

THES/BID: un «thesaurus» per l'informatica giuridica e il diritto dell'informatica, redatto con l'ausilio dell'elaboratore elettronico

Costantino Ciampi
Elio Fameli

Sandro Ricci
Giuseppe Trivisonno

SOMMARIO: 1. *Gli scopi e le caratteristiche generali dell'opera realizzata.* — 2. *Le fasi operative per la redazione di THES/BID;* 2.1. *La raccolta e la selezione del materiale lessicale;* 2.2. *La normalizzazione;* 2.3. *La classificazione e la strutturazione;* 2.4. *L'organizzazione e la stampa automatica degli indici.* — 3. *I programmi elettronici per il controllo e l'espansione della struttura del Thesaurus e la redazione automatica degli indici;* 3.1. *Le liste di lavoro;* 3.2. *L'espansione automatica delle relazioni fra i termini e il controllo delle regole di costruzione del Thesaurus;* 3.3. *Le liste di stampa.*

La relazione qui di seguito pubblicata ha per oggetto la descrizione della metodologia e delle procedure di lavoro seguite nella costruzione d'un *thesaurus* settoriale, realizzato da un gruppo di ricercatori dell'Istituto per la documentazione giuridica del CNR e recentemente pubblicato dalla Casa editrice Le Monnier col titolo *THES/BID: A Computer-based Thesaurus of Terminology in Computers and the Law*; non verrà pertanto fatto riferimento né al dibattito teorico sui sistemi di classificazione e indicizzazione con riguardo all'uso dei *thesauri* (sui quali esiste ormai una vasta letteratura cui in questa sede si rinvia), né alle iniziative dei numerosi organismi internazionali operanti nel campo della gestione delle informazioni e della standardizzazione, che hanno portato all'individuazione di norme tecniche precise per la redazione di questi ormai indispensabili strumenti documentari.

1. GLI SCOPI E LE CARATTERISTICHE GENERALI DELL'OPERA REALIZZATA

THES/BID è un *thesaurus* strutturato dei descrittori (parole chiave e locuzioni) in lingua inglese impiegati nell'ambito della «Bibliografia internazionale d'informatica e diritto» (BID) e del corrispondente archivio interrogabile *on line*¹ per indicizzare e reperire il materiale bibliografico relativo al-

Gli Autori sono ricercatori del CNR presso l'Istituto per la documentazione giuridica di Firenze.

1. L'Istituto per la documentazione giuridica ha cominciato a raccogliere e pubblicare materiale bibliografico nel campo dell'informatica giuridica e del diritto dell'informatica fin dal 1971; l'at-

l'informatica, alle sue applicazioni nel campo del diritto e ai problemi giuridici connessi. L'opera rappresenta il primo tentativo organico di sistematizzazione generale dei concetti ricorrenti nel campo dell'informatica giuridica e del diritto dell'informatica e si propone quindi all'attenzione non solo dei documentalisti, dei lettori della «Bibliografia» e degli utenti di BID, ma anche di tutti coloro che sono comunque interessati a questa nuova disciplina.

THES/BID è stato compilato con l'ausilio d'un elaboratore elettronico, impiegando procedure automatiche di controllo e di stampa appositamente predisposte per agevolare il lavoro dei redattori. La descrizione delle procedure e dei programmi, concepiti secondo metodologie originali, è svolta in maniera analitica e dettagliata all'interno del § 3 di questa relazione; gli Autori hanno ritenuto infatti che l'esperienza maturata nella costruzione di THES/BID potesse essere proficuamente utilizzata nella redazione di altri *thesauri*, in settori disciplinari diversi.

Nella prima fase del lavoro, dalla letteratura specialistica analizzata per la «Bibliografia internazionale» in un arco di tempo ormai quasi decennale, sono stati selezionati, non soltanto sulla base della loro frequenza d'uso ma anche su quella della loro significatività, circa 2.000 descrittori; una volta normalizzati (secondo i criteri esposti nel § 2), questi sono stati numerati progressivamente e quindi classificati con l'apposizione di codici derivati da un apposito schema (*Classification Table*), utilizzato anche per la classificazione delle unità bibliografiche di BID. Questi dati sono stati raccolti in un archivio elettronico denominato MASTER, mentre in un archivio separato,

tuale «Bibliografia internazionale d'informatica e diritto» — che, come parte integrante della rivista «Informatica e diritto», esce dal 1975, a cura dell'Istituto e con la collaborazione di numerosi centri stranieri, presso la Casa editrice Le Monnier di Firenze — ha infatti raccolto l'esperienza di un'altra pubblicazione del genere, il «Bollettino bibliografico d'informatica generale e applicata al diritto», edito con periodicità trimestrale, sempre a cura dell'Istituto, negli anni 1972 e '73. Dal 1976, in base a precisi accordi di collaborazione, stipulati tra l'Istituto per la documentazione giuridica da una parte e la Corte di Cassazione e la Camera dei deputati dall'altra, i documenti raccolti nel «Bollettino» e nella «Bibliografia» sono reperibili *on line* sulla rete di terminali del Sistema Italgire e in quella collegata col Centro per la documentazione automatica della Camera. Attualmente la consistenza dell'archivio è di oltre 10.000 documenti, con un incremento medio annuo di 2.000 unità bibliografiche.

Sul «Bollettino bibliografico d'informatica generale e applicata al diritto» cfr. l'*Introduzione* di C. CIAMPI al n. 2, 1972 (I), del «Bollettino» stesso; sull'archivio BID cfr. C. CIAMPI, *La ricerca automatica dei documenti nell'archivio BID*, in «Informatica e diritto», 1977 (III), n. 1, pp. 168-190; per una presentazione aggiornata e analitica della «Bibliografia internazionale» e dello stesso archivio BID si veda E. FAMELI, R. NANNUCCI e R.M. DI GIORGI, *La documentazione nel campo dell'informatica giuridica e del diritto dell'informatica e la «Bibliografia internazionale d'informatica e diritto»*, relazione presentata al 3° Congresso internazionale organizzato dalla Corte di Cassazione sul tema «L'informatica giuridica e le comunità nazionali ed internazionali», Roma, 9-14 maggio 1983.

L'attuale edizione del *Thesaurus* è in lingua inglese. La favorevole accoglienza riscossa da THES/BID ha indotto gli Autori a curarne personalmente la traduzione italiana (che apparirà nel corso del 1983 nella collana della Giuffrè «Informatica e ordinamento giuridico») e a collaborare con alcuni gruppi di ricerca stranieri che si sono impegnati a tradurlo in altre lingue (francese, spagnolo, tedesco). Gli Autori attendono inoltre all'aggiornamento del testo già pubblicato, in modo da realizzarne una seconda edizione che sarà insieme plurilingue e adeguata al continuo sviluppo della terminologia del settore.

chiamato NOTE, sono stati via via registrati i termini non accettati e le annotazioni particolari (*Scope Notes*) relative all'uso di alcuni descrittori.

In una seconda fase del lavoro sono stati individuati i principali rapporti, di tipo verticale (gerarchie concettuali) e di tipo orizzontale (sinonimia, quasi-sinonimia e altre relazioni di tipo associativo), esistenti tra i diversi descrittori selezionati; sono stati inoltre fissati i rinvii dai termini esclusi a quelli accolti nel *Thesaurus* e i collegamenti tra alcune annotazioni particolari registrate in NOTE e i descrittori cui esse si riferiscono. Le relazioni tra i descrittori (compresi i rinvii e i collegamenti), dopo essere state codificate, sono state inserite in un archivio elettronico chiamato RELAT.

L'elaboratore elettronico, sulla base d'un insieme di programmi appositamente redatti², ha effettuato una prima serie di controlli sui tre archivi (MASTER, NOTE, RELAT), verificando l'esattezza formale delle classificazioni e delle codificazioni ed eliminando le eventuali duplicazioni di stringhe.

Successivamente, sulla base delle proprietà preventivamente definite per gli operatori relazionali, l'elaboratore ha sviluppato la rete di relazioni registrata in RELAT. Lo sviluppo delle proprietà di reciprocità, simmetria e transitività di alcuni operatori, da una parte, e il controllo, anch'esso automatico, della proprietà di irreflessività e di altre proprietà rilevanti delle strutture algebriche generate dalle reti di relazioni, dall'altra, hanno permesso d'ampliare notevolmente il numero dei rapporti originariamente definiti e, nel contempo, di depurare il *Thesaurus* dagli inevitabili errori. Particolarmente apprezzabile è stata l'espansione automatica delle relazioni: le oltre 4.000 relazioni inizialmente registrate in RELAT si sono più che raddoppiate mediante l'applicazione della macchina e i dati così ottenuti sono stati organizzati in un nuovo archivio, chiamato RELAT FILE ESPANSO.

Le relazioni gerarchiche verticali tra i termini (individuate dall'operatore BT, *Broader Term*, e dal suo reciproco NT, *Narrower Term*) sono state concepite come strutture algebriche alberiformi, le cui proprietà formali sono state sviluppate e controllate automaticamente. Per ogni insieme di catene gerarchiche aventi una stessa origine concettuale è stato costruito un albero con un'unica radice e numerosi nodi intermedi e finali; all'interno d'ogni albero s'è poi provveduto a controllare, con l'ausilio della macchina, che non vi fossero salti di gerarchia nel passaggio da un nodo all'altro (nelle catene BT/NT) e che ogni nodo avesse un solo superiore gerarchico.

Le relazioni orizzontali di sinonimia, quasi-sinonimia e associazione tra i termini (individuate dall'operatore RT, *Related Term*, e sviluppate automaticamente secondo le proprietà di simmetria e transitività unidirezionale) sono state invece concepite come strutture algebriche reticolari, non cicliche; anch'esse sono state controllate automaticamente, onde evitare errori logici e

2. Cfr. *infra*, § 3.

ridondanze e, sempre con l'ausilio della macchina, è stata verificata l'osservanza della regola d'incompatibilità tra la relazione di BT e quella di RT per coppie uguali di termini.

Dell'elaboratore ci s'è avvalsi, infine, anche per organizzare graficamente la stampa dei numerosi indici e liste di cui il *Thesaurus* è composto nell'attuale edizione.

Da quanto fin qui esposto — e si chiarirà meglio nel seguito — risulta che THES/BID ha caratteristiche particolari che lo distinguono da altri *thesauri*. In primo luogo, esso è costruito *a posteriori* sulla base dell'analisi lessicale svolta sui titoli e sugli *abstracts* delle unità bibliografiche selezionate per la «Bibliografia»; questa caratteristica lo oppone alla categoria dei *thesauri* costruiti *a priori*, prima cioè dell'organizzazione d'un sistema di documentazione nel particolare settore applicativo. In secondo luogo THES/BID è composto, oltre che da singole parole chiave, anche da descrittori pluritermini che individuano concetti complessi e articolati; questa scelta, se da un lato ha incrementato notevolmente il numero dei termini accolti, dall'altro ha permesso una maggiore precisione nella descrizione del contenuto dei documenti, limitando notevolmente i rischi legati alla post-coordinazione dei descrittori monotermini. In terzo luogo, la scelta metodologica fondamentale d'associare a ogni descrittore un numero d'identificazione ha consentito di rendere le varie procedure di controllo, elaborazione ed espansione indipendenti dalla forma linguistica impiegata; i vantaggi principali che ne derivano attongono al risparmio di memoria, alla velocità d'esecuzione dei programmi e alla possibilità di tradurre facilmente il *Thesaurus* in altre lingue.

Quanto alla struttura, THES/BID si presenta non come una semplice lista di termini ammessi (*authority list*), ma come un vocabolario strutturato di descrittori in cui sono individuate, oltre alle relazioni di preferenza, le relazioni gerarchiche e quelle di sinonimia, quasi-sinonimia e associazione. I descrittori così strutturati, poi, vengono presentati non solo in liste alfabetiche e in liste gerarchiche (ordinate dal concetto più ampio a quello più specifico), ma anche in liste organizzate per classi secondo l'ordinamento sistematico delle voci e sottovoci del particolare schema di classificazione adottato per BID. Quest'ultimo tipo di ordinamento rappresenta un'utile sintesi tra le tecniche dell'indicizzazione e quelle della classificazione, mettendo a disposizione dell'utente i vantaggi di entrambe e riducendo nel contempo i limiti intrinseci di ognuna.

Infine, l'impiego dell'elaboratore, liberando i redattori del *Thesaurus* da una serie di compiti puramente meccanici e ripetitivi (quali gli ordinamenti dei descrittori, i controlli formali nell'applicazione delle regole di costruzione ed espansione delle relazioni e l'organizzazione grafica degli indici), ha consentito loro di concentrarsi quasi esclusivamente sulla definizione dei criteri da seguire nella realizzazione dell'opera (selezione, normalizzazione e strutturazione dei termini) e sull'applicazione delle regole prefissate (proprietà delle

relazioni e delle strutture algebriche utilizzate, individuazione delle incompatibilità e delle ridondanze all'interno di queste ultime). Tale applicazione è stata particolarmente rigorosa e coerente in quanto s'è provveduto a formalizzare le regole e a tradurle in istruzioni di programma; l'elaboratore è stato così messo in grado di svolgere un vero e proprio lavoro di sviluppo delle relazioni, di controllo su di esse e di supporto nella costruzione delle strutture concettuali.

Considerati nel loro insieme, i numerosi programmi realizzati (22 programmi in linguaggio PL/1, per un totale di 2.440 istruzioni, cui vanno aggiunte 450 istruzioni raccolte in 14 «exec» funzionanti nell'ambito del sistema operativo CMS) sulla base delle regole formalizzate (40 regole, la cui applicazione è stata poi verificata su quasi 10.000 relazioni) possono essere valutati come il nucleo centrale d'un vero e proprio «sistema esperto» nel campo della costruzione di *thesauri*.

Intorno a questo nucleo, fin d'ora si può prevedere di sviluppare nuovi programmi destinati all'analisi semantica dei descrittori, al controllo delle ridondanze nell'uso della relazione associativa tra termini appartenenti ad alberi concettuali diversi, alla gestione delle varie fasi relative all'incremento della base lessicale e alla sua traduzione in altre lingue.

2. LE FASI OPERATIVE PER LA REDAZIONE DI THES/BID

2.1. La raccolta e la selezione del materiale lessicale

Il materiale lessicale di base utilizzato per la preparazione del *Thesaurus* è tratto, oltre che dalle voci dello schema di classificazione predisposto per BID, dalle parole e locuzioni utilizzate nei titoli e negli *abstracts* delle unità documentarie raccolte nelle bibliografie specialistiche curate dall'Istituto per la documentazione giuridica, in particolare nel «Bollettino bibliografico d'informatica generale e applicata al diritto» (1972, 1973) e nella «Bibliografia internazionale d'informatica e diritto» (1975-1982)³.

S'è provveduto, inoltre, a integrare questo materiale lessicale consultando una serie di *thesauri*, dizionari, vocabolari, glossari e soggetti relativi alle materie di specifico interesse (informatica, diritto, scienza dell'informazione e della documentazione, biblioteconomia, filosofia, cibernetica, logica, linguistica, matematica, scienze politiche, economiche e sociali, scienza dell'amministrazione, ecc.)⁴.

3. Sull'evoluzione delle tecniche bibliografiche utilizzate nel «Bollettino» prima e nella «Bibliografia internazionale» poi, cfr. la relazione di E. FAMELI, R. NANNUCCI e R.M. DI GIORGI, *La documentazione nel campo dell'informatica giuridica...*, cit., pp. 24-31.

4. Dei numerosi *thesauri* consultati si citano qui: *Thesaurus of Computing Terms*, 8ª ed., Manchester, National Computing Centre, 1977, 310 pp.; C.K. SCHULTZ, *Thesaurus of Information Science Terminology*, Metuchen (N.J.), London, The Scarecrow Press, 1978, 288 pp.; *A Thesau-*

S'è in tal modo ottenuta una base lessicale composta da alcune migliaia di termini; su di essa è stata svolta un'attenta selezione in base a criteri non soltanto sostanziali, attinenti al contenuto concettuale dei descrittivi stessi, ma anche tecnici o formali, in quanto collegati alle caratteristiche strutturali del *Thesaurus* come strumento documentario. Tra i primi rientrano il grado di significatività del termine nell'ambito della materia considerata, la sua esattezza terminologica, la relativa frequenza d'uso, la previsione sull'accettabilità da parte dell'utenza e la sua collegabilità con altri termini precedentemente inclusi nel *Thesaurus*.

Criteri di selezione formale sono stati applicati, invece, eliminando i descrittivi ridondanti in quanto esprimenti lo stesso concetto con forme solo morfologicamente o sintatticamente diverse (per esempio, è stato eliminato «*Information access*», mentre è stato conservato «*Access to information*»), individuando i sinonimi in senso stretto e rinviando dal termine non preferito a quello accolto (per esempio, individuati i due sinonimi «*Magnetic memory*» e «*Magnetic storage*», si è scelto d'indirizzare il lettore dal primo termine, non preferito, al secondo, invece accolto nel *Thesaurus*) e, infine, sostituendo — quando lo si è ritenuto opportuno — due o più termini con una locuzione che ne rappresenta il coordinamento (ad esempio, «*Antitrust regulation*» e «*Trade regulation*» sono stati coordinati nell'unico descrittore «*Antitrust and trade regulation*»).

Sempre nella fase di selezione sono stati poi isolati i nomi propri, distinguendo quelli geografici da quelli individuanti enti, istituti, associazioni, progetti, macchine, programmi, sistemi informativi, ecc.: i primi, dopo essere stati normalizzati, sono stati raccolti nell'apposito «*Elenco dei nomi geografici*», mentre i secondi sono stati registrati nella «*Lista delle sigle e abbreviazioni*» solo quando per essi era nota la sigla o l'abbreviazione corrispondente. Analogamente, in questa stessa lista, sono stati registrati i con-

rus for Informatics, Roma, Intergovernmental Bureau for Informatics (IBI), June, 1980, pp. xiv + 167 + 75, AIDS/8801. Per un'aggiornata bibliografia generale sui *thesauri* cfr. la *Bibliografia internazionale di Thesauri (International Bibliography of Thesauri)*, a cura di S. RIZZO, Roma, Camera dei deputati, novembre 1979, pp. VIII + 64.

Tra i vocabolari e dizionari d'informatica, particolarmente utile è risultata la consultazione del *Dizionario di informatica* di A. CHANDOR, Bologna, Zanichelli, 1972, 354 pp., del *Computer Dictionary Handbook* di C.J. SIPPL e C.P. SIPPL, Indianapolis (Indiana), Howard W. Sams & Co., 1972, 778 pp. e del *Dictionary of Data Processing (Including Applications in Industry, Administration and Business. English, Deutsch, Français)* di A. WITTMANN e J. KLOS, Amsterdam, Elsevier Scientific Publishing Co., 3^a ed. riv., 1977, 1^a rist. 1979, 347 pp.

Sulla terminologia giuridica, in particolare, si sono consultate le seguenti opere: E. JOWITT e C. WALSH, *Jowitt's Dictionary of English Law*, 2^a ed. a cura di J. Burke, voll. 2, London, Sweet and Maxwell, 1978; J.S. JAMES, *Stroud's Judicial Dictionary of Words and Phrases*, 4^a ed., voll. 5, London, Sweet and Maxwell, 1971-1974; *Words and Phrases Legally Defined*, 2^a ed. a cura di J.B. Saunders, voll. 6, London, Butterworths, 1978; R. HERBST, *Dictionary of Commercial, Financial and Legal Terms Comprising Trade and Industry (English, Deutsch, Français)*, voll. 3, Zug, Translegal Publishers, 1966-1975.

Per la terminologia relativa alle varie discipline considerate in THES/BID, oltre alle opere citate, sono stati utilizzati in generale i numerosi testi disponibili presso la biblioteca dell'Istituto per la documentazione giuridica.

cetti che nella letteratura tecnica si è soliti esprimere con le corrispondenti sigle o abbreviazioni (per esempio: KWOC = *Key Words Out Context*).

Al termine della selezione i descrittori si sono ridotti a circa duemila.

2.2. La normalizzazione

Una volta selezionato, il materiale lessicale raccolto è stato normalizzato in ordine sia alla registrazione nel *Thesaurus* (e quindi alle varie elaborazioni necessarie per la preparazione e la stampa dei numerosi indici che lo compongono), sia alle esigenze di reperimento automatico, mediante i descrittori accolti, delle unità bibliografiche nell'archivio in linea.

Il primo tipo di normalizzazione effettuato riguarda la scelta delle varianti grafiche delle parole, dipendenti dalle aree geografiche in cui esse sono usate: il criterio generalmente seguito è stato quello di preferire la forma inglese a quella americana (ad esempio, è stata preferita la forma inglese «*Labour law*» a quella americana «*Labor law*»).

Quanto all'uso dei segni d'interpunzione e dei segni speciali, questo è stato limitato al massimo al fine d'evitare problemi nell'ordinamento alfabetico delle liste: gli unici segni utilizzati sono la virgola, l'apostrofo (per il genitivo sassone), il trattino e le parentesi tonde. Di queste ultime, in particolare, ci si è avvalsi per specificare l'ambito concettuale del descrittore (per esempio: «*Data protection (International law aspects)*»).

Per la scelta del numero grammaticale (singolare o plurale) sono stati adottati criteri indicativi e orientamenti di massima desunti dagli usi linguistici della letteratura specialistica. Così, la forma singolare è stata preferita per i descrittori che indicano concetti astratti, come qualità («*Economic efficiency of computer systems*»), facoltà, attività («*Administration*», «*Automation*»), scienze, discipline e campi di studio («*Philosophy*», «*Logic*», «*Architecture*»). Ugualmente è stato usato il singolare per i nomi collettivi, le qualifiche professionali («*Programmer*», «*Operator*») e gli organi costituzionali dello Stato («*Parliament*», «*Government*»), oltre che per designare un'intera classe di elementi («*Hospital*», «*School*», «*Income tax*»).

S'è invece preferito il plurale per i termini che individuano macchine, strumenti, oggetti fisici («*Computers*», «*Terminals*»), che designano i soggetti esercitanti le professioni giuridiche («*Judges*», «*Lawyers*») o che riguardano le ripartizioni amministrative («*Regions*», «*Counties*»).

Come descrittori, inoltre, sono stati usati solo sostantivi e locuzioni sostantivali; non sono stati quindi utilizzati isolatamente verbi, aggettivi e avverbi.

Nei descrittori pluritermini, infine, s'è provveduto a normalizzare la sequenza delle parole e l'uso di congiunzioni e preposizioni.

2.3. La classificazione e la strutturazione

Una volta effettuata la normalizzazione, a ogni descrittore è stato assegnato un numero progressivo d'identificazione, la cui prima cifra indica la provenienza del descrittore stesso: la cifra '0' rivela che si tratta d'un termine non accettato nel *Thesaurus*, da cui vien fatto rinvio a un descrittore accolto, di significato affine; la cifra '1' segnala che la parola chiave è stata estratta dallo schema di classificazione; infine, la cifra '2' individua i termini provenienti da materiale lessicale registrato nella «Bibliografia» o acquisito dalle opere di consultazione utilizzate. L'attribuzione del numero d'identificazione a ciascun descrittore ha consentito d'effettuare tutte le elaborazioni relative al controllo, alla strutturazione e agli ordinamenti dei termini, operando su numeri e non su stringhe alfabetiche, col notevole vantaggio d'occupare minor spazio in memoria e di rendere più veloce l'esecuzione dei programmi.

Successivamente, i descrittori sono stati classificati in base ai codici dello schema di classificazione adottato per BID. Sovente è stato necessario attribuire più codici al medesimo descrittore, sia per coprire interamente la sua area semantica, sia per esprimere i molteplici aspetti sotto cui esso può essere esaminato. Il raggruppamento dei descrittori per classi ha reso più agevole il successivo lavoro di strutturazione, cioè d'individuazione dei rapporti semantici fra i termini, in quanto la creazione di classi di descrittori concettualmente omogenei è valsa di per sé a delimitare l'ambito di ricerca delle relazioni.

Com'è ormai prassi generalmente seguita nella realizzazione dei *thesauri*, la strutturazione dei descrittori è stata basata su tre tipi di relazioni, cioè sulle relazioni di preferenza, su quelle di gerarchia e su quelle di associazione.

Le relazioni di preferenza sono indicate con l'operatore US (*Use* = usa) e col suo reciproco UF (*Used For* = usato per) e servono a rinviare da un non-descrittore a un descrittore e viceversa (per esempio: «*Cadastre*» US «*Real property registers*»; «*Real property registers*» UF «*Cadastre*»). Le relazioni di gerarchia vengono segnalate dall'operatore BT (*Broader Term* = termine più ampio) e dal suo reciproco NT (*Narrower Term* = termine più ristretto) e collegano verticalmente tra loro i termini appartenenti alla stessa famiglia semantica. Infine, le relazioni di associazione o affinità sono espresse dal simbolo RT (*Related Term* = termine associato) ed evidenziano i rapporti di sinonimia, equivalenza ed associazione d'idee tra i descrittori.

In particolare, le relazioni gerarchiche sono state individuate con riferimento a regole logiche generali, a regole sistematiche relative allo specifico sistema di classificazione adoperato e a regole grammaticali.

Tra le prime hanno trovato ampio impiego: 1) la regola in base alla quale un termine esprime un concetto concreto ha come superiore gerarchico il termine che designa il concetto astratto corrispondente (per esempio:

«Adaptive systems» BT «Learning and adaptation»); 2) il principio per cui il termine specifico è considerato di significato meno ampio rispetto al termine che individua il genere (per esempio: «Appeal» BT «Actions and defenses»); 3) il criterio secondo il quale sono considerati in relazione di BT due termini concettualmente collegati tra loro da un rapporto d'appartenenza (per esempio: «Fundamental rights» BT «Constitutional law»); infine, 4) la norma per cui il termine che designa il prodotto o l'oggetto d'una attività rimanda, come suo superiore gerarchico, al termine che indica l'attività corrispondente e questo, a sua volta, al termine che individua il soggetto: così, «Abstracts» (prodotto dell'attività) ha per BT «Abstracting» (attività), mentre «Criminal investigation» (attività) ha per BT «Criminal Police» (agente).

Per quanto riguarda le regole sistematiche, ai termini appartenenti a un livello inferiore dello schema di classificazione sono stati assegnati come BT i termini appartenenti ai livelli superiori dello schema stesso: per esempio, «Computers in sports» (voce dello schema di classificazione cui corrisponde il codice 5.1.4) ha per BT «Computers in selected areas» (voce 5.1.).

In molti altri casi hanno trovato invece applicazione alcune semplici regole grammaticali. Così, ai descrittori pluritermini costituiti da due sostantivi collegati tra loro con una preposizione sono stati assegnati come superiori gerarchici tali sostantivi o termini a essi affini (per esempio: «Sociology of law» BT «Sociology» e BT «Legal science»).

Le relazioni di tipo associativo tra i descrittori sono state individuate facendo riferimento non solo alla vera e propria associazione semantica (sinonimia ed equivalenza) tra i termini, ma anche ai criteri, più discrezionali, dell'associazione d'idee. La scelta d'estendere nel senso indicato lo sviluppo dei rapporti di RT è stata motivata, da un lato, dall'esigenza di conferire al *Thesaurus* una funzionalità precipua in ordine alla maggiore reperibilità possibile dei dati pubblicati nella «Bibliografia» o registrati nell'archivio in linea, dall'altro, dall'opportunità di contemperare il rigido verticalismo concettuale delle liste gerarchiche e sistematiche con le possibilità d'estensione orizzontale della ricerca documentaria che sono offerte da questo tipo di collegamento tra i descrittori.

L'associazione tra sinonimi in senso stretto è stata drasticamente limitata ai soli casi in cui l'accoglimento d'entrambi i termini come descrittori risultasse giustificata dalla presenza di precise condizioni oggettive (in particolare, l'uguale rilevanza e diffusione nell'uso, pur nel rispetto delle diverse sfumature di significato). Diversamente s'è fatto ricorso alle relazioni di preferenza, rinviando dal termine non accettato al descrittore registrato nel *Thesaurus* (US) e, reciprocamente, richiamando sotto quest'ultimo il sinonimo respinto (UF). Esemplicando, sono stati collegati in RT «Computer instructions» e «Statements», ma da «Memory» s'è opportunamente rinvio (US) a «Storage».

L'associazione tra quasi-sinonimi o equivalenti (per esempio: «EDP-supported planning processes» RT «Planning process automation») è stata invece

impiegata largamente; tra l'altro, in fase di preparazione dell'opera, alla più agevole ed estesa individuazione di questo particolare tipo di relazioni è risultata funzionale la preventiva attribuzione dei codici di classificazione ai descrittori.

Infine, rapporti di tipo associativo non semantico (e quindi non riconducibili né alla sinonimia, né all'equivalenza) sono stati costruiti tanto all'interno d'uno stesso albero concettuale quanto tra nodi d'alberi concettuali distinti. Un esempio del primo tipo è il rapporto di RT tra «*Control of citizens*» e «*Control of individuals*», descrittori entrambi appartenenti all'albero concettuale che ha per radice il termine «*Sociology*». Il rapporto associativo tra «*Communication media*» e «*Information dissemination*» (il primo *Top Term*, il secondo appartenente alla catena gerarchica che ha come suo vertice il termine «*Information*») può essere invece considerato un esempio del secondo tipo.

2.3.1. *Gli archivi di lavoro*

L'impiego dell'elaboratore elettronico per redigere il *Thesaurus* richiede la predisposizione di appositi archivi di lavoro e di particolari procedure che, dal punto di vista tecnico, formano oggetto d'analitica descrizione nel terzo paragrafo della presente relazione; qui di seguito ci si limita invece a fornirne una sintetica illustrazione, accessibile anche al lettore non esperto di programmazione.

Nel primo dei *files* realizzati – denominato MASTER FILE – vengono registrate insieme le parole e le locuzioni estratte dallo *Schema di classificazione* e quelle costituenti le parole chiave inserite nei volumi bibliografici finora editi a cura dell'Istituto; di questo archivio vengono poi elaborate elettronicamente la versione in ordine alfabetico (vedi fig. 1) e quella in ordine numerico (vedi fig. 2).

In un *file* distinto – chiamato MASTER PERMUTATO – sono contenuti gli stessi termini immessi nel *file* precedente, sottoposti però a permutazione, cioè fatti ruotare con l'impiego d'una *routine* di tipo KWOC (cfr. *KWOC Index* nel § 2.4.).

Altro archivio di lavoro è il RELAT FILE (vedi fig. 3), in cui sono individuati alcuni tipi di relazioni tra i termini del MASTER FILE e tra questi e quelli del NOTE FILE. Gli operatori di relazione utilizzati nella registrazione dei dati sono espressi con le sigle BT, RT, US, TT, HT, SC, DL. Gli operatori NT e UF vengono invece introdotti automaticamente dalla macchina che li ricava, per reciprocità, dalle relazioni di BT e US.

Un ultimo archivio di lavoro, denominato NOTE FILE (vedi figg. 4 e 5), è destinato a contenere due liste distinte. La prima di esse (*Neg List*) riporta i termini non accolti nel *Thesaurus*, che rinviano – mediante l'operatore re-

MASTER ALFA

| | |
|-------|--|
| 20018 | ABSOLUTE ADDRESS = 7120 |
| 20001 | ABSTRACT = 65 |
| 20002 | ABSTRACT BULLETINS = 82 |
| 20003 | ABSTRACT JOURNALS = 82 |
| 10001 | ABSTRACT MACHINES = 721 |
| 10002 | ABSTRACTING = 65 |
| 20004 | ABUSES = 33 |
| 10003 | ACCEPTANCE ON EXAMINATION = 3413 |
| 29943 | ACCESS TIME = 7110, 7121 |
| 20005 | ACCESS TO DATA = 3743, 64, 7110 |
| 20006 | ACCESS TO INFORMATION = 311 |
| 20007 | ACCESSION = 144 |
| 20011 | ACCOUNTING = 141, 532 |
| 20009 | ACCOUNTING MANAGEMENT = 532 |
| 20010 | ACCOUNTING SECURITY = 532 |
| 20012 | ACCURACY = 7120 |
| 20013 | ACRONYM = 64 |
| 20014 | ACT OF PARLIAMENT = 114 |
| 20863 | ACTIONS AND DEFENCES = 120, 211 |
| 20016 | ADAPTIVE SYSTEMS = 722 |
| 20017 | ADDRESS = 7120 |
| 20019 | ADDRESSING SYSTEMS = 7120 |
| 20020 | ADMINISTRATION = 10, 530 |
| 20022 | ADMINISTRATION SCIENCE = 10 |
| 20021 | ADMINISTRATIVE AUTHORITIES = 10 |
| 10006 | ADMINISTRATIVE AUTOMATION (ORGANIZATION) = 133 |
| 10007 | ADMINISTRATIVE AUTOMATION (STRUCTURE) = 133 |
| 10005 | ADMINISTRATIVE AUTOMATION = 131, 133 |
| 20025 | ADMINISTRATIVE BRANCHES = 10 |
| 20026 | ADMINISTRATIVE CONTROL = 10 |

Fig. 1

lazionale US – ai termini ammessi; questa lista consente d'ampliare le chiavi d'accesso al *Thesaurus* nei confronti dell'utente, ma insieme fornisce al documentalista criteri non casuali e progressivamente precisabili per una formazione tendenzialmente controllata dal patrimonio lessicale specialistico. In una seconda lista (*Scope Note List*) vengono registrate le definizioni in senso lato di quelle parole o espressioni che non risultino di per se stesse sufficientemente chiare o che siano state accolte in accezioni particolari o parziali: essa quindi fornisce delucidazioni sul significato di certi termini, indica in quale senso essi sono usati e può anche segnalare relazioni con termini simili. La prima cifra dei numeri d'identificazione delle due liste descritte è costituita, rispettivamente, dallo 0 (*Neg List*) e dal 3 (*Scope Note List*).

La fusione degli archivi MASTER, PERMUTATO, RELAT e NOTE è resa possibile dall'esistenza, come elemento di raccordo tra i *files*, del numero d'identificazione dei termini.

| MASTER NUMER | |
|--------------|---|
| 10001 | ABSTRACT MACHINES = 721 |
| 10002 | ABSTRACTING = 65 |
| 10003 | ACCEPTANCE ON EXAMINATION = 3413 |
| 10004 | PROCEEDINGS OF CONFERENCES AND CONGRESSES ON INFORMATICS = 83 |
| 10005 | ADMINISTRATIVE AUTOMATION = 131, 133 |
| 10006 | ADMINISTRATIVE AUTOMATION (ORGANIZATION) = 133 |
| 10007 | ADMINISTRATIVE AUTOMATION (STRUCTURE) = 133 |
| 10008 | ADMINISTRATIVE COOPERATION = 1332 |
| 10009 | ADMINISTRATIVE DECISION PREDICTION = 024 |
| 10010 | ADMINISTRATIVE LAW ASPECTS OF COMPUTER LAW = 32 |
| 10011 | ALLOCATION OF HEARINGS = 121 |
| 10012 | ANALOG COMPUTERS = 7110 |
| 10013 | ANALYSIS AND FORECASTING OF JUDICIAL DECISIONS = 124 |
| 10014 | ANTITRUST AND TRADE REGULATION (INTERNATIONAL LAW ASPECTS) = 304 |
| 10015 | ARCHIVES = 144 |
| 10016 | ARTIFICIAL INTELLIGENCE = 722 |
| 10017 | ARTIFICIAL INTELLIGENCE AND LEGAL REASONING = 023 |
| 10018 | ARTIFICIAL LANGUAGES = 513 |
| 10019 | ASYNCHRONOUS SWITCHING FUNCTIONS = 721 |
| 10020 | AUTOMATED DRAFTING OF ADMINISTRATIVE MEASURES = 135 |
| 10021 | AUTOMATED FORMS = 221 |
| 10022 | AUTOMATED TRANSLATION = 513 |
| 10023 | AUTOMATION IN PARLIAMENT = 110 |
| 10024 | AUTOMATION IN THE PUBLIC ADMINISTRATION = 130 |
| 10025 | AUTOMATION IN COURTS AND TRIBUNALS = 120 |
| 10026 | AUTOMATION-ORIENTED LEGISLATION = 39 |
| 10027 | FUNDAMENTAL RIGHTS = 311 |
| 10028 | BEHAVIOURAL SCIENCES AND LEGAL INFORMATICS = 036 |
| 10029 | BIBLIOGRAPHIES OF INFORMATICS = 82 |
| 10030 | BIONICS = 722 |

Fig. 2

La stampa del *Thesaurus* secondo le procedure di lavoro illustrate consente d'ottenere le liste finali descritte nel § 2.4.

2.3.2. Il controllo e l'espansione automatica delle relazioni

In questa fase l'elaboratore, opportunamente programmato secondo quanto esposto nel § 3.2., analizza le relazioni memorizzate nel RELAT FILE ed effettua una serie di controlli ed espansioni in base alle proprietà definite dagli Autori del *Thesaurus* per le relazioni stesse. Qui di seguito si presentano le principali proprietà delle relazioni e delle strutture algebriche utilizzate.

RELAT FILE

| FILE: | RELAT | SORT1 | ISTITUTO DOCUMENTAZIONE GIURIDICA - CNR |
|-------|-------|-------|---|
| 10002 | SC | 30004 | |
| 10003 | BT | 20217 | |
| 10003 | BT | 21300 | |
| 10004 | BT | 21243 | |
| 10005 | BT | 10109 | |
| 10005 | BT | 20020 | |
| 10005 | HT | | |
| 10005 | RT | 10084 | |
| 10005 | RT | 10310 | |
| 10005 | RT | 20032 | |
| 10005 | SC | 30004 | |
| 10006 | BT | 10005 | |
| 10006 | RT | 10007 | |
| 10007 | BT | 10005 | |
| 10008 | BT | 10006 | |
| 10009 | BT | 20030 | |
| 10009 | RT | 10251 | |
| 10010 | BT | 19987 | |
| 10010 | HT | | |
| 10010 | RT | 29916 | |
| 10010 | SC | 30004 | |
| 10011 | BT | 10150 | |
| 10011 | RT | 10250 | |
| 10012 | RT | 10216 | |
| 10012 | RT | 20075 | |
| 10013 | BT | 10251 | |
| 10014 | BT | 10245 | |
| 10014 | RT | 19986 | |
| 10015 | BT | 20738 | |
| 10015 | RT | 20565 | |
| 10016 | BT | 20339 | |
| 10016 | HT | | |
| 10016 | RT | 20775 | |
| 10016 | RT | 20794 | |
| 10016 | SC | 30004 | |
| 10017 | BT | 10016 | |
| 10017 | BT | 10267 | |
| 10018 | BT | 20829 | |
| 10018 | HT | | |
| 10018 | RT | 20427 | |
| 10018 | RT | 20589 | |
| 10018 | SC | 30004 | |
| 10019 | BT | 10356 | |
| 10020 | BT | 10024 | |
| 10020 | RT | 10263 | |
| 10020 | RT | 20023 | |
| 10021 | BT | 10114 | |
| 10022 | BT | 21654 | |
| 10022 | RT | 29987 | |
| 10023 | BT | 10109 | |
| 10023 | DL | 10092 | |
| 10023 | DL | 10142 | |

Fig. 3

NOTE NUMER

| | |
|-------|------------------------------|
| 00002 | ADMINISTRATION AND EDP |
| 00003 | ADMINISTRATION OF JUSTICE |
| 00004 | ADMINISTRATION OF THE COURTS |
| 00005 | ADMINISTRATIVE REFORM |
| 00006 | ADP |
| 00007 | AUTOMATED MANAGEMENT |
| 00011 | BIBLIOGRAPHIC DOCUMENTATION |
| 00012 | BOOK-KEEPING |
| 00013 | BOUND FORM |
| 00014 | BUREAU AUTOMATION |
| 00015 | BUREAUCRACY |
| 00016 | CENTRAL PROCESSOR |
| 00018 | CIVIC RIGHTS |
| 00020 | CLASS TERM |
| 00021 | COMPANY |
| 00022 | COMPILING PROGRAMS |
| 00023 | COMPUTER-AIDED TEACHING |
| 00024 | COMPUTER-ASSISTED LEARNING |
| 00027 | COMPUTER CRIMES |
| 00028 | COMPUTER EDUCATION |
| 00029 | COMPUTER ENGINEER |
| 00030 | COMPUTER ERA |
| 00031 | COMPUTER FRAUD |
| 00032 | COMPUTER INPUT |
| 00034 | COMPUTER MAKER |
| 00035 | COMPUTER OUTPUT |
| 00036 | COMPUTER PRODUCER |
| 00037 | COMPUTER PROFESSIONAL |
| 00038 | COMPUTER REVOLUTION |
| 00039 | COMPUTER SCIENTIST |

Fig. 4

Proprietà delle relazioni

US = Use

A US B: il termine A non è accettato nel *Thesaurus*; in luogo di esso si consiglia d'adoperare il termine B. La relazione di US – definita (e controllata dalla macchina) come irriflessiva, non simmetrica e intransitiva – ha come relazione inversa (o reciproca) la relazione di UF, che presenta le stesse proprietà di US.

UF = Used for

B UF U: il termine B è stato impiegato in sostituzione del termine A, non accettato. La relazione di UF è sviluppata automaticamente, per reciprocità, dalla relazione di US ($x \text{ US } y \Rightarrow y \text{ UF } x$).

NOTE ALFA

| | |
|-------|--|
| 30007 | ADD TO THIS THESAURUS THE NAMES OF DOCUMENTATION RETRIEVAL SYSTEMS AS NEEDED |
| 30002 | ADD TO THIS THESAURUS THE NAMES OF INDIVIDUAL COMPUTER MANUFACTURERS AS NEEDED |
| 30001 | ADD TO THIS THESAURUS THE NAMES OF INDIVIDUAL COMPUTERS AS NEEDED |
| 30003 | ADD TO THIS THESAURUS THE NAMES OF INDIVIDUAL INSTITUTIONS AS NEEDED |
| 30005 | ADD TO THIS THESAURUS THE NAMES OF INFORMATION RETRIEVAL LANGUAGES AS NEEDED |
| 30010 | ADD TO THIS THESAURUS THE NAMES OF INTERNATIONAL INFORMATION SYSTEMS AS NEEDED |
| 30008 | ADD TO THIS THESAURUS THE NAMES OF JUDICIAL INFORMATION RETRIEVAL SYSTEMS AS NEEDED |
| 30006 | ADD TO THIS THESAURUS THE NAMES OF LEGAL DOCUMENTATION RETRIEVAL SYSTEMS AS NEEDED |
| 30009 | ADD TO THIS THESAURUS THE NAMES OF PARLIAMENTARY INFORMATION RETRIEVAL SYSTEMS AS NEEDED |
| 00002 | ADMINISTRATION AND EDP |
| 00003 | ADMINISTRATION OF JUSTICE |
| 00004 | ADMINISTRATION OF THE COURTS |
| 00005 | ADMINISTRATIVE REFORM |
| 00006 | ADP |
| 09999 | ASSEMBLY LANGUAGES |
| 00007 | AUTOMATED MANAGEMENT |
| 09992 | AUXILIARY STORAGE |
| 00011 | BIBLIOGRAPHIC DOCUMENTATION |
| 00012 | BOOK-KEEPING |
| 00013 | BOUND FORM |
| 00014 | BUREAU AUTOMATION |
| 00015 | BUREAUCRACY |
| 09965 | CADASTRE |
| 09968 | CAR REGISTRATION |
| 00016 | CENTRAL PROCESSOR |
| 00018 | CIVIC RIGHTS |
| 00020 | CLASS TERM |
| 09969 | COMMERCIAL MANAGEMENT |
| 00021 | COMPANY |
| 09966 | COMPANY LAW |

Fig. 5

TT = Top Term

A *TT*: il termine A viene definito come *Top Term*, qualora costituisca il vertice d'una piramide concettuale o — impiegando una diversa terminologia — la radice d'una struttura algebrica alberiforme. L'elaboratore ha consentito di verificare che fossero definiti *TT* soltanto i descrittori posti alla radice degli alberi e, corrispondentemente, che nessuno di tali descrittori fosse rimasto sprovvisto della definizione di *TT*. La relazione di *TT* è monadica, ha cioè un solo termine alla sua sinistra.

HT = Heading Term

B HT: il termine B viene definito *Nodal Heading*, a condizione che esso rappresenti un nodo concettualmente rilevante cui si colleghi un gruppo consistente d'altri termini (non meno di 5). Anche la relazione di HT, come quella di TT, è monadica.

SC = Scope Note

A SC n: al termine A segue una nota (progressivamente numerata nell'archivio NOTE, richiamata dal numero 'n', composto di cinque cifre, a destra di SC) che ne chiarisce il significato e l'uso. La relazione di SC è biadica, irriflessiva, non simmetrica e intransitiva.

BT = Broader Term

A BT B: il termine A ha come concetto gerarchicamente superiore il termine B. Le proprietà della relazione di BT — anch'esse verificate automaticamente — sono l'irriflessività, la non-simmetria e l'intransitività. La relazione inversa (o reciproca) è la relazione di NT che ha le stesse proprietà di BT.

NT = Narrower Term

B NT A: il termine B ha come inferiore gerarchico il termine A. La relazione di NT è sviluppata automaticamente, per reciprocità, dalla relazione di BT ($x BT y \Rightarrow y NT x$).

RT = Related Term

A RT B: il termine B è associato al termine A (e, simmetricamente, A è associato a B) perché i due termini sono in un rapporto di sinonimia, di quasi-sinonimia o di generica associazione. Alla relazione di RT (anch'essa irriflessiva) vengono applicate automaticamente la proprietà simmetrica e quella transitiva unidirezionale, per cui:

$$\begin{aligned} & x RT y \Rightarrow x RT y, y RT x \quad (\text{simmetria}) \\ & x RT y \wedge y RT z \Rightarrow x RT y, y RT x, y RT z, z RT y, \\ & \qquad \qquad \qquad x RT z, z RT x \quad (\text{transitività unidirezionale}) \\ \text{ma } & x RT y \wedge z RT y \Rightarrow x RT y, y RT x, z RT y, y RT z \end{aligned}$$

Vengono inoltre effettuati dei controlli automatici sull'incompatibilità fra la relazione di RT e quella di BT e controlli «antiridondanza» (vedi § 3.2.2., II parte).

Per la redazione del *Thesaurus*, infine, ci si è avvalsi d'un ulteriore operatore, denominato DL (*Delete*); esso ha consentito d'eliminare quelle relazioni

di tipo RT che, prodotte dall'espansione automatica, non sono state però giudicate concettualmente accettabili. La relazione di DL – che ha le proprietà dell'irriflessività, della simmetria e dell'intransitività – è stata volta a volta registrata nell'archivio RELAT, ma, avendo solo funzione strumentale nella fase di compilazione della *Lista alfabetica strutturata*, non è stata ovviamente riprodotta nella stampa del *Thesaurus*.

Proprietà degli alberi

Negli alberi concettuali costruiti per il *Thesaurus* sono individuabili le seguenti proprietà: 1) sono monogerarchici ($x \text{ BT } y \wedge x \text{ BT } z \wedge (x, y, z) \in \mathcal{A}_1 \Rightarrow$ 'errore di ramificazione'); 2) hanno una radice (TT) e una soltanto; 3) hanno almeno tre nodi; inoltre in ciascun albero: 4) gli HT sono al vertice di almeno 5 nodi e 5) è osservata la regola dell'incompatibilità tra la relazione di BT e quella di RT ($x \text{ BT } y \wedge x \text{ RT } y \Rightarrow$ 'errore di incompatibilità diretta'; $x \text{ BT } y \wedge y \text{ BT } z \wedge x \text{ RT } z \Rightarrow$ 'errore di incompatibilità indiretta').

Per quanto riguarda le relazioni fra gli alberi, nel *Thesaurus* realizzato sono state ammesse la congiunzione e la disgiunzione, ma non anche l'inclusione (vedi fig. 6).

2.4. L'organizzazione e la stampa automatica degli indici

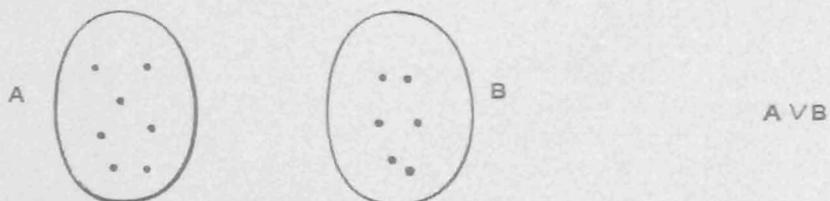
Oltre che nelle fasi della costruzione, del controllo e dello sviluppo della rete di rapporti concettuali fra i descrittori, l'elaboratore è stato utilizzato anche per organizzare graficamente la stampa delle numerose liste che compongono il *Thesaurus* e la cui consultazione coordinata permette tre tipi fondamentali d'accesso ai termini: quello alfabetico, quello gerarchico e quello sistematico (cfr. fig. 7).

Le dieci liste che compongono THES/BID nell'attuale edizione possono essere stampate su tabulato oppure registrate su supporto magnetico per essere utilizzate sia da un sistema elettronico interrogabile in linea per mezzo di terminali video-stampanti, sia da una macchina fotocompositrice automatica che, opportunamente predisposta, integra i dati memorizzati ricavandone delle matrici su pellicola, pronte per riprodurre l'opera in *offset*.

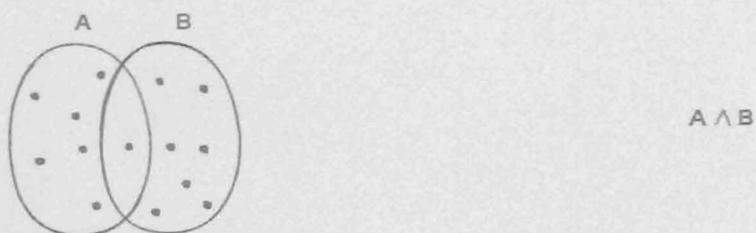
Per la consultazione di THES/BID sono state adottate tutte e tre le alternative: la stampa su tabulato viene prodotta presso il Centro elettronico dell'Istituto per la documentazione giuridica; la consultazione in linea è possibile sia collegandosi al Centro elettronico dell'Istituto che nell'ambito del Sistema Italgire della Corte di Cassazione (archivio THBID); la fotocomposizione automatica è realizzata in collaborazione con gli stabilimenti tipografici della ILTE di Moncalieri (Torino).

LE RELAZIONI TRA GLI ALBERI

1) ALBERI DISGIUNTI



2) ALBERI CONGIUNTI



3) UN ALBERO NON PUO' ESSERE INCLUSO IN UN ALTRO ALBERO

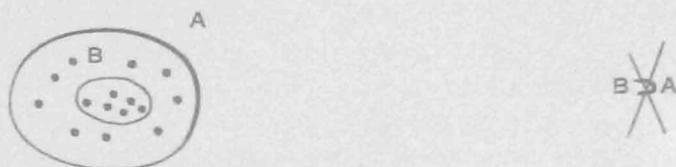


Fig. 6

Nelle pagine seguenti sono illustrati gli indici riprodotti nell'edizione fotocomposta presso l'ILTE e pubblicata dalla Casa editrice Le Monnier di Firenze.

| <u>GLI INDICI DEL THESAURUS</u> | | |
|---|---|---|
| Nell'attuale edizione THES/BID è composto da dieci liste: | | |
| <u>Descrittori concettuali</u> | | |
| 1) | L'elenco alfabetico permutato dei termini usati nei descrittori (Indice KWOC) | } consentono di consultare il <u>Thesaurus</u> secondo una strategia di ricerca alfabetica |
| 2) | La lista alfabetica strutturata dei descrittori | |
| 3) | L'indice alla lista gerarchica dei descrittori | } consentono di consultare il <u>Thesaurus</u> secondo una strategia di ricerca gerarchica |
| 4) | L'elenco dei <u>Top Terms</u> | |
| 5) | L'elenco degli <u>Heading Terms</u> | |
| 6) | La lista gerarchica dei descrittori | |
| 7) | Lo schema di classificazione di BID | } consentono di consultare il <u>Thesaurus</u> secondo una strategia di ricerca sistematica |
| 8) | La lista dei descrittori ordinata per classi | |
| <u>Nomenclatura</u> | | |
| 9) | L'elenco dei termini geografici | |
| 10) | L'elenco delle sigle e abbreviazioni | |
| | a) con ingresso alfabetico dalle sigle e abbreviazioni | |
| | b) con ingresso alfabetico dalle espressioni corrispondenti | |

Fig. 7

KWOC INDEX OF PERMITTED TERMS
(Elenco alfabetico dei termini usati nei descrittori)

Il *KWOC Index of Permitted Terms* consente di risalire dalle singole parole componenti le locuzioni accolte nel *Thesaurus* alle locuzioni di cui esse fanno parte. Si tratta d'un indice alfabetico costruito secondo la tecnica degli indici KWOC: nella colonna di sinistra, in esponente, sono elencati in ordine alfabetico i termini singoli, mentre nella colonna di destra sono riportati i contesti (cioè le locuzioni) da cui i termini stessi sono tratti.

La consultazione del *KWOC Index* costituisce una necessaria chiave d'accesso alla *Structured Alphabetic List* per chi non conosca già la sequenza delle parole all'interno dei descrittori pluritermini accolti nel *Thesaurus*.

| <i>Kwoc index</i> | ACQUISITION |
|-------------------|---|
| | A |
| ABSOLUTE | ABSOLUTE ADDRESS |
| ABSTRACT | ABSTRACT ABSTRACT BULLETINS ABSTRACT JOURNALS ABSTRACT MACHINES |
| ABSTRACTING | ABSTRACTING COMPUTERIZED ABSTRACTING |
| ABSTRACTS | ENCODED ABSTRACTS INFORMATIVE ABSTRACTS |
| ABUSES | ABUSES COMPUTER ABUSES DATA PROCESSING ABUSES |
| ACCEPTANCE | ACCEPTANCE ON EXAMINATION |
| ACCESS | ACCESS TIME ACCESS TO DATA ACCESS TO INFORMATION ADMINISTRATIVE FILE ACCESS DIRECT ACCESS STORAGE DOCUMENT CONTENT ACCESS FREE ACCESS MULTIPLE ACCESS PRIORITY (TIME-SHARING ACCESS) REMOTE ACCESS RIGHT OF ACCESS SELECTIVE ACCESS SEQUENTIAL ACCESS STORAGE SYSTEM ACCESS TIME-SHARING ACCESS |

STRUCTURED ALPHABETIC LIST
(Lista alfabetica strutturata dei descrittori)

La *Structured Alphabetic List* rappresenta la lista principale del *Thesaurus* in quanto contiene, in ordine alfabetico, tutte le parole chiave e le locuzioni esaminate (anche quelle non accettate), con l'indicazione dei vari tipi di rapporti tra esse intercorrenti. I descrittori sono seguiti, tra parentesi, dai rispettivi codici di classificazione (in base ai quali sono ordinati nella *Class List*) e i rapporti tra loro sono indicati mediante operatori relazionali codificati secondo l'uso corrente, illustrato nel § 2.3.

| <i>Alphabetic list</i> | ACRONYM |
|---|--|
| ABSOLUTE ADDRESS (7120) | ABUSES (33) |
| BT ADDRESS (7120) | BT COMPUTER CRIMINALITY (331) |
| RT STORAGE ALLOCATION (7111) | NT COMPUTER ABUSES (331) |
| STORAGE LOCATIONS (7111) | DATA PROCESSING ABUSES (331) |
| ABSTRACT (65) | RT CRIMES (122,151,33) |
| BT ABSTRACTING (65) | CRIMINAL RECORDS (122) |
| NT ABSTRACT BULLETINS (82) | OFFENDERS (122,151) |
| ABSTRACT JOURNALS (82) | ACCEPTANCE ON EXAMINATION |
| ENCODED ABSTRACTS (65) | (3413) |
| INFORMATIVE ABSTRACTS (65) | BT COMPUTER CONTRACTS (341) |
| RT DOCUMENT CONTENT ACCESS (64) | PURCHASE (536) |
| SUMMARY (65) | ACCESS TIME (7110, 7121) |
| ABSTRACT BULLETINS (82) | BT ACCESS TO DATA (3743,64,7110) |
| BT ABSTRACT (65) | SYSTEM ACCESS (383,7121) |
| RT ABSTRACT JOURNALS (82) | ACCESS TO DATA (3743, 64, 7110) |
| SECONDARY PUBLICATIONS (61) | BT DATA BANKS (1321,1322,66) |
| ABSTRACT JOURNALS (82) | NT ACCESS TIME (7110,7121) |
| BT ABSTRACT (65) | ADMINISTRATIVE FILE ACCESS |
| RT ABSTRACT BULLETINS (82) | (10,3743)- |
| SECONDARY PUBLICATIONS (61) | FREE ACCESS (3743) |
| ABSTRACT MACHINES (721) | SELECTIVE ACCESS (7120) |
| BT THEORY OF INFORMATICS (721) | RT ACCESS TO INFORMATION (311) |
| NT TURING MACHINES (721) | SYSTEM ACCESS (383,7121) |
| ABSTRACTING (65) | ACCESS TO INFORMATION (311) |
| <i>This is a nodal heading</i> | BT INFORMATION LAW (3) |
| <i>Note: This term is a nodal heading which is at the</i> | RT ACCESS TO DATA (3743,64,7110) |
| <i>vertex of a large group of other terms. Use</i> | DUTY TO GIVE INFORMATION |
| <i>a more specific term, where possible.</i> | (311,371) |
| BT DOCUMENTATION SCIENCE (032,60) | FREEDOM OF INFORMATION (311) |
| NT ABSTRACT (65) | FREEDOM OF THE PRESS (311,3734) |
| COMPUTERIZED ABSTRACTING | RIGHT OF ACCESS (3741) |
| (65,723) | RIGHT TO INFORMATION (311,371) |
| SUMMARY (65) | SYSTEM ACCESS (383,7121) |
| RT SEMANTIC COMPRESSION | ACCESSION (144) |
| (035,513,61) | BT FACILITIES MANAGEMENT (144) |
| TEXT COMPRESSION (035,513,61) | |

INDEX TO THE HIERARCHY
(Indice alla lista gerarchica dei descrittori)

Nell'*Index to the Hierarchy* sono disposti in successione alfabetica i singoli descrittori (parole o locuzioni) del *Thesaurus*, integrati dall'indicazione dei codici di classificazione ad essi corrispondenti; per ciascuno è segnalata la radice dell'albero concettuale (o degli alberi) in cui è inserito.

La consultazione dell'*Index to the Hierarchy* consente pertanto di risalire dal singolo descrittore al *Top Term* della catena gerarchica cui esso appartiene.

| <i>Index to the hierarchy</i> | ADMINISTRATIVE AUTHORITIES |
|---|--|
| ABSOLUTE ADDRESS (7120) <i>see</i> PROGRAMMING (7120) | ACCEPTANCE ON EXAMINATION (3413) <i>see</i> COMPUTER LAW (3) <i>or</i> LAW (022, 038, 114, 123, 136, 211) |
| ABSTRACT (65) <i>see</i> DOCUMENTATION (60) | ACCESS TIME (7110, 7121) <i>see</i> AUTOMATION (10, 40) <i>or</i> COMPUTER SCIENCE (031, 710) <i>or</i> DOCUMENTATION (60) <i>or</i> INFORMATICS (031, 710, 720) |
| ABSTRACT BULLETINS (82) <i>see</i> DOCUMENTATION (60) | ACCESS TO DATA (3743, 64, 7110) <i>see</i> DOCUMENTATION (60) |
| ABSTRACT JOURNALS (82) <i>see</i> DOCUMENTATION (60) | ACCESS TO INFORMATION (311) <i>see</i> COMPUTER LAW (3) <i>or</i> LAW (022, 038, 114, 123, 136, 211) |
| ABSTRACT MACHINES (721) <i>see</i> INFORMATICS (031, 710, 720) | ACCESSION (144) <i>see</i> INFORMATION (032, 311, 61, 724) |
| ABSTRACTING (65) <i>see</i> DOCUMENTATION (60) | ACCOUNTING (141, 532) |
| ABUSES (33) <i>see</i> COMPUTER LAW (3) <i>or</i> CRIMINALITY (122, 151, 302, 33) | |

LIST OF TOP TERMS (TT)
(Elenco dei *Top Terms*)

La *List of Top Terms (TT)* contiene l'elenco dei descrittori che sono alla radice dei 52 alberi concettuali in cui il *Thesaurus* è strutturato.

| <i>List of top terms (TT)</i> | TRANSPORT |
|-------------------------------|---------------------|
| ADMINISTRATION | LEGAL DECISIONS |
| AUTOMATION | LEGAL INFORMATICS |
| COMMERCE | LINGUISTICS |
| COMMUNICATION MEDIA | MACRO-ORGANIZATIONS |
| COMPUTER LAW | MANAGEMENT |
| COMPUTER SCIENCE | MATHEMATICS |
| CORPORATIONS | MONEY AND CREDIT |
| CRIMINALITY | PHILOSOPHY |
| CRIMINOLOGY | PLANNING |
| CYBERNETICS | POLICE |
| DATA PROCESSING | POLITICAL SCIENCE |
| DOCUMENTATION | POLITICS |
| ECONOMICS | PROFESSIONS |
| EDUCATION | PROGRAMMING |
| ELECTROTECHNICS | PSYCHOLOGY |
| ENVIRONMENT | PUBLISHING HOUSES |
| ERGONOMICS | RESEARCH |
| EXPERT | SOCIAL SCIENCES |
| FIRMS | SOCIETY |
| HEALTH | SOCIOLOGY |
| INDUSTRY | STATE |
| INFORMATICS | STATISTICS |
| INFORMATION | TECHNOLOGY |
| INSURANCE | TELECOMMUNICATIONS |
| LANGUAGE | TRAFFIC |
| LAW | TRANSPORT |

LIST OF HEADING TERMS (HT)

(Elenco degli Heading Terms)

La *List of Heading Terms (HT)* è la lista dei 156 descrittori che, nella strutturazione degli alberi, rappresentano nodi concettualmente rilevanti per il numero delle relazioni che da essi dipendono (non meno di 5).

I descrittori sono raggruppati in due elenchi: nel primo sono disposti secondo l'ordine alfabetico generale, con l'aggiunta – tra parentesi – dell'indicazione del numero dei nodi che dipendono da ciascuno; nel secondo sono divisi per gruppi in base al numero delle relazioni ad essi connesse e all'interno d'ogni gruppo sono ordinati alfabeticamente.

| <i>List of heading terms (HT)</i> | NETWORKS |
|--|---|
| ABSTRACTING (7) | CENTRAL GOVERNMENT (30) |
| ACCOUNTING (14) | CIVIL LAW (28) |
| ADMINISTRATIVE AUTOMATION (18) | CIVIL LAW ASPECTS OF COMPUTER LAW (22) |
| ADMINISTRATIVE LAW (8) | CLASSIFICATION (13) |
| ADMINISTRATIVE LAW ASPECTS OF COMPUTER LAW (7) | CODING (9) |
| ARTIFICIAL INTELLIGENCE (27) | COMMERCIAL LAW (22) |
| ARTIFICIAL LANGUAGES (9) | COMMERCIAL LAW ASPECTS OF COMPUTER LAW (16) |
| ARTIFICIAL RECOGNITION (11) | COMPUTER CONTRACTS (16) |
| AUTOMATION IN PARLIAMENT (7) | COMPUTER CRIMINALITY (10) |
| AUTOMATION IN THE PUBLIC ADMINISTRATION (52) | COMPUTER SYSTEM EQUIPMENT (16) |
| AUTOMATION IN THE STATE (86) | COMPUTER SYSTEMS (101) |
| BANKS (7) | COMPUTERIZED LEGAL INFORMATION RESEARCH AND RETRIEVAL (7) |
| BIBLIOGRAPHIES (5) | |
| BUSINESS MANAGEMENT (24) | |

| <i>List of heading terms (HT)</i> | LIBRARIES |
|-----------------------------------|---|
| 5 BIBLIOGRAPHIES | ADMINISTRATIVE LAW ASPECTS OF COMPUTER LAW |
| HARDWARE | AUTOMATION IN PARLIAMENT |
| MICROFILM | BANKS |
| SCHOOL | COMPUTERIZED LEGAL INFORMATION RESEARCH AND RETRIEVAL |
| SIMULATION | CRIMINAL INVESTIGATION |
| 6 DATA BASES | DEBUGGING |
| DICTIONARIES | ECONOMIC ISSUES IN INFORMATICS |
| INDEX | FUNDAMENTAL RIGHTS |
| MODELS | INFORMATION DISSEMINATION |
| PUBLIC AUTHORITIES | INTERESTED PARTY'S RIGHTS (DATA PROTECTION) |
| STANDARDIZATION | INTERNATIONAL LAW |
| SYSTEM ACCESS | |
| 7 ABSTRACTING | |

HIERARCHY LIST
(Lista gerarchica dei descrittori)

La *Hierarchy List* contiene la descrizione della struttura gerarchica generale del *Thesaurus* nella sua attuale edizione. In questa lista il materiale lessicale complessivo è raggruppato e organizzato in 52 alberi concettuali aventi per radici altrettanti descrittori, considerati di rilevanza semantica generale (vedi *List of Top Terms*).

All'interno d'ogni albero il livello gerarchico è rappresentato graficamente dal diverso numero di punti che precedono la parola o l'espressione; pertanto, nella ricerca del superiore o dell'inferiore gerarchico d'un determinato concetto, occorrerà scorrere la lista fino alla riga individuata da un numero di punti rispettivamente minore o maggiore d'una sola unità. In base allo stesso criterio, al medesimo livello di gerarchia concettuale sono da considerare i descrittori preceduti da un ugual numero di punti.

La consultazione della *Hierarchy List* fornisce al documentalista che deve indicizzare una data unità bibliografica e allo stesso utente della «Bibliografia» e dell'archivio BID una specie di mappa del campo di conoscenza in cui si colloca l'argomento oggetto d'esame o di ricerca; in base alle gerarchie di concetti indicate nella lista è quindi possibile – a seconda delle esigenze – ampliare o restringere l'analisi in modo sistematico, controllando validamente la scelta dei termini per l'indicizzazione o il reperimento delle informazioni.

Hierarchy list

DOCUMENTATION

DOCUMENTATION (60)

- AUTOMATED DOCUMENTATION (60)
- BACKLOG DOCUMENTATION (66, 68)
- BIBLIOGRAPHIES (66)
 - • AUTOMATED BIBLIOGRAPHIES (66)
 - • BIBLIOGRAPHIC RESEARCH (144, 66)
 - • BIBLIOGRAPHIC UPDATING (66)
 - • LEGAL BIBLIOGRAPHIES (042, 66)
 - • RETROSPECTIVE BIBLIOGRAPHIES (66)
- COSTS OF DOCUMENTATION (144, 69)
- DATA BANKS (1321, 1322, 66)
 - • ACCESS TO DATA (3743, 64, 7110)
 - • • ACCESS TIME (7110, 7121)
 - • • ADMINISTRATIVE FILE ACCESS (10, 3743)
 - • • FREE ACCESS (3743)
 - • • SELECTIVE ACCESS (7120)
 - • COMMERCIAL DATA BANKS (1322, 512, 536)
 - • DATA BANK MANAGEMENT (1321, 1322, 66)
 - • DATA BANK NETWORKS (1321, 1322, 66, 68)
 - • DATA BANK OPERATOR (361, 415)
 - • DATA RETRIEVAL (66, 728)

CLASSIFICATION TABLE
(Schema di classificazione di BID)

La *Classification Table* è uno schema di classificazione di tipo decimale che presenta il quadro di riferimento sistematico delle materie e degli argomenti rilevanti nel campo dell'informatica giuridica e del diritto dell'informatica. Ciascuna delle nove classi generali di cui si compone – numerate dallo 0 all'8 – è strutturata al suo interno in sottoclassi, a loro volta gerarchicamente organizzate secondo livelli di specificità gradualmente crescenti. A ogni voce della *Classification Table* corrisponde un numero di codice a una o più cifre, costruito in dipendenza della posizione gerarchica da essa occupata nel sistema; il contenuto semantico relativo è definibile però solo mediante un processo d'interpretazione sistematica che colleghi la voce presa in considerazione con quelle via via più generali ad essa connesse.

Gli elementi costitutivi della *Classification Table* (codici di classificazione e voci corrispondenti), riferiti a una specifica unità documentaria, consentono di collocarne il contenuto all'interno d'un sistema organizzato di conoscenze. Corrispondentemente, dal momento che alla struttura di questo sistema di classificazione sono stati collegati anche i singoli descrittori del *Thesaurus* mediante l'apposizione dei relativi codici, sono resi possibili il rinvio biunivoco tra parole chiave e codici di classificazione e una scomposizione analitica dell'informazione tanto nella fase dell'indicizzazione quanto in quella del reperimento.

Classification table

1. AUTOMATION IN THE STATE AND INTERNATIONAL...

- 0. LEGAL INFORMATICS: THEORY, SYSTEMATICS, APPLICATIONS AND RELATIONS
 - 0.1. LEGAL INFORMATICS IN GENERAL
 - 0.1.1. Theory
 - 0.1.2. Systematics
 - 0.1.3. Terminology
 - 0.1.4. Others
 - 0.2. LEGAL INFORMATICS: SPECIAL QUESTIONS
 - 0.2.1. Analysis of legal language by computers
 - 0.2.2. Legal information and legal information retrieval systems
 - 0.2.3. Artificial intelligence and legal reasoning
 - 0.2.4. Decision theory. Judicial and administrative decision prediction, and judicial behaviour
 - 0.2.5. Formalization of legal language. Symbolic logic and law
 - 0.2.6. Others
 - 0.3. LEGAL INFORMATICS RELATED TO OTHER DISCIPLINES
 - 0.3.1. Computer science. Informatics
 - 0.3.2. Information science. Documentation science
 - 0.3.3. Cybernetics
 - 0.3.4. Philosophy. Logic. Mathematics
 - 0.3.5. Linguistics. Semiology
 - 0.3.6. Social, human and behavioural sciences (psychology, sociology, statistics, etc.)
 - 0.3.7. Economics
 - 0.3.8. Jurisprudence

CLASS LIST
(Lista dei descrittori ordinata per classi)

Nella *Class List* sono riprodotte le voci della *Classification Table*. Sotto ogni voce sono elencati i termini del *Thesaurus* ad essa attinenti, eventualmente seguiti dall'indicazione tra parentesi degli altri codici di classificazione sotto cui sono allo stesso modo raggruppati. I termini sono elencati in due gruppi distinti, al loro interno alfabeticamente ordinati: il primo gruppo comprende i descrittori estratti dalla *Classification Table*, mentre nel secondo sono contenute tutte le parole o espressioni, registrate nel *Thesaurus*, che non provengono dalla *Classification Table*, bensì direttamente dalla letteratura specialistica esaminata.

Class list 0.24. DECISION THEORY. JUDICIAL AND ADMINISTRATIVE...

- 0. **LEGAL INFORMATICS: THEORY, SYSTEMATICS, APPLICATIONS AND RELATIONS**
- 0.1. **LEGAL INFORMATICS IN GENERAL**
 - 1) LEGAL INFORMATICS
 - 2) JURIMETRICS
- 0.1.1. **THEORY**
 - 1) THEORY OF LEGAL INFORMATICS
- 0.1.2. **SYSTEMATICS**
 - 1) SYSTEMATICS OF LEGAL INFORMATICS
- 0.1.3. **TERMINOLOGY**
 - 1) TERMINOLOGY OF LEGAL INFORMATICS
- 0.1.4. **OTHERS**
- 0.2. **LEGAL INFORMATICS: SPECIAL QUESTIONS**
- 0.2.1. **ANALYSIS OF LEGAL LANGUAGE BY COMPUTERS**
 - 1) ANALYSIS OF LEGAL LANGUAGE BY COMPUTERS
 - 2) INTERPRETATION OF LAW (0381)
 - LEGAL DICTIONARY (043, 513)
 - LEGAL LANGUAGE (025, 0382, 513)
 - LEGAL TEXTS (022, 0382)
 - STRUCTURAL ANALYSIS OF LANGUAGE (025, 035, 513)
- 0.2.2. **LEGAL INFORMATION AND LEGAL INFORMATION RETRIEVAL SYSTEMS**
 - 1) LEGAL INFORMATION RETRIEVAL SYSTEMS (123)
 - 2) LAW (038, 114, 123, 136, 211)
 - LEGAL AUTOMATED DOCUMENTATION (032, 114, 123, 136, 211)
 - LEGAL DATA BANKS (032, 114, 123, 136, 211)
 - LEGAL DATA BASES (032, 114, 123, 136, 211)
 - LEGAL DATA RETRIEVAL (032, 114, 123, 136, 211)
 - LEGAL DOCUMENTATION (032, 114, 123, 136, 211)
 - LEGAL DOCUMENTATION RETRIEVAL SYSTEMS (032, 114, 123, 136, 211)
 - LEGAL DOCUMENTS (032, 114, 123, 136, 211)
 - LEGAL FILES (032, 114, 123, 136, 211)
 - LEGAL INFORMATION (114, 123, 136, 211)

GEOGRAPHIC LIST
(Elenco dei termini geografici)

La *Geographic List* contiene l'elenco alfabetico dei nomi di Paesi (o di gruppi di Paesi politicamente significativi) utilizzati nell'ambito della documentazione esaminata per individuare le aree geografiche in cui si svolgono determinate attività di ricerca o esperienze applicative; tali nomi sono riprodotti nella forma normalizzata, che è quindi vincolante per l'indicizzatore e per chi ricerca le informazioni.

Quando il nome d'uno Stato è seguito da quello d'una regione ad esso appartenente, i due nomi sono separati tra loro mediante un trattino; i nomi di città sono invece indicati tra parentesi, dopo il nome dello Stato o della regione cui si riferiscono.

| <i>Geographic list</i> | LAOS |
|-----------------------------|--|
| EGYPT | HUNGARY |
| EL SALVADOR | ICELAND |
| ETHIOPIA | INDIA |
| EUROPE | INDONESIA |
| EUROPE, EASTERN | IRAN |
| EUROPE, WESTERN | IRAQ |
| FINLAND | IRELAND |
| FRANCE | ISRAEL |
| FRENCH GUYANA | ITALY |
| FRENCH OVERSEAS TERRITORIES | ITALY (MILAN) |
| GABON | ITALY - LAZIO (ROME) |
| GAMBIA | ITALY - LOMBARDIA |
| GERMAN DEMOCRATIC REPUBLIC | IVORY COAST |
| GERMANY | JAMAICA |
| GERMANY, FEDERAL REPUBLIC | JAPAN |
| GHANA | JORDAN |
| GREECE | KENYA |
| GREENLAND | KMER REPUBLIC |
| GUATEMALA | KOREA |
| GUINEA | KOREA, DEMOCRATIC PEOPLE'S REPUBLIC |
| GUYANA | KOREA, REPUBLIC |
| HAITI | KUWAIT |
| HONDURAS | LAOS |

ACRONYM LIST (1 e 2)
(Elenco delle sigle e abbreviazioni)

L'*Acronym List* è costituita dalle sigle ricorrenti nel campo dell'informatica giuridica e del diritto dell'informatica e dalle espressioni che di quelle sigle rappresentano lo scioglimento; dei due elenchi che compongono la lista, il primo è basato sull'ordine alfabetico delle sigle e il secondo su quello delle espressioni ad esse relative.

Le abbreviazioni contenute tra parentesi, quando seguono espressioni corrispondenti a nomi di istituti, enti, associazioni, ecc., indicano il Paese (secondo la sigla automobilistica internazionale) in cui essi hanno sede o, eventualmente, il loro carattere internazionale (INT.); quando, invece, sono posposte a titoli di riviste, guide, repertori e simili, indicano che si tratta di pubblicazioni (PUB.).

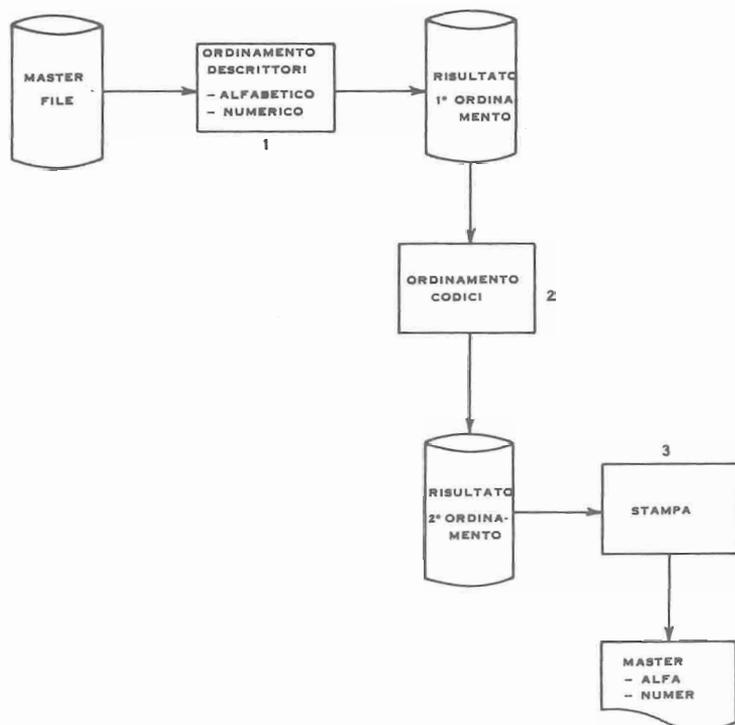
| <i>Acronym list 1</i> | ACORN |
|-----------------------|---|
| A&I | ABSTRACTING AND INDEXING |
| A/D | ANALOG DIGITAL |
| AAA | AMERICAN AUTOMOBILE ASSOCIATION (USA) |
| AAAC | AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF CRIMINOLOGY (USA) |
| AAAS | AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE (USA) |
| AAC | AMERICAN ASSOCIATION OF CRIMINOLOGY (USA) |
| AACI | AMERICAN ASSOCIATION FOR CONSERVATION INFORMATION (USA) |
| AACR | ANGLO-AMERICAN CATALOGING RULES (PUB.) |
| AACSL | AMERICAN ASSOCIATION FOR THE COMPARATIVE STUDY OF LAW (USA) |
| AAEDC | AMERICAN AGRICULTURAL ECONOMICS DOCUMENTATION CENTER (USA) |
| AAFS | AMERICAN ACADEMY OF FORENSIC SCIENCES (USA) |
| AAHA | AMERICAN ACADEMY OF HEALTH ADMINISTRATION (USA) |

| <i>Acronym list 2</i> | ALGORITHMIC REMOTE MANIPULATION |
|--|---------------------------------|
| A PROGRAMMING LANGUAGE | APL |
| A SYSTEM FOR NOTARIES (F) | ALVAN |
| ABNORMAL END | ABEND |
| ABSOLUTE VALUE BIT SYNCHRONIZER | AVBS |
| ABSTRACTING AND INDEXING | A&I |
| ACCOUNTING AND FINANCE OFFICE | AFO |
| ACCOUNTING INFORMATION SYSTEM | AIS |
| ACRONYMS, INITIALISMS, AND ABBREVIATIONS DICTIONARY (PUB.) | AIAD |
| ACTION INFORMATION ORGANIZATION | AIO |
| ACTIVE FILTER NETWORK | AFN |
| ACTIVELY SHARED KNOWLEDGE | ASK |
| ADAPTIVE MODE PLANNING SYSTEM INPUT | AMPSIN |
| ADDRESS INDICATING GROUP | AIG |
| ADJUTANT GENERAL MANAGEMENT INFORMATION SYSTEM | AGMIS |

3. I PROGRAMMI ELETTRONICI PER IL CONTROLLO E L'ESPANSIONE DELLA STRUTTURA DEL «THESAURUS» E LA REDAZIONE AUTOMATICA DEGLI INDICI

3.1. Le liste di lavoro

3.1.1. Stampa del MASTER FILE



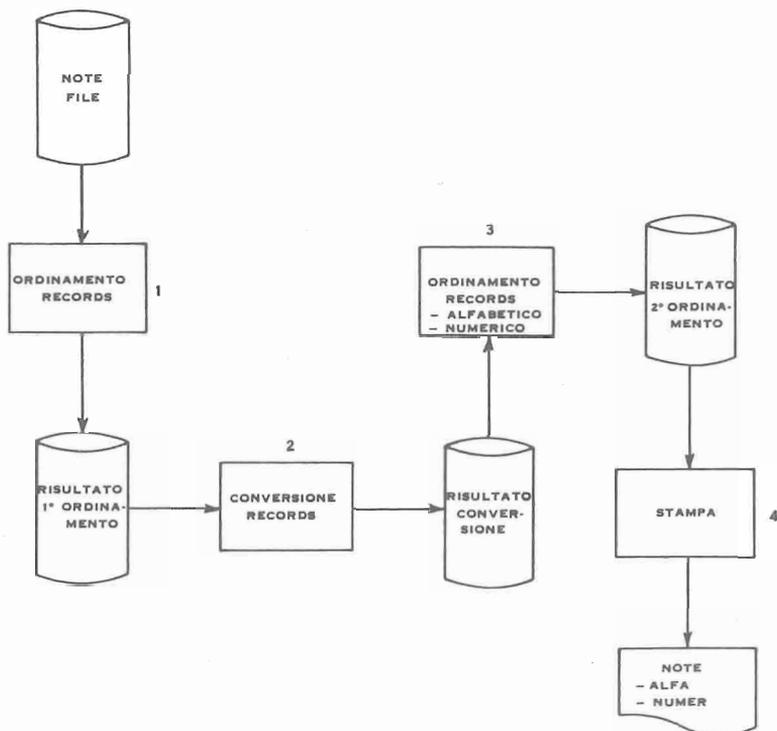
Descrizione (3.1.1.)

Tracciato *records* del MASTER FILE: il formato del *file* è variabile e la lunghezza massima di ogni *record* è di 80 caratteri. I primi 5 caratteri sono destinati al numero d'identificazione. Il resto del *record* ha un formato variabile e i descrittori sono distinti dai codici di classificazione mediante il segno di '='; nel caso in cui ad un descrittore siano stati assegnati più codici, questi sono separati tra loro con una virgola.



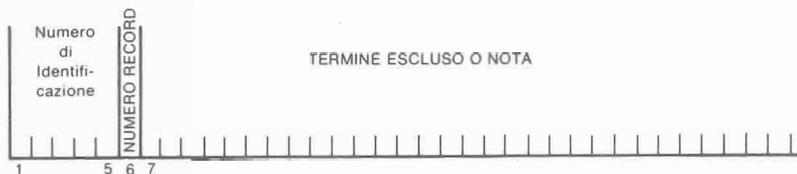
- (1) I *records* del MASTER FILE vengono ordinati alfabeticamente per descrittore o numericamente per numero d'identificazione.
- (2) I codici di classificazione relativi ad ogni descrittore vengono ordinati secondo la sequenza dei codici dello schema di classificazione.
- (3) Viene eseguita la stampa dei dati secondo le modalità grafiche stabilite e il tipo di ordinamento prescelto.

3.1.2. Stampa del NOTE FILE



Descrizione (3.1.2.)

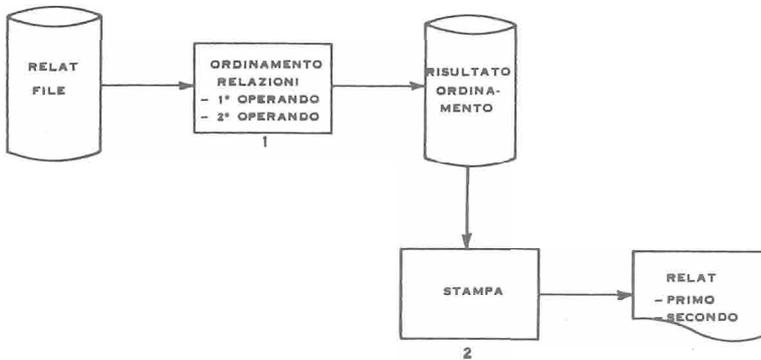
Tracciato *records* del NOTE FILE: il formato del *file* è variabile e la lunghezza massima di ogni *record* è di 80 caratteri. I primi 5 caratteri sono destinati al numero d'identificazione; il sesto al numero del *record*, dato che le espressioni registrate in questo *file* possono occupare più di una riga; i restanti caratteri del *record* sono destinati a contenere il testo con formato variabile.



→ 80

- (1) I *records* del NOTE FILE vengono ordinati in base ai numeri d'identificazione e, al loro interno, per numero di *records*.
- (2) I *records* del NOTE FILE vengono convertiti in un diverso formato, in cui viene eliminato il numero di *record* dato che le espressioni su più righe vengono concatenate tra loro.
- (3) I *records* del NOTE FILE vengono ordinati alfabeticamente per espressione o numericamente in base al numero d'identificazione.
- (4) Viene eseguita la stampa dei dati secondo le modalità grafiche stabilite e il tipo di ordinamento prescelto.

3.1.3. Stampa del RELAT FILE



Descrizione (3.1.3.)

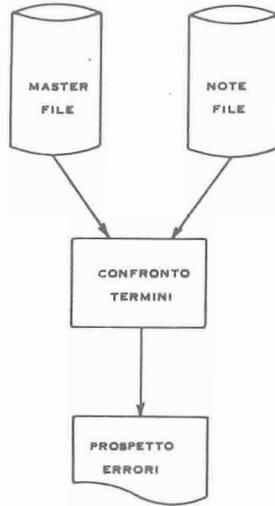
Tracciato *records* del RELAT FILE: il formato del *file* è fisso e la lunghezza di ogni *record* è di 14 caratteri. Ogni *record* contiene, nelle prime 5 e nelle ultime 5 posizioni, i numeri d'identificazione dei due operandi e, al centro (posizioni 7 e 8), la sigla corrispondente all'operatore relazionale (BT, RT, US, TT, HT, SC, DL).



- (1) I *records* del RELAT FILE vengono ordinati in base al numero d'identificazione del primo o del secondo operando.
- (2) Viene eseguita la stampa dei dati secondo le modalità grafiche stabilite e il tipo di ordinamento prescelto.

3.2. L'espansione automatica delle relazioni fra i termini e il controllo delle regole di costruzione del Thesaurus

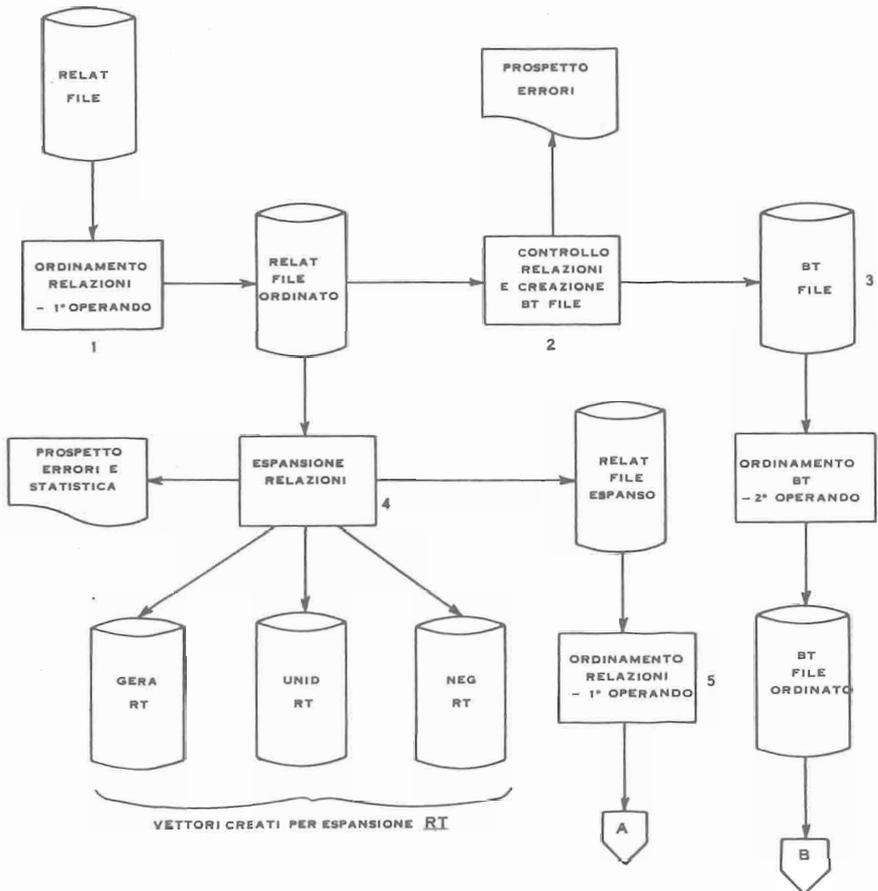
3.2.1. Confronto tra MASTER FILE e NOTE FILE



Descrizione (3.2.1.)

I termini contenuti nei due *files* vengono confrontati tra loro e sono segnalati i casi in cui lo stesso termine compare in entrambi i *files*.

3.2.2. Controllo ed espansione delle relazioni (I parte)



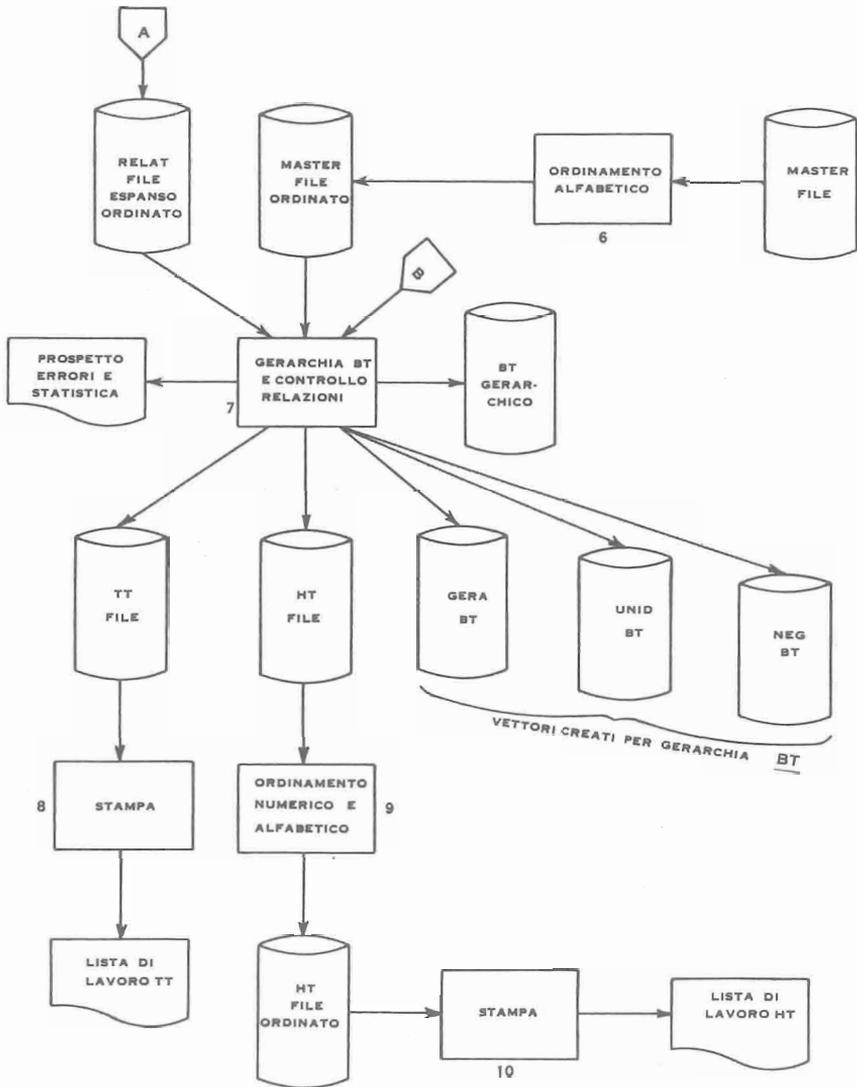
Descrizione (3.2.2.) (I parte)

- (1) I records del RELAT FILE vengono ordinati in base al numero d'identificazione del primo operando (riprodotto a sinistra dell'operatore relazionale).
- (2) Viene eseguita una serie di controlli. Innanzitutto si verifica che gli operatori memorizzati corrispondano a quelli previsti e che, per ciascuno di essi, vengano rispettate le relative proprietà (per es. il carattere monadico o biadico dell'operatore); in secondo luogo si controlla che all'interno d'una stessa relazione non compaia due volte il medesimo operando (es.: 21125 BT 21125 => 'errore'); infine viene accertato che una data coppia di operandi non sia collegata da uno stesso operatore una volta in un ordine e una seconda volta nell'ordine inverso (es.: 01118 US 12143 \wedge 12143 US 01118 => 'errore'). Nello stesso passo, mediante l'estrazione delle relazioni di BT dal RELAT FILE, viene creato il BT FILE.
- (3) I records del BT FILE vengono ordinati in base al numero d'identificazione del secondo operando (riprodotto a destra dell'operatore relazionale).

(4) Viene eseguita l'espansione automatica delle relazioni memorizzate nel RELAT FILE, costruendo le relazioni reciproche per i BT e gli US e le relazioni simmetriche e transitive (unidirezionali) per gli RT (es. di relazioni reciproche: 12125 BT 13148 => 13148 NT 12125; 01218 US 13148 => 13148 UF 01218; es. di relazioni simmetriche e transitive: 12451 RT 15621 \wedge 15621 RT 22404 => 15621 RT 12451 \wedge 22404 RT 15621 \wedge 12451 RT 22404 \wedge 22404 RT 12451). I vettori utilizzati per l'applicazione alle relazioni di RT della proprietà transitiva unidirezionale vengono poi conservati su disco (*files* UNID RT e GERA RT) per consentire ai redattori del *Thesaurus* d'individuare in essi gli eventuali errori logici segnalati dal programma nell'apposito prospetto. Egualmente vengono conservati su disco e stampati in un prospetto separato i vettori di quelle relazioni di RT che vengono ignorati per un errore logico di costruzione nell'espansione della proprietà transitiva (*file* NEG RT).

(5) I *records* del RELAT FILE ESPANSO vengono ordinati in base al numero d'identificazione del primo operando.

3.2.2. Controllo ed espansione delle relazioni (II parte)



Descrizione (3.2.2.) (II parte)

(6) I records del MASTER FILE vengono ordinati alfabeticamente.

(7) Con le relazioni di BT ordinate sul secondo operando (v. punto 3) vengono costruite le catene gerarchiche degli alberi concettuali del *Thesaurus*, che sono memorizzate nel file BT GERARCHICO secondo l'ordine alfabetico dei descrittori all'interno dello stesso livello gerarchico. I vettori utilizzati per la costruzione delle catene gerarchiche vengono poi conservati su disco (files UNID BT e GERA BT) per consentire ai

redattori del *Thesaurus* d'individuare in essi gli eventuali errori logici, di salto di gerarchia e di ramificazione, segnalati dal programma nell'apposito prospetto. Egualmente vengono conservati su disco e stampati in un prospetto separato i vettori di quelle relazioni di BT che vengono ignorati per un errore logico di costruzione delle catene gerarchiche (*file* NEG BT).

Viene poi eseguita una serie di controlli. Innanzitutto viene verificato che una data coppia di operandi non compaia una volta con l'operatore BT e una seconda volta con l'operatore RT (es.: 12148 BT 13126 \wedge 12148 RT 13126 \Rightarrow 'errore d'incompatibilità diretta'; oppure 12148 BT 13126 \wedge 13126 BT 14158 \wedge 12148 RT 14158 \Rightarrow 'errore d'incompatibilità indiretta').

In secondo luogo vengono eseguiti dei controlli di ridondanza, all'interno dello stesso albero concettuale, relativamente alle relazioni di RT; in particolare viene segnalato come ridondante il fatto che da un nodo partano più relazioni di RT verso nodi legati tra loro da un rapporto di gerarchia (es.: 15156 RT 18143 \wedge 15156 RT 13148 \wedge 18143 BT 13148 \Rightarrow 'errore di ridondanza').

In terzo luogo vengono controllati i *Top Terms* per verificare che quelli definiti dai redattori come tali siano effettivamente al vertice degli alberi concettuali e, viceversa, che non sia stato omissso di definire come *Top Term* un termine al vertice delle catene gerarchiche. Per quanto riguarda gli *Heading Terms* i termini vengono pesati in base al numero di sottonodi da essi dipendenti nelle catene gerarchiche: il programma controlla che i nodi da cui dipendono almeno 5 sottonodi siano stati definiti *Heading Terms* (HT) e, viceversa, che nessuno degli altri sia stato erroneamente definito HT.

Al termine dei controlli i *Top* e gli *Heading Terms* vengono memorizzati su disco, rispettivamente nel TT FILE e nell'HT FILE. Viene infine controllato che nessun termine del MASTER FILE sia stato omissso nella costruzione degli alberi concettuali.

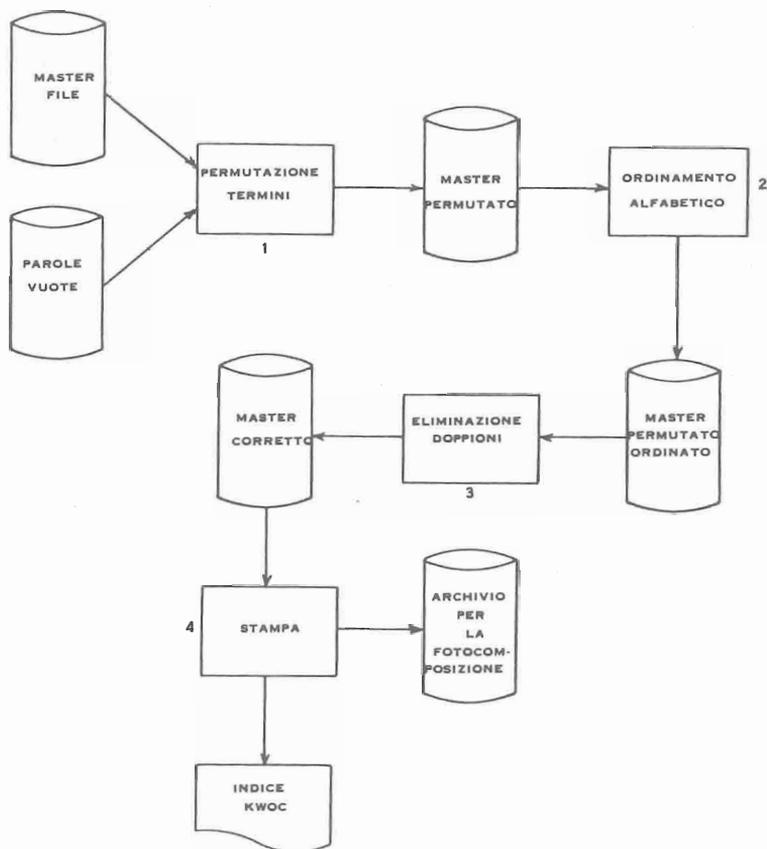
(8) Viene eseguita la stampa di lavoro dei *Top Terms*, in ordine alfabetico e secondo le modalità grafiche stabilite.

(9) I *records* dell'HT FILE vengono raggruppati secondo l'ordine crescente del numero di sottonodi e ordinati alfabeticamente all'interno di ciascun gruppo.

(10) Viene eseguita la stampa degli *Heading Terms* secondo l'ordine indicato nel punto (9) e le modalità grafiche stabilite.

3.3. Le liste di stampa

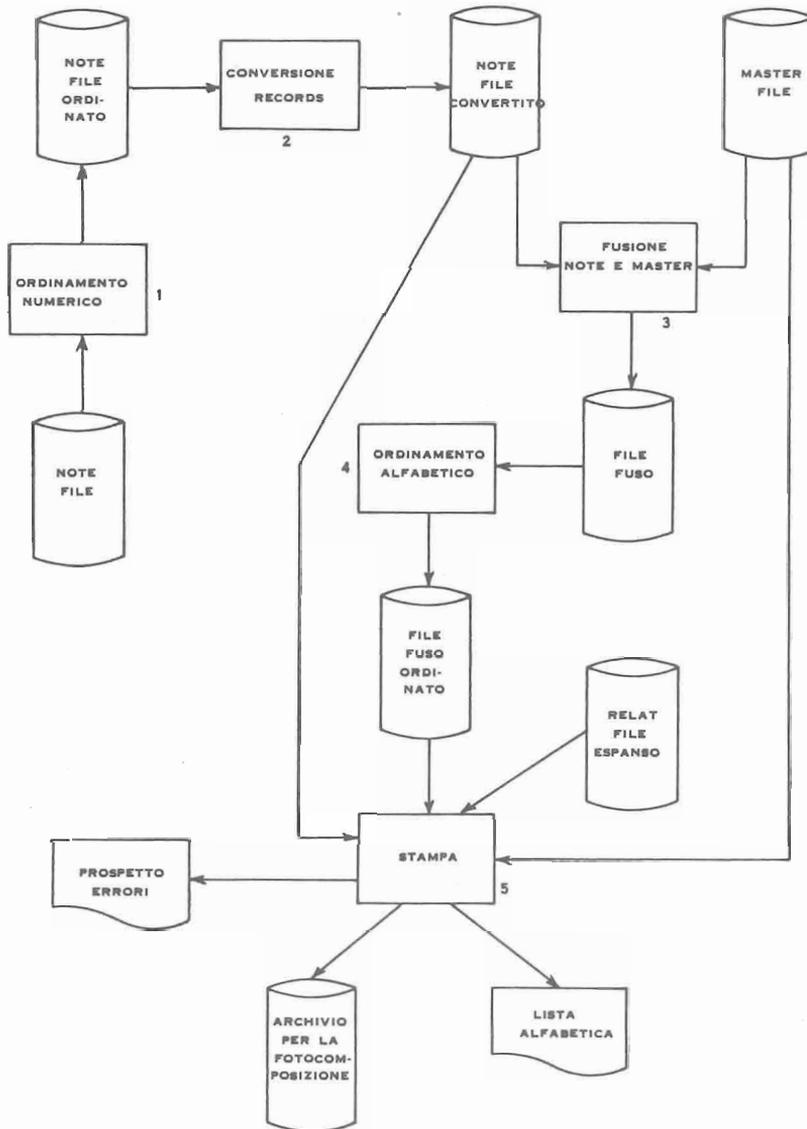
3.3.1. Indice KWOC



Descrizione (3.3.1.)

- (1) Vengono permutati i termini componenti ciascun descrittore, escludendo le parole grammaticali e quelle considerate dai redattori prive di significato autonomo.
- (2) Vengono ordinati alfabeticamente i *records* del MASTER PERMUTATO.
- (3) Le eventuali stringhe ripetute sono eliminate.
- (4) Viene eseguita, secondo le modalità grafiche stabilite, la stampa su tabulato dell'indice e viene creato un archivio corrispondente che contiene anche i codici speciali per la fotocomposizione automatica.

3.3.2. Lista alfabetica strutturata



Descrizione (3.3.2.)

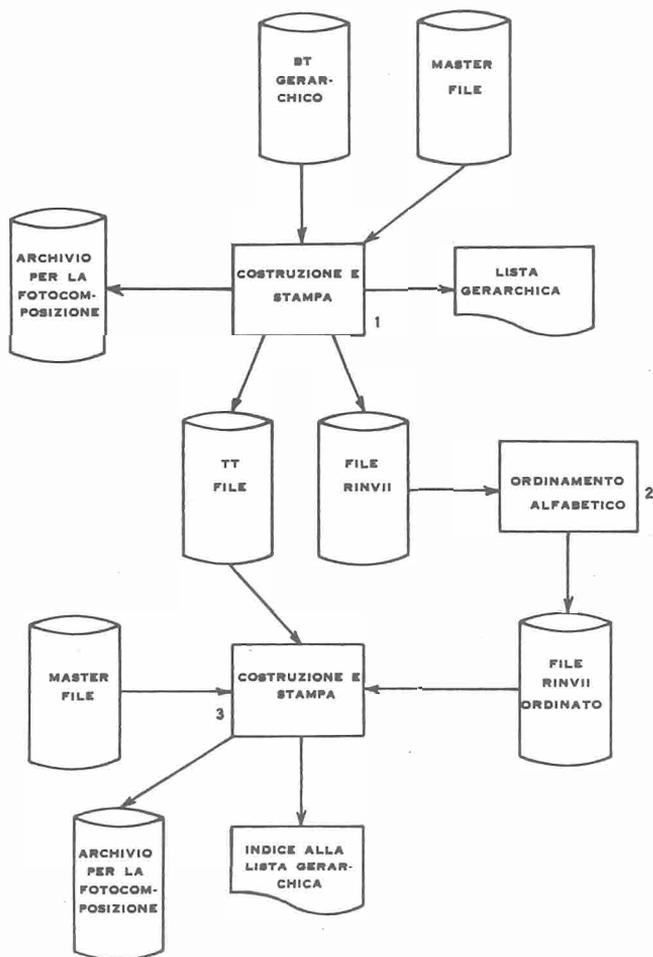
- (1) I records del NOTE FILE vengono ordinati in base ai numeri d'identificazione e, al loro interno, per numero di records.
- (2) I records del NOTE FILE vengono convertiti in un diverso formato, in cui viene eliminato il numero di record, dato che le espressioni su più righe sono concatenate tra loro.

(3) I *records* del NOTE FILE e del MASTER FILE vengono riuniti nel FILE FUSO, con l'esclusione di quelle espressioni il cui numero d'identificazione cominci col 3 (NOTE).

(4) I *records* del FILE FUSO vengono ordinati alfabeticamente.

(5) Viene eseguita la stampa dell'indice su tabulato, secondo le modalità grafiche stabilite; inoltre viene creato un archivio corrispondente, che contiene anche i codici speciali per la fotocomposizione automatica, e viene eventualmente stampato un prospetto-errori relativo a quei numeri d'identificazione del RELAT FILE che non compaiono né nel MASTER FILE né nel NOTE FILE.

3.3.3. Lista gerarchica e Indice alla lista gerarchica



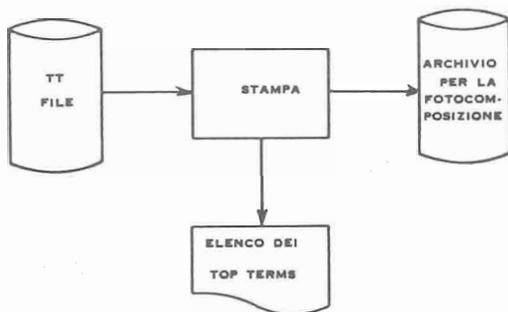
Descrizione (3.3.3.)

(1) Utilizzando le catene gerarchiche memorizzate nell'archivio BT GERA e i descrittori corrispondenti ai numeri d'identificazione registrati nel MASTER FILE, viene costruita la *Lista gerarchica dei descrittori* per ciascun albero concettuale. Il programma provvede inoltre a stampare la lista gerarchica su tabulato, secondo le modalità grafiche stabilite, e a creare un archivio corrispondente, contenente anche i codici speciali per la fotocomposizione automatica. Vengono infine memorizzati su disco il *file dei Top Terms* e un secondo *file* contenente i rinvii dai termini del MASTER FILE ai *Top Terms*.

(4) I *records* del FILE RINVII vengono ordinati alfabeticamente.

(3) Utilizzando i *records* del FILE RINVII, del TT FILE e del MASTER FILE, il programma costruisce l'*Indice alla lista dei descrittori*, provvedendo a stamparlo su tabulato, secondo le modalità grafiche stabilite, e a creare un archivio corrispondente, che contiene anche i codici speciali per la fotocomposizione automatica.

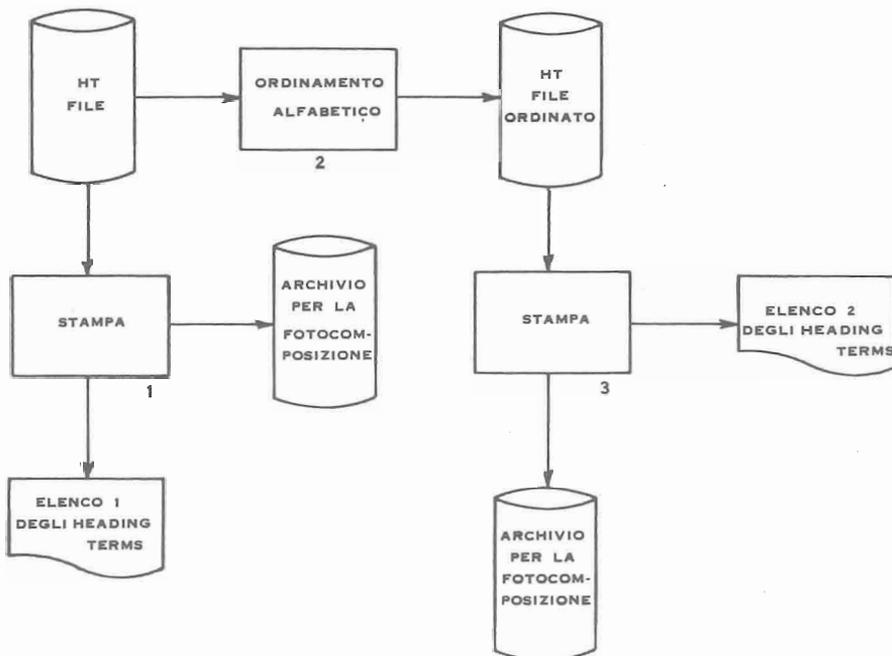
3.3.4. Elenco dei Top Terms



Descrizione (3.3.4.)

Viene eseguita la stampa dell'*Elenco dei Top Terms* su tabulato, secondo le modalità grafiche stabilite. Il programma provvede inoltre a creare un archivio corrispondente, che contiene anche i codici speciali per la fotocomposizione automatica.

3.3.5. Elenco degli Heading Terms



Descrizione (3.3.5.)

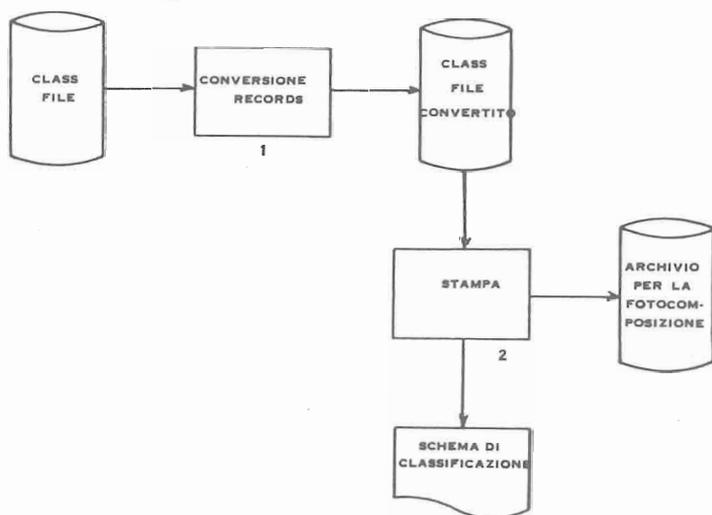
(1) Viene eseguita, secondo le modalità grafiche stabilite, la stampa dell'*Elenco 1 degli Heading Terms*, raggruppando i termini in base al numero delle relazioni ad essi con-

nesse e ordinandoli alfabeticamente all'interno d'ogni gruppo. Viene inoltre creato un archivio corrispondente, che contiene anche i codici speciali per la fotocomposizione automatica.

(2) I *records* dell'HT FILE vengono ordinati alfabeticamente.

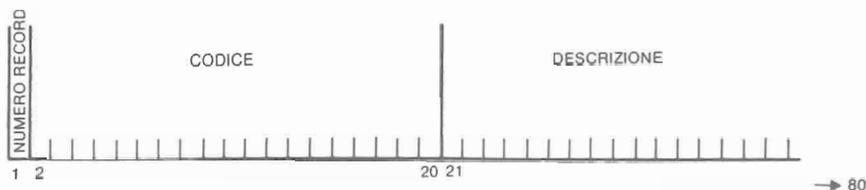
(3) Viene eseguita la stampa su tabulato dell'*Elenco 2 degli Heading Terms*, in ordine alfabetico e secondo le modalità grafiche stabilite. Viene inoltre creato un archivio corrispondente, che contiene anche i codici speciali per la fotocomposizione automatica.

3.3.6. Schema di classificazione



Descrizione (3.3.6.)

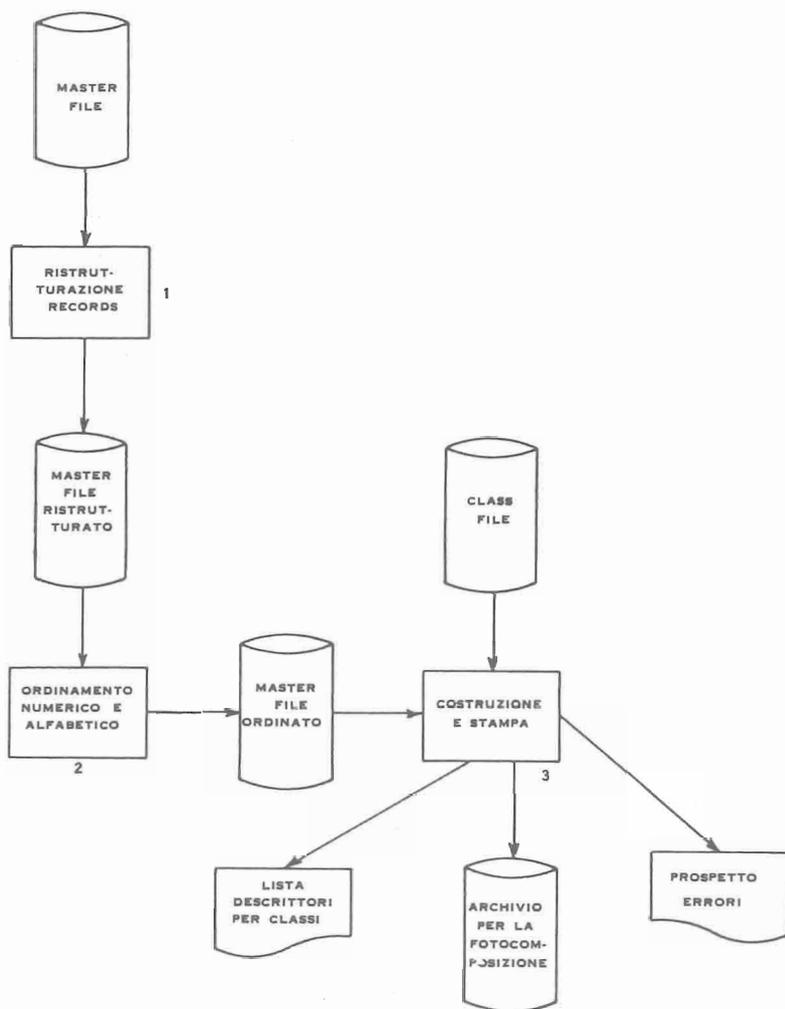
Tracciato *records* del CLASS FILE: il formato del *file* è variabile e la lunghezza massima di ogni *record* è di 80 caratteri. Nel primo carattere viene indicato il numero del *record*, dato che le informazioni registrate in questo *file* possono occupare più d'una riga; nel campo da 2 a 20 viene registrato il codice di classificazione; i restanti caratteri del *record* sono destinati a contenere le voci di classificazione corrispondenti ai codici.



(1) I *records* del CLASS FILE vengono convertiti dal formato fisso di 80 caratteri a un formato variabile, concatenando i *records* di quelle voci dello *Schema di classificazione* che occupano più di 80 caratteri.

(2) Viene eseguita la stampa dello *Schema di classificazione* su tabulato, secondo le modalità grafiche stabilite, e viene creato un archivio corrispondente, che contiene anche i codici speciali per la fotocomposizione automatica.

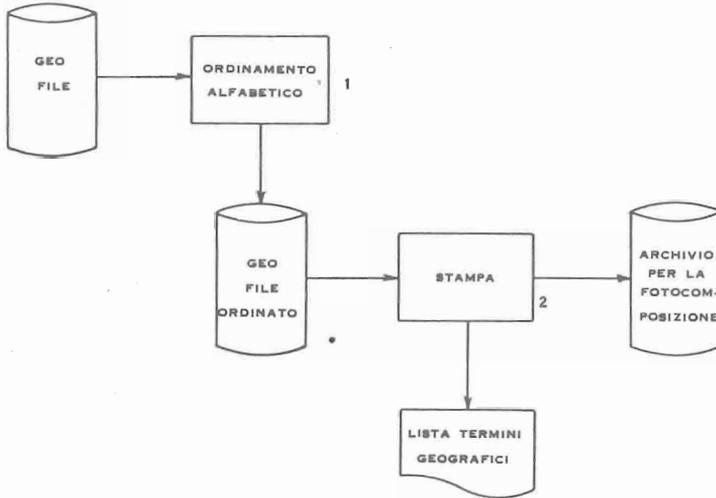
3.3.7. Lista dei descrittori ordinati per classi



Descrizione (3.3.7.)

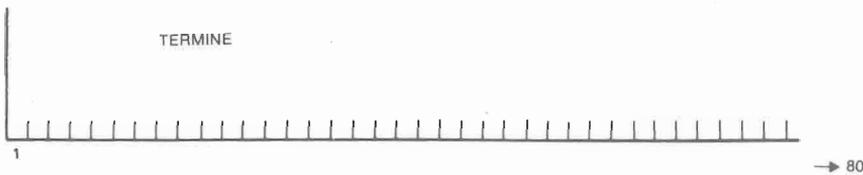
- (1) I records del MASTER FILE vengono riprodotti tante volte quanti sono i codici di classificazione assegnati e ogni record riprodotto viene fatto precedere, a turno, da uno dei codici di classificazione.
- (2) I records del MASTER FILE, ristrutturato secondo le modalità indicate nel punto (1), vengono ordinati in base al codice di classificazione, alla prima cifra del numero d'identificazione e, al loro interno, alfabeticamente.
- (3) Viene stampata su tabulato la *Lista dei descrittori ordinata per classi*, utilizzando il MASTER FILE ORDINATO e il CLASS FILE. Inoltre viene creato un archivio corrispondente, che contiene anche i codici speciali per la fotocomposizione automatica, ed eventualmente stampato un prospetto dei codici di classificazione errati.

3.3.8. Lista dei termini geografici



Descrizione (3.3.8.)

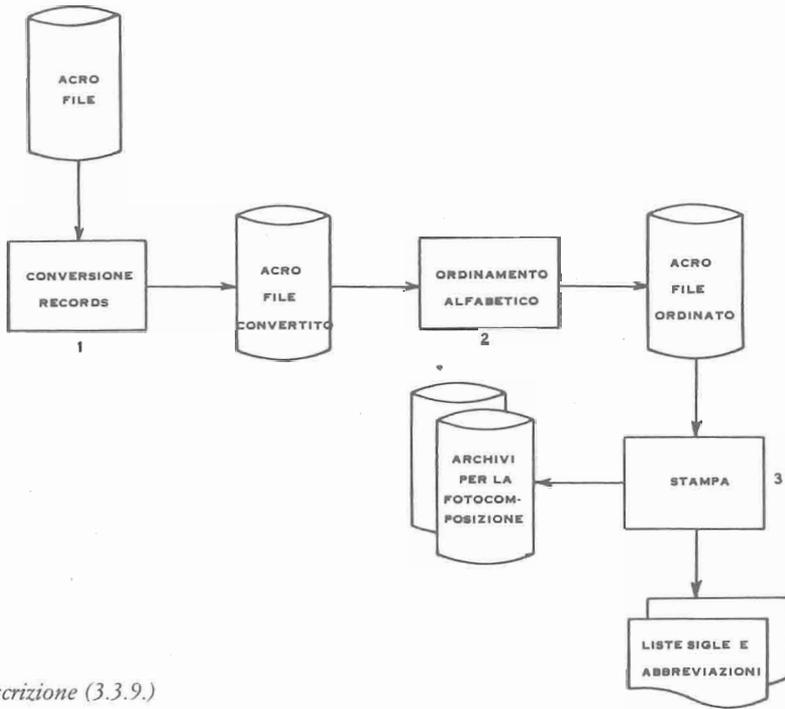
Tracciato *records* del GEO FILE: il formato del *file* è variabile e la lunghezza massima del *record* è di 80 caratteri. Ogni *record* è destinato a contenere la versione normalizzata di un termine geografico.



(1) I *records* del GEO FILE vengono ordinati alfabeticamente.

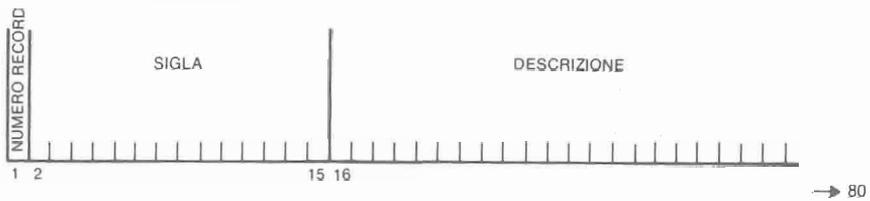
(2) Viene eseguita, secondo le modalità grafiche stabilite, la stampa su tabulato della *Lista dei termini geografici* e viene creato un archivio corrispondente, che contiene anche i codici speciali per la fotocomposizione automatica.

3.3.9. Lista delle sigle ed abbreviazioni



Descrizione (3.3.9.)

Tracciato *records* dell'ACRO FILE: il formato del *file* è variabile e la lunghezza massima d'ogni *record* è di 80 caratteri. Nel primo carattere viene indicato il numero del *record*, dato che le informazioni registrate in questo *file* possono occupare più d'una riga; nel campo da 2 a 15 viene registrata la sigla; i restanti caratteri del *record* sono destinati a contenere le espressioni che rappresentano lo scioglimento delle sigle.



- (1) I *records* dell'ACRO FILE vengono convertiti da un formato fisso di 80 caratteri a un formato variabile, concatenando i *records* corrispondenti alle espressioni che occupano più di 80 caratteri.
- (2) I *records* dell'ACRO FILE CONVERTITO vengono ordinati alfabeticamente in base alla sigla o all'espressione corrispondente, in relazione alla lista che s'intende produrre.
- (3) Viene eseguita, secondo le modalità grafiche stabilite, la stampa della lista delle sigle e abbreviazioni nelle due versioni, con ingresso alfabetico una volta dalle sigle e una volta dalle relative espressioni. Vengono inoltre creati su disco i due archivi corrispondenti, che contengono anche i codici speciali per la fotocomposizione automatica.