1

Capitolo A05
procedure di supporto alla esposizione

Contenuti delle sezioni

a. opportunità di automatismi di supporto a esposizioni-ST p. 4
b. schematizzazione delle esposizioni-ST p. 6
c. struttura dei kapitoli p. 8
d. compiti degli automatismi di supporto p. 10
e. prestazioni di controllo e operazioni diagnostiche p. 14
f. componenti di riferimento derivabili di una esposizione p. 17
g. sostenere la controllabilità del testo p. 19
i. scelte per i programmi di supporto p. 23
j. strutture di dati per il testo-M p. 25
k. manovre per costruire il testo-M p. 29
m. prospettive di ulteriori automatismi di supporto ai testi ST p. 18

33 pagine

A05 0.01 In questo capitolo vengono presi in considerazione i programmi in corso di sviluppo o solo auspicati che consentono di supportare l'evoluzione dei files dei contenuti di MATeXp e in prospettiva l'evoluzione di esposizioni simili.

Per questo ampliamento di obiettivi potenziali si introduce la nozione di **esposizione-ST**; con questo termine intendiamo una esposizione la cui stesura è in divenire e che presenta nozioni scientifico-tecnologiche, in particolare nozioni afferenti all'area STEM, ossia all'area comprendente Science, Technology, Engineering and Mathematics.
Ogni esposizione-ST deve presentare nozioni della matematica e dei modelli di genere fisico-matematico delle entità (oggetti e processi) del mondo reale che vengono indagati. Quindi MATeXp ha titolo per essere considerata un esempio di esposizione-ST.

A05 0.02 Nel seguito per sostenere i testi del genere accennato presenteremo dei programmi redatti nel linguaggio di programmazione C++ relativamente semplici e utilizzabili con computers comuni, cioè con apparecchiature facilmente disponibili.
Si tratta di programmi messi a punto con modalità artigianali con i quali si sperimentano le operazioni di sostegno annunciate sui files costituenti MATeXp.

Nelle pagine che seguono, per maggiore chiarezza, useremo varie convenzioni apposite.

Chiamiamo "redattore" di una esposizione-ST la persona o il gruppo di persone che fanno evolvere tale documento, sia come documento originale, sia come rielaborazione di uno o più documenti preesistenti, da presentare a lettori con esigenze confrontabili con quelle dei destinatari dei documenti da rielaborare.

Diciamo **testo-ST** la rappresentazione mediante files di un prodotto della redazione di una esposizione-ST.

Come vedremo, di una esposizione-ST **T** si devono considerare tre versioni.
Useremo "testo-S" per il complesso dei files sorgente costituente **T** redatti nel cosiddetto linguaggio-S che nel caso di MATeXp è il linguaggio Plain TEX.
Useremo il termine "testo-M" per denotare le rappresentazioni nella memoria del computer dei files costituenti il testo-S.
Ci serviremo di testo-R per la versione leggibile del testo **T**; tale oggetto, se si utilizza Plain TEX, può essere fornito da files con le estensioni .dvi, .ps, .pdf e da altre derivabili da queste.

Ci serviremo della sigla MTM, interpretabile come acronimo "Mathematical Text Manipulation" o come "Manipolazione di Testi Matematici", per caratterizzare i metodi, i procedimenti e i programmi che hanno lo scopo di facilitare e rendere più controllata l'evoluzione delle esposizioni-ST.

A05 0.03 Conviene ribadire che i programmi-MTM che vengono messi a punto e sperimentati con MATeXp si servono di strumenti software comuni e sostanzialmente semplici che possono essere utilizzati con modalità artigianali.
Per i metodi e i procedimenti che possono essere prospettati in generale per le esposizioni-ST si possono invece prevedere prestazioni anche molto diverse.
Le esigenze che si potrebbero prendere in considerazione sono molteplici, gli strumenti pensabili possono dispiegarsi in varie direzioni; possono venire prospettate attività di lunga durata che richiederebbero notevole impegno, soprattutto per la loro validazione in relazione a sperimentazioni accurate.

Quindi nel seguito distingueremo tra generalizzazioni graduali delle procedure-MTM per MATeXp in corso di completamento e già in parte utilizzate per recenti fasi di crescita di MATeXp e prospettive di maggiore ambizione che richiedono più competenze e più risorse complessive.

Alle procedure più ambiziose si possono assegnare compiti facilmente definibili a grandi linee, ma sensibilmente più impegnative da realizzare.
Queste richiedono strumenti software industriali piuttosto che artigianali, sperimentazioni con contenuti effettivi che dovrebbero essere resi ampiamente utilizzabili da autori ai quali non ha senso chiedere di addentrarsi nei loro dettagli tecnici.
Pensiamo soprattutto alle prestazioni di sistemi per la codifica resi disponibili dall'inizio del 2026 come Claude Mythos 5.

Queste procedure più ambiziose pensiamo che vadano applicate in particolare a esposizioni di nozioni utili ad applicazioni circoscritte della matematica che si ritiene opportuno tenere collegate a precise nozioni generali.

Per le procedure avanzate si prospettano attività per gran parte delle quali ora sono solo prevedibili alcune finalità e si può presumere che richiederanno studi e sperimentazioni piuttosto impegnativi.

A05 0.04 Vediamo per sommi capi le finalità delle procedure-MTM disponibili o di prossima realizzazione.
Esse intendono sostenere la coerenza lessicale e strutturale del testo corrente, facilitare la rilevazione di alcuni tipi di mancanze del testo, e agevolare alcuni tipi di modifiche, di unificazioni e di differenziazioni di brani del testo.
Si tratta di operazioni sul testo per le quali un redattore dovrebbe tenere sotto controllo centinaia o migliaia di pagine, e quindi dovrebbe spendere molto tempo in controlli solo minuziosi.

Un'altro compito assegnato alle prime procedure MTM consiste nel generare automaticamente parti del testo logicamente ricavabili dalle rimanenti, in primis indici delle occorrenze di termini e di simboli specifici.

Le procedure MTM concretamente consistono in programmi scritti in un sottoinsieme ben delimitato del linguaggio C++, che in buona parte si trovano nella tradizionale area del linguaggio C.
Esse quindi richiedono competenze di programmazione piuttosto limitate, contenute in quelle presentate nei capitoli J10, J11, J12, J14 e J18.

Questi programmi vengono messi a punto con il sistema di sviluppo MicroSoft Visual Studio e possono essere utilizzati su un qualsiasi personal computer.

A05 0.05 Presentiamo brevemente i generi delle manovre in corso di implementazione.

Analisi della struttura dei kapitoli-files (kapf) e relativi controlli di correttezza.

Analisi statistica dei kapf e presentazioni dei risultati.

Trasformazioni di files `.tex` in nuovi files `.tex` per migliorarne alcune caratteristiche di prima evidenza, come organizzazione formale delle righe e numerazioni delle ripartizioni.

Produzione files di supporto `.txt` e `.htm`.

Generazione di indici definitivi o da sottoporre a una ridotta postproduzione, in particolare indici pubblicabili sul Web.

A05 0.06 Aggiungiamo qualche parola sulle procedure che intendono sostenere lo sviluppo di presentazioni di argomenti scientifico-tecnologici diverse da MATeXp, ma con caratteristiche simili.

Si possono prospettare presentazioni redatte anche in linguaggi diversi da Plain TEX, ma in ogni caso contenenti formule matematiche e diagrammi schematici riguardanti andamenti quantitativi e comportamenti riconducibili a proprietà strutturali, che si servono di termini introdotti da definizioni accurate e coerenti.
A queste presentazioni inoltre si chiede abbiano una ben definita organizzazione che prevede ripartizioni organizzate in modo gerarchico, come quella di MATeXp riguardante tomi, kapitoli, sezioni e paragrafi e con sistemi di collegamenti interni che seguono regole di completezza verificabili da automatismi.

Qualcuna delle considerazioni che seguono potrebbe servire anche a testi redatti in linguaggi diversi da Plain TEX, purché preveda una versione sorgente esprimibile con caratteri ASCII o Unicode e una versione leggibile derivabile dalla precedente mediante automatismi ampiamente disponibili.
Hanno queste caratteristiche i linguaggi HTML, XML e il linguaggio dei sistemi Wikimedia.

A05 a. opportunità di automatismi di supporto a esposizioni-ST

A05 a.01 Come per la massima parte delle attività umane, anche la redazione di testi può trarre vantaggio dalla disponibilità di automatismi in grado di aiutarle.

Questo genere di supporto a testi in corso di sviluppo si può muovere in più direzioni. Qui cominciamo con presentare a grandi linee le operazioni di supporto che vengono concretamente sperimentate su MATeXp, operazioni che corrispondono alle esigenze redazionali più evidenti e che si possono considerare le più semplici da soddisfare.

Conviene tuttavia premettere alcune considerazioni generali sui testi la cui redazione si intende sostenere con automatismi, quale che sia il generi dei testi.

A05 a.02 Per avere automatismi di sostegno più efficaci, più efficienti e di uso più sicuro è opportuno che i testi soddisfino regole formali piuttosto ben definite.
Più in particolare, soprattutto per test-ST, servono regole lessicali e sintattiche; a queste ultime si possono aggiungere regole della struttura complessiva per ciascuno dei documenti. Per queste regole useremo i termini **regole-LS** e **regole-LSS**.
Come esempi di regole-LS ricordiamo le regole dettate dal Sistema Internazionale per la scrittura delle grandezze fisiche, ben precise anche per quanto riguarda caratteristiche tipografiche. Ricordiamo anche le loro varianti definite da importanti associazioni scientifiche e professionali che devono rispettare tradizioni e opportunità molto sentite nelle rispettive aree di attività dovute a esigenze pratiche di misurazione, di movimentazione e di trasformazione degli oggetti di loro interesse.

Nel caso di testi trattati con sistemi tipografici come Plain TEX, vi sono precise regole-LSS riguardanti le forme delle macro, le regole sintattiche delle espressioni matematiche e la strutturazione complessiva dei documenti.
Il loro rispetto tuttavia qui non costituisce un problema in quanto si suppone di operare solo su files .tex che hanno superato il vaglio della compilazione e della produzione di testi-R leggibili di estensione .dvi .

A05 a.03 Si osserva che se per un testo si stabiliscono regole-LS più vincolanti si possono proporre automatismi più efficaci, efficienti e più affidabilmente controllabili. Questi però in genere riducono la libertà espressiva del redattore.
Serve quindi un certo equilibrio tra regole vincolanti e possibilità espressive.

Si osserva anche che aumentando la qualità e la varietà delle prestazioni degli automatismi si possono far crescere sia il rigore effettivo delle regole, che le possibilità espressive delredattore.

Va quindi decisamente auspicata la crescita nel tempo degli automatismi disponibili, nel senso della loro portata.
Questa crescita evidentemente può svilupparsi anche grazie all'aumento delle esperienze dovuto all'effettivo utilizzo dei programmi di supporto.

A questaauspicta crescita possono contribuire in misura rilevente anche i più recenti sistemi AI per il controllo delle esposizioni discorsive e per la messa a punto assistita del software.

A05 a.04 Torniamo alle indicazioni a grandi linee sulle esigenze dei programmi in corso di sperimentazione su MATeXp.

Controllare che il testo rispetti i requisiti che sono imposti dalle sue finalità e suggerire interventi correttivi ai redattori.

Ricavare dal testo che viene via via scritto dal suo redattore dei dati di riferimento che possono migliorare la sua fruizione da parte dei lettori ai quali è rivolto; esempi: tavola dei contenuti, indice di termini ed elenchi di nomi di persone, iniziative e prodotti citati.

Effettuare trasformazioni formali per adeguare l'organizzazione del testo ad aggiunte e a cambiamenti di una certa estensione; due esempi di cambiamenti: eliminazione, inserimento o scambio di posto di alcune parti del testo; adozione di nuovi termini o di nuovi simboli per una parte definita dell'esposizione.

Bisogna osservare che le aggiunte e i cambiamenti di contenuti implicanola necessità di adeguare tutti i complessi di dati di riferimento, indici necessariamente ordinati, tabelle indicazioni statistiche, ...
Questi adeguamenti se fatti manualmente sono evidentemente faticosi e lenti; è decisamente opportuno che siano automatizzati.

A05 a.05 Le precedenti opportunità sono particolarmente rilevanti per tutti i testi-ST per due motivi.
(1) Ai contenuti dei testi-ST si richiede accuratezza e coerenza espositiva più di quanto accada per i testi degli altri generi, ad esempio alla narrativa di evasione.
(2) L'importanza della efficienza nello sviluppo di questi testi che deriva dalla attuale rapida evoluzione delle conoscenze in tanti settori disciplinari; degli avanzamenti continui si riscontrano sostanzialmente in tutto l'ecosistema STEM a causa della crescita generale delle interazioni fra le diverse discipline.

Crescono quindi le occasioni e le iniziative nelle quali risulta opportuno e spesso necessario avvalersi di conoscenze, di competenze, di materiali e di interi dispositivi sviluppati in ambienti STEM specialistici.

La problematica della rapidità e della accuratezza degli aggiornamenti ad ampio raggio attualmente costituisce un importante fattore di crescita complessiva. Infatti va osservato che molte iniziative innovative, non necessariamente di area STEM, ma con forti conseguenze economiche e sociali, richiedono di essere sviluppate in ambienti multidisciplinari.

Inoltre un atteggiamento di multidisciplinarità risulta sempre più necessari anche sul piano dell'orientamento di tutte le iniziative che vogliono essere innovative, che si pongono obiettivi accuratamente mirati e che cercano di avere ricadute di medio-lungo termine.

A05 a.06 Cerchiamo ora di capire quali tipi di azioni a favore dello sviluppo di MATeXp e di testi simili si possono chiedere ad automatismi realizzabili con risorse limitate e in tempi contenuti.

Innanzi tutto è necessari che il testo da supportare abbia una struttura organizzativa chiaramente definita che sia il più possibile accuratamente rispettata in tutte le sue parti.

Risultano allora importanti i programmi diagnostici in grado di controllare il rispetto delle regole strutturali. Infatti nel corso di una lunga attività di redazione le possibilità di derogare alle regole sono concrete; anche a causa della tendenza dei redattori di migliorare continuamente i contenuti, sia nelle fasi nelle quali si rivedono parti scritte in precedenza, che alla conclusione di una fase nella quale si sono effettuate rilevanti ampliamenti del testo.
Questi miglioramenti spesso si effettuano in fasi successive e riferendosi a diverse parti della normativa e a diversi livelli di accuratezza. I programmi quindi devono facilitare gli interventi graduali dei redattori.

Monitoraggio del procedere degli sviluppi, in particolare con indicazioni statistiche sugli aggiornamenti stessi.
Controllo e monitoraggio dei cambiamenti apportati riguardo al rispetto delle regole strutturali, al mantenimento della omogeneità delle scelte stilistiche e al rispetto della coerenza terminologica.

TOC, indici di termini, elementi di notazioni, persone, opere, pubblicazioni.

Indicazioni di carenze di definizioni e di sviluppi

A05 b. schematizzazione delle esposizioni-ST

A05 b.01 Per poter sostenere efficacemente mediante automatismi l'evoluzione di testi-ST è necessario definire in modo chiaramente condivisibile e algoritmicamente gestibile uno schema di organizzazione strutturale per la lunga sequenza di caratteri che si intende tenere sotto controllo.

Questo schema deve consentire di monitorare i vari requisiti che la redazione deve sforzarsi di rispettare nel corso della sua evoluzione.

Un requisito primario per un testo che si intende sostenere è l'essere costituito da un complesso di parti ben definite che chiamiamo kapitoli, ciascuno dei quali viene registrato in un file sequenziale simbolico dotato di opportune proprietà volte a sostenere leggibilità, controllabilità algoritmica e gestibilità nelle varie fasi dello sviluppo. Ricordiamo che un kapitolo può essere un capitolo o un fascicolo.
Per sottolineare i ruoli dei kapitoli, soprattutto in relazione al loro controllo con programmi effettivi (i programmi MTM) ad essi associamo anche il termine **kapf**.

A05 b.02 Presentiamo esempi di altri requisiti che si vogliono rispettati dalle esposizioni-ST; per maggiore precisione conviene fare riferimento ai *testi*-S dei suoi kapitoli, ovvero dei suoi capitoli e dei suoi fascicoli.

Ciascun kapitolo deve essere facilmente identificabile (e quindi indirizzabile) e il complesso dei kapitoli si può considerare un insieme totalmente ordinato. Più precisamente ogni kapitolo è indentificabile con un trigramma di lettere e cifre decimali (sulla cui composizione ritorneremo); e l'ordine lessicografico dei trigrammi rispecchie l'ordine totale dei kapitoli.

Ciascun kapitolo è costituito da ripartizioni di due livelli ben identificabili; più precisamente ogni kapitolo si può considerare una sequenza di sottosequenze che chiamiamo "sezioni" e ogni sezione si può considerare una sequenza di sottosequenze che chiamiamo "paragrafi";

Ogni sezione è identificata da un "tetragramma" costituito dal trigramma del corrispondente kapitolo e da un quarto carattere che può essere una cifra decimale o una lettera; la progressione delle sezioni nel kapitolo è rispecchiata dall'ordine delle cifre e delle lettere, convenendo che le cifre precedono le lettere.

Ogni paragrafo è identificato da un "exagramma", ossia da una sequenza di 6 caratteri, costituita dal tetragramma della corrisponente sezione seguita da una coppia di cifre decimali con esclusione di "00"; la progressione delle sezioni nel kapitolo è rispecchiata dall'ordine delle coppie di cifre decimali.
Per facilitare la riorganizzazione delle parti di una *esposizione* non si richiede la precisa progressione dei trigrammi, dei tetragrammi e degli exagrammi identificativi. Inoltre la prima sezione di ogni kapitolo e alcuni suoi paragrafi non sono presentati esplicitamente, ma si possono individuare algoritmicamente.

A05 b.03 Vediamo ora cosa occorre chiedere alla schematizzazione dei testi da supportare.

Risulta opportuno fare riferimento a tre diverse versioni per ciascuno dei kapf di questi testi.

La prima versione costituisce il **testo-S** già introdotto: è il testo sorgente prodotto e rielaborato dal redattore, scritto in caratteri trattabili senza ambiguità con comuni text editors. In particolare prendiamo in considerazione files costituiti da caratteri ASCII primari.

Dobbiamo poi avere presente il cosiddetto **testo M**, il testo ricavato dal testo-S e organizzato nella memoria del computer incaricato di eseguire le elaborazioni sui kapitoli; conviene osservare subito che le elaborazioni saranno piuttosto diverse nei casi dei capitoli e dei fascicoli, ma che in ogni caso saranno impegnative e questo implica che è necessario che il testo-M contenga molti dati non esplicitati nel

testo-A, ma derivabili dalla sua analisi, sottoposti a controlli di coerenza e completezza e in grado di sostenere le trasformazioni che vengono richieste ai programmi MTM.

Si deve tener conto in modo ben preciso di quello che chiamiamo **testo-R**, testo leggibile normalmente su carta o su schermo che viene ottenuto come prodotto per il rendering dal testo-A (possibilmente sostenuto da un testo-M controllato positivamente) sottoposto al programma traduttore WinShell.

Va ricordato che le esigenze della leggibilità sono determinanti per l'efficacia di tutti i possibili automatismi per il supporto di testi sorgente renderizzabili.

A05 b.04 Un testo-M conviene considerarlo organizzato da una gerarchia di ripartizioni rappresentabile da una arborescenza distesa a più livelli che distinguiamo con interi naturali.
Abbiamo già visto i kapitoli, che consideriamo le componenti di una ripartizione del *testo*-S alla quale assegnamo il livello 2.
Ogni kapitolo va visto come sequenza di sezioni, componenti di una ripartizione che poniamo al livello 3.
Ogni sezione va vista come sequenza di paragrafi, le ripartizioni poste al livello 1), le sezioni componenti di una ripartizione che poniamo al livello 4.

Si possono inoltre considerare le righe dei files come componenti di una ripartizione al livello 5 e i caratteri delle righe come componenti di una ripartizione al livello 6.

Inoltre l'intero *testo*-S si può collocare a un formale livello 0, in quanto contenitore complessivo ripartito dalla sequenza dei tomi.

I kapitoli si possono significativamene raggruppare in successivi "tomi" caratterizzati dalla prima lettera maiuscola dei trigrammi dei loro kapitoli; i tomi vanno visti come componenti di una ripartizione al livello 1; va detto questa ripartizione interessa solo in relazione ad attività complessive di redazione e che sono caratterizzate
Nel caso di MATeXp abbiamo un tomo A contenente l'avvio della presentazione, tomi B, C, D, G, I, J, K, P e T riguardanti macroaree della matematica, tomo W comprendente i capitoli del prontuario e tomo X comprendente i fascicoli dei riferimenti, in gran parte ricavabili aotomaticamente dai *testi*-A dei vari capitoli.

La suddetta gerarchia di ripartizioni fornisce a un testo una struttura piuttosto utile che chiamiamo **struttura kspr**, termine che segnala la importanza prevalente delle ripatizioni in kapitoli, sezioni, paragrafi e righe.

Come vedremo, è opportuno anche distinguere le **sezioni le esplicite** formalmente ben distinguibili all'interno di ogni kapitolo, dalle sezioni nonesplicite e distinguere i **paragrafi espliciti** formalmente ben delimitati all'interno di ogni sezione dai nonespliciti.

Nelle elaborazioni, ossia nelle esecuzioni dei programmi MTM, le componenti di ciascuna delle ripartizioni dei vari livelli vengono convenientemente caratterizzate da interi progressivi; per questi si possono adottare o i numeri naturali (0, 1, 2, ...) oppure i positivi (1, 2, 3, ...).

Vedremo anche che conviene servirsi di due numerazioni: la prima la chiamiamo **numerazione sommersa**, serve per le ripartizioni nonesplicite e viene utilizzata nelle elaborazioni concernenti il testo-M; la seconda, detta **numerazione emersa**, concerne le ripartizioni esplicite, è rilevabile interamente nel testo-S e nel testo-R fino al livello 4 (non ai livelli delle righe sorgente e dei caratteri).

A05 c. struttura dei kapitoli

A05 c.01 Descriviamo qui la struttura dei kapitoli, ossia dei files che registrano i loro *testi*-A scritti nel linguaggio Plain TEX.

Innanzi tutto occorre segnalare che la maggior parte dei kapitoli seguono quella che chiamiamo "struttura prevalente", mentre i restanti seguono quella che chiamiamo "struttura ridotta".

Qui iniziamo a presentare la struttura prevalente e successivamente otterremo la ridotta attraverso mere eliminazione di elementi della struttura prevalente; procederemo dalle ripartizioni più comprensive, le sezioni, a quelle via via meno estese.
Per gli elementi della struttura ci serviremo spesso degli identificatori usati nei programmi MTM per tali entità, in modo di facilitare il collegamento fra le pagine che seguono e i testi dei programmi.

A05 c.02 La sequenza delle sezioni di ogni kapitolo ha le componenti identificate da una cosiddetta matricola trattata dall'identificatore sma con i possibili valori sma = 0, 1, 2 , ...
I successivi paragrafi di ciascuna sezione sono trattati da una loro matricola trattata dall'identificatore pma che assume i successivi valori interi naturali pma = 0, 1, 2 , ...

Nella sezione 0 si distinguono un paragrafo iniziale con pma=0, un paragrafo dedicato alla TOS, tavola delle sezioni con pma=1 ed eventuali paragrafi espositivi dedicati alla presentazione del kapitolo i quali, rispettivamente, hanno pma = 2, 3, ...
Questi paragrafi di presentazione, non sempre presenti e in genere non più di 3, iniziano con righe di intestazione aventi le forme

```
\prg{0.01}  ,  \prg{0.02}  ,  \prg{0.03}  ,  ...
```

A05 c.03 Il paragrafo iniziale di ogni kapitolo è costituito da 7 righe che seguono schemi abbastanza rigidi esprimibili con le espressioni che seguono.

`%^kI` *specifica del kapf, spazio, eventuale commento*
`\input mn001fop \input mn000MCR %\input pictex`
`\def\titq{`*titolo del tomo*`}`
`\def\trgq{`*trigramma del kapitolo*`}  \def\data{`*DATA*`}`
`\def\tikq{`*titolo del kapitolo*`} %^k \prg %\sct{ npg{`*NPG*`}` *trigramma del kapitolo*
`\input mn002hfp \input mn000Hyph.txt`
`\def\sct#1{{\FbSd \trgq~#1}} \def\prg#1{{\FaSd \trgq~#1}~} %\prg`

Nelle forme precedenti la sigla segnaposto *DATA* rappresenta la data nella quale il kapitolo ha ricevuto l'ultimo aggiornamento che il redattore considera significativo, data nella forma "ordinata" yyyy-mm-dd ; la sigla *NPG* invece rappresenta il numero delle pagine del kapitolo in esame.

Molti elementi delle 7 righe precedenti sono stringhe con il compito di rendere facilmente analizzabile il testo sorgente del kapitolo singolo e i sorgenti di più kapitoli utilizzando semplici comandi "find" disponibili in molti text editors, ad esempio in PSPad e in Notepad++.

Nel paragrafo iniziale le 7 righe iniziali del kapf possono essere seguite da righe di commento utili per il redattore; queste hanno tipicamente lo scopo di ricordagli parti del kapitolo che vanno riviste, ad esempio parti che sono state redatte frettolosamente, oppure che si considerano in attesa di una modifica che le adegua a nuove scelte di contrenuto o di forma di una consistente portata; queste scelte potrebbero essere desiderate ma ancora poco definite e le modifiche sopra accennate potrebbero richiedere tempo.

A05 c.04 Il secondo paragrafo della sezione iniziale, con pma=1, inizia con la prima riga della tavola dei titoli delle sezioni esplicite del kapitolo (in sigla TOS), riga avente la forma

`\TOS%\prg`

Seguono le righe dedicate alle successive sezioni espositive del kapitolo.
Ciascuna riga ha la forma

`{`*LdS*`.}` *titolo della sezione* `p.` *PIdS* `\f \p %\sct{`

La sigla *LdS* sta per lettera caratterizzante la sezione; la sigla *PIdS* sta per pagina del testo-R, de testo leggibile, nella quale inizia la sezione.

La sequenza delle sezioni è terminata dalla riga della forma

`\npg{`*NPG*`}%k̂ \prg`

Il segnaposto *NPG* deve fornire lo stesso numero di pagine presentato nella quinta riga del primo paragrafo (con pma=0) 0 della prima sezione del kapitolo.

La coerenza richiede che il trigramma del kapitolo che compare nella specificazione del kapitolo presente nella prima riga della prima sezione coincida con quello presente nella quarta riga (rma=3) dopo la scrittura `trgq`.
Il secondo paragrafo e la prima sezione si concludono con la riga, in genere bianca, che precede la prima riga del terzo paragrafo oppure con la prima riga della seconda sezione del kapitolo, ossia della prima sezione esplicita (ed espositiva).

A05 c.05 Le sezioni esplicite presentano un primo paragrafo (pma=0) riguardante l'intestazione della sezione e una sequenza di paragrafi espliciti che inizia con la riga di intestazione del paragrafo.

L'intestazione della sezione prevede tre righe (rma=0, 1, 2). Le prime due hanno la forma
`\npg{sct{ {`*lettera*`.}` *titolo della sezione*`} %\prg`
`\JV`
L'ultimaè una semplice riga bianca.

Seguono i paragrafi espliciti; ciascuno di essi inizia con la riga di intestazione della forma
`\prg{`*LdS*`.nn}` *possibile commento o nido-dn*
Può seguire una richiesta di a capo `\JP` e quindi si hanno le righe del paragrafo.
Nelle sezioni si possonob avere linee di commento caratterizzate da un carattere `%` iniziale.

A05 c.06 L'ultima sezione di un kapitolo in genere si conclude con il suo ultimo paragrafo.
Questo si conclude con la riga, spesso riga bianca, che precede l'ultima sezione del kapitolo.
Questa è costituita dalla ultima riga del file presa in considerazione dal traduttore WinShall che deve avere la forma
`\endkap\bye %^k \sct{ \prg{`
Questa riga può essere seguita da righe che vengono ignorate dal traduttore per il rendering, ma possono essere utili per il redattore.

A05 d. compiti degli automatismi di supporto

A05 d.01 Il testo di MATeXp o di un'altra esposizione-ST che cerca di essere considerata ampiamente attendibile e proficuamente utilizzabile per applicazioni concrete, si può considerare una struttura informativa costituita da parti tendenzialmente discorsive e e da parti più tecniche e operative con contenuti che dipendono da un complesso di convenzioni specialistiche che devono essere presentate con definizioni coerenti.

A questo proposito è opportuno distinguere due versioni di un testo-ST: la versione sorgente o testo-ST-S da pensare curata da un redattore, costituita da file di caratteri ASCII scritti nel linguaggio Plain TEX, e la versione leggibile o testo-ST-R, ottenuta come trasformazione della precedente fornita dal traduttore WinShell.

Occorre osservare che il traduttore davanti a una riga del testo-ST-S ignora tutti i caratteri che seguono un carattere "**%**" che non sia preceduto da \ ; questi caratteri costituiscono quelle che chiamiamo **zone invisibili** del testo-S, zone che possono essere utilizzate dal redattore per formulare commenti utilizzabili nel procedere dello sviluppo dell'esposizione.

Ogni testo-ST-S può essere analizzato e rielaborato da programmi che intendono sostenere lo sviluppo della esposizione; questi strumenti li chiamiamo **programmi MTM**.

Al testo-ST-S sorgente, per essere efficacemente controllabile, si chiede di osservare un complesso di regole piuttosto ampio e articolato, che innanzi tutto deve fornire indicazioni interpretabili senza ambiguità dai programmi MTM.

A05 d.02 Il contenuto di un testo-ST si suppone che sia in progressivo sviluppo e in genere si richiede che sia facilmente adattabile a nuove esigenze; questa adattabilità va considerata un requisito di una gran parte dei documenti prodotti dagli studi nelle discipline scientifiche e tecnologiche.

È inoltre auspicabile che anche le regole di una struttura informativa del genere exposizione possano essere facilmente adattate al presentarsi di nuovi tipi di richieste che possono essere avanzate in relazione alle nuove esigenze che possono provenire da varianti dei loro contenuti.

Queste varianti si può immaginare possano derivare dai vari generi di cambiamenti di scenari che si riscontrano nell'evolversi delle attività dell'attuale mondo della ricerca e dello sviluppo.

Assistiamo continuamente al rendersi disponibili di nuovi strumenti di indagine (sistemi e metodi computazionali, strumenti di misura, materiali costitutivi, processi produttivi, ...), al presentarsi di nuove opportunità applicative, al rendersi evidenti difficoltà precedentemente sottovalutate nelle previsioni riguardanti la scalabilità delle realizzazioni.
Questi cambiamenti di scenari sono spesso poco prevedibili, ma è ragionevole prevedere che sia altamente probabile incontrarne più d'uno.

A05 d.03 Tornando alle regole imposte alle esposizioni-ST, c'è da aspettarsi che in molte circostanze risulti opportuno definirle con maggiore precisione per riuscire a realizzare procedimenti di monitoraggio più tempestivi o più approfonditi.

C'è da aspettarsi anche che le regole di una esposizione-ST si debbano adattare a richieste di aggiunta di nuovi tipi di contenuti oppure a richieste di adattamenti a nuovi destinatari con caratteristiche di formazione e di interessi diverse da quelle dei destinatari iniziali.
Gli adattamenti delle regole risultano necessari in particolare quando si cerca di rendere una esposizione più facilmente fruibile, oppure più adattabie a varianti di contenuti, oppure di più agevole manutenzione.

Potrebbe anche rendersi necessario soddisfare nuove esigenze riguardanti obiettivi più aggiornati, o quando si volessero aggiungere obiettivi che risultano ragionevolmente raggiungibili dopo che sono state rese superabili difficoltà precedentemente giudicate fuori portata.

A05 d.04 Solo a un testo-ST che viene redatto rispettando precise regole lessicali, sintattiche e strutturali risulta agevole applicare procedure computazionali in grado di analizzare le sue pagine per sostenere il suo sviluppo in varie direzioni.

Innanzi tutto ogni testo di una certa complessità richiede di essere corredato di indici che facilitino la sua fruizione.
Tali indici vanno ricavati da determinate parti del testo stesso; in assenza di automatismi è necessario un attento lavore dell'autore o di un suo incaricato; nell'ambito di una casa editrice di consistenti dimensioni questo lavoro è affidato ad un cosiddetto "editor".

Questo lavoro viene spesso sostenuto ricorrendo a procedure semiautomatiche; se si riesce a disporre di procedure di supporto più automatiche e più versatili (ovviamente più impegnative) si possono avere vantaggi molto più rilevanti.

Gli indici di un testo-ST in genere sono più necessari e più impegantivi di quelli che servono ad un testo prevalentemente discorsivo, in quanto il suo linguaggio (a partire dal lessico) deve essere molto disciplinato.

Un importante supporto automatico a un testo riguarda la correzione ortografica e la ricerca dei possibili refusi ed è ormai prassi comune.
Nel caso di testi-ST anche questa ricerca risulta particolarmente impegnativa a causa della maggior cura richiesta per la precisione lessicale, e sintattica; questa in particolare riguarda le formule matematiche, le definizioni di algoritmi e i testi sorgente di programmi, generi di testi che devono rispettare regole assai precise.

A05 d.05 Sensibilmente più impegnativo è il controllo della coerenza semantica che riguarda innanzi tutto la completezza e la consequenzialità delle argomentazioni.
Bisogna inoltre considerare che spesso si rendono opportune delle deroghe alla coerenza del linguaggio, in quanto una coerenza integrale può portare a definizioni ed enunciati molto pesanti alla lettura.
Questa possibile pesantezza Nei testi descrittivi e narrativi, evidentemente, viene accuratamente evitata; anzi spesso il ricorso alla ambiguità delle situazioni e delle metafore, per la possibilità di suscitare emozioni costituisce un pregio letterario.
Al contrario l'ambiguità e la vaghezza del linguaggio nelle argomentazioni delle discipline ST possono essere assai dannose.
Per i testi-ST-S quindi si prospettano dei compromessi equilibrati fra rispetto rigoroso della coerenza e leggibilità delle pagine.
In effetti la pesantezza delle esposizioni ligie alle definizioni viene ampiamente evitata ricorrendo a semplificazioni, a espressioni figurate, ad analogie e a metafore.
Ad esempio in molte presentazioni si confondono oggetti geometrici con loro lunghezze, aree o volumi; si confondono insiemi finiti con i numeri dei loro elementi; si confondono funzioni con valori da loro assunti.
Inoltre trattando argomenti impegnativi accade di confondere una entità piuttosto astratta con una delle sue rappresentazioni, in genere con quella che l'autore pensa essere la più adottata.

Va anche segnalato che i supporti computazionali a un corpo informativo disponibile in forma digitale e rispettoso di regole chiare, come cerca di essere MATeXp, consente di facilitare in misura rilevante la sua riutilizzazione per varianti dei fini, in particolare per specifici scopi didattici e divulgativi.

Questo ultimo genere di prestazioni riguarda MATeXp solo per le sue possibili riprese per finalità ben delimitate, ma conviene prospettarle, in quanto la loro implementazione può avvalersi di parecchi moduli di programma e di parecchi schemi operativi che devono far parte della strumentazione computazionale riguardante le accennate destinazioni didattiche e divulgative.

A05 d.06 Per lo sviluppo di MATeXp inizialmente si è proceduto nel controllo del testo sorgente servendosi di alcuni source editors liberamente disponibili, PSpad e Notepad++.
Successivamente si sono utilizzati i primi programmi MTM messi a punto specificamente per MATeXp utilizzato sostanzialmente solo in modo procedurale al livello della applicazione desktop e avvalendosi solo di librerie di comune disponibilità.

Vediamo in breve le regole che vengono fatte rispettare e gli arricchimenti prodotti automaticamente.

Il testo di MATeXp è organizzato in tomi, kapitoli, sezioni, paragrafi e righe secondo un ordine semirigido, ossia rappresentabile attraverso sigle che seguono precisamente un ordinamento e lessicografico, rendendo il testo-S totalmente ordinato e dotato di ripartizioni totalmente ordinate.

Tuttavia le sigle presentano frequenti lacune motivate dalla opportunità di favorire le possibili e prevedibili modifiche future, soprattutto le possibili aggiunte integrative.

A05 d.07 In ogni paragrafo si distinguono sequenze di righe discorsive e sequenze di righe dedicate a formule matematiche e a tabelle.

Le prime sono scritte in un italiano ricco di tecnicismi piuttosto specifici e che tiene conto della innegabile importanza delle espressioni in inglese e del lessico, spesso gergale, degli addetti ai lavori.
Le seconde riguardano formule racchiuse tra $ e $, oppure tra $$ e $$, costituenti nidi specifici come `{\dn` *termine* `}`, `\YA{`*autore*`}`, `\Yw{`*opera*`}`, `\IL{`*rinvio interno*`}`... .

Le parti discorsive, oltre alle normali regole, devono rispettare attentamente la precisione lessicale.

Nella presentazione devono essere seguite regole di consequenzialità.
In particolare ogni nozione specifica che non si trovi tra quelle che si presumono ben riconosciute dai lettori previsti, dovrebbe essere usata solo dopo essere stata ben definita.
A questa precedenza della definizione spesso si deve derogare per evitare eccessive pesantezze; ciascuna di queste deroghe tuttavia dovrebbe essere esplicitata e motivata e dovrebbero essere segnalati i punti del testo sorgente o i passi della letteratura specialistica nei quali si possono trovare le indicazioni in grado di chiarire ogni possibile dubbio interpretativo.

A05 d.08 Vediamo rapidamente come possano essere utilizzate proficuamente le zone invisibili del testo sorgente.

I brani che restano da sviluppare, da confermare, da ricollocare o da precisare in momenti successivi possono essere marcati mediante stringhe di marcatura che hanno il ruolo di marcatori di problemi lasciati in sospeso.
Queste srtringhe vanno scelte con attenzione per poter essere chiaramente individuate (singolarmente e collettivamente) insieme alla loro finalità quando il redattore intenderà risolvere uno dei problemi in sospeso.
In particolare servono stringhe di marcatura per passi del testo che si prevede debbano essere aggiornati periodicamente.

Oltre alle segnalazioni utili alla gestione del procedere della redazione, nelle zone invisibili, nel testo-S si possono inserire stringhe che vengono utilizzate da procedure che scorrono il testo sorgente per

ricavare indicazioni di vari generi, in particolare rilevazioni statistiche sui contenuti presenti ed elenchi di problemi dello sviluppo lasciati in sospeso.

Attraverso zone invisibili del testo-S si possono inserire in vari brani dell'esposizione macro TEX con nomi per i quali si chiede soltanto che non interferiscano con le altre macro in uso; queste macro consentono di individuare elementi sui quali possono agire precise procedure di supporto.

segnaliamo anche la possibilità di inserire in gran parte dei punti del testo-S macro TEX invisibili, che non hanno alcun effetto sul testo leggibile, ma che possono segnalare con precisione elementi di contenuto considerati incerti o provvisori.

A05 e. prestazioni di controllo e operazioni diagnostiche

A05 e.01 Le prestazioni di controllo dei programmi MTM si basano sulla individuazione nel testo sorgente di elementi qualificanti e sulla verifica della loro accettabilità lessicale, sintattica e della struttura delle ripartizioni tcsp.

Le verifiche sui diversi tipi di elementi che si trovano nel testo-S evidentemente vanno effettuate con manovre che dipendono dalle caratteristiche dei vari tipi di elementi.

Alcune verifiche si effettuano con algoritmi tendenzialmente semplici riguardanti stringhe di lunghezza contenuta; questo accade ad esempio per gli esagrammi che fungono da coordinate dei paragrafi del testo aventi la forma "Tnnlmm" e quindi facilmente riconoscibili.

Per altri tipi di elementi la verifica della accettabilità si realizza stabilendo se la stringa che costituisce l'elemento compare o meno in un determinato elenco di parole o di locuzioni; esempi locuzioni del linguaggio della matematica, elenco di identificatori riservati di un linguaggio di programmazione, macro del linguaggio Plain TEX.
Questi elenchi di stringhe di riferimento vanno definiti con accuratezza; taluni devono essere aggiornati in accordo con le modifiche che si vanno effettuando nel testo-A della esposizione, soprattutto in seguito ad arricchimenti dei contenuti.

Si può pensare che per sviluppare esposizioni controllabili con le procedure del genere MTM convenga partire da un insieme collauato di elementi e di tipi di elementi che si possa ritoccare e far crescere con una certa facilità mentre si procede con le modifiche e con l'ampliamento del testo, avendo cura che risulti sempre relativamente agevole far entrare nel testo elementi di tipi nuovi, precedentemente non previsti, e formalmente trattabili con nuovi algoritmi su stringhe brevi.

In questi ampliamenti, evidentemente, si deve curare la compatibilità con il lessico e con gli algoritmi precedenti.

Occorre segnalare che l'introduzione di un nuovo elemento spesso rende necessaria la revisione di non pochi elementi precedenti.
Questo accade soprattutto quando si introducono generalizzazioni e differenziazioni di entità precedentemente definite in modi che in un successivo momento si possa decidere di precisarli e di raffinarli.

A05 e.02 Presentiamo le manovre riconducibili al controllo della correttezza che sono state prese in considerazione, cominciando con quelle che chiamiamo "individuazioni", manovre che consistono nello scorrimento e nell'analisi dei files `.tex` nei quali è registrato il testo sorgente.

Individuazione dei titoli di capitoli e sezioni al fine di verificare il loro ordinamento e di tenere sotto controllo il loro complesso.

Individuazione dei nidi della forma `{\dn`*termine*`}` , e suo accompagnamento con l'esagramma del paragrafo nel quale si trova.

Una tale operazione la chiamiamo anche "individuazione e localizzazione" del nido dn e serve a tenere sotto controllo il complesso delle occorrenze dei termini più significativi e a evitare che tra di essi si abbiano conflitti di omonimia, ambiguità di sinonimia e varianti usate senza esplicitazione che possano generare confusione o fraintendimenti.

Il termine "individuazione e localizzazione" viene usato anche per denotare altre operazioni simili che stiamo per incontrare.

Individuazione delle occorrenze di notazioni che il redattore ha opportunamente marcato in parti invisibili, anch'esse da accompagnare con gli esagrammi che fanno da coordinate dei relativi paragrafi; servono a controlli simili a quelle sui termini giudicati significativi.

Individuazione e localizzazione delle citazioni di autori o di altre personalità in quanto elementi del testo della forma `\Ya{`*nome*`}` .
Di questi elementi si intende controllare l'ortografia e la reperibilità di informazioni che di essi si trovano sul Web; questi tipi di dati nel caso di MATeXp servono alla redazione e all'aggiornameto del fascicolo X15.

Individuazione e localizzazione delle citazioni di opere fornite da elementi della forma `\Yw{`*titolo*`}` ; questi dati servono solo a tener sotto controllo la loro presenza nel capitolo X21.

A05 e.03 Risulta molto utile anche l'estrazione dei rinvii della forma `\IL{`*coordinata*`}` seguita dalla registrazione della relativa coppia

$$\langle \textit{paragrafo richiamante, ripartizione richiamata} \rangle .$$

Qui il paragrafo richiamante è rappresentato dalla codifica dell'esagramma del corrispondente paragrafo; la ripartizione richiamata, che diciamo anche "bersaglio", può essere un paragrafo, una sezione o un intero kapitolo del testo-S e viene rappresentata, rispettivamente, da un exagramma, da un tetragramma della forma `\IL{`*tetragramma*`}` e da un trigramma della forma `\IL{`*trigramma*`}` .

Serve inoltre l'individuazione delle stringhe alfanumeriche esterne ai nidi delle formule matematiche aventi la forma $...$, la forma $$...$$ o forme simili; tutte queste stringhe devono essere considerate lessicalmente accettabili in ogni punto del testo-S.

La correttezza formale delle espressioni matematiche viene controllata dalle esecuzioni del traduttore WinShell; la loro individuazione può servire al controllo della loro correttezza in quanto elementi delle argomentazioni riguardanti nozioni, definizioni, teoremi e di altre affermazioni matematiche.
Questo genere di correttezza in definitiva è a carico del redattore del *testo*; i programmi di supporto servono a sostenere il redattore fornendoli brani del *testo* riguardanti oggetti simili a quelli che sta attualmente indagando.
Per questo genere di supporto possono rendersi utili appositi elenchi di stringhe controllate o comunque giudicate accettabili che trattano entità, trasformazioni o dimostrazioni collegate in qualche modo ai temi che il redattore sta esaminando. ‘ riesempi con laQuesto si dovrà ottenere con la verifica che ciascuna di esse è presente in un apposito elenco generale delle stringhe .

Questi elenchi devono essere mantenuti con controlli che forse dovrebbero essere effettuari dopo ogni ampliamento dell'esposizione; evidentemente a chi cura questi elenchi si chiede una buona consapevolezza del lessico, della sintassi e dei significati che si stanno adottando nella intera esposizione corrente.

Evidentemente gli impegnativi elenchi sopra accennati sono molto utili a mantenere la coerenza lessicale e sintattica complessiva; inoltre possono contribuire a un elenco che intende costituire una buona base per la individuazione dei possibili refusi nel testo.

A05 e.04 È necessaria anche l'individuazione delle macro della forma `\`*stringa alfabetica* e per ciascuna di esse la verifica se sia presente o meno in uno dei due elenchi disgiunti delle macro del sistema TEX e delle macro del redattore.

Questi due elenchi vengono gestiti diversamente. Mentre il primo elenco è sostanzialmente fissato dal linguaggio, per le macro definite dal redattore si deve garantire che non si abbiano conflitti di omonimia tra di esse e con macro TEX e si deve assicurare che le sinonimie interne siano contenute e giustificate.

Si vuole inoltre che le macro del redattore possano essere ricordate facilmente e questo richiede che abbiano nomi significativi, ossia si trovino in sintonia con i termini che esprimono le entità che le macro coinvolgono.
Si richiede inoltre che il complesso dei loro nomi risulti coerente, e in particolare che risulti abbastanza facile intuire quando ci si trova davanti a due macro tali che una abbia portata più generale dell'altra, oppure tali che una sia più incisiva dell'altra.

Si è trovato che l'efficacia dei nomi delle macro va perseguita soprattutto attraverso successivi adattamenti, ciascuno dei quali può anche riguardare gruppi di macro piuttosto estesi.

A05 e.05 Passiamo ora in rapida rassegna le operazioni di diagnosi dei files sui quali è registrato il testo, operazioni finalizzate a verificare e a mantenere il rispetto della loro coerenza e della loro organizzazione, almeno sul piano formale.

Scorrimento dei files-capitoli forniti da un elenco dato.

Scorrimento di una data selezione delle righe porzione di un dato file.

Scorrimento di una riga di un file fino alla prima occorrenza del segno "%" oppure fino alla sua fine.

Controllo in un kapitolo della coerenza fra indice collocato al suo inizio e sequenza delle sezioni, compresa la verifica delle coincidenze dei titoli.

Controllo dell'ordinamento delle sezioni e dei paragrafi all'interno di un kapitolo con la segnalazione delle eventuali scorrettezze, oppure con il loro aggiustamento attraverso modifiche delle etichette.

Estrazione da alcuni kapitoli dei titoli del kapitolo stesso e delle sue sezioni al fine di assicurare l'aggiornamento dei contenuti degli indici nei capitoli X10, X11 e X14.

Individuazione in alcuni capitoli dei rinvii TnnSmm, verifica della presenza nel testo-S del bersaglio, ossia del corrispondente paragrafo (entro data sezione e dato kapitolo) e raccolta delle coppie $\langle$ paragrafo rinviante , ripartizione rinviata $\rangle$.

Infine segnaliamo la prospettiva di effettuare esami complessivi della consistenza dei vari argomenti trattati e della adeguatezza dei collegamenti tra i vari argomenti basati su occorrenze di termini, su correlazioni fra termini e su collegamenti fra le parti.

Mentre le prestazioni precedenti possono essere realizzate con algoritmi relativamente semplici, queste ultime prestazioni richiederanno un complesso lavoro attraverso tentativi, verifiche e aggiustamenti sensibilmente dipendenti dai contenuti trattati.
==========

A05 f. componenti di riferimento derivabili di una esposizione

A05 f.01 In un testo-ST si possono distinguere componenti derivabili da gran parte del testo rimanente che possono costituire utili **elementi di riferimento** per il testo stesso.
Con questo termine denotiamo elementi del testo che possono facilitare la fruizione del testo stesso, in quanto consentono al lettore di orientarsi rapidamente quando ricerca passi, termini e notazioni che gli permettono di soddisfare sue curiosità.

Elenchiamo rapidamente le operazioni che servono alla generazione di componenti dell'*esposizione*.

Dal testo di ogni capitolo si può determinare la struttura informativa che consente di associare a ciascuna delle sue righe il paragrafo e la sezione di apppartenenza.

Estrazione dalle righe contenenti stringhe dalle forma `%;T` che segnalano la definizione di un termine e controllo della loro correttezza.

Generazione dell'indice in X11 che raccoglie le informazioni precedenti.

Estrazione dalle righe contenenti stringhe delle seguenti forme

`%;Na` , `%;Nc` , `%;Nd` , `%;Ne` , `%;Ni` , `%;Nm`

riguardanti informazioni su una notazione e controllo della loro correttezza.

Generazione dell'indice in X12 che raccoglie le informazioni precedenti.

Ricodifica delle etichette di sezioni e di paragrafi dovute a manovre di inserimento o di eliminazione.

A05 f.02 Per concludere la presente sezione accenniamo ad altre due prospettive per le attività di sviluppo di testi-ST-S.

La prima riguarda la traduzione del suo testo in altra lingua.

La traduzione di materiali-ST di una certa ampiezza può essere significativa, soprattutto in questo periodo di rapida crescita degli interessi nei confronto delle svariate innovazioni scientifiche e tecnologiche e delle loro conseguenze socio-economiche.

Inoltre le recenti innovazioni nel settore della traduzione automatica hanno resi facilmente disponibili strumenti di sostegno per le traduzioni auspicabili tali da renderle fattibili molto più agevolmente e più rapidamente che nel passato.

Si osserva inoltre che molti componenti di una esposizione possono essere inseriti in una sua traduzione senza cambiamenti o in seguito a trasformazioni che possono essere effettuate con specifici procedimenti automatici relativamente semplici.

Questi procedimenti si devono servire di dati lessicali che devono essere messi a punto nell'ambito di una visione complessiva delle componenti (essenzialmente termini ed espressioni) delle formulazioni delle nozioni matematiche e delle operazioni di gestione e trasformazione dei testi.

Qualche tentativo è stato avviato, sempre a un modesto livello artigianale, per trasformare MATeXP nella lingua inglese.
Si è osservato che gli attuali traduttori automatici per essere utilizzati efficacemente richiedono operazioni preparatorie e di finitura assai attente, ma anch'esse facilitabili da meccanismi di elaborazioni di testi realizzabili senza eccessive difficoltà e con l'impegno di risorse piuttosto contenute.

Inoltre emerge l'opportunità di disporre di dizionari specialistici (riguardanti matematica, fisica, informatica, ...) su più lingue che cerchino di rispettare stili lessicali e sintatici di vasta portata.

Per la costruzione e l'aggiornamento di strutture di dati testuali sarebbe decisamente auspicabile il contributo di consistenti competenze.

A05 f.03 Una seconda prospettiva riguarda l'adattamento del testo ad altri linguaggi di supporto alla comunicazione che prevedano la agevole disponibilità di un linguaggio sorgente trattabile con programmi procedurali consolidati, come C++, Java, Python, PHP, ...

Una prima alternativa, riguardante un cambiamento relativamente piccolo, consiste nell'uso di LaTeX invece di Plain TEX.

Una seconda alternativa per il rimpiazzamento di Plain TEX sono i linguaggi per il Web come HTML ed XML.
Questi linguaggi consentono una organizzazione della ipertestualità dei contenuti che può essere molto efficace; in questa direzione si aprono nuove possibilità, accompagnate dai relativi problemi.
Per passare da TEX ad HTML si pongono problemi riguardanti principalmente le formule matematiche, sia quelle tra $ e $ all'interno dei periodi discorsivi che all'interno delle formule autonome tra $$ e $$.

In questa direzione sembra siano richiesti interventi, anche impegnativi, del redattore e di una persona con buona conoscenza sia dei linguaggi di comunicazione coinvolti che dei contenuti da tradurre.

Conviene osservare che le operazioni di traduzione di una esposizione potrebbero essere accompagnte da modifiche di parti dello stesso testo originale che possono costituire miglioramenti nelle formulazioni suggerite da un confronto attento fra le loro versioni in lingue diverse.

Questi miglioramenti di ritorno fanno parte dei miglioramenti che possono provenire da ogni ampliamento di una esposizione di contenuti di buon livello informativo in relazione alle necessarie revisioni dei materiali che possono essere richieste dagli ampliamenti della visione delle nozioni che presentano molteplici collegamenti con altre nozioni rilevanti, ovvero dei materiali che possono essere esaminati fruttosamente da nuovi punti di vista.

A05 f.04 Altri strumenti di comunicazione meritano di essere presi in considerazione.

Uno molto interessante è costituito dai sistemi Wiki: anch'essi consentono di esporre nozioni, idee e punti di vista con prestazioni ipertestuali versatili che presentano indubbi vantaggi, ma per contro richiedono un lavoro non trascurabile per la precisazione dei collegamenti fra nozioni diverse richiesti dai meccanismi di organizzazione dei contenuti che consentono di adattarli alle esigenze del fruitore nel corso delle sue operazioni di lettura e di progressiva riorganizzazione del suo inquadramento degli argomenti che va affrontando.

In una direzione ben diversa si possono cercare versioni di un testo-ST in sistemi di comunicazione di larghissima diffusione, ad esempio in un word processor come Word di MS Office o simili; maggiori indicazioni su questi strumenti non siamo in grado di fornire.

Tutte le varie alternative indicate sembrano avere caratteristiche interessanti per ottenere da un testo esteso e disciplinato estratti circoscritti ragionevolmente controllabili da singoli operatori e con impiego relativamente ridotto di tempi e di risorse computazionali.

A05 g. sostenere la controllabilità del testo

A05 g.01 Attualmente per rendere il contenuto di MATeXp controllabile con efficacia e poco impegno umano si agisce sul *testo* sorgente in due modi.

(1) Si mantiene il *testo* rispettoso di precise regole riguardanti la sua organizzazione complessiva mediante le ripartizioni dei vari livelli e riguardanti l'impiego di un lessico controllato e di notazioni adeguate e coordinate.

(2) Alle parti del *testo* da rendere visualizzabili in sequenze di pagine attraverso files con le estensioni `.dvi` e `.pdf`, si aggiungono **commenti finali** nelle parti delle righe che seguono la prima occorrenza del carattere `%` non preceduto da `\` e si inseriscono macro TeX della forma `\yna` e simili con il ruolo di marcatori di precisi punti del *testo*.
I commenti finali è opportuno siano anch'essi regolati da criteri di omogeneità e chiarezza formale, in quanto devono essere interpretabili senza ambiguità dai redattori nei tempi lunghi richiesti dalla stesura di presentazioni molto meno scorrevoli di quelle impiegate per le narrazioni.

Gli **inserti invisibili** del redattore costituiti dai commenti finali e dai marcatori invisibili servono soprattutto a facilitare gli interventi automatici sul testo sorgente eseguibili mediante comuni editori di testi (PsPAD, Notepad++) e mediante altri programmi di ampia disponibilità (Excel, Regex, ...), in grado di effettuare generici controlli di omogeneità e coerenza dei contenuti.
Controlli automatici è opportuno che siano esercitati anche sugli stessi inserti invisibili.

Essi inoltre contribuiscono a stabilire quali occorrenze di termini, di notazioni e di altre componenti minute con ruoli specifici devono comparire negli indici dei riferimenti interni e contribuiscono a consentire loro elaborazioni piuttosto agevoli. I controlli automatici, soprattutto, sono alla base della generazione automatica dei fascicoli dell'*esposizione*.

A05 g.02 Prima di occuparci delle manovre di supporto allo sviluppo dei testi-ST conviene presentare gli aspetti generali delle loro strutture.

Per questo conviene introdurre tre diverse versioni di ciascuna di queste strutture informative.

La prima versione la chiamiamo testo-S, testo sorgente: si tratta del testo sviluppato dall'agente che chiamiamo "redattore", contenuto in una sequenza di files sequenziali costituiti da caratteri ASCII, ciascun file costituente un capitolo del testo oppure un suo fascicolo ausiliario destinato a contenere informazioni di supporto ai capitoli, soprattutto formate da elenchi di unità informative (titoli di ripartizioni, termini, persone, ...).

Come già accennato ogni capitolo del testo-S contiene contenuti dell'*esposizione* che diventeranno leggibili e commenti trascurati nel rendering curati dal redattore con due generi di scopi: facilitare il proprio lavoro di redazione (che può svolgersi in successivi intervalli temporali anche distanti), e per contribuire allo stessa impianto strutturato dei contenuti.

Ogni file del testo-S va sottoposto al traduttore WinShell che lo trasforma biunivocamente in un capitolo o in un fascicolo costituente il testo-R, testo leggibile in formato inizialmente `.dvi` e successivamente `.pdf`, da presentare ai lettori e da far circolare attraverso canali telematici o attraverso stampati.

Per poter far intervenire procedimenti automatici sui contenuti occorre agire sulla versione costituente il cosiddetto testo-M, testo organizzato nella memoria del computer in modo da poter essere elaborato con i previsti programmi MTM scritti nel linguaggio C++.

Questa versione contiene tutte le informazioni registrate nel testo-S e in più informazioni derivate che hanno lo scopo di rendere effettive e relativamente agevoli le azioni dei sopra accennati programmi.

A05 g.03 Veniamo ora alla struttura prevista per ciascuno dei testi-ST, facendo riferimento sia alla versione-S che alla versione-M e facendo riferimenti approssimati alla più recente disponibilità di MATeXp.

Il testo va considerato come una sequenza di caratteri ASCII che risulta ripartita a vari livelli che presentiamo procedendo dalle ripartizioni più comprensive alle più fini.

La prima ripartizione prevede successivi tomi caratterizzati da lettere maiuscole da considerare in ordine alfabetico crescente, attualmente "ABCDGIPTWX".

Ciascuno dei tomi è costituito da una sequenza di capitoli caratterizzati da trigrammi della forma Tnn, con T lettera maiuscola del tomo ed nn coppia di decimali da 01 a 98 (in MATeXp il numero dei capitoli di un tomo va da 5 a una quarantina); anche i trigrammi procedono in ordine crescente, ma presentano evidenti lacune.
Ogni capitolo è contenuto in un file `.tex` e per ciascuno di essi usiamo anche il termine gergale "kapf".

Ogni capitolo è costituito da una sequenza di sezioni (in MATeXp grosso modo una decina per capitolo). Le successive sezioni sono caratterizzate da lettere mminuscole da considerare nell'ordine alfabetico e con possibili lacune.
Ogni sezione è quindi individuata da un trigramma della forma Tnnl , dove l intende rappresentare la minuscola caratterizzante la sezione nel capitolo.

Ogni sezione è formata da una sequenza di paragrafi, ciascuno caratterizzato da un esagramma della forma Tnnlmm, con Tnnl teragramma della sezione e mm rappresentante una coppia di cifre decimali che vanno da 01 a 99, anch'essi da considerare in ordine crescente e con possibili lacune; in MATeXp grasso modo una sezione contiene da 5 a 10 paragrafi.

Ogni paragrafo, ovviamente. è costituito da una sequenza di righe caratterizzate dai progressivi all'interno di un kapf che conviene esprimenre con interi da 0 in su; amche le linee bianche vanno contate e non ha senso parlare di lacune tra le righe del testo-S.
Con MATeXp si hanno paragrafi con numeri di righe che grosso modo vanno da 10 a 50; infatti si cerca di avere paragrafi che possono essere presentati interamente in una schermata video.

Ogni riga è costituita da una sequenza di caratteri ASCII da considerare numerati da 0 in su.

A05 g.04 Oltre alle sezioni e ai paragrafi ben evidenti nelle versioni-S e nelle versioni-R, nalla versione-M si sono introdotte sezioni e paragrafi rilevanti alla lettura, ma necessarie per avere che per ogni coppia di ripartizioni, la seconda immediatamente più fine della prima, ogni parte della prima (la cosiddetta "partizione over") è costituita esattamente dalla giustapposizione di parti della seconda (la cosiddetta "ripartizione under").

Abbiamo quindi che ogni kapf presenta come sezione 0 le righe dalla iniziale a quella che precede la riga con `\csin`; essa viene considerata costituita da un unico paragrafo.

Nel kapf segue una sezione 1 comprendente le righe da quella con `\csin` nella riga che precede quella contenente la prima riga di una sezione evidente, ossia la riga contenente `\sct`.
Questa sezione presenta come primo paragrafo esplicito costituito dalle linee che forniscono i contenuti delle pagine del capitolo; si possono poi avere successivi paragrafi inizianti con la riga che inizia con `\prg` e prosegue con la caratterizzazione del paragrafo della forma 0.01, 0.02, 0.03 e simili.

Seguono le sezioni esplicite caratterizzate dalle minuscole che rappresentiamo con "S" e quindi identificabili nell'intero testo con i tetragrammi della forma TnnS.

Ogni sezione presenta un paragrafo implicito costituito dalle prime righe della sezione fino alla pro=ima riga del parimo paragrafo esplicito che inizia con la scritture `\prg` e prosegue con l'identificativo del paragrafnel capitolo della forma Smm e che fa parte dell'esagramma della forma TnnSmm che indentifica il paragrafo nel testo.

A05 g.06 Sono previste manovre dei tipi seguenti.

Controllo della organizzazione complessiva dei files-capitoli; prevede la verifica della non ambiguità e dell'ordinamento delle etichette che caratterizzano le ripartizioni costituenti capitoli, sezioni e paragrafi.

Controllo dei tokens, grosso modo stringhe alfanumeriche che possono contenere i segni - e ' , che si trovano nei testi, con la segnalazione dei tokens escluse dal previsto elenco di riferimento.
Per ciascuno dei tokens segnalati sarà deciso dal redattore l'inserimento tra quelli da considerare leciti, oppure la correzione della occorrenza trovata.

Controllo delle occorrense delle macro e verifica della appartenenza di ciascuna di esse all'elenzo delle macro TEX, oppure all'elenco delle macro del redattore; nel caso di assenza si ha può avere la correzione dell'occorrenza, oppure l'inserimento del token nell'elenco che lo riguarda.

Controllo dei termini contenuti nei nidi `tGa`*termine* `}`consistente nella verifica della loro appartenenza al relativo elenco, che nel caso di termine nuovo viene seguita dalla sua accettazione doppure dalla correzione della occorrenza.

Controllo delle notazioni evidenziate invisibilmente come significative; per ciascuna notazione viene precisato il tipo e verificata la sua presenza nell'elenco di quelle previste e nel caso di assenza viene decisa la correzione o l'inclusione nell'elenco.

Controllo delle occorrenze delle stringhe dei tipi previsti nell'elenco comprendente abbreviazioni, specificazioni e variatori contenuto nel file X13 seguito, nel caso di assenza, dalla correzione dell'occorrenza o dell'inclusione della stringa nel fascicolo.

Controllo delle occorrenze di nomi di persone che compaiono nei nidi `\Ya{`*nome*`}` s e, nel caso di assenza di ciascuno di essi, dalla correzione del nome o dalla sua inclusione nell'elenco X15.

Controllo delle citazioni delle opere che compaiono nei nidi `\Yw{`*citazione*`}` seguito, nel caso di assenza, dalla correzione o dell'inclusione nell'elenco X16.

A05 g.07 Presentiamo le operazioni che toccano le strutture informative di riferimento, relativamente semplici, che vemgono richiamate da uno o più tipi di manovre.

Implementazione di un accettatore-trasduttore a stati finiti (AccTrdF) generico fornito da una descrizione equivalente a una matrice di transizione e trasduzione.

Operazioni per la gestione di un elenco che segue un ordinamento di tipo lessicografico basato su una operazione di confronto fra due stringhe controllata da un AccTrdF; le operazioni consistono in inizializzazione dell'elenco, search di una stringa con segnalazione di presenza o assenza, inserimento di una stringa, eliminazione di una stringa.

Operazioni più complete delle precedenti riguardanti stringhe di items accompagnati da numeri e da localizzazioni di occorrenze.

A05 g.08 Operazioni di scorrimento di una lista di files-capitoli o di una parte di sezioni e paragafi di un file-capitolo e delle corrispondenti righe.

Accettazione, qualificazione, eventuale segnalazione diagnostica di una riga di un file sorgente mediante un meccanismo del livello di un AccTrdF.

Correzione mediante un AccTrdF di una linea, di un periodo o di un paragrafo di un file sorgente.

Controllo della definitezza delle ripartizioni in sezioni e paragrafi di capitoli di un testo sorgente, del rispetto dell'ordinamento delle sezioni nei capitoli e del rispetto dell'ordinamento dei paragrafi nelle sezioni.

Conviene osservare esplicitamente che sono esclusi i trigrammi dei capitoli della forma Tnn con nn = 00 e gli esagrammi dei paragrafi della forma TnnSmm con mm = 00.

Gestione di elenchi di stringhe corredate da numeri o localizzazioni di occorrenze, elenchi che seguono ordinamenti lessicografici definiti da una operazione doi confronto basata su un automa AccTrdf.

Esame di un elenco di items che seguono un ordinamento lessicografico, in ciascuno dei quali può essere marcata o comunque individuabile algoritmicamente, qualche parola da evidenziare e sua trasformazione in un più completo elenco KWIC, ossia in un elenco ordinato di items che possono essere permutati circolarmente.

A05 i. scelte per i programmi e per le strutture di dati

A05 i.01 Per denotare i programmi che si stanno mettendo a punto per il controllo del testo di MATeXp useremo anche, come già enunciato, la sigla MTM, acronimo di Mathematical Text Manipulation.
Prospettando automatismi più generali e più ambiziosi useremo anche il più comprensivo MMMC, acronimo di "Man Machine Math control" nel quale man machine math va intesa come "matematica trattabile sia da persone umane che da macchine".

I programmi MTM, come già accennato, sono scritti in un sottoinsieme ridotto del linguaggio C++ come "desktop application", ossia come software applicativo che come piattaforma per la sua esecuzione richiede un semplice personal computer di tipo desktop o del tipo portatile.

Si tratta di programmi poco interattivi (anche se una maggiore interattività sarebbe auspicabile), per ciascuno dei quali si prevede un file di governo che precisa le finalità di una esecuzione e i parametri che servono a precisarle, uno o più files per i testi da elaborare e i files per i testi da generare; si possono avere anche files da aggiornare da sottoporre a letture iniziali e a riscritture successive.

A05 i.02 Tra i files da leggere in genere si trova un file `.txt` contenente un elenco dei files di specificazione `.tex` che si devono esaminare o modificare.

Tra i files da generare si trova normalmente un file di log di tipo `.txt` nel quale, per ciascuna delle esecuzioni del programma in esame, vengono registrati i contenuti del file di governo e gli esiti di rilievo della esecuzione effettuata.

Per molte manovre tra i files da generare si trova un file `.txt` destinato a contenere le segnalazioni delle porzioni di testo da modificare o da riesaminare; le porzioni da modificare contengono configurazioni erronee, da dichiarare illegali; le porzioni da riesaminare contengono elementi giudicati ambigui, poco chiari o in disaccordo con il contesto da segnalare al redattore affinché intervenga per chiarirle, per armonizzarle con il contesto o per migliorarle in qualche modo.

I testi da esaminare riguardano sia kapf, che testi ausiliari (elenchi, dati statistici, ...) che si intendono certificare come adeguati o come aggiornati.

A05 i.03 Per avere programmi MTM che possono operare con una buona efficienza anche sui computers più comuni, si prevede di tenere in memoria uno o più kapf.
Per ciascun kapf può essere utile sia l'equivalente del testo-S che una rappresentazione del testo-M o di una sua parte.
Segnaliamo che un testo-M comprende sistemi i puntatori che rendono il contenuto del kapf una struttura di dati gestibile con efficienza e abbastanza facilmente.
Questa efficienza può riguardare sia la velocità esecutiva, che la disponibilità di collegamenti interni al testo in grado di fare da riferimento per procedure progettabili e collaudabili con relativa semplicità.

Conviene quindi avere chiari i collegamenti tra un testo-S registrato su un file `.tex` leggibile e il corrispondente testo-M consistente in una struttura di dati che nel seguito chiameremo anche "struttura-M" del kapf.

Per ottenere la sopra auspicata efficienza di programmazione sono previsti anche data sets ausiliari che possono riguardare l'intera *esposizione*organizzati in modo da essere utilizzabili in sintonia con strutture-M.

Va anche segnalato che questi data sets, che chiamiamo "data sets di riferimento", richiedono di essere accuratamente aggiornati dopo ogni modifica dei contenuti.

A05 i.04 Buona parte dei data sets di riferimento sono previsti lentamente variabili, in quanto tengono conto di scelte organizzative complessive delle nozioni che vengono trattate (e studiate), scelte che vengono modificate solo in seguito a decisioni che possono avere portata strategica per gli sviluppi della *esposizione*.

Presentiamo a grandi linee questi data sets.

Elenco dei files `.tex` contenenti il testo-S; in esso conviene distinguere i files ritenuti stabili da quelli che, nella attuale fase di sviluppo dell'*esposizione* ci si aspetta che possono essere sensibilmente modificati.

Elenco delle macro TEX utilizzate.

Elenco delle macro definite specificamente per la redazione del testo-S.

Elenco dei files ausiliari previsti.

Struttura informativa riguardante le varie manovre eseguibili con le relative specifiche che per ogni esecuzione devono essere precisate ed essere inserite mediante loro opportune codifiche nei files di governo; le specifiche delle varie manovre comprendono le specificazioni dei files di ingresso, di uscita e da aggiornare che le manovre prevedono.

La struttura informativa delle manovre serve per la interpretazione sistematica delle richieste contenute nei files di governo, richieste che devono essere esaminate e giudicate legali prima di diventare esecutive.

Repertorio dei messaggi diagnostici prevedibili; attualmente si tratta di segnalazioni in italiano, ma si vuole che la loro trauzione in altre lingue sia del tutto agevole; a questo contribuisce la attuale disponibilità di buoni traduttori automatici.

A05 i.05 Presentiamo, ancora a grandi linee, i data sets di riferimento che devono essere aggiornati in seguito ad ogni esecuzione che viene resa necessaria da una sessione di modifiche di un testo-ST.

Elenco delle macro, distinguendo le macro TEX dalle macro del redattore.

Elenco di lessemi del testo-S; questi comprendono di parole significative, nelle lingue naturali, nidi-$-$ esprimenti semplici formule delimitate da due singoli dollari, tokens previsti per i testi-S; l'insieme di questi ultimi lessemi deve essere definito e aggiornato con cura attraverso algoritmi di scansione relativamente semplici.

Elenco dei termini che compaiono nei nidi-dn

Elenco delle notazioni dei diversi tipi segnalate dai marcatori-N.

Elenco delle cosiddette "stringhe AbbrSpecVar" raccolte nel file X14.

Elenco delle persone citate nei nidi-Ya.

Elenco dei lavori citati nei nidi-Yw.

Codifiche dei classificatori delle linee del testo-S.

Codifiche degli accettatori-trasduttori riguardanti la individuazione nelle linee del testo-S dei vari tipi dei brani contenuti in queste linee e di interesse per il supporto allo sviluppo.

Codifiche degli accettatori-trasuttori (AccTrdF) elementari agenti sui lessemi.

Elenco dei paragrafi del testo-S contenuti in una lista di kapf con le relative posizioni iniziale e finale.

A05 j. strutture di dati per il testo-M

A05 j.01 Presentiamo ora le strutture informative, le operazioni e i procedimenti che compaiono nei programmi MTM procedendo in un modo che cerca di essere schematico e uniforme.

Per questa presentazione e per i testi dei programmi MTM introduciamo vari termini esplicativi specifici e vari identificatori che sono stati scelti con dichiarata pignoleria che riteniamo indispensabile per sviluppare discorsi e programmi sufficientemente precisi ed esaurienti.
Riteniamo infatti che sia importante la sintonia tra i termini usati nelle motivazioni e nelle descrizioni dei programmi e gli identificatori più significativi che compaiono nei testi sorgente dei programmi.

Per ogni array che incontreremo in seguito, se il suo identificatore è `arr`, denoteremo:
con `arr[i]` il suo componente nella posizione `i` ,
con `arr[<arrL]` la sequenza dei suoi componenti dalla posizione iniziale 0 alla posizione ultima `arrL-1`
con `arr[<=arrL]` la sequenza dei suoi componenti dalla posizione 0 iniziale alla posizione `arrL` ,
con `arr[i..s]` la sequenza dei componenti di `arr` nelle posizioni dalla `i`alla `s-1` .
con `arr[i..<=s]` la sequenza dei componenti di `arr` nelle posizioni dalla `i`alla `s`.

A05 j.02 Vediamo le abbreviazioni più generiche e semplici.

char = carattere che occupa un byte, otto bits;
int = intero che occupa 4 bytes, 32 bits;
seq-c = sequenza di caratteri, ossia di bytes che possono riguardare caratteri leggibili, blank, `\0` e `\n`;
seq-i = sequenza di interi da 32 bits;
seq-b = sequenza di bits;
array-char = array costituito da caratteri;
array-int = array costituito da interi da 32 bits;
spis = riguarda una sequenza nella quale si distinguono successive sottosequenze e consiste nella sequenza dei puntatori alle posizioni iniziali delle sottosequenze ai quali si aggiunge il puntatore alla posizione successiva all'ultimo componente della sequenza in causa; una spis richiede una seq-i e un int che fornisce la sua lunghezza;
spis-c = sequenza di caratteri munita di una sua spis; richiede una seq-i per i puntatori, un int per la sua lunghezza, oltre alla seq-c nella quale si intendono distinguere le sottosequenze;
spis-i = sequenza di interi da 32 bits munita da una sua spis; richiede una seq-i per i puntatori, un int per la sua lunghezza, oltre alla seq-i le cui sottosequenze si vogliono distinguere;
spis-ccc sequenza di terne di caratteri e relativa spis; riguarda una seq-ccc, ossia una sequenza di terne di caratteri alfanumerici, e richiede una seq-i per i puntatori che distinguono le sue sottosequenze e un int per la sua lunghezza.

A05 j.03 I nomi e i richiami delle variabili e degli arrays collegati sono scelti seguendo i criteri di sistematicità lessicale che seguono.

Se un array ha nome `arr` , alla sua lunghezza diamo il nome `arrL`.
Se come nome esemplare per una spis-c o per una spis-i assumiamo `seq`, scriviamo `seqL` per la sua lunghezza, `seqJ[]` per la sequenza dei puntatori agli inizi delle sottosequenze da distinguere ed `seqN`per il numero delle sottosequenze.

Conviene osservare che per ciascuna di tali spis abbiamo `seqL = seqJ[seqN]` e che la sottosequenza contraddistinta dal progressivo `j` che può assumere i valori `0, 1, ... , seqN-1` occupa le componenti dalla iniziale `seq[seqJ[j]]` fino all'ultima con valore `seqJ[j+1]-1`.

È opportuno che chi si occupa di programmi MTM abbia presenti delle valutazioni realistiche per le estensioni degli arrays che consentono di elaborare un particolare testo-M.
I valori massimi prevedibili con una attendibilità considerata sufficiente per una lunghezza di nome `seqL` e per un numero di oggetti di nome `seqN` li identifichiamo, risp., con `seqLmx` e con `seqNmx`.

A05 j.04 Un dato di ingresso presente in tutti i programmi e è la specificazione del file di governo nell'array-char `fgov[..fgovL]` ; in generale si può prevedere `fgovLmx = 30` .

Per molte manovre serve un file contenente, una in ciascuna linea, le specificazioni dei files .tex da elaborare, ciascuna in una linea; le specificazioni sono registrate come stringhe nullterminated nella spis-c ⟨ `feN , feJ[<=feN] , fe[<feL]` ⟩ .
Previsioni: `feNmx = 200` e `feLmx = 3500` .

I titoli dei kapitoli e delle sezioni dell'*esposizione* vengono trattati in memoria da due spis-c.

Per il numero dei kapitoli da trattare con una esecuzione di un proramma abbiamo gli identificatori equivalenti `kN` e `feN`; per il mumero delle sezioni `sN` e per il numero dei paragrafi `pN`; per il numero dei tomi `tN`.
Per i titoli dei kapitoli usiamo: array-char `tika[..tikaL]` organizzato dai puntatori in `tikaJ[..kN]` ; in altri termini: i titoli dei kapitoli sono registrati nella spis-c `[[ tikaN , tikaJ[<=kN] , tika[<tikaL] ]]` .
Per MATeXp sono prevedibili `tikaLmx = 4000` e `kNmx = 200` .

I titoli delle sezioni vanno nella spis-c `[[ sN , tiszJ[<=tiszN] , tisz[<tiszL] ]]` .
Per MATeXp sono prevedibili `sezNmx = 1500` e `tiszLmx = 50000` .

Per il collegamento tra kapitoli e sezioni: `smFkm[kma]` fornisce il progressivo (tipicamente identificato da `sma`) della prima sezione del kapitolo `kma`.

sequenza delle lettere dei tomi: `cdtm[..tN] = "ABCDGIJKPTWX "`
sequenza dei codici dei kapitoli: `kcd[..kN] , kNmx = 200` ,
`kcd[ik] =` *prgr. del tomo* `*100 +` *intero del suo identificativo*; ad esempio il codice del capitolo G52 è 452 .

A05 j.05 Consideriamo il testo-M, ossia la struttura che organizza un testo-S in memoria servendosi del complesso di puntatori descritto in precedenza.

Come si è detto per un *testo*-M occorre fare riferimento a vari livelli di ripartizioni; questi livelli vengono rappresentato dalle lettere segnalate nella tabella che segue.

livello 1 tomi, lettera t;
livello 2 kapitoli, lettera k;
livello 3 sezioni, lettera s;
livello 4 paragrafi, lettera p;
livello 5 righe, lettera r;
livello 6 caratteri, lettera c.

Aggiungiamo che per talune manovre converrebbe utilizzare anche un livello dei periodi da collocare tra il livello dei paragrafi e quello delle righe; se introdurremo questo livello che caratterizzeremo con la lettera P.

A05 j.06 Ciascuna delle ripartizioni di un testo viene rappresentata da interi naturali progressivi (0, 1, 2, ...) che chiamiamo **matricole delle ripartizioni**.

Vengono utilizzate matricole relative all'intero testo e matricole relative a ripartizioni più comprensive, ossia di livello inferiore.
Procedendo dai livelli si maggiore dettaglio adottiamo gli identificatori

`cma, rma, (Pma,) pma, sma, kma` e `tma` .

Dunque per l'intero testo sutilizziamo le seguenti matricole.

Vengono utilizzati sistematicamente:
`cma = 0,1,...,cN-1` con `cN <= cNmx` per i caratteri;
`rma = 0,1,...,rN-1` con `rN <= rNmx` per le righe;
`pma = 0,1,...,pN-1` con `pN <= pNmx` per i paragrafi;
`sma = 0,1,...,sN-1` con `sN <= sNmx` per le seziono;
`kma = 0,1,...,kN-1` con `kN <= kNmx` per i kapitoli;
`tma = 0,1,...,tN-1` con `tN <= tNmx` per i tomi;

Previsione delle estensioni per un testo-SP non dissimile da MATeXp:
tomi tN, tNmx = 12-16
kapitoli kN, kNmx = 160 - 200
sezioni sN, sNmx = 1396 - 1500
paragrafi pN, pNmx = 13972 - 15000
pagine pagN pagNmx = 180 x 25 = 4500
righe rN, rNmx = 400 000
caratteri cN, cNmx = 20 000 000

A05 j.07 La sequenza di tutti i caratteri di un testo-M viene trattata nell'array-char di base `c[<cN]` .

Per i collegamenti tra le ripartizioni di un testo-M abbiamo:
lista delle coppie (paragrafo rinviante , paragrafo/sezione/kapitolo rinviato nell'array int p2psk[p2pskN] con p2psk[ipp] = ip + 10000*ipsk ove
ip = indice paragrafo rinviante, ipsk = indice del p/s/k rinviato e p2pskN ¡ p2pskNmx = 5000 (ora 2856 per 158 kapf);
relazione inversa psk2p[p2pskN] con psk2p[ipsk] = ipsk + ip*10000 array ottenuto ordinando in situ gli interi ottenuti dalla sequenza dei `p2psk[ip]` .

Segue il complesso degli arrays di collegamento tra le ripartizioni del *testo*-M, ossia dei puntatori ai primi caratteri di ogni ripartizione.

```
cmFrm[rNmx]   rmFpm[pNmx]   pmFsm[sNmx]   smFkm[kNmx]
      cmFpm[pNmx]   rmFsm[sNmx]   pmFkm[kNmx]
            cmFsm[sNmx]   rmFkm[kNmx]
                  cmFkm[kNmx]
```

Può essere utile controllare anche i primi caratteri inutili di ogni linea con l'array di puntatori `cmuFrm[<rNmx]` che potrebbe riguardare le posizioni nelle singole righe oppure le posizioni nell'array `c[<cN]` .

Segnaliamo che la lettera "`F`" che compare all'interno dei nomi di arrays precedenti si può leggere come "from": ad esempio `rmFsm[isma]` andrebbe qualificato affermando: "fornisce la matricola della prima riga assegnata alla sezione avente la matricola `isma`".

A05 j.08 Si passa dal trigramma di un kapitolo “Tnn” alla sua matricola `kma` mediante la sequenza dei trigrammi `yk[<3][<kN]` individuando in essa l’intero naturale `kma` tale che sia `yk[1..3][kma] == "Tnn"` .

Si passa dal tetragramma di sezione “Tnnl” alla sua matricola `sma` ricavando `kma` da “Tnn” come sopra e cercando la posizione della sezione caratterizzata dalla lettera “l” nella spis-c delle lettere delle sezioni nei successivi kapitoli `ys[..sN]` ripartita per kapitoli dalla spis-c `smFkm[<=kN]`; `sma` è l’unico indice che soddisfa la relazione `smFkm[kma] <= sma < smFkm[kma+1]` e tale che sia `ys[isma] == "l"` .

Si passa dall’exagramma di paragrafo “TnnSmm” alla matricola di tale paragrafo `pma` ricavando `kma` e `sma` da “TnnS” come sopra e ottenendo `pma` considerando la sequenza delle matricole “mm” dei successivi paragrafi immagazzinabile nell’array-int yp[¡pN], array- ripartito per sezioni successive dalla spis-i `pmFsm[<=sN]`: `pma` è l’unico intero naturale per cui `yp[pma]` assume il valore val(”mm”) e soddisfa il vincolo `pmFsm[sma] <= pma < pmFsm[sma+1]` .

A05 k. manovre per costruire il testo-M

A05 k.01 accettazione delle ripartizioni

.. legge da file di governo e sistema nella spis-c tt [[kN, feJ[¡=Nk] , fe[¡feJ[feN]]]] la sequenza delle specificazioni dei kapitoli da trattare.

.. ottiene la seq-c delle lettere maiuscole dei tomi `ctom[tN]` contenute con ripetizioni nel caratteri `spekm[Fkm[kma]+2]` per `kma = 0..kN-1`; ottiene sequenza trigrammi dei kapitoli tgk[..3][..kN] e la suddivide per kapitoli con la spis-ccc `[[ tN, tgkJ[<=tN], tgk[<tgkL] ]]`

.. ottiene la seq-i dei codici dei kapitoli `Cdkapf[<kN]`, anch'essa suddivisa da spis `kmFtm[Nk+1]`, a partire dal relativo trigramma `tgk[1:3][kma]` definendo

```
Cdkapf[kma] = itk[ik]*1000 + ival(tgk[2:3][ik],
```

dove `itk[ik]` è tale che `codk[itk[kma]] == tgk[1][kma]`.

.. potendo assumere la correttezza delle sezioni e dei paragrafi dei kapitoli (v.o.), scorre i files per definire sequenza dei trigrammi di sezioni e paragrafi e costruire la seq-c dei caratteri delle sezioni entro i kapitoli `csek[sN]`, anch'essa organizzata in spis-c da `jtgk[..Nk+1]`, e la seq-i degli interi specificazioni dei paragrafi entro le sezioni `ispes[pN]` organizzata in una spis da `ispesJ[<=sN]`.

.. risulta stabilito il collegamento tra indici e specificazioni dei paragrafi, nonchè in seguito a semplificazione, il collegamento tra sezioni e kapitoli.

`spec(ik,is,ip) = tgk[1:3][ik] + csec[is] + decod(ispes[ip]` e
`iiicode(TnnSmm) = (ik,is,ip)`,
dove `ik = icode(Tnn), is = icode(L,ik), ip = decod(is,mm)`;
Qui abbiamo usato le functions `decod, iiicode` e `iicode` che effettuano le transcodifiche necessarie che preciseremo più avanti.

A05 k.02 estrazioni da file .tex

Per elaborare il testo sorgente servono le seguenti varie operazioni di estrazione dal testo-M di suoi elementi caratterizzanti.

Estrazione di rinvii interni a kapitoli, a sezioni e a paragrafi;
estrazione di nididn aventi la forma
`{\dn` *termine*`]tGc`
con i relative exagrammi TnnSmm;
estrazione delle occorrenze di notazioni marcate con i relativi exagrammi;
estrazione delle citazioni di personalità con i relativi esagrammi;
estrazione delle citazioni di opere con i relativi exagrammi;
estrazione dei rinvii coordinata, loro validazione e registrazione di coppie da paragrafo sorgente a coordinata bersaglio;
individuazione stringhe alfanumeriche estratte dai nidi-$...$, dai nidi-$:.$$ e da nidi simili e verifica di assenza dal corrispondente elenco generale;
individuazione di macro aventi forma del tipo `\abcd` e verifica della loro presenza nell'elenco delle macro di TEX o delle macro del redattore.

A05 k.03 diagnosi dei files per il rispetto di struttura e coerenza

Corre sui files forniti da un elenco predisposto nel file di governo;
corre su un file fino alla prima occorrenza di `\bye`;

corre su ogni linea fino a new line o fino alla prima occorrenza di "%" ;
esamina l'indice iniziale della forma `\csin` *sezioni, titoli e pagine* `\nop` ;
verifica la sequenza delle sezioni e i loro titoli;
esamina la sequenza dei paragrafi con eventuale diagnosi;
estrae titoli di kapitoli e di sezioni che potrebbero essere emessi nel relativo indice KWIC;
estrae coordinate bersagli di rinvii `\IL` per controllarne la presenza nel testo;
sse richiesto le registra insieme alle linee del rinvio.

A05 k.04 costruzione dei collegamenti entro il testo-M

Costruzione dei rinvii righe di kapitolo –¿ paragrafi di capitolo.
estrazione dei termini dalle linee dei files con `%;T TnnSmm` e verifica della legalità dell'exagramma;
estrazione delle linee dei files con `%;Na TnnSmm` , `%;Nc Ynnymm` `%;Nd Ynnymm` `%;Ne TnnSmm` `%;Ni TnnSmm` `%;Nm TnnSmm` e verifica della legalità dell'exagramma; estrazione dalle linee dei files che iniziano con `%;A TnnSmm` e verifica della legalità dell'exagramma;

costruzione dell'indice dei termini X11;
costruzione dell'indice delle notazioni X12;
costruzione di parte dell'indice AbbSpecVar X14;

A05 k.05 estrazioni da stringhe

Molte operazioni consistono nell'esame di una stringa da una riga, da un periodo o da un paragrafo altro allo scopo di individuare elementi qualificanti dei contenuti.

Oggetti che possono essere ricercati:
una data stringa null terminated;
una data sottostringa di un array-char determinata dalla posizione iniziale e dalla posizione che segue la sua ultima;
nidi della forma `{\dn` *termine*`}`;
nidi aventi una delle forme `\IL{`*TnnSmm*`}` , `\IL{`*TnnS*`}` , `\IL{`*Smm*`}` , `\IL{`*:Smm*`}` o `\IL{`*:S*`}`;
nidi della forma `\Ya{`it nome di personalità`}`;
nidi della forma `\Yw{`it nome di opera`}`.

Sono prevedibili le ricerche di altri nidi di forme simili alle precedenti, che evidentemente sono facilemente individuabili.

In prospettiva va esaminata la possibilità di ricercare sottostringhe caratterizzate da un riconoscitore a stati finiti definito mediante una sua codifica.

A05 k.06

Vengono studiate anche procedure aventi lo scopo di aggiornare l'intero testo-M in seguito a richieste dei tipi seguenti:

spostamenti di kapitoli, sezioni o paragrafi;
eliminazioni di kapitoli, sezioni o paragrafi;
inserimenti di kapitoli, sezioni o paragrafi;
modifiche di elementi individuabili con algoritmi come quelli accennati in k.05.

A05 m. prospettive di ulteriori automatismi di supporto ai testi ST

A05 m.01 Sono molte le considerazioni che si possono sostenere sulla opportunità di rendere disponibili strumenti software che consentano di supportare le attività per la pubblicistica di argomenti matematici e di più in generale di argomenti scientifico-tecnologici, strumenti con finalità e prestazioni ben maggiori dei programmi di livello artigianale come i programmi MTM.

Ovviamente questi auspicati strumenti richiedono risorse rilevanti in varie direzioni:
direzione della progettazione in relazione ad esigenze culturali e didattiche ampiamente riconosciute;
direzione della programmazione di strumenti interattivi, multimediali, operativi sul Web e in grado di interagire con basi dati e piattaforme AI dedicate alla matematica e alle sue applicazioni;
direzione del monitoraggio delle prestazioni e delle ricadute.

A05 m.02 Un primo ordine di considerazioni sull'importanza del software di supporto ai testi-ST riguarda la necessità di produrre tempestivamente presentazioni didattiche e divulgative riguardanti argomenti di ricerca e sviluppo in rapida evoluzione.

Altre considerazioni possono riguardare la necessità di presentare a settori di attività applicative di elevata importanza aggiornamenti tempestivi sopra risultati recenti progressi che si ritiene possano influenzare gli accennati settori attraverso l'adozione di nuovi strumenti di nuovi metodi e di nuovi approcci, con la conseguente apertura di nuove prospettive sui piani commerciale ed economico, forse anche sul piano sociale e sul piano culturale.

Questi aggiornamenti è opportuno siano presentati nelle forme più adatte ai singoli settori applicativi, anche attraverso la introduzione di termini di portata settoriale da proporre attraverso definizioni necessariamente ben fondate e ben comprensibili.

Queste ultime considerazioni si possono integrare con argomentazioni riguardanti la possibilità che nel prossimo futuro si abbia una crescita rapida e diffusa della domanda di conoscenze di afferenti alle discipline STEM.

A05 m.03 Si tratterebbe di un cambiamento rilevante, in particolare in Italia, stante il limitato interesse attuale per questi problemi, soprattutto se rapportato al diffuso mantenimento di interesse verso attività e prospettive dotate di tradizione ma destinate ad esaurirsi in poco tempo in quanto obiettivamente svantaggiose se confrontate con pratiche sostenute dalle innovazioni che vengono continuamente immesse nel mercato e nell'immaginario collettivo dalle imprese in grado di sviluppare le iniziative più avanzate.

A questo proposito si possono citare in particolare le realizzazioni del mondo dell'intelligenza artificiale che sono state presentate negli anni recenti.
Queste realizzazioni sono balzate agli onori della cronaca soprattutto a partire dal lancio del sistema generativo ChatGPTavvenuto nel 2022.
Da allora continua a giungerci dalle imprese del settore AI un flusso continuo di nuovi strumenti e di nuove metodiche; spesso questi strumenti sono offerti in uso gratuito sia per aumentare la prospettiva della loro efficacia e della loro desiderabilità sia per ricavare dal loro utilizzo diffuso indicazioni sulla loro efficacia e nuovi dati che i produttori possono usare per migliorare quasi con continuità le loro prestazioni.

Questi meccanismi di advertising e di ampliamento delle prestazioni dei prodotti agenti nell'infosfera sono evidentemente favoriti dalle esistenti reti telematiche e sono rese estesamente disponibili da dispositivi sempre più onnipresenti come smart phones, tablets e assistenti digitali personali (orologi, occhiali,...) .
Tutto questo fa sì che le innovazioni di frontiera stanno avendo una crescente influenza sulle vite di tutte le persone che vivono in paesi dotati di infrastrutture digitali.
Inoltre da questo discende un potenziale maggiore interesse nei confronti di iniziative divulgative concernenti le nozioni attribuibili al mondo STEM.

A05 m.04 Un terzo ordine di considerazioni tocca la crescente disponibilità di strumenti e di infrastrutture in grado di sostenere l'efficacia delle iniziative divulgative e didattiche nella direzione STEM.

A questo proposito ci limitiamo a segnalare la crescente diffusione di ambienti Web dedicati alle innovazioni tecnologiche e alle relative discussioni critiche. Citiamo come esempi le iniziative dell'MIT, della Simons Foundation, di Springer Nature, di Scientific American e di WIRED.

Ricordiamo in particolare la possibilità di trovare in questi ambienti collezioni molto notevoli di fotografie, di diagrammi e di tabelle di interesse scientifico e tecnologico.

Conviene anche ricordare la accessibilità in rete di tutti i maggiori periodici scientifici e di buon livello divulgativo e di un gran numero di blog dedicate ad analisi critiche dei nuovi prodotti e delle recenti iniziative.

Rispetto alle tradizionali pubblicazioni cartacee si tratta di prodotti che comportano costi che possono essere giudicati accessibili da gran parte delle persone seriamente interessate.

Inoltre le informazioni e i discorsi critici ottenibili dal Web hanno il vantaggio del tempestivo aggiornamento delle notizie.

Va anche osservato che la difficoltà della lingua straniera, prevalentemente dell'inglese, attualmente può essere tranquillamente superata ricorrendo ai servizi di traduzione automatica sostanzialmente immediata ampiamente disponibili; questi sono gratuiti fino a un primo livello di prestazioni e a un livello professionale presentano costi decisamente inferiori a quelli dei canali più tradizionali che qualche decennio fa avevano ben poche alternative.

Aggiungiamo anche la disponibilità in rete di strumenti della AI che consentono di orientarsi su ampie gamme di strumenti e di tendenze innovative quasi in tempo reale.
Per molti documenti reperiti e leggibili in rete e di rilevanti estensioni si possono ottenere con pochi clik anche dei concisi riassunti orientativi.

A05 m.05 Torniamo alle procedure di supporto ai testi-ST-M e alle prospettive di allargare la loro portata, di aumentare la loro incisività e di cercare anche di allargare lo sguardo per comprendere il possibile supporto a simili iniziative riguardanti nozioni di altri settori di studio che danno primaria importanza alla **razionalità inclusiva** e alle analisi quantitative.

Si possono individuare e cominciare ad esaminare due generi di prospettive.

Al primo genere appartengono le prospettive della generazione automatica di files leggibili nei linguaggi HTML, XML e WikiMedia, nonché nel formato `.pdf`.

Al secondo genere appartengono le prospettive di generazione di files nella lingua inglese o in altre lingue che possono avere un ruolo di rilievo per la divulgazione e la didattica della matematica e le sue applicazioni.

Va anche auspicata la possibilità di elaborare stringhe per ricercare o individuare frammenti significativi, come detto in [:g05], o per molti altri scopi immaginabili, servendosi di riconoscitore a stati finiti o di altre macchine sequenziali caratterizzate attraveso loro codifiche.

A05 m.06 Aggiungiamo qualche parola sulle procedure che intendono sostenere lo sviluppo di presentazioni di argomenti scientifico-tecnologici diverse da MATeXp, ma con caratteristiche simili.

Si possono prospettare presentazioni redatte anche in linguaggi diversi da Plain TEX, ma in ogni caso contenenti formule matematiche e diagrammi schematici riguardanti andamenti quantitativi e comportamenti riconducibili a proprietà strutturali, che si servono di termini introdotti da definizioni accurate e coerenti; a queste presentazioni inoltre si chiede abbiano una ben definita organizzazione che prevede con ripartizioni organizzate in modo gerarchico, come quella di MATeXp riguardante tomi, kapitoli, sezioni e paragrafi e con sistemi di collegamenti interni che seguono regole di completezza verificabili da automatismi.

Qualcuna delle considerazioni che seguono potrebbe servire anche a testi redatti in un altro linguaggio, purché preveda una versione sorgente esprimibile con caratteri ASCII o UNICODE e una versione leggibile derivabile dalla precedente mediante automatismi condivisibilmente comprensibili.
Hanno queste caratteristiche i linguaggi HTML, XML e il linguaggio dei sistemi Wikimedia.

`Testo fruibile in https://www.mi.imati.cnr.it/alberto/ e https://arm.mi.imati.cnr.it/Matexp/matexp_main.php`