

Una interfaccia per il browsing grafico di thesauri strutturati

R.Bartoli, A.Garibaldi*, M.Greco*, R.Potenza**, O.Signore*, B.Testa***

* CNUCE - CNR - via S. Maria, 36 - 56126 Pisa

** Centro di Studio per la Stratigrafia e Petrografia delle Alpi Centrali - via Botticelli, 23 - 20123 Milano

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 - Introduzione

La crescita del mercato dell'informazione online, per la quale erano state formulate ottimistiche previsioni all'inizio degli anni '80, è stata di fatto molto rallentata da vari fattori, alcuni di natura essenzialmente tecnologica (velocità e affidabilità delle reti di trasmissione dati, complessità delle procedure di logon, etc.), ma altri legati alle caratteristiche intrinseche dei sistemi di Information Retrieval.

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.

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. 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.

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 base di dati, in cui si definisce uno schema concettuale dei dati e una serie di operazioni su tale schema.

Come puntualizzato in [Smith89], alcuni utenti dei sistemi di reperimento delle informazioni hanno spesso bisogno di supporto per definire o raffinare gli argomenti di loro interesse, altri devono migliorare le loro conoscenze sull'argomento per decidere esattamente quali sono gli elementi di interesse, altri ancora hanno una idea chiara dei concetti da ricercare, ma hanno delle difficoltà ad esprimerli (in qualche caso anche in linguaggio naturale a persone che si incaricano di svolgere la ricerca). Altri, infine, 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).

In tutti questi casi, si rende necessario un raffinamento 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 tutti i sistemi di Information Retrieval più diffusi 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. Inoltre, si può ingenerare disorientamento nell' utente, che potrebbe aver bisogno di ritornare al punto iniziale per poter seguire un diverso cammino tra i concetti rappresentati nel thesaurus. Questo tipo di fenomeno è ben noto anche in altri campi, come gli ipertesti.

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

L' importanza di una appropriata interfaccia utente è stata sottolineata da parte di diversi autori nel corso degli ultimi anni. Una buona interfaccia utente deve mettere l' utente in grado di formulare ricerche significative mediante descrittori appropriati. Per realizzare un tale tipo di interfacce, deve essere possibile poter creare un meccanismo di comunicazione per i concetti chiave. Uno degli approcci possibili è quello di 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.

L' idea di realizzare il sistema Proteus è basata sul fatto che, come notato 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, ostano due fattori:

- le modalità richieste dai sistemi per formulare la query (la maggior parte dei sistemi richiede la formulazione della query mediante una notazione formale basata sulla logica booleana, meno efficace dei sistemi statistici, come riportato in [Salton83a]);
- la difficoltà per l'utente di definire in modo preciso il contenuto della query.

A questi due fattori, si aggiungono gli elementi già sottolineati in precedenza per quanto concerne l'utilizzo di sistemi tradizionali, disponibili sulle reti pubbliche:

- difficoltà di connessione;
- eterogeneità del linguaggio.

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. 1), 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.

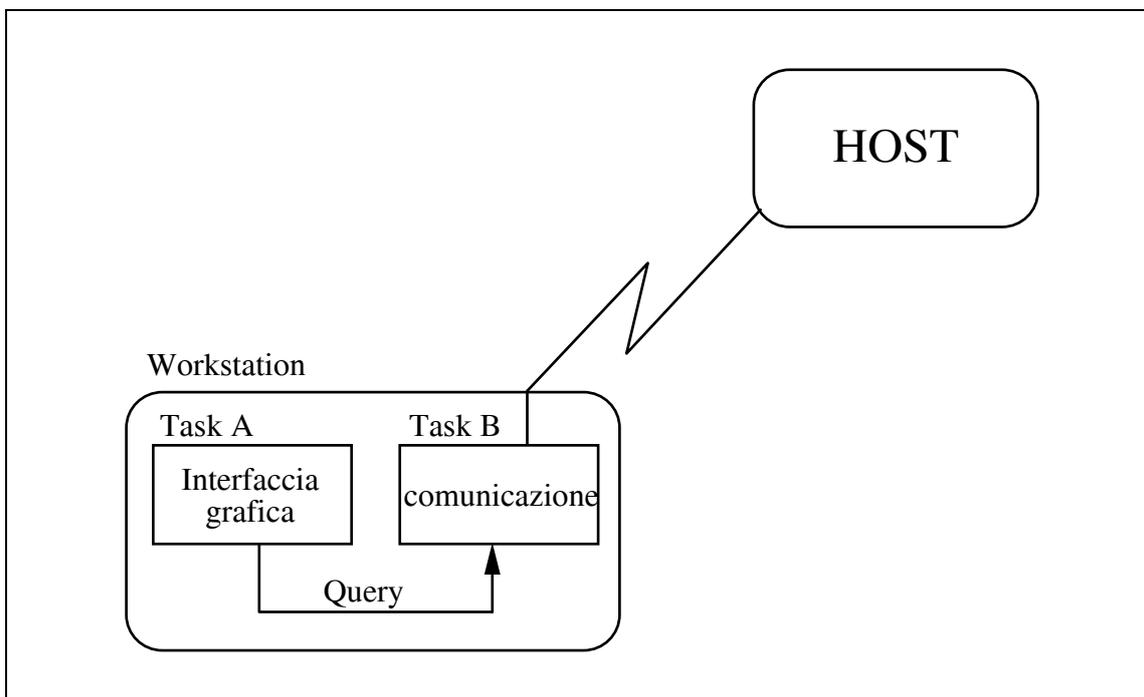


Fig. 1 - L'architettura generale di Proteus

Il primo passo è stato quindi la realizzazione di una interfaccia di tipo generalizzabile, adattabile a qualunque sistema di Information Retrieval (da qui il nome Proteus), 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.

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.

2 - L' architettura di Proteus

In questa sezione verranno descritte le caratteristiche implementative dell' applicazione realizzata. In particolare, verranno descritti il database di supporto utilizzato, e il database costituito dall' insieme di tabelle necessarie a modellare il thesaurus su cui si è realizzato un Thesaurus Manager.

2.1 - Il database di supporto

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.

2.2 - Il database che modella il thesaurus GEODOC

La struttura che definisce un thesaurus è traducibile in una base dati relazionale. Descrivere un thesaurus, quindi, è equivalente a mostrare il database che definisce i suoi termini e la relazione che li lega.

GEODOC è il thesaurus italiano corrispondente al sistema di Lessici e Thesauri prodotti nei Paesi della Commissione COGEODOC dell' IUGS (International Union of Geological Sciences) per la gestione e l' analisi dei documenti nell' ambito della Rete Internazionale di Documentazione Geologica([BRGM-CNRS-BGR74], [VSEGEI82], [CNR-BRGM-CNRS-BGR77], [Carimati84], [CNRS-BRGM86], [Shimomura89]). Un sottoinsieme di questi è stato armonizzato nelle sei lingue principali della Commissione, per costituire il Thesaurus multilingue IUGS, destinato a sviluppare un vocabolario geologico che possa essere utilizzato come riferimento terminologico comune da differenti centri di documentazione (Rassam88).

La struttura logica del thesaurus GEODOC è di tipo misto, poiché i termini sono raggruppati gerarchicamente in un limitato numero di temi (36), ma le parole chiave, divise in *Descrittore* e *Non Descrittore*, sono legate da relazioni diverse.

La base dati relazionale in cui è stato tradotto GEODOC è costituita dalle tabelle KEYWORDS e TEMI.

KEYWORDS	(T1, Rel, T2)
T1	parola chiave che identifica il termine.
Rel	relatore che lega due termini.
T2	parola chiave che identifica il termine legato (tramite il relatore riportato nel campo Rel) al termine rappresentato dalla parola chiave in T1.

TEMI	(Tema, Kw)
Tema	identifica il tema.
Kw	contiene la parola chiave.

Nella tabella KEYWORDS è implementata la relazione n:m tra T1 e T2, poiché ogni termine può essere legato, tramite un relatore, a termini diversi.

Esiste una relazione 1:n fra Tema e Kw nella tabella TEMI, nel senso che ad un tema possono essere associate più parole chiave, mentre una parola chiave è legata ad un solo tema.

Il thesaurus GEODOC è strutturato a multi-albero, per cui un singolo termine può avere come BT (Broader Term) più termini.

I relatori definiti sono: BT (e l' inverso NT), USE (e l' inverso UF), SA.

3 - Il funzionamento dell' interfaccia

L'insieme di funzionalità messe a disposizione può essere sostanzialmente suddiviso in tre gruppi, ognuno dei quali riassume una ben precisa fase dell' interazione con la base dati.

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

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.

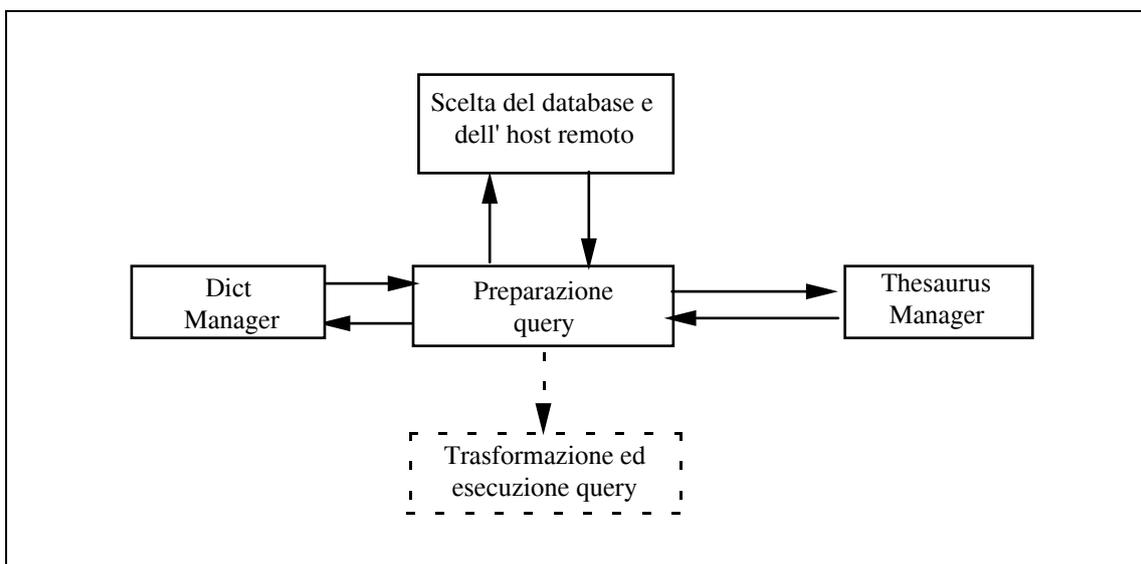


Fig. 2 - I gruppi di funzionalità messi a disposizione dall' interfaccia

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.

3.1 - La scelta del database e dell' host remoto

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

The image shows a dialog box with a title bar containing the text "Presentatevi!". Below the title bar, the word "Password" is displayed on the left side. To the right of "Password" is a rectangular text input field. At the bottom of the dialog box, there are two rounded rectangular buttons: "Conferma" on the left and "Termina" on the right.

Fig. 3 - La box di presentazione: l'utente digita la propria password

The image shows a dialog box with a title bar containing the text "Selezione basata su". Below the title bar, the word "SELEZIONARE" is displayed on the left side. To the right of "SELEZIONARE" is a rectangular area containing two radio button options: "Database" (with a filled radio button) and "Host" (with an empty radio button). To the right of this area, there are two rounded rectangular buttons: "Conferma" on top and "Annulla" on the bottom.

Fig. 4 - La box per la scelta della priorità: o il particolare database, o il particolare host

Selezionare i termini per la connessione

DB accessibili

Host accessibili

Sistemi di gestione

Scelta corrente

DB

Host

Sistema

Proseguì

Descrizione

Annulla

Termina

Fig. 5 - La box 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).

Lista dei campi del documento

Nome DB

- Tema
- Tipo documento
- Autore
- Sede autore
- Titolo originale
- Lingua originale
- Titolo tradotto
- Riferimenti bibliografici
- Parole chiave

Termina Selezione

Esegui

Annulla

Fig. 7 - La box con il "documento virtuale"

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 box (Fig. 8).

Valore per query

Autore

Conferma

Salva e continua

Annulla

Fig. 8 - La box di query

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.

Query costruita

Query : 1 AND 2 AND 3

- 1 Autore : Bigazzi OR Bonadonna
- 2 Tema: 227.A.03
- 3 Tipo documento: Congresso

Esegui Annulla Ritorna alla selezione

Fig. 9 - La box 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 verificarsi 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.

3.3 - Il Dict Manager

Uno schema del sottoinsieme di funzionalità fornite dal Dict Mgr è riportato in figura 10.

Il Dict Mgr mette a disposizione dell' utente una box con una lista dei possibili valori che il campo può assumere e gestisce le selezioni effettuate da tale lista. L'utente potrà

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

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.

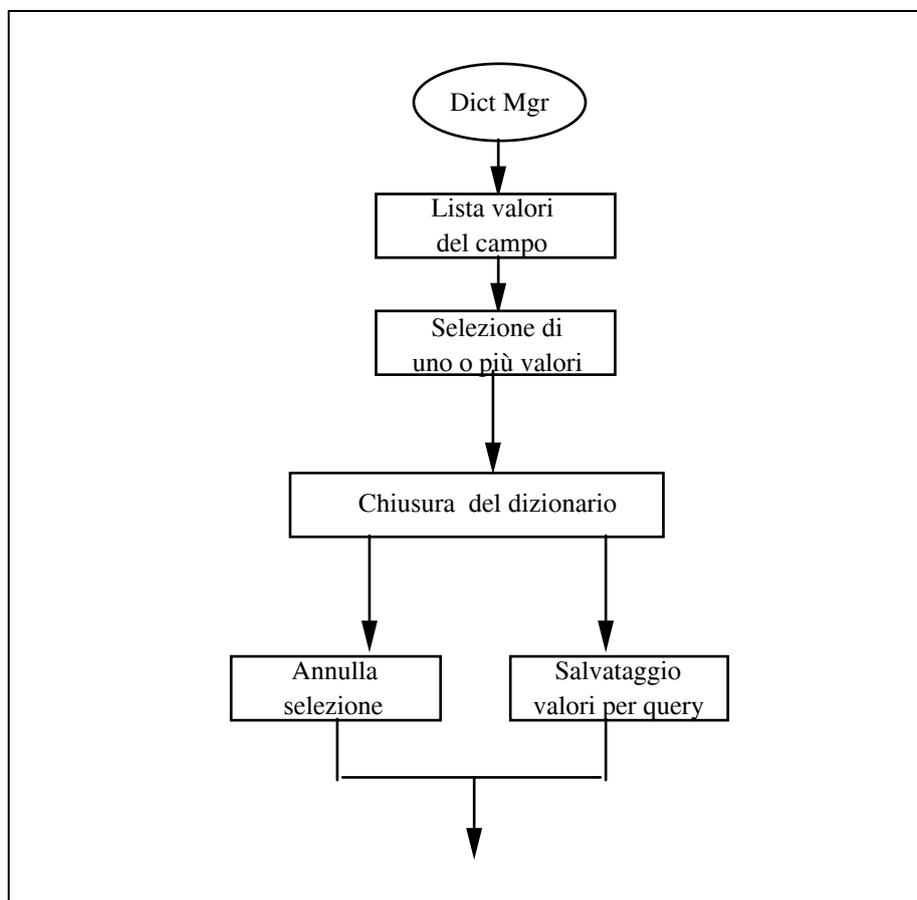


Fig. 10 - Le funzionalità del Dict Mgr

3.4 - 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.

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.

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

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

Nella fase in cui si sceglie il termine iniziale della navigazione, l'utente troverà le opzioni mostrate in figura:

File	Contenuti Navigazione	Help
	Parola Iniziale...	
	Tema Iniziale...	

L'opzione "**Parola Iniziale**" consente di effettuare la ricerca del termine iniziale (una parola chiave) partendo da una parola troncata o di inserire direttamente la parola iniziale.

L'opzione "**Tema Iniziale**" invece mostra una dialog box in cui va inserito un tema o un prefisso per tema cui fa seguito, rispettivamente, una lista di parole chiave associate al tema scelto o una lista di temi col prefisso editato in precedenza. Dalla lista di temi se ne sceglie poi uno con cui partire per la ricerca delle parole chiave.

La scelta dei relatori limita il numero "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).

Dato che GEODOC ha una struttura multialbero, si ha un *butterfly display* (display a farfalla), come quello mostrato in figura 11.

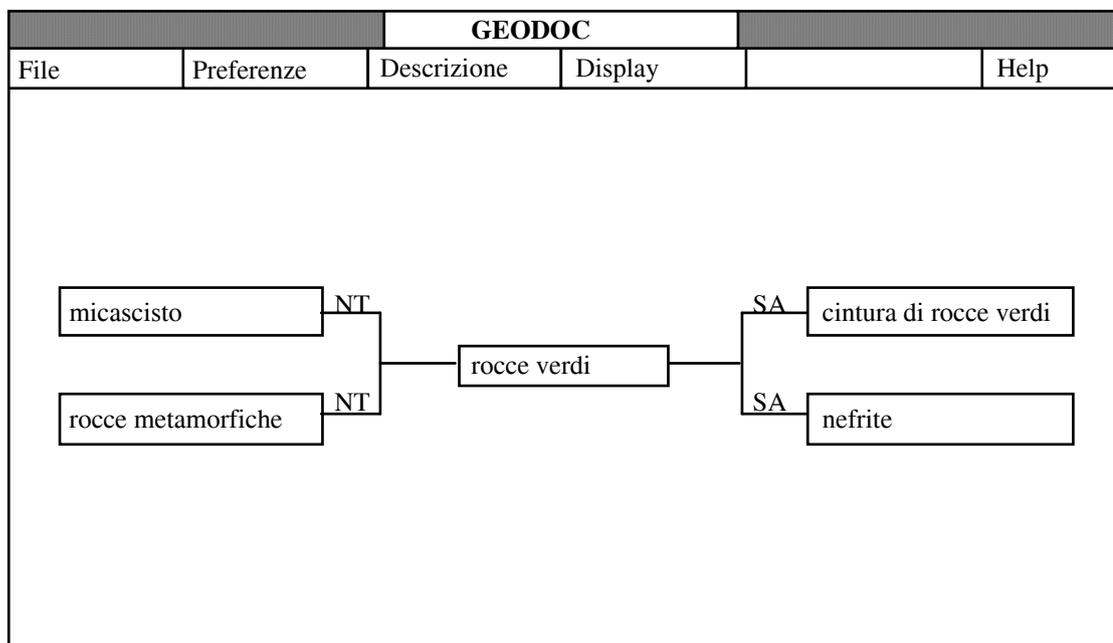


Fig. 11 - Un esempio di grafo per GEODOC

Tale figura mostra un intorno di tre livelli della parola chiave iniziale "rocce verdi".

Il livello centrale conterrà sempre uno ed un solo un 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.

Sulla struttura visualizzata sono 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.

4 - 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 dialog box, 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. Ciò potrà essere realizzato dando la possibilità, all'utente, di selezionare dalla lista il token da correggere e disporlo in un apposito edit control, per permettere di compiersi le opportune manipolazioni.
- **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. L' utente non dovrà così formulare nuovamente la query, ricordando le scelte fatte in precedenza, oppure potrà 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, cambio dei relatori durante la navigazione e modifica delle relazioni tra i termini, spostando 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
- [BRGM-CNRS-BGR74] *Sciences de la Terre - Thesaurus Franco-Allemand*, BRGM, Orléans, 1974, 126 pp.
- [Carimati84] carimati R., Potenza R., Testa B.: *Lessico Internazionale di Scienze della Terra, edizione italiana 1984*, CNR, Roma, (1984), 232 pp.
- [Chang88] Chang P.Y., E.F., Myre W.W.: *OS/2 EE Database Manager overview and technical highlights*, IBM Systems Journal, Vol. 27, No.2, 1988

- [CNR-BRGM-CNRS-BGR77] Manzoni M (ed.): *Thesaurus dei termini-chiave*, Patron Editore, Bologna (1977), 16 pp.
- [CNRS-BRGM86] *Pascal Geode Lexique français-anglais. Science de la Terre*, CNRS-CDST, Paris, 289 pp.
- [Croft89] Croft W.B.: *Advanced Information retrieval techniques*, HYPERTEXT'89 Course Notes, Pittsburgh, 1989.
- [Dror90] Dror A., Lafore R.: *OS/2™ Presentation Manager Programming Primer*, Osborne McGraw Hill (1990).ISBN 0-07-881467-7
- [Duncan89] Duncan R.: *Il multitasking di OS/2: sfruttare il protected mode dell'80286*. Microsoft Systems Journal Ed.It. (Gennaio 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
- [IBMSJ272] IBM Systems Journal, Vol. 27, No.2, 1988
- [IBMSJ273] IBM Systems Journal, Vol. 27, No.3, 1988
- [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)
- [Nierstrasz] Nierstrasz O.: *A Survey of Object-Oriented Concepts in Object-Oriented Concepts, Databases, and Applications*. (Kim W., Lochowsky F. H. Eds.), Addison Wesley, ISBN 0-201-14410-7
- [Petzold89] Petzold C.: *La tecnica multithread di OS/2 nelle applicazioni del Presentation Manager*. Microsoft Systems Journal Ed.It. (Febbraio 1989)
- [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

- [Provasi90] Provasi G., Pelagatti G., Colombo F.: *La nuova informatica per l'impresa degli anni '90*, Allegato al n. 101 di Data Manager, Fratelli Pini Editori, Marzo 1990
- [Rada87] Rada R., Martin B.K.: *Augmenting Thesauri for Information Systems*, ACM TOIS Vol.5, N. 4 (October 1987) pp. 378-392
- [Rassam88] Rassam G.N., Gravesteijn J., Potenza R.: *Multilingual Thesaurus of geosciences*, Pergamon Press, N.Y., LIII + 516 pp.
- [Salton83a] Salton G., Fox E., Wu U.: *Extended Boolean information retrieval*, Communications of the ACM, Vol. 26, N. 11 (Nov. 1983), pp.1022-1036
- [Salton83] Salton G., McGill M.J.: *Introduction to modern Information Retrieval*, McGraw-Hill, New York (1983)
- [Salton86] Salton G.: *Recent Trends in Automatic Information Retrieval*, 1986 ACM Conf. in Information Retrieval, Pisa
- [Salton89] Salton G.: *Automatic Text processing*, Addison-Wesley (1989), ISBN 0-201-12227-8
- [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
- [Shimomura89] Shimomura R.H.: *GeoRef Thesaurus and Guide to Indexing V Edition*, A.G.I., Alexandria VA, XII + 732 pp.
- [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)

- [VSEGEI82] *Mnogoyazychnyj Tezaurus po geologii (Russkij variant)*, VSEGEI, Leningrado, Vol. I 248 pp., Vol. II 274 pp., Vol. III 102 pp.
- [Wheeler88] Wheeler E.F., Ganek A.G.: *Introduction to System Application Architecture*, IBM Systems Journal, Vol. 27, No.3, 1988

Manuali di riferimento

- [API1] IBM Operating System/2™ Extended Edition Version 1.1 Technical Reference *Application Programming Interface Volume 1*, IBM Programming Family (Dicembre 1988)
- [API2] IBM Operating System/2™ Extended Edition Version 1.1 Technical Reference *Application Programming Interface Volume 2*, IBM Programming Family (Dicembre 1988)
- [API3] IBM Operating System/2™ Extended Edition Version 1.1 Technical Reference *Application Programming Interface Volume 3*, IBM Programming Family (Dicembre 1988).
- [CASEPM] *CASE:PM User Manual*, CASEWORKS inc., Atlanta, GA (1989)
- [C2CLR] IBM C/2™ Version1.1 *Compile, Link and Run*, IBM Programming Family (Settembre 1988), N.15F0383
- [C2Debug] IBM C2™ Version1.1 *Debug*, IBM Programming Family (Settembre 1988), N. 15F0377
- [C2Fund] IBM C/2™ Version1.1 *Fundamentals*, IBM Programming Family (Settembre 1988), N. 15F0384
- [C2LR] IBM C/2™ Version1.1 *Language Reference*, IBM Programming Family (Settembre 1988), N.15F0382
- [C2LS] IBM Operating System/2™ Extended Edition Version 1.1 Technical Reference *C/2 Language Syntax*, IBM Programming Family (Dicembre 1988).
- [CR] IBM Operating System/2™ Extended Edition Version 1.1 *Command Reference*, IBM Programming Family (Novembre 1988) N. 90X7664..
- [DMPG] IBM Operating System/2™ Extended Edition Version 1.1 *Database Manager Programming Guide and Reference*, IBM Programming Family (Novembre 1988).
- [DMSR] IBM Operating System/2™ Extended Edition Version 1.1 *Database Manager SQL Reference*, IBM Programming Family (Novembre 1988).
- [INF53] *Informatica Oggi*, No.53 (Settembre 1989)

- [**TKBEP**] IBM Operating System/2™ 1.1 Toolkit *Application Development : Building Executable Programs*, IBM Programming Family (Dicembre 1988), N. 00F8835
- [**TKPG**] IBM Operating System/2™ 1.1 Toolkit *Application Development : Programming Guide*, IBM Programming Family (Dicembre 1988), N. 00F8833
- [**TKPO**] IBM Operating System/2™ 1.1 Toolkit *Application Development : Programming Overview*, IBM Programming Family (Dicembre 1988), N. 00F8834
- [**UG**] IBM Operating System/2™ Extended Edition Version 1.1 *User's Guide and Reference*, IBM Programming Family (Novembre 1988) N. 90X7658