

Mathematics Subject Classification e schemi correlati nell'ambito di OAI¹

ANTONELLA DE ROBBIO² – DARIO MAGUOLO – ALBERTO MARINI

Questo lavoro vuole dare un'idea dei ruoli che le classificazioni di soggetto specifiche di area disciplinare possono giocare nel movimento Open Archives per la libera diffusione delle informazioni nelle attività di ricerca. La matematica e la Mathematics Subject Classification sono i fulcri attorno ai quali ci muoveremo per scoprire una varietà di modalità di presentazione, protocolli e strumenti per l'interoperabilità tra umani e macchine. La Open Archives Initiative (OAI) va intesa come un quadro di riferimento efficace per questo gioco. Nella prima parte del lavoro descriviamo gli schemi di classificazione tematici più importanti in matematica e nelle discipline collegate. Successivamente, accenniamo alla struttura degli schemi specifici di disciplina in funzione di un'agile consultazione (browsing) e illustriamo differenti modalità di browsing, implementate negli strumenti che abbiamo prodotto e raccolto ne La pagina delle classificazioni scientifiche. Infine, daremo uno sguardo alla definizione, all'implementazione e all'uso di un linguaggio di programmazione per la generazione di presentazioni ipertestuali per strutture complesse di dati. Nella seconda parte, elenchiamo differenti strategie per la comunicazione tramite e-print nella ricerca scientifica, per giungere alle definizioni di base dell'Open Archives Initiative. Una panoramica delle funzionalità attualmente utilizzate dagli archivi OAI compatibili, gestiti dal software Eprints, ci porterà poi a formulare alcune ipotesi di lavoro circa i ruoli che le classificazioni di soggetto per la matematica e le discipline collegate possono svolgere negli scenari del movimento Open Archives.

Parole chiave: Schemi di classificazione per la matematica – Open Archive – MSC Mathematics Subject Classification – Classificazioni di soggetto specifiche

Classificazioni di soggetto

Le classificazioni di soggetto sono strumenti primari per l'organizzazione della conoscenza e della terminologia nelle discipline scientifiche.

Sono prodotte principalmente da società professionali, da istituzioni accademiche o di ricerca, spesso per essere impiegate nelle basi di dati bibliografiche sviluppate dagli stessi enti. Anche se molti fra questi enti sono di ambito nazionale o regionale, le classificazioni di soggetto hanno generalmente un interesse internazionale e sono considerate un mezzo di comunicazione per la comunità scientifica internazionale.

Schemi per la matematica

Mathematics Subject Classification (MSC)³ è lo schema sviluppato dagli uffici editoriali delle due basi di dati bibliografiche più importanti del mondo per la ricerca matematica:

- *MathSci*, che è prodotto dall'American Mathematical Society;
- *Zentralblatt MATH*, che è prodotto dalla European Mathematical Society, dal Fachinformationszentrum (FIZ) di Karlsruhe (Germania) e da altre unità editoriali decentrate in Europa.

MSC copre tutti i rami della matematica pura ed applicata, compresi calcolo delle probabilità e statistica, analisi numerica, informatica, fisica ed economia matematica, la teoria dei sistemi e del controllo, le teorie dell'informazione e della comunicazione.

MSC ha subito nel tempo un certo numero di revisioni; l'ultima versione è entrata in vigore nel gennaio del 2000, ed è perciò denominata **MSC2000**.

Sul fronte della didattica della matematica, lo **Zentralblatt für Didaktik der Mathematik Classification Scheme**⁴ è usato per la base di dati bibliografica *MATHDI*, che è pubblicata dalla Società Matematica Europea, dal FIZ di Karlsruhe e dal Zentrum für Didaktik der Mathematik presso la Karlsruhe Universität, in collaborazione con Math Doc Cell (Francia).

Schemi per l'informatica, la fisica, il controllo e la tecnologia dell'informazione

Nel campo dell'informatica, compresi hardware, software, reti, la teoria, metodologie e applicazioni, lo strumento più importante è il **Computing Classification System (CCS)**⁵.

È sviluppato dall'Association for Computing Machinery (USA) per classificare gli articoli nei repertori *Computing Reviews* e *Guide to Computing Literature*, che sono pubblicati dallo stesso ente.

La sezione *68 – Computer Science* di MSC è stata definita in corrispondenza con una grande parte di CCS.

Nei campi della fisica teorica, sperimentale e applicata e in astronomia abbiamo **Physics and Astronomy Classification Scheme (PACS)**⁶. La sezione *02 – Metodi matematici nella fisica* di PACS assomiglia molto ai codici del primo livello di MSC relativi a matematica pura, calcolo delle probabilità e statistica. PACS è sviluppato e sottoposto a revisione, almeno ogni due anni, dall'American Institute of Physics.

Una versione di PACS viene adottata come Sezione A dell'**INSPEC Classification**⁷.

INSPEC è un servizio d'informazione bibliografica fornito dall'Institution of Electrical Engineers (Regno Unito). Riguarda fisica, ingegneria elettrica, elettronica, comunicazioni, ingegneria dei controlli, computer, informatica e tecnologia dell'informazione.

La **INSPEC Classification** ha altre tre sezioni importanti:

- Sezione B: Ingegneria elettrica e elettronica
- Sezione C: Calcolatori e controlli
- Sezione D: Tecnologia dell'informazione.

Schemi per l'economia

L'economia, nei suoi vari settori di attività, sia teorica che pratica, fa un uso sempre più esteso di strumenti matematici; le scienze matematiche, per converso, mutuano da quest'area elementi per la formulazione di problemi e teorie nuove.

Ciò risulta chiaro se si guarda al ruolo che la matematica ha assunto nel "**Journal of Economic Literature Classification System**"⁸, prodotto dall'American Economics Association, per quanto la sua rivista di indicizzazione e la corrispondente base di dati *EconLit* presentano.

Tali argomenti sono situati soprattutto nelle sezioni *62 – Statistica*, *90 – Ricerca operativa, programmazione matematica* e *91 – Teoria dei giochi, economia, scienze sociali e del comportamento* di MSC2000.

Schemi specifici di disciplina e schemi generali

Oltre a questi, esistono molti altri schemi di classificazione di soggetto utilizzabili in ogni disciplina o campo di discipline scientifiche.

Ulteriori schemi di classificazione sono quelli generali, non orientati verso discipline specifiche, come la **Classificazione Decimale Dewey**⁹.

Schemi di classificazione: dalla struttura al *browsing*

La struttura comune degli schemi di classificazione di soggetto

La struttura degli schemi di classificazione di soggetto, siano essi specifici di disciplina o generali, è essenzialmente la stessa: un sistema relazionale di *categorie*, identificate da codici alfanumerici, il cui significato è specificato da *descrizioni* o note esplicative in un certo linguaggio naturale (soprattutto, per la ricerca scientifica, l'inglese, anche se spesso sono disponibili traduzioni ed edizioni multilingui).

In generale, c'è una relazione principale, che nella maggior parte dei casi è a

forma di albero (relazione monogerarchica o, semplicemente, gerarchica) e le categorie sono denominate *nodi*. A volte, tuttavia, la relazione principale è un tipo più rilassato di ordine parziale, che permette a un nodo di avere più *padri*, ossia antecedenti immediati (perciò la relazione è chiamata multigerarchica).

Altre relazioni sono considerate come riferimenti incrociati, che rendono possibili collegamenti fra percorsi divergenti della relazione principale dello schema.

Gli schemi di classificazione di soggetto variano nel tempo attraverso versioni successive; generalmente una versione mantiene la sua validità per l'indicizzazione e la ricerca in una base di dati bibliografica per un periodo più o meno lungo di anni.

Due versioni successive possono essere messe in relazione collegando le categorie della versione più vecchia con quelle della versione più recente che mantengano una certa corrispondenza nel significato, anche se il collegamento può non essere di uno a uno, o non conservare la struttura nel passaggio fra le versioni, a causa di divisioni, fusioni, riorganizzazioni, morti e nascite dei soggetti, così come vengono rappresentati dalle categorie nelle due versioni.

Per esempio, **Mathematics Subject Classification** ha 5531 categorie in una gerarchia a tre livelli. Il livello superiore conta 63 nodi. I riferimenti incrociati, dotati spesso di testo esplicativo (“For...”), sono dei seguenti tipi: *vedi anche* – *vedi principalmente* – *vedi*.

Sono inoltre presenti alcune note per l'indicizzazione e la ricerca post-coordinata.

Physics and Astronomy Classification Scheme ha una gerarchia a quattro livelli. Il livello superiore conta 10 nodi, il secondo livello 66.

Dalla struttura al *browsing*

Per le loro caratteristiche strutturali, le classificazioni di soggetto sono strumenti efficienti per il *browsing* e la ricerca nei database bibliografici, nei cataloghi e in altri contenitori di metadati.

Inoltre, le classificazioni di soggetto possono dar vita a strumenti per l'organizzazione della conoscenza: per collezioni lessicali estratte da database di metadati o di documenti a testo completo; per terminologie, glossari, dizionari o enciclopedie, rassegne; per biblioteche distribuite di documenti digitali nativi o digitalizzati da supporto analogico.

L'insieme delle descrizioni di uno schema di classificazione costituisce già di per sé una importante risorsa terminologica.

Gli H-volumi ne *La pagina delle classificazioni scientifiche*

Differenti modalità di *browsing* possono essere ottenute attraverso diverse tecniche ipertestuali; noi abbiamo sviluppato degli strumenti con lo scopo di dimostrare alcune di queste modalità.

*La pagina delle classificazioni scientifiche*¹⁰ illustra e riunisce questi strumenti. È presentata sia in inglese che in italiano e comprende le seguenti sezioni:

- *La pagina della classificazione matematica*
- *Mathematics Subject Classification (MSC) e Classificazione Decimale Dewey (CDD)*
- *Classificazioni scientifiche di soggetto in relazione.*

Gli strumenti che abbiamo prodotto consistono di sistemi di pagine HTML sintatticamente semplici, ma altamente connesse e coordinate, chiamate *H-volumi*.

Gli H-volumi possono ammontare anche a migliaia di file, scritti in linguaggio HTML con semplici *routine* Javascript. Nel nostro ambiente di lavoro sono generati da un gruppo di programmi in linguaggio C standard, partendo dai file ASCII che presentano liste di record senza ridondanze e glossari riguardanti i valori degli attributi.

Gli H-volumi possono essere usati per visualizzare un complesso strutturato di informazioni di qualsiasi tipo, come repertori, collezioni biografiche, collezioni di metadati, database, glossari, dizionari, enciclopedie, ecc.

L'effettiva produzione degli H-volumi parte dai file ASCII ottenuti manipolando insieme di dati e testi esistenti, in particolare pagine Web già disponibili. Quest'attività di preparazione è svolta parzialmente "a mano" (cioè usando interattivamente un qualsiasi *editor* di testi flessibile), e in parte usando procedure di *text processing* sviluppate contestualmente alle procedure per la generazione delle pagine HTML.

Ritorniamo ora a esaminare in dettaglio le sezioni de *La pagina delle classificazioni scientifiche*.

La pagina della classificazione matematica

La pagina della classificazione matematica raccoglie sei presentazioni ipertestuali a *frame* dell'ultima versione della Mathematics Subject Classification (MSC2000).

Da un file ASCII contenente l'intera MSC2000, sono state ottenute una presentazione a *frame* semplice e una presentazione a doppia visione.

Il primo processo, per la generazione di una presentazione a *frame* semplice, è stato applicato a un file contenente una traduzione italiana di MSC2000, mentre, lavorando sui due file in combinazione, abbiamo ottenuto una presentazione a

frame semplice che mostra per ogni categoria le descrizioni in entrambe le lingue.

Da un file risultante dal confronto tra la versione di MSC del 2000 e quella del 1991 abbiamo ottenuto una presentazione a *frame* semplice che mostra i cambiamenti intervenuti.

Infine, dalla combinazione del primo file ASCII con un file che contiene dati su pagine specifiche per soggetto di siti Web rilevanti, abbiamo ottenuto una presentazione a *frame* semplice di pagine guida con *link* alle pagine specifiche individuate nei siti Web.

Nella Fig. 1 abbiamo innanzitutto un esempio di presentazione a *frame* semplice.

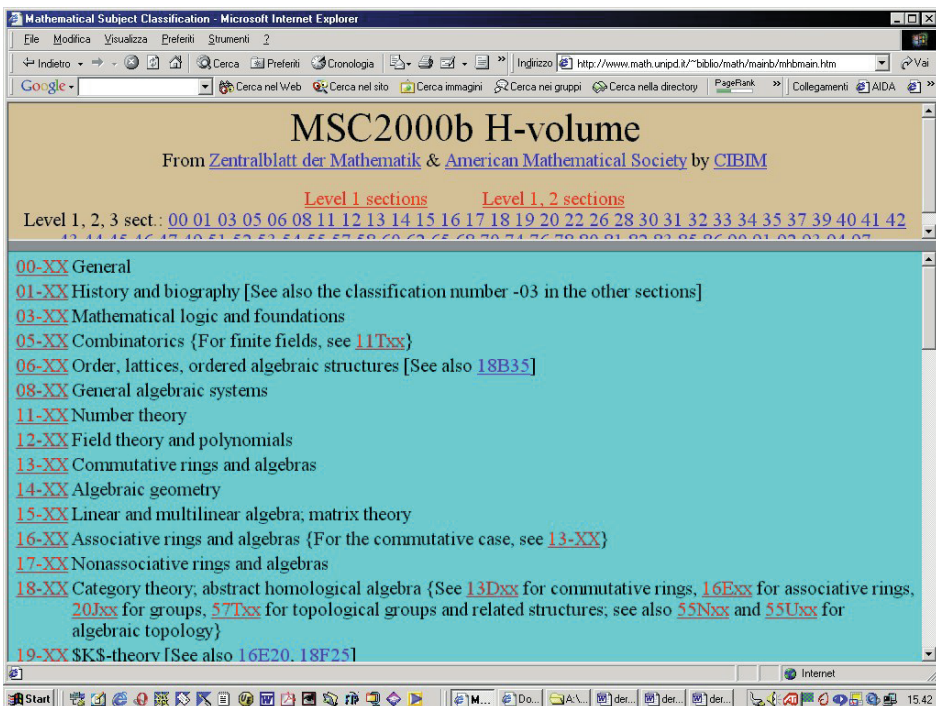


Fig. 1 – Presentazione a *frame* semplice

Le finestra superiore contiene una sorta di indice, che dà accesso a differenti tagli dello schema di classificazione: presentazioni a lista unica delle categorie di livello 1 e 1 – 2 e un insieme indicizzato di presentazioni a lista che copre l'intero schema. Per questo insieme, il rettangolo superiore mostra la lista delle prime due cifre dei codici delle 63 categorie del livello 1; ogni elemento nella lista punta ad una pagina che è mostrata nel riquadro sottostante, il quale contiene a sua volta una presentazione a lista del sottoalbero che soggiace all'indicata categoria di livello 1.

In tal modo, la lunga lista di tutte le categorie della classificazione è divisa in un certo numero di sottoliste, cosicché si possa scorrere velocemente lo schema delle classificazioni trasferendo soltanto file di dimensioni moderate.

D'altro canto, presentazioni a visione doppia o multipla possono essere utilizzate per navigare tra collegamenti trasversali, sia all'interno di una stessa versione dello schema di classificazione che attraverso schemi di diverse versioni. È quindi possibile muoversi avanti e indietro su tali versioni attraverso finestre parallele.

Ecco qui (Fig. 2) un esempio di presentazione a doppia visione che mostra le connessioni tra le categorie della Classificazione Decimale Dewey, 21^{ma} edizione, e di MSC2000.

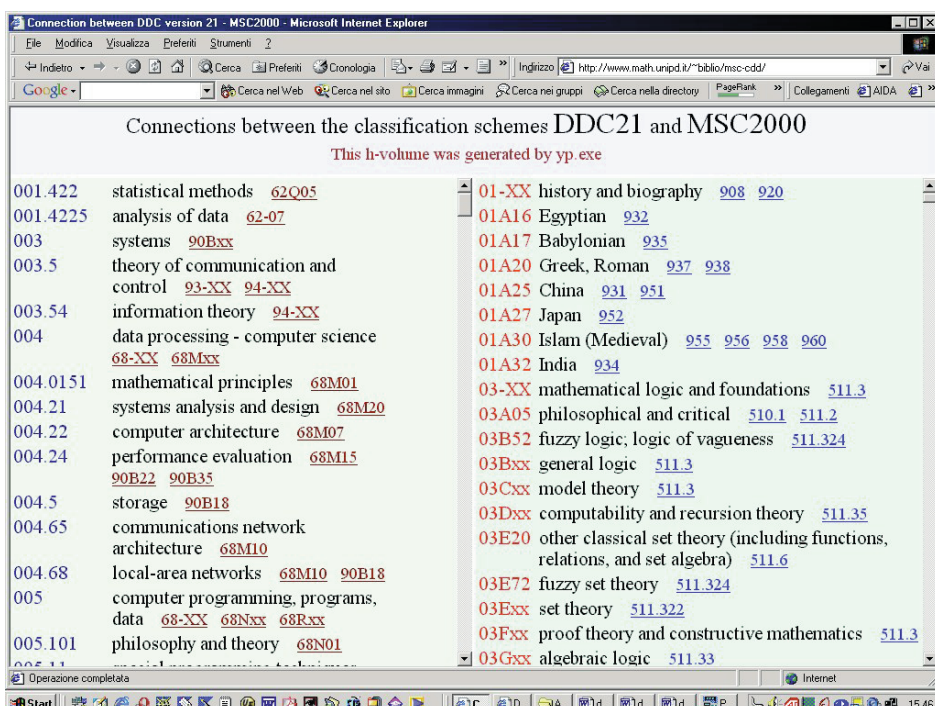


Fig. 2 – Presentazione a doppia visione

La presentazione a doppia visione contenuta ne *La pagina della classificazione matematica* è in effetti la duplicazione di una presentazione a *frame* semplice dell'albero di MSC2000 (con le descrizioni in inglese), che permette di navigare attraverso i riferimenti incrociati mantenendo la visibilità simultanea dei contesti in cui i riferimenti sono inseriti.

Mathematics Subject Classification (MSC) e Classificazione Decimale Dewey (CDD)

La sezione *Mathematics Subject Classification (MSC) e Classificazione Decimale Dewey (CDD)* de *La pagina delle classificazioni scientifiche* comprende due presentazioni in inglese:

- la pagina appena mostrata di connessioni tra categorie appartenenti alla CDD, 21^{ma} edizione, e a MSC2000
- un H-volume che presenta una lista KWIC ricavata dall'insieme combinato delle descrizioni di:
 - una proposta di revisione della sezione 510 – *Matematica* della CDD, presentata nel gennaio del 2001
 - MSC2000
 - Le sezioni E – N della classificazione ZDM, codificate come 97E – 97N nello stile di MSC.

Gli H-volumi di liste KWIC (come nella Fig. 3) sono definiti in funzione della scoperta di somiglianze testuali fra le descrizioni dei soggetti in uno o più schemi o versioni di schema di classificazione, per ottenere suggerimenti circa possibili affinità tra contenuti.

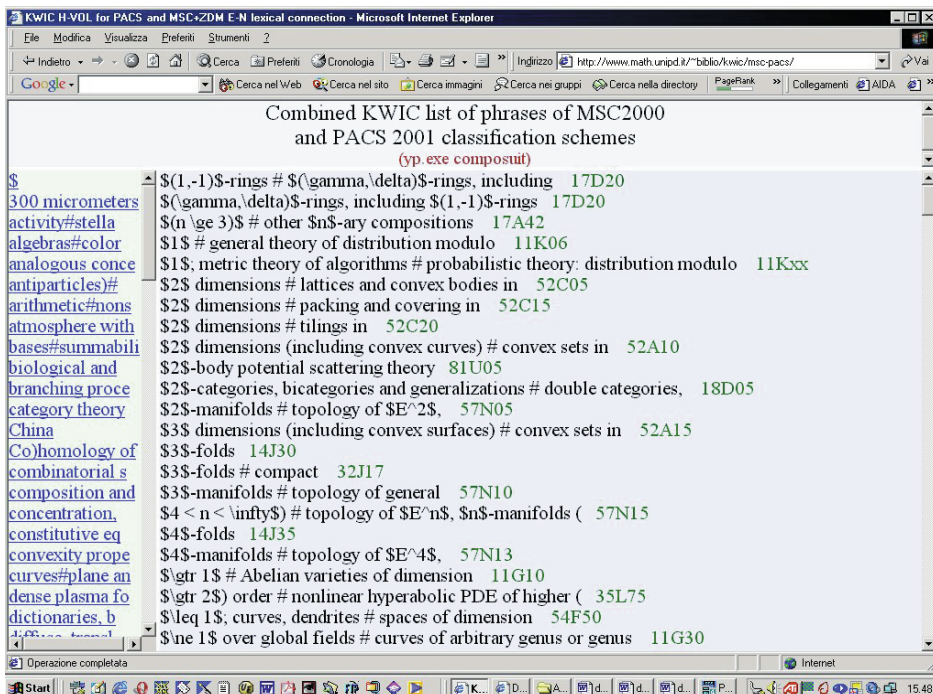


Fig. 3 – H-volumi delle liste KWIC

Una lista KWIC (KWIC è un acronimo per *KeyWords In Context*) presenta ogni descrizione mediante le sue permutazioni circolari che cominciano con una parola o con una sequenza di parole significativa; la lista di tutte queste permutazioni è ordinata secondo la lista delle parole o sequenze significative.

Con un metodo simile a quello impiegato per la presentazione a *frame* semplice, una lunga lista ordinata, com'è generalmente una lista KWIC, può essere dotata di una specie di distributore che permette di raggiungere rapidamente determinati punti o sezioni della sequenza. Un distributore può essere costruito a puntatori a lettere iniziali, a parole iniziali di sezioni impaginate, a sottoliste relative a particolari categorie di entità. La lista delle descrizioni permutate, suddivisa in liste più piccole e più gestibili, è visualizzata a destra, mentre il distributore compare nel *frame* di sinistra.

Le liste KWIC possono essere pensate e costruite non tanto ad uso dell'utente finale, quanto piuttosto come strumento ausiliario per il riconoscimento e la creazione giustificata di collegamenti strutturati all'interno di o fra gli schemi di classificazione. Questa attività, anche se può trarre notevole beneficio dalle tecniche automatizzate, richiede una quantità di conoscenza specifica che non può essere automatizzata, almeno con le attuali tecnologie. I collegamenti così individuati possono successivamente essere visualizzati attraverso presentazioni a visione multipla degli schemi di classificazione interessati.

Classificazioni scientifiche di soggetto in relazione

La sezione *Classificazioni scientifiche di soggetto in relazione* de *La pagina delle classificazioni scientifiche* contiene un insieme di presentazioni in lingua inglese (in un caso bilingue):

- una presentazione a doppia visione, che mostra collegamenti fra le categorie di ACM Computing Classification System (1998) e MSC2000;
- liste KWIC separate delle descrizioni di MSC2000, di PACS 2001, di ACM Computing Classification System (1998);
- liste KWIC combinate delle descrizioni di MSC2000 e PACS 2001, e di MSC2000 e ACM Computing Classification System (1998).

Verso un linguaggio per la generazione di presentazioni

Gli H-volumi che abbiamo creato sono considerati non già prodotti finali di riferimento, ma prototipi capaci di mettere in luce i problemi reali che si profilano nella costruzione di H-volumi più completi e professionali; essi permettono altresì, essendo strumenti di documentazione, verifiche di efficacia.

Lo sviluppo di tali prototipi, infatti, ha prodotto la specifica di meccanismi di parametrizzazione, strutture di dati e modi di elaborazione che hanno portato alla

definizione di un linguaggio di programmazione orientato alla manipolazione di presentazioni ipertestuali e alla visualizzazione di strutture matematiche.

Si è in effetti dato inizio alla definizione e all'implementazione di un linguaggio sperimentale denominato TAMP (Text Analysis, Manipulation and Presentation). TAMP è mirato all'analisi di file di testo di formato specificato (TeX, HTML, XML, ecc.), all'organizzazione di basi di conoscenza specifiche dotate di collegamenti ad altre risorse di Internet e alla loro presentazione attraverso pagine HTML.

Il linguaggio è implementato per mezzo di un unico programma in C, denominato YP, che legge e genera solo file di tipo *plain* ASCII.

Il primo file di *input*, caratterizzato dall'estensione .yppg, è il file sorgente del programma che dovrà eseguire l'elaborazione dei dati testuali. Successivamente vengono letti e scritti molti altri file a cui questo programma fa riferimento.

Tali file contengono dati o i sorgenti di programmi specificamente dedicati alla generazione di file HTML o di altri file pubblicabili (per esempio, file TeX), alla preparazione di file intermedi – per esempio, liste ordinate in base a criteri definiti e che raccolgono oggetti resi disponibili da file parziali non ordinati (in particolare, file estratti da pagine Web) – o al controllo di manipolazioni di alcuni tipi (in realtà, pochi) di strutture matematiche a partire da espressioni relativamente semplici di strutture di base, allo scopo di produrre presentazioni leggibili di strutture significative, possibilmente in un buon contesto di consultazione. L'implementazione è solo a una “versione minore di 1” ed è povera per molti aspetti, ma ha alcune caratteristiche che hanno permesso la produzione di pratiche pagine Web e che sembrano interessanti per la ricerca di ulteriori sviluppi.

Il linguaggio può controllare parecchi tipi di dati: quelli di base sono numeri interi (ma non ancora i reali) e stringhe; controlla complessi di dati di base come sequenze, tabelle e sequenze di sequenze. Inoltre è possibile manipolare alcune strutture specifiche di presentazione (liste KWIC di indicizzazione, glossari, ecc.) e le rappresentazioni di particolari strutture matematiche (permutazioni, partizioni, grafi, alberi, percorsi sul piano combinatorio, ecc.).

Mentre è fornita una buona scelta di operatori sui tipi di dati di base e sui loro aggregati, in realtà sono stati implementati solo pochi operatori che agiscono su strutture specifiche. D'altra parte, il programma interprete YP ha buone caratteristiche di estensibilità: i tipi di dati sono parametrizzati, semplici schemi permettono l'introduzione di identificatori e di caratteristiche funzionali generali di nuovi operatori. Le loro azioni possono essere implementate in procedure la cui collocazione e il cui ruolo non sono difficili da sintonizzare con le caratteristiche degli operatori esistenti.

Fra i più ricchi tipi di dati il linguaggio fornisce alcuni speciali generi di costrut-

tori, entità composite destinate alla costruzione di strutture di presentazione. Un esempio emblematico è fornito dal cosiddetto motore KWIC: la sua definizione richiede di specificare i campi di un file piatto, le procedure catalogate incaricate di distinguere ed accettare questi campi, le procedure catalogate incaricate di sviluppare i campi differenti di elementi KWIC in forma definitiva e i parametri richiesti da alcune procedure.

Dichiarazioni specifiche permettono di attivare i costruttori, dando modo di scegliere per essi parametri come schemi di controllo per la generazione dei file da generare e i prefissi dei loro nomi.

Una caratteristica importante del linguaggio è la possibilità di definire automi a livelli differenti di generalità. Gli automi del tipo più generale possono essere definiti da un ricco gergo specifico, che offre l'opportunità di determinare modelli efficaci degli accettori, trasduttori, analizzatori e generatori del testo, attraverso raffinamenti successivi.

Inoltre, il traduttore del linguaggio proposto può essere utilizzato con un pre-processore versatile che permette una buona estensione di sostituzioni, inserzioni, selezioni e iterazioni: le sue strutture di controllo possono agire su variabili che vengono assegnate a stringhe, numeri interi e file.

Questo pre-processore limita il reale svantaggio principale del linguaggio, vale a dire la mancanza di modularità. Un gruppo di dichiarazioni che sarebbe naturale incapsulare in un modulo può essere registrato in un file dotato di stringhe fittizie: questo file può così essere incluso in altri file sorgente, in quello principale o in un file che a sua volta può essere incluso in un altro file.

Il quadro di riferimento di OAI

La comunicazione tramite *E-print*: strumenti e architetture in rete

La produttività della ricerca scientifica è strettamente connessa alla veloce diffusione dei suoi risultati: perciò al lento processo formale della sottomissione di articoli a riviste si sono aggiunti altri, più veloci, metodi di disseminazione.

In un primo momento la disseminazione ha coinvolto documenti stampati, come i rapporti tecnici e gli atti informali dei congressi. Poi i ricercatori hanno iniziato a servirsi di Internet, mettendo i loro *paper* in siti ftp e, dall'apparire del Web, in siti Web. Questi contenitori sono comunque assai frammentati, con l'effetto che una ricerca sui loro contenuti risulta molto difficile; inoltre, non c'è garanzia che l'informazione resti archiviata dopo la conclusione del progetto scientifico che l'ha generata.

Con il passare del tempo sono state sviluppate diverse strategie per la comunica-

zione nella ricerca scientifica mediante documenti elettronici, o *e-print*, basate su:

- piccoli archivi specializzati;
- archivi centralizzati quali: arXiv¹¹ per fisica e discipline collegate, matematica, scienze non lineari, informatica; CogPrints¹² per scienze cognitive, intelligenza artificiale, linguistica computazionale, neuroscienze, ecc.;
- archivi istituzionali singoli o collegati, quali NCSTRL¹³ ed ERCIM Technical Reference Digital Library¹⁴ per informatica e matematica;
- reti distribuite collegate da un protocollo per l'interoperabilità, quali RePEc¹⁵ per l'economia e DoIS¹⁶ per la biblioteconomia e la scienza dell'informazione;
- server ad ombrello, quale MPRESS¹⁷ per la matematica;
- server collegati a gruppi di riviste o promossi da editori commerciali, ecc.

I servizi per la ricerca nel Web e l'immagazzinamento interno dei documenti trovati, come Researchindex (in precedenza CiteSeer)¹⁸, forniscono una soluzione che è stata apprezzata in special modo dai ricercatori nel campo dell'informatica. Gli *e-print* collocati in *homepage* personali vengono raccolti e immagazzinati senza alcuna particolare preoccupazione da parte degli autori circa i metadati; il servizio comprende altresì il collegamento fra i riferimenti bibliografici.

L'Open Archives Initiative

L'**Open Archives Initiative (OAI)**¹⁹ è uno sforzo internazionale per sviluppare standard di interoperabilità per la disseminazione di contenuti nella Rete.

OAI sottolinea la diversità tra l'essere un fornitore di dati (cioè, in sostanza, un editore) e l'essere un fornitore di servizi (cioè, interfaccia per ricerca, *browsing*, collegamenti di e da riferimenti bibliografici). D'altra parte, nulla impedisce che uno stesso sistema comprenda ed integri entrambe le funzioni. Per i ricercatori è persino possibile sviluppare archivi aperti personali, cui accedere per sviluppare siti Web personali e altri servizi, oppure da sottoporre a raccolta verso archivi di Dipartimento.

Il concetto base di OAI è la raccolta dei metadati, definita con l'*OAI Protocol for Metadata Harvesting*²⁰. In questo scenario non importa più *dove* i documenti sono archiviati; i metadati – compreso l'indirizzo in Internet – dei documenti presenti in qualsiasi archivio compatibile con OAI e registrato possono essere raccolti dai fornitori di servizi usando il protocollo di OAI in un "archivio virtuale" globale.

Gli archivi di autodeposito ad accettazione controllata compatibili con OAI: il software EPrints 2

EPrints è un software libero (General Public License) per la gestione degli archi-

vi di *e-print*, sviluppato all'Electronics and Computer Science Department dell'Università di Southampton (UK).

È rivolto ad organizzazioni e comunità piuttosto che a singoli individui. Fornisce un'interfaccia per gli amministratori di sistema, per i redattori incaricati di gestire l'accettazione di *e-print* nell'archivio, per gli autori che depositano i loro documenti e per gli utenti che vogliono accedere ai documenti mediante una ricerca sui metadati.

Il sistema viene configurato di *default* per gestire un archivio istituzionale di *e-print*, ma può essere riconfigurato con etichette e contenuti di campo per metadati completamente differenti.

Qualsiasi versione di EPrints è completamente interoperabile con la contemporanea versione dell'*OAI Protocol for Metadata Harvesting*.

Conclusioni

Il nostro lavoro è stato finalizzato alla definizione di metodologie per l'elaborazione di testi, funzionali allo sviluppo di presentazioni ipertestuali per strutture di documentazione complesse. Tali modalità di presentazione possono arricchire le funzionalità di *browsing* di archivi e fornitori di servizi nel quadro operativo di OAI, facendo in modo che una rete completa di collegamenti fra settori specifici possa guidare attività avanzate nella comunicazione di ricerca.

In particolare, stiamo studiando la possibilità di sviluppare il software EPrints con strumenti modellati su quelli sperimentali che abbiamo implementato per *La pagina delle classificazioni scientifiche*.

Facendo perno sulla **Mathematics Subject Classification**, si possono lanciare e attraversare ponti all'interno della matematica e fra le discipline che vivono e si sviluppano con la matematica. Ciò equivale a dire che si possono lanciare ponti tra tutti i rami della conoscenza scientifica e tecnologica, se siamo consapevoli delle dinamiche sempre più potenti e interconnesse con cui le discipline matematiche stanno facendo evolvere le attività modellistiche e computazionali in ogni campo della conoscenza umana.

Note

¹ Lavoro presentato all'International Congress of Mathematicians 2002 – Beijing (China), August 28-31, 2002 *EIC-Satellite Conference at Tsinghua University – Electronic Information and Communication in Mathematics*.

- 2 Per un contributo precedente nello stesso filone di ricerca, pubblicato su “AIDAinformazioni”, si veda Antonella De Robbio, *Banche dati per la matematica, cooperazione per il 2000*. “AIDAinformazioni”, 17 (1999), n. 3-4, p. 42-55 (liberamente accessibile su Web: <<http://www.aidainformazioni.it/pub/derobbio3499.html>>).
- 3 <<http://www.ams.org/msc/>>.
- 4 <<http://www.mathematik.uni-osnabrueck.de/projects/zdm/>>.
- 5 <<http://www.acm.org/class/1998/>>.
- 6 <<http://www.aip.org/pubservs/pacs.html>>.
- 7 <<http://www.iee.org.uk/publish/inspec/docs/classif.html>>.
- 8 <<http://www.aeaweb.org/journal/elclasjn.html>>.
- 9 <<http://www.oclc.org/dewey/products/index.htm>>.
- 10 <<http://www.math.unipd.it/~biblio/math/eng.htm>>.
- 11 <<http://arXiv.org>>.
- 12 <<http://cogprints.soton.ac.uk>>.
- 13 <<http://www.ncstrl.org>>.
- 14 <http://www.iei.pi.cnr.it/DELOS/EDL/ETRD_L_Con/>.
- 15 <<http://www.repec.org>>.
- 16 <<http://dois.mimas.ac.uk/>>.
- 17 <<http://mathnet.preprints.org>>.
- 18 <<http://citeseer.nj.nec.com>>.
- 19 <<http://www.openarchives.org>>.
- 20 <<http://www.openarchives.org/OAI/openarchivesprotocol.htm>>.

Riferimenti bibliografici

- Antonella De Robbio – Dario Maguolo – Alberto Marini, *Scientific and general subject classifications in the digital world*. “High Energy Physics Libraries Webzine”, issue 5, November 2001 <<http://doc.cern.ch/heplw/5/papers/4/>>
- Alberto Marini, *Text processing for presentation and manipulation of mathematical resources* presentato al Workshop *Electronic media in mathematics*, Coimbra (Portogallo), 13-15 settembre 2001 <<http://www.mat.uc.pt/EMM/index.html>>.