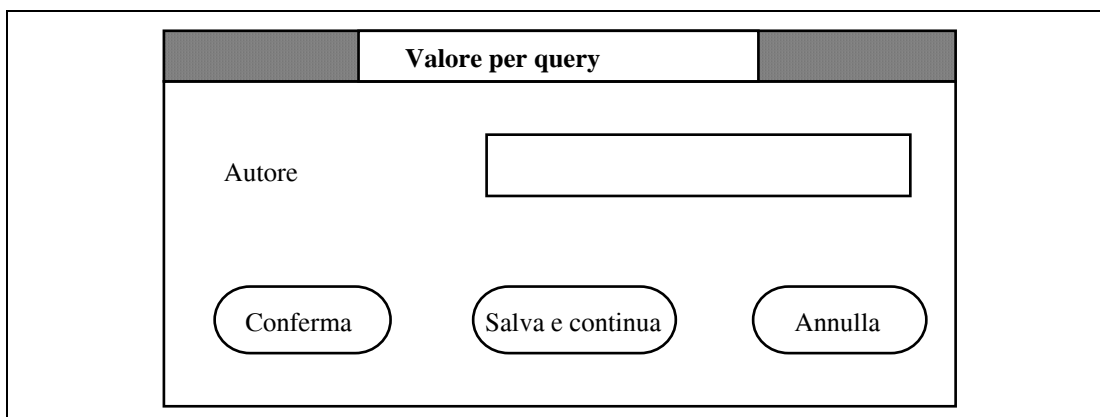


Fig. 13 - La finestra di query

Se il campo selezionato è un campo con dizionario o con thesaurus, potranno essere richiesti, da menù, rispettivamente i servizi del Dict Mgr, o del Thesaurus Manager. L'utente può sempre, comunque, formulare direttamente una query, immettendo direttamente i valori richiesti nella apposita finestra (Fig. 13).

In questo caso, i servizi offerti da Dict Mgr e Thesaurus Manager saranno semplicemente quelli di verificare la correttezza dei termini immessi.

La query costruita è sempre disponibile per verifica ed eventuali modifiche, prima di essere effettivamente inviata all' host per essere eseguita.

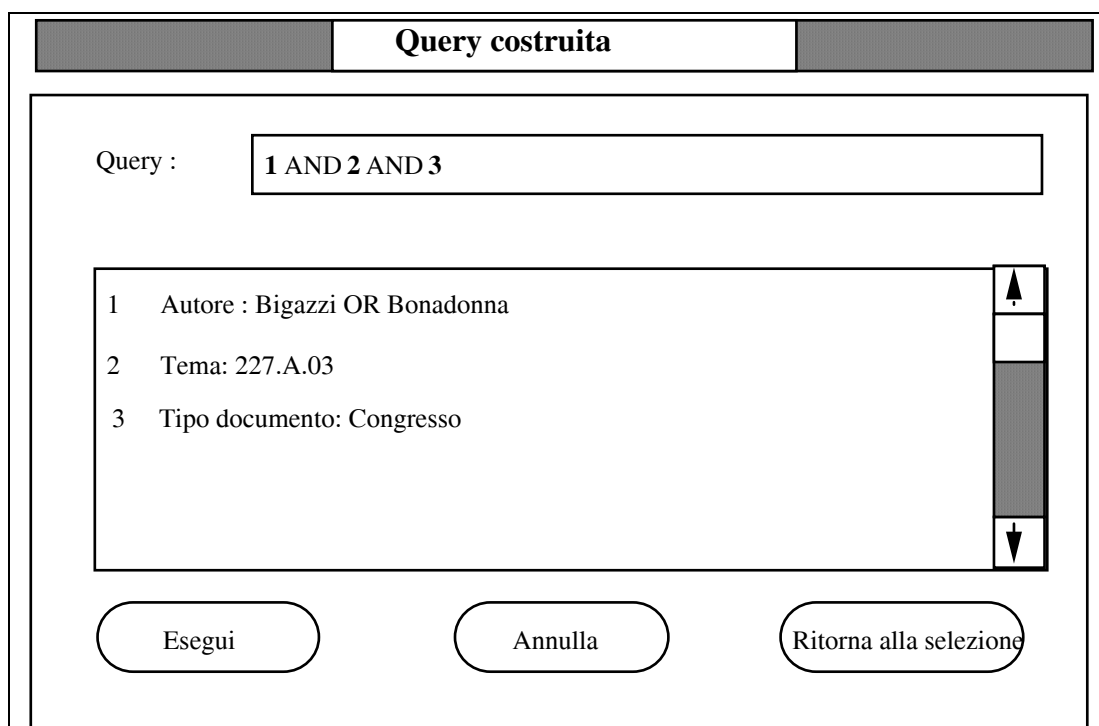


Fig. 14 - La finestra che mostra la query formulata

Se l'utente è soddisfatto della query finora formulata potrà chiederne l'esecuzione sull'host remoto¹, attraverso un click su "Esegui"; se invece ritiene che la query non sia ancora completa, ma necessita di termini aggiuntivi per esserlo, può premere su "Ritorna alla selezione" con cui riappare la lista dei campi del documento per dare l'opportunità di compiere le ulteriori selezioni. Può anche accadere che, pur essendo soddisfatto dei termini inseriti nella query, l'utente non lo sia del modo in cui essi sono combinati logicamente. Egli potrà allora editare direttamente nell'edit control la query che con precisione desidera compiere sul database remoto, ed uscire poi con la selezione di uno dei due push-button di cui sopra.

5.3 - Il Dict Manager

Uno schema del sottoinsieme di funzionalità fornite dal Dict Mgr è riportato in fig. 15.

¹ Questa funzionalità non è ancora stata implementata in modo completo.

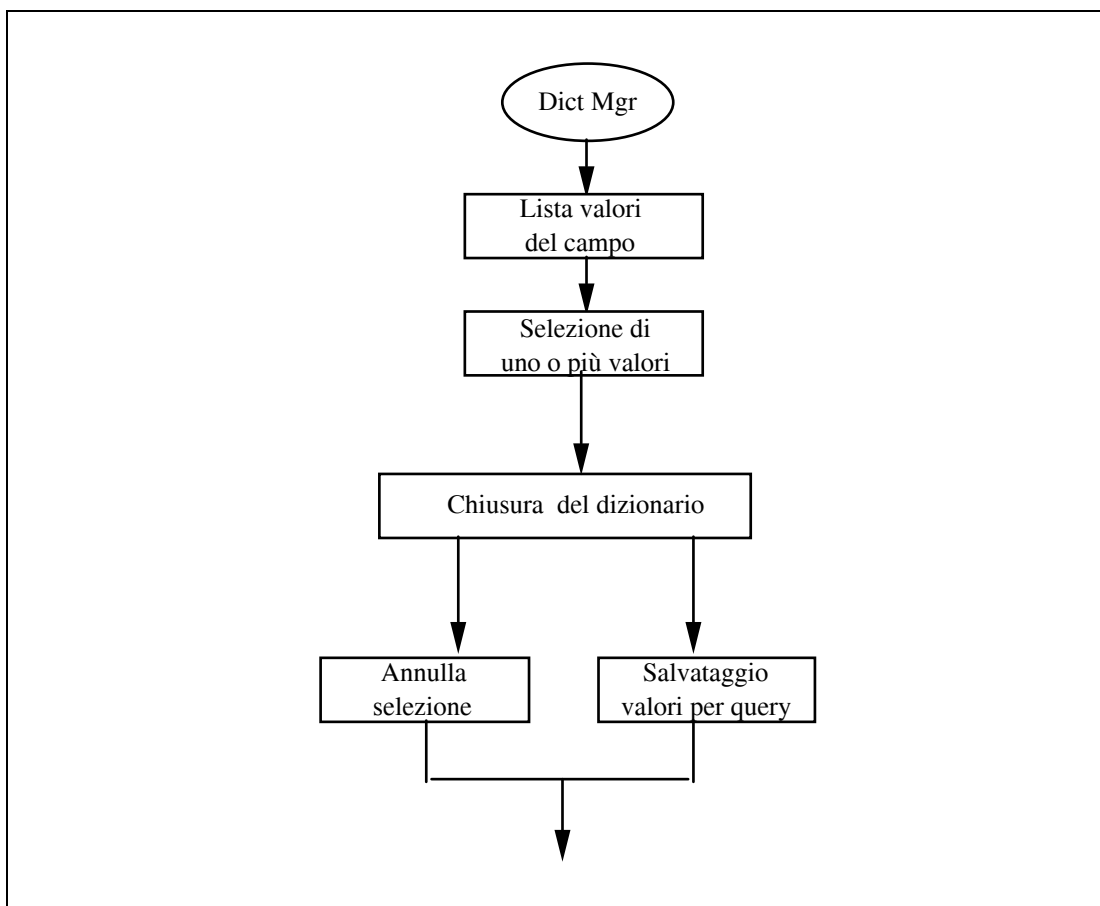


Fig. 15 - Le funzionalità del Dict Mgr

Il Dict Mgr mette a disposizione dell' utente una finestra con una lista dei possibili valori che il campo può assumere e gestisce le selezioni effettuate da tale lista. L'utente potrà anche effettuare una selezione multipla ed i termini correntemente selezionati sono quelli evidenziati in nero nella list box.

Ogni termine selezionato è messo in "OR" con quelli già scelti dalla lista.

5.4 - Il Thesaurus Manager

È il gruppo di funzionalità con cui si gestisce la ricerca dei termini di un thesaurus permettendo la navigazione sulla rappresentazione grafica della sua struttura logica. Le funzionalità realizzate per il Thesaurus Manager sono assolutamente generali, mentre la rappresentazione grafica ed il modo in cui è gestita la navigazione sono peculiari al particolare thesaurus.

Possiamo suddividere il dialogo tramite il Thesaurus Manager in tre fasi distinte:

- scelta dei relatori;
- scelta del termine iniziale;
- navigazione sul thesaurus.

La scelta dei relatori limita il numero dei legami "visibili" nel corso della particolare navigazione, in quanto i termini evidenziati saranno solo quelli connessi tra di loro dai relatori selezionati (e dai loro inversi).

In generale possiamo dire che un thesaurus è rappresentato graficamente da un grafo in cui i nodi sono termini del thesaurus e gli archi la relazione che li lega. Nel caso di thesauri a struttura gerarchica, la rappresentazione grafica sarà quella di un *albero* (fig.16), mentre nel caso di thesauri multi-albero si ha un display a *farfalla*.(fig. 17).

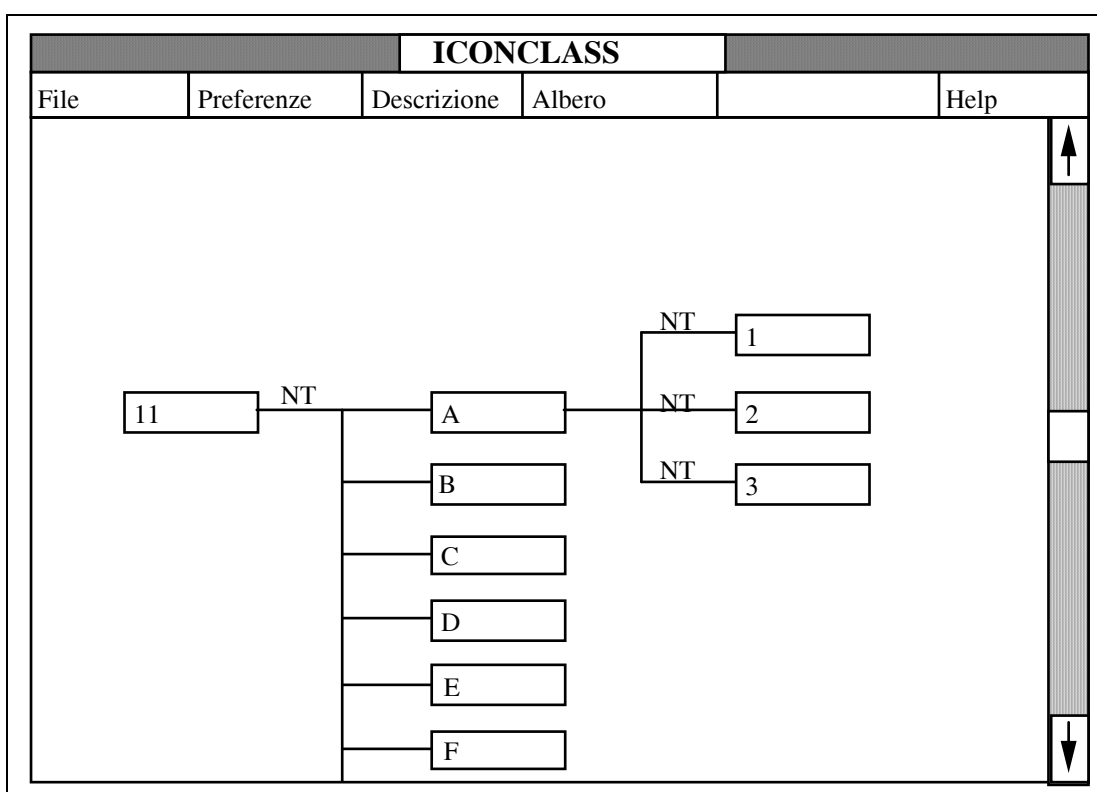


Fig. 16 - Un esempio di display ad albero

Nel caso del display ad albero, verranno evidenziati tre livelli: quello del nodo corrente, tutti i suoi "figli", il "padre".

Nel caso del display a farfalla, il livello centrale conterrà sempre uno ed un solo nodo, quello rispetto a cui si è realizzato il display. Il livello a sinistra contiene tutti i termini che hanno il termine centrale nel codominio, calcolato rispetto ai relatori selezionati. Sul livello di destra sono disposti tutti i termini che stanno nel codominio del termine centrale rispetto agli stessi relatori.

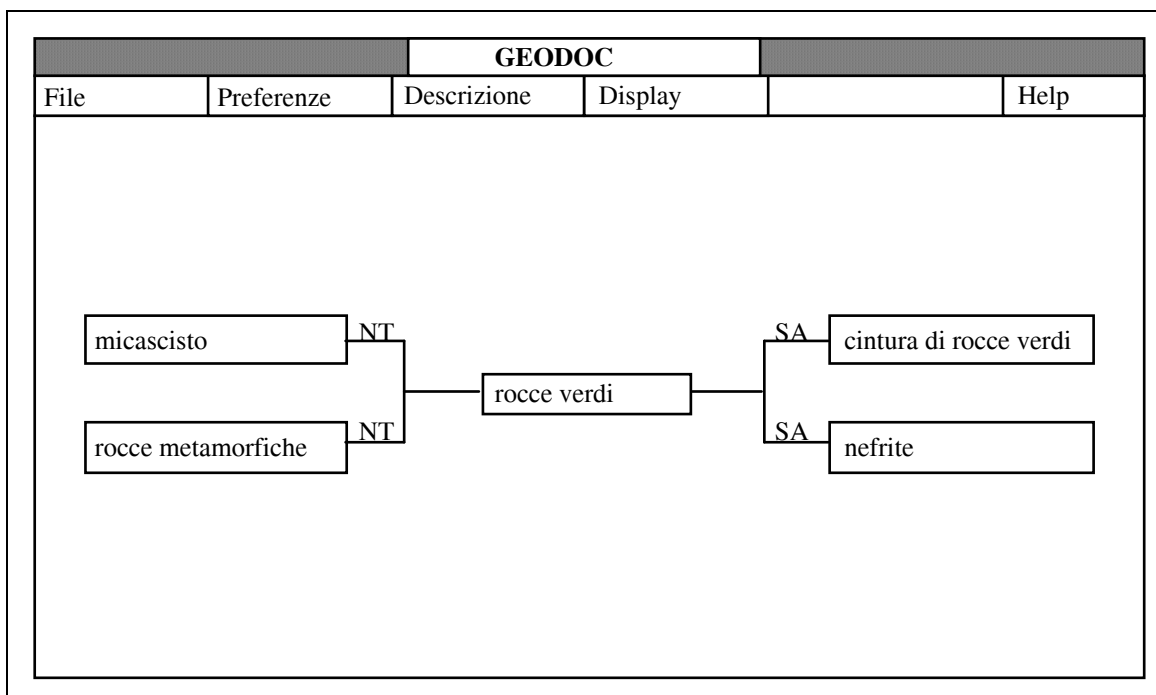


Fig. 17 - Un esempio di display “a farfalla”

In entrambi i casi, la modalità di interazione è del tipo “*point and click*” e le operazioni consentite sono le stesse. Sulla struttura visualizzata sono infatti consentite operazioni di scrolling verticale, di estensione verso termini più generici o, nel verso opposto, verso termini più specifici, e ritorno al display di partenza.

L' uso del colore rende più esplicito il significato dei vari nodi: il nodo corrente (rosso) è distinguibile da quello scelto come nodo iniziale (ciclamineo) e da tutti gli altri (verdi).

Alcune opzioni permettono di centrare il nodo attivo, o di ritornare al display iniziale dopo una navigazione complessa.

6 - Conclusioni

L' interfaccia realizzata permette all' utente non ancora esperto di ricevere informazioni sui campi del documento e sui termini del thesaurus prima di procedere alla formulazione della query, mentre l' utente esperto, che ha maggiore familiarità con la struttura dei campi, può evitare il passo preliminare. In ogni caso le interrogazioni, guidate dai menu e dalle finestre di dialogo, che contengono già tutte

le informazioni e necessitano di un input minimo da parte dell' utente, risultano assai semplificate.

Anche se gli strumenti utilizzati si sono dimostrati di notevole complessità (in particolare l' ambiente OS/2 e PM), sono stati conseguiti gli obiettivi che ci eravamo inizialmente posti nella realizzazione dell' interfaccia, e cioè:

- Fornire all' utente uno strumento che facilitasse le interrogazioni su un documento visualizzandone la rappresentazione concettuale.
- Consentire all' utente l' utilizzo di strumenti di consultazione al fine di facilitare la formulazione della query. Sono stati realizzati, a tale proposito, un gestore di dizionario e ed un gestore di thesaurus.
- Dare la possibilità di eseguire la navigazione sul thesaurus in parallelo con la fase di costruzione della query sugli altri campi. Ciò consente di sfruttare i tempi morti in cui l' utente dovrebbe attendere una risposta dal sistema.
- Realizzare il tutto conservando una totale indipendenza dai dati, tranne per il gestore del thesaurus, per il quale la dipendenza dai dati è intrinseca.
- Fornire una certa facilità d' uso, nonostante l' elevato numero di opzioni messe a disposizione, ed impedire all' utente di eseguire azioni da cui seguirebbe un comportamento dell' interfaccia diverso dalle aspettative.
- Far sì che l' applicazione realizzata rispettasse gli standard stabiliti da SAA per la portabilità del software, e avesse un aspetto esteriore conforme alla CUA.

I possibili sviluppi del lavoro possono orientarsi in diverse direzioni:

- **Correzione della query formulata.** È possibile per ora cambiare solo il modo in cui i token della query sono connessi logicamente. Si prevede però di consentire la correzione ad un livello più basso, al fine di poter modificare ognuno dei token che compongono la query.
- **Salvataggio della query per usi futuri.** Qualora l' host remoto sia momentaneamente non accessibile, poiché non attivo o per problemi nella rete, può tornare utile un salvataggio della query per dare la possibilità di eseguirla non appena siano soddisfatte le condizioni per la comunicazione, oppure per poter sottomettere la query periodicamente o in batch.
- **Modifica delle relazioni fra i termini del thesaurus.** Questo discorso entra a far parte di un discorso ancora più vasto, e cioè quello della manipolazione. In tale ambito rientrano le operazioni di aggiunta, cancellazione e modifica dei termini, modifica delle relazioni tra i termini semplicemente modificando la loro posizione nel grafo che visualizza la struttura logica del thesaurus.

Bibliografia

- [ANSI] American National Standard Institute : *Relational Database Language (SQL)*, (ANS X3.135 - 1986)
- [Bartoli82] Bartoli R., Romano G.A., Signore O.: *Implementation of Common Command Language on STAIRS/VS-TLS*, in *Deontic Logic, Computational Linguistics and Legal Information Systems* (A.A. Martino Ed.), North Holland (1982)
- [Blair85] Blair D.C., Maron M.E.: *An Evaluation of Retrieval Effectiveness for a Full-Text Document Retrieval System*, Comm. of ACM, Vol. 28, N. 3 (March 1985), pp. 289-299
- [Croft89] Croft W.B.: *Advanced Information retrieval techniques*, HYPERTEXT'89 Course Notes, Pittsburgh, 1989.
- [Frei83] Frei H.P., Jauslin J.F.: *Graphical presentation of information and services: a user-oriented interface*, Information Technology: Research and Development, N. 2, pp.23-42
- [Harris85] Harris S., Brightman H.: *Design implications of a task-driven approach to unstructured cognitive tasks in office work*, ACM TOIS, Vol. 3, N. 3 (July 1985), pp. 292-306
- [ISO86] International Standard ISO 2788, *Documentation Guidelines for the establishment and development of monolingual thesauri*, International Organization for Standardization, Svizzera (1986)
- [Monarch87] Monarch I, Carbonell J.: *CoalSORT: A Knowledge-Based Interface*, IEEE Expert (Spring 1987), pp.39-53
- [Negus77] Negus A.E.: *Euronet Guideline: Standard Commands for Retrieval Systems*, The Institution of Electrical Engineers, London (1977)
- [Negus80] Negus A.E., Snowden A.E.: *EURONET-DIANE User's Guide (Common Command Language)*, Scicon Consultancy International Ltd. (1980)
- [Pollitt84] Pollitt A.S.: *End user touch searching for cancer therapy literature-a rule based approach*, in *Proceedings of the Sixth Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, Vol.17, N.4, (June 1983), pp.136-145

- [Rada87] Rada R., Martin B.K.: *Augmenting Thesauri for Information Systems*, ACM TOIS Vol.5, N. 4 (October 1987) pp. 378-392
- [Salton83] Salton G., McGill M.J.: *Introduction to modern Information Retrieval*, McGraw-Hill, New York (1983)
- [Salton89] Salton G.: *Automatic Text processing*, Addison-Wesley (1989), ISBN 0-201-12227-8
- [Schreiber80] Schreiber F.A., Difilippo C., Zagolin M.: *Un preprocessor per l' interfacciamento di sistemi eterogenei di information retrieval su una rete di calcolatori*, Proceedings del convegno annuale AICA '80 (29-31 Ottobre 1980), pp.1245-1263
- [Signore90] Signore O., Bartoli R. : *Implementation of a historical-geographical database with support of imprecise dates*, DEXA 90: Proceedings of the International Conference in Vienna, Austria, 1990 (Tjoa A.M., Wagner R., Eds.), Springer Verlag, Wien-New York
- [Smith89] Smith P.J., Shute S.J., Galdes D., Chignell M.H.: *Knowledge-Based Search Tactics for an Intelligent Intermediary System*, ACM TOIS, Vol. 7, N. 3 (July 1989)
- [Thompson71] Thompson D.: *Interface design for an interactive information retrieval system: A literature survey and a research system description*, J. Am. Soc. Inf. Sci. (1971), pp. 361-373
- [Trigg86] Trigg R.H., Weiser M.: *TEXTNET: A network-based approach to text handling*, ACM TOIS, Vol. 4, N. 1 (January 1986), pp. 1-23
- [VanRijsbergen79] Van Rijsbergen C.J.: *Information retrieval*, Second edition, Butterwoths, London(1979)