### ACCADEMIA NAZIONALE DEI LINCEI

ANNO CCCLXXVIII - 1981

# CONTRIBUTI DEL CENTRO LINCEO INTERDISCIPLINARE DI SCIENZE MATEMATICHE E LORO APPLICAZIONI N. 56

CLAUDIO SAPORETTI

## RISULTATI E PROSPETTIVE DELL'ANALISI DEI TESTI ACCADICI MEDIANTE IL CALCOLATORE ELETTRONICO

Conferenza tenuta il 14 dicembre 1979



ROMA ACCADEMIA NAZIONALE DEI LINCEI 1981

I quasi seicento segni, usati in forma a volte ideografica a volte sillabica, un vocabolario ricchissimo (piu' del doppio, e' stato calcolato, di quello biblico), una evoluzione della lingua legata ad un'area vastissima che dalla Mesopotamia giunge all'Anatolia e all'Egitto e ad un periodo che va ben oltre un millennio di storia, una ricchezza incredibile di documenti di ogni tipo, tutti di prima mano, in parte editi ed in parte ancora sepolti sotto la sabbia onei Musei; sono dati sufficienti, questi, per comprendere immediatamente perche'si sia pensato all'utilizzazione del computer anche per l'Accadico.La conoscenza dei dati dell'Accadico (l'"Assiro-Babilonese") elegata ad una schedatura che certamente e' molto avanzata sia da parte di privati che a livello di compilazione di lessici e di manuali epigrafici. Ma va da se'che la schedatura privata conosce un campo di utilizzazione molto limitato, mentre quella dei Vocabolari e dei Manuali, da una parte e' portata alla conoscenza ufficiale degli studiosi con lentezza, e dall'altra si trova spesso nella difficolta' di offrire aggiornamenti tempestivi. Come e' noto agli Orientalisti, i grandi Dizionari di Assiro (a parte quelli ormai superatissimi quali il Delitzsch e il Bezold) sono l'Akkadisches Handwörterbuch del von Soden (iniziato nel 1959 e giunto ora, senza aggiornamenti, alla parola uzuzzu), ed il Dizionario di Chicago (CAD) (iniziato nel 1956 ed in cui sono state pubblicate finoratredici lettere dell'alfabeto, senza aggiornamenti se si vuole escludere un vecchio aggiornamento della H in appendice alla G). Anche se gli giornamenti non tarderanno a venire, é comunque evidente che, una volta pubblicati, si dovra' attendere anccra molto per avere i successivi; e d'altronde, aggiornati o no, i lessici riportano sempre delle esemplificazioni, mai quella totalita' di dati indispensabile allo studioso. Cosi dicasi dei repertori quali il Manuale di René Labat (edito nel 1948 e recentemente

aggiornato nel 1976) o il Das Akkadische Syllabar (del 1948, aggiornato nel '67). Una schedatura elettronica a disposizione di tutti puo' invece risolvere i problemi rimasti ed integrare le molte lacune, senza pretendere tuttavia di essere sostitutiva dei Lessici, che continueranno ovviamente ad avere sempre la loro insostituibile funzione. Anche su altre lingue del Vicino Oriente Antico si stanno d'altronde eseguendo, in altre sedi, ricerche di questo tipo: basti citare il lavoro di Whitaker sulle concordanze dell'Ugaritico (1972) (1) e quello degli studiosi dell'Universita di Friburgo sui testi sumeri, con cui sono in contatto anche illustri studiosi italiani.

Scopo di questa relazione e' di esplicare lo stato dei lavori in Italia sulla memorizzazione dell'Accadico, con aspetti utili anche ad altre lingue, mentre per quanto riguarda l'aspetto piu tecnico relativo alla metodologia ed ai sistemi usati per l'elaborazione elettronica rimando alle descrizioni dei singoli programmi esposti da John Settles, Arthur A. Sorkin, Sal J. Fallone (2).

La lingua accadica, e' noto, e' scritta in cuneiforme, cioe' mediante una grande quantita' di segni incisi o scolpiti, che i Semiti hanno adottato dalla scrittura sumera applicando all'Accadico il loro valore fonetico senza considerare il significato dato loro dai Sumeri. Ne risulta una frequente differenziazione di valore fonetico per ogni segno, legato ai differenti significati, e dunque letture, dati dai Sumeri al singolo ideogramma. Tale differenza di lettura puo essere legata a diversi periodi storici ed anche a diversi ambienti geografici. Ad esempio il segno con valore NI puo avere anche un valore DIK in epoca neoassira e tardobabilonese. Molto spesso coesistono poi valori diversi in quasi tutti i periodi e le aree. Cosi' il valore RID/T/T convive con il valore MIS, NE con BIL, e cosi' via. Succede di conseguenza, che nei testi di un particolare periodo ci siano anche molti valori indifferentemente usati per il medesimo segno. Cosi'ad esempio il segno 449 ha, in epoca

tardobabilonese, il valore ŠI e Š $E_{20}$ , IGI, LIM/LEM, LÌ, PÀN e forse anche LAM $_5$ , accanto ad altri valori attestati in ambienti differenti, quali SI $_{17}$ , GI $_8$ , LIMI, LÚM, BAT $_5$ , ÍNA/I/U.

Va da se' che questo fenomeno complica l'analisi, non tanto a livello di memorizzazione ai fini di una lemmatizzazione da usarsi per compilare lessici oper una analisi di tipo grammaticale (e' memorizzata infatti la traslitterazione del testo, dove il valore esatto del segno e' gia' stato studiato e scelto dall'Assiriologo), ma piuttosto quando si voglia considerare l'uso e la frequenza di un segno in rapporto ad altri determinati segni che lo seguano o lo precedano. Ne' basterebbe creare un algoritmo che tenga semplicemente conto dei valori sillabici diversi per ogni segno perche', come conseguenza del fenomeno appena descritto e della grande quantità dei segni stessi, si verifica un secondo fenomeno non meno caratteristico, e cioe' la coesistenza di un medesimo valore in più segni diversi. Può essere che questa coesistenza fosse legata magari a sfumature di antica pronuncia, che nella differenziazione dei segni aventi il medesimo valore hanno lasciato un indizio, o meglio un accenno di se'; spesso tuttavia troviamo che i segni aventi il medesimo valore sono usati indifferentemente nello stesso preciso contesto, e cio' porta ad una ulteriore difficolta per l'analisi.

E' nota altresi' la presenza di un altro fenomeno caratteristico del Cuneiforme, cioè il frequente uso degli ideogrammi
sumerici riportati con il loro significato, ma letti con il corrispondente valore accadico. L'evoluzione grafica di ogni segno e' stata generalmente la stessa sia che il segno fosse usato nel suo valore sillabico, sia che avesse mantenuto il suo
originale significato sumerico, e dunque nel testo puo' esserci
una coesistenza di valori totalmente diversi in un segno ripetuto nella identica forma (3); e puo' addirittura accadere che
il medesimo segno coesista in un contesto che ne richiede l'uso in immediata successione (4).

Alla presenza, talvolta notevole, di segni,o di un gruppo di segni, usati con il loro originale valore ideografico, sono legati altri due aspetti della scrittura dell'Accadico. Il primo e'l'uso di determinativi, premessi ad una parola per determinarne appunto, con immediato anticipo rispetto alla sua lettura, la natura e la specie. Sono circa una trentina gli ideogrammi usati a questo scopo, e vanno dal determinativo immediatamente preposto al nome di divinita'a quelli che indicano nomi geografici, titoli o mestieri, oggetti in cuoio o legno. stoffe, vasi, mesi e cosi' via. Poiche' tali segni erano muti segnali di preavviso, intimamente legati alla parola seguente, e' ovvio che devono esserne considerati parte integrante; ma ne complicano tuttavia l'analisi con la loro presenza. Ugualmente dobbiamo dire di quegli altri determinativi, meno numerosi, non gia anteposti ma posposti al nome, quali KAM/KÁM posposto ai numeri, KI ai nomi geografici, per non parlare del frequentissimo MES o del meno frequente HI.A posposti ad indicare il plurale.

Un breve accenno, infine, al secondo problema legato alla presenza degli ideogrammi; parlo cioe' dei complementi fonetici. Gli ideogrammi sumerici, ho gia'avuto modo di ricordare, erano letti in lingua accadica anche se erano rappresentati con il loro originario significato; la loro presenza, cioe', era ben diversa da quella, ad esempio, di motti latini intercalati in uno scritto in lingua italiana. Da cio il fenomeno, ben noto, dei complementi fonetici applicati all'ideogramma per favorire la lettura piu esatta. Totalmente assente nei testi che sono oggetto di memorizzazione in Italia e'la trascrizione completa o quasi, in Accadico, dell'ideogramma sumerico; frequentissimo invece e' l'uso del complemento fonetico corrispondente alla prima od all'ultima parte della parola (di solito una sillaba, talvolta la sola vocale) in cui l'ideogramma sumerico deve essere svolto. Il complemento fonetico aiutava gli antichi lettori, ed a maggior ragione aiuta oggi noi, sia per una giusta interpre-

tazione lessicale dell'ideogramma, sia per una giusta interpretazione grammaticale della parola (5). Cosi' come gli ideogrammi-determinativi andranno considerati al computer come legati alla parola successiva, anche le sillabe-complementi fonetici andranno dunque strettamente rapportate alla parola precedente o seguente, cui esse si riferiscono. Il problema e complicato dalle varianti di questo fenomeno, che in fase di studio per una migliore codificazione dei testi si e'dovuto seriamente considerare. Ad esempio ci sono complementi fonetici posposti o preposti non gia' agli ideogrammi, ma ai segni sillabici stessi, quando sono rari o usati con un valore insolito, tale da potere creare confusione (6). E c'e' di piu: trattandosi di chiarificazioni della lettura di un segno che può generare confusione, e'evidente che questo fenomeno non si limita, come per gli ideogrammi, all'aggiunta di un complemento fonetico alla fine o al principio della parola, ma puo verificarsi anche all'interno di questa (7).

Mi sono dilungato sui problemi di trascrizione perche' lo stato attuale del lavoro in Italia vede proprio l'applicazione di programmi legati da una parte alle concordanze dei segni e alla loro distinzione di valori, dall'altra alla stampa automatica del testo in una trascrizione che possa essere accettabile a qualunque Assiriòlogo, e dunque il piu' possibile vicina al tipo di trascrizione convenzionale ormai universalmente accettato.

L'esigenza di approntare un nuovo tipo di schedatura dell'Assiro-Babilonese, e di conseguenza di affrontare i problemi che ho finora esposto, e'nata in Italia in seguito a due corsi di analisi elettronica, particolarmente stimolanti, che l'Ingegner Enrico Maretti, nel 1970, ha tenuto nella sede dell'Istituto Miceneo ed Egeo-Anatolico (Consiglio Nazionale delle Ricerche, Roma). La data in cui si sono svolte rappresenta pero'solo un riferimento simbolico, perche'il periodo immediatamente successivo ha visto un impegno, non una realizzazione. Come risultato di vari contatti e degli studi effettuati si era arrivati a concepire, anche se non a realizzare, un sistema di programmazione relativo all'analisi dei segni che poteva dare risultati soddisfacenti. Ma a livello di lemmatizzazione i risultati erano stati pochi.

Il contatto con Giorgio Buccellati, Professore di Accadico presso l'Universita di California (UCLA) di Los Angeles e gia compagno di studi all'Universita' Cattolica di Milano, ed il successivo approdo del Progetto alla Divisione Linguistica CNUCE (Consiglio Nazionale delle Ricerche) (8) hanno portato alla soluzione di molti problemi, suggerito la radicale rinuncia ad alcuni sistemi gia' considerati, e soprattutto favorito il nascere e la realizzazione di un progetto di collaborazione. Giorgio Buccellati aveva infatti da parte sua iniziato gia' da prima una ricerca analoga, con la collaborazione di esperti programmatori, o convertiti alla problematica della linguistica, come John Settles, o linguisti essi stessi, come David Packard, cui si devono interessanti lavori di carattere statistico (Lineare A), bibliografico (bibliografia omerica) e lessicale (concordanze dell'Ugaritico). La ricerca era, allora, dedicata e limitata all'analisi delle lettere antico-babilonesi, in vista dell'analisi di tutto il corpus antico-Babilonese, come suggeriva anche il titolo del Progetto che aveva iniziato a realizzare (OBLAP: "Old Babylonian Linguistic Analysis Project"). Il Progetto americano, di cui si era gia avuta una prima notizia alla XVII "Rencontre Assyriologique Internationale" di Bruxelles, é stato presentato dallo stesso Buccellati al 183° "Meeting of American Oriental Society", tenuto a Washington nel 1973 e, nel medesimo anno, al convegno "Computational and Mathematical Linguistics", tenuto a Pisa, i cui atti sono stati pubblicati in seguito, a cura di Antonio Zampolli e Nicoletta Calzolari (9). Nel corso del medesimo Convegno e' stato annunciato anche il Progetto ALTAN (Analisi Linguistica dei testi assiri e nuziani) relativo alla parte italiana, con la specificazione che i due progetti sarebbero stati paralleli ecomplementari (10).

L'accordo internazionale, contratto tra i due responsabili con l'avallo del Rettore dell'Universita' di California e dell'allora Direttore dell'I.M.E., Professor Carlo Gallavotti, ha stabilito infatti l'uso dei medesimi sistemi di codificazione, con il mutuo scambio di programmi e di dati. Comuni, oltre ai sistemi, anche i fini che i due Progetti si sono prefissi: anzitutto quello di costituire una "banca dei dati" a disposizione di tutti gli studiosi, e fornire poi i risultati degli studi e delle analisi mediante una serie di pubblicazioni particolari. Una simile ripartizione di compiti tra parti che si prefiggono gli stessi fini ed usano gli stessi sistemi ci e'sembrato rispondere ad una corretta impostazione della programmazione. Questa ripartizione porta naturalmente a specializzazioni, che non restano separate, ne tanto meno si escludono; tutt'altro: si integrano e sono complementari l'una con l'altra. Su questa linea sono nati altri progetti che hanno accolto la medesima metodologia contemporaneamente ad un allargamento e potenziamento del Progetto americano (11).

E'necessario chiarire, a questo punto, quale sia il campo di interesse perseguito in Italia, e quali le ragioni che hanno portato a scegliere questo campo, piuttosto che un altro. La scelta e' intimamente legata agli studi ed all'insegnamento del Professor Karlheinz Deller, allievo del von Soden, che ha insegnato presso la Facolta' Orientale del Pontificio Istituto Biblico, intorno agli anni '60. A lui si devono nuovi impulsi in piu' campi (12). Soprattutto egli ha dato nuovo impulso allo studio dell'Assiro, additando l'importanza dei documenti economici e giuridici, specie privati, scritti in questa lingua. Dopo il suo passaggio a Professore presso il "Seminar für Sprachen und Kulturen des Vorderen Orients" dell'Universita' di Heidelberg, Karlheinz Deller ha allargato i suoi studi ai testi di Nuzi. Sulla scia di questi interessi, che ha saputo comunicare

ai suoi allievi, e'nata la rivista "Assur" di cui e' "editor" per il neo-Assiro e il Nuziano e che, in una serie a parte ("Nuzi Studies"), accoglie ora anche gli studi sulla civilta' e sulla lingua di Nuzi. E'evidente, da quanto detto, che il progetto di analisi dei testi assiri e nuziani si inserisce coerentemente in un solco gia' tracciato, lungo un filone gia perseguito. I vantaggi di una tale scelta sono ovvî: anzitutto la collaborazione di quegli studiosi in qualche modo legati all'insegnamento di Karlheinz Deller, una loro più approfondita conoscenza dei testi e, importantissima, la disponibilita' di corpora gia' studiati ed elaborati a fondo. Il "seme" del computer e' caduto cosi' in un terreno ampiamente fecondo. In altre parole, non si e' partiti dal nulla per verificare la possibilita' di applicazione dell'elettronica, ma tale applicazione non e' stata che un ulteriore passo in avanti di una impresa gia' in corso.

Devo tuttavia chiarire meglio l'argomento quando parlo di Assiro e Nuziano. L'Assiro, e' noto, e' quella lingua parlata al Nord, e che ebbe come fulcro l'Assiria, la cui capitale fu (ma non sempre) Assur, sul luogo dell'attuale Qala'at Shergat. I testi assiri, o comunque raccolti in collezioni assire (si pensi alla biblioteca di Assurbanipal) sono migliaia, e coprono un'area assai vasta, specie se vogliamo includere le iscrizioni reali che i re lasciarono talvolta nei paesi conquistati. Pensare ad una memorizzazione di tutto l'Assiro sarebbe stata, ed e'ancora oggi, idea eccessivamente ambiziosa. Era meglio incominciare da un nucleo di testi, per allargarsi via via. Questa prudente decisione e' stata facilitata dal fatto che l'Assiro e' divisibile in tre grandi gruppi, chiamati, secondo le fasi cronologiche, antico (o paleo)-Assiro, medio e neo. La divisione non e'solo convenzionale; rispecchia infatti effettive realta' storico-linguistiche, ed anche geografiche. Significativo e' il caso del corpus antico-assiro, costituito di tavolette in gran parte non gia' ritrovate in Assiria, ma a centinaia di Km. di distanza in varie localita' dell'Anatolia, soprattutto a Kültepe (antica Kaniš).

Dal punto di vista linguistico, i tre aspetti dell'Assiro presentano diversita' notevoli, ed anzi alcune differenze si notano addirittura anche all'interno di ciascun gruppo, legate o a intervalli di tempo che separano la compilazione delle tavolette (ad esempio i due gruppi dei testi di Kültepe, separati da una soluzione di continuita' cronologica), o a ragioni geografiche (differenze dialettali del linguaggio di citta' periferiche rispetto a quello della capitale), o a diversa tipologia di testi (lingua della cancelleria statale e linguaggio dei documenti fatti redigere da privati). Questa divisione era gia' di per se' una traccia di lavoro. Bastava incominciare dal corpus posseduto in maniera piu' completa (nel nostro caso, il corpus medio-assiro) e verificare qui l'uso e la convenienza del computer; e poi, nel caso di risultati positivi, allargarsi secondo il criterio ritenuto piu' valido al momento opportuno.

Le difficolta' iniziali, inevitabili, causa di tempi dilavoro talvolta molto piu lunghi del previsto, si sono verificate dunque nel contesto di un corpus noto, ed il loro superamento e' stato enormemente facilitato da questo fatto. La varieta' del corpus medio-assiro ha permesso inoltre di verificare l'applicazione del computer in contesti differenti, con gradi crescenti di difficolta'. Nei 1000 e piu testi, oggi pressocche tutti memorizzati, troviamo infatti una notevole varieta' di tipologie. Si va da una ricca quantita' di documenti economici e giuridici, in ognuno dei quali troviamo un aspetto diverso di un sistema commerciale e bancario particolarmente interessante, a raccolte di leggi divise per argomento; da lettere di ogni tipo (private e reali) a trattati di medicina e di cosmetica; da rituali talvolta oscuri e complicati a manuali per l'allevamento dei cavalli; da disposizioni di sovrani relative all'ordinamento del loro palazzo e del loro harem ad elenchi dei tributi ricevuti da feudatari, con la specificazione del loro uso; ne' mancano i documenti più ufficiali, quali le iscrizioni rea-

li (scritte pero' in Babilonese) e più privati, come i contratti di adozione e di matrimonio. Particolarmente interessanti sono poi gli elenchi di persone e delle loro famiglie, che registrano deportazioni di popoli di paesi circostanti; sono di questo tipo alcune grandi tavole provenienti da Kār-Tukulti-Ninurta, ancora inedite ma che ho avuto la fortuna di vedere e di studiare recentemente (13). Tutto questo insieme di testi ha costituito il corpus-base, il gruppo-cavia, la prima pietra della banca dei dati. Diro' poi dei risultati che sono stati ottenuti, o ci si appresta ad ottenere, da questa prima "tesaurizzazione" di patrimonio linguistico. Qui voglio per ora sottolineare soltanto il crescente interesse che altri studiosi stranieri, che si erano applicati in precedenza o per ora si stanno applicando a questi testi, portano verso questo nuovo tipo di analisi. C'e' fra essi chi ha avuto modo di collazionare documenti noti solo in copia, e chi ne possiede di inediti; questo loro interesse, che si sta concretizzando in una fattiva collaborazione da intravedere ormai non lontana, e' una ulteriore garanzia di serieta e completezza del lavoro per la parte strettamente assiriologica.

La fase in cui, dopo le prime applicazioni e verifiche, si poteva passare ad un allargamento del lavoro ad altri corpora, e' da tempo venuta. C'erano tre prospettive diverse: il corpus dei testi neo-assiri, quello dei paleo-assiri, quello nuziano. Motivate ragioni hanno subito indotto a posticipare l'interesse per i testi neo-assiri, nonostante la grande ricchezza di materiale che il Professor Deller, forte della sua alta specializzazione proprio nel neo-Assiro, poteva (come può sempre) fornire. La notizia di tentativi ed esperimenti fatti a Chicago dal collega Simo Parpola su un gruppo di testi appunto neo-assiri, il desiderio di costruire una linea "verticale" di storia della lingua assira, e quindi la necessita di affrontare subito i testi più antichi, o l'alternativa di allargare "in orizzontale" lo studio della civilta mesopotamica contempora-

nea al medio-assiro, affrontando dunque l'analisi del Nuziano, hanno portato alla esclusione del neo-Assiro dalla scaletta degli interessi piu' immediati; ovviamente, questa esclusione non implica la rinuncia. A questa decisione non e' stato d'altronde estraneo il desiderio personale di perseguire studi che avessero un addentellato con l'Anatolia, essendo compiuti in un Istituto per gli studi micenei ed egeo-anatolici; anche in tal senso si prestavano meglio i testi antico-assiri, ritrovati appunto in Anatolia, ed anche i testi nuziani, legati al mondo della civilta hurrica. Come e noto, i testi rinvenuti a Nuzi e nell'area circostante sono stati scritti infatti da popolazioni hurrite che hanno usato (fortunatamente, date le difficolta' che si incontrano per la comprensione del Hurrico) la lingua accadica, sia pure con numerosi anacoluti ed errori. L'interesse che ad Heidelberg si era elevato vivissimo per i testi nuziani in quel momento, ha fatto propendere per la scelta di questo corpus. Cio' permetteva di stringere ulteriormente i contatti ed intensificare la collaborazione con Karlheinz Deller. Ne' era trascurabile, ovviamente, il fatto che il corpus nuziano, la cui trascrizione veniva pian piano elaborata, costituiva una materia prima di valore assai considerevole. Tutte queste ragioni, soprattutto di ordine pratico, sono state determinanti per la scelta del corpus nuziano, cioè per quella scelta "orizzontale" di allargamento allo studio della civilta nordmesopotamica. Cio' implica, dico subito, un altro campo di analisi di estremo interesse; quello dell'intera area mesopotamica nella seconda meta del 2º millennio, che richiederebbe l'analisi linguistica e storica anche dei testi medio-babilonesi di eta cassita.

I testi nuziani sono numerosi, ma gli argomenti trattati sono più ridotti rispetto a quelli attestati nei documenti medio-assiri: si tratta di testi della cancelleria reale e di cancellerie private, questi ultimi soprattutto documenti di adozioni, autentiche e false, di matrimonio, di trasferimento di beni, di dichiarazioni e processi, oltre ad un cospicuo numero di documenti amministrativi. La loro analisi e'avvenuta ed avviene ad opera dei collaboratori del Professor Deller, che rivede il testo definitivo della traslitterazione. Dopo questo controllo, una copia del testo e'passata al Progetto ALTAN per la codificazione e la memorizzazione. Finora sono pervenuti al Progetto testi nuziani trascritti dallo studioso irakeno Abdulillah Fadhil e dalla tedesca Gudrun Dosch. Collabora con i Nuziologhi di Heidelberg la Dottoressa italiana Paola Negri, che ho avuto il piacere di avviare allo studio ed alla memorizzazione dei testi e che lavora attivamente al Progetto. La fase attuale del lavoro sul corpus di Nuzi vede memorizzati c. 1400 testi su 2500 pervenuti da Heidelberg.

Il procedimento per l'analisi dei testi sia assiri che nuziani segue, naturalmente, le varie fasi cui va soggetto ogni corpus, e rimando per questo al lavoro di Antonio Zampolli, L'elaborazione elettronica dei dati linguistici: stato delle ricerche e prospettive, pubblicato proprio a cura di questo Centro Linceo Interdisciplinare. L'input richiede una codificazione particolare del materiale; anzi, le codificazioni sono di due tipi, una per il testo vero e proprio, l'altra relativa ad una cosiddetta "scheda di testa" che accompagna, quale vera e propria carta d'identita', il testo stesso. La codificazione avviene sulla traslitterazione definitiva: il testo cioe'e' trasportato in scrittura fonetica, che riproduce segno per segno, unendo le sillabe di una stessa parola con un trattino, ed indicando gli ideogrammi, come d'uso, con la scrittura maiuscola. Una leggera variante consiste nel dare al cuneo verticale spesso anteposto al nome di persona la forma del N.1 (che d'altronde in Accadico si scrive proprio con questo cuneo) al posto della m minuscola ("maschile" o "mister" che si voglia chiamare) o della barra in alto, di solito espressa convenzionalmente con una I maiuscola. Per il resto, minime sono le varianti al testo nella fase di preparazione dell'input, riducendosi solo a

leggeri interventi; essi sono di una certa rilevanza praticamente solo nei nomi di persona, e sono due: il primo e' per tutti i nomi, e consiste in un segnale simbolico che serve a distinguerli da qualsiasi altra parola e dunque, volendo, a crearne automaticamente una raccolta; il secondo intervento consiste invece in un + al posto del normale trattino (-) nel punto di separazione tra gli elementi formativi del nome, quando naturalmente questi elementi sono piu'di uno; manchera' cioè nei nomi del tipo Ubru, Šelubu, Gazru, ma dividera gli elementi dei nomi del tipo Assur+uballit, Eriba+Adad, Assur+resa+isi. Questo +, che nella stampa del testo compare invece sotto la veste del normale trattino, e' un avvertimento per il computer affinche' tenga presente, all'atto della elaborazione, che perquanto uniti perche' formativi di un nome personale gli elementi vanno in realta' separati ai fini dell'analisi. I problemi relativi alla resa di consonanti con segni diacritici nella tastiera della perforatrice sono stati risolti invece con la semplice sostituzione di altri segni (ad esempio la X per la T enfatica) che lasciano il posto a quelli regolari all'atto della conversione a stampa. La distinzione tra ideogrammi e valori fonetici e' espressa da un asterisco anteposto agli ideogrammi, sufficiente perche' essi vengano considerati tali al momento dell'analisi e perche vengano trascritti in maiuscolo, secondo la convenzione in uso, all'atto della stampa. Ai determinativi viene anteposta una sigla-simbolo, sufficiente non solo per distinguerli ai fini di una collezione a parte qualora si volesse realizzarla, ma soprattutto perche'l'analisi della parola cui sono legati non incontri difficolta dalla loro presenza. Sigle-simbolo di questo tipo sono anteposte anche ai nomi geografici e divini, affinche' non sfuggano ad una eventuale raccolta quelli che non hanno l'usuale determinativo, come URU, ÍD, KUR ecc. per i primi, ed il DINGIR per i secondi, come ad es. Sîn e Kūbu). I complementi fonetici degli ideogrammi sono considerati intimamente legati alla parola che l'ideogramma

esprimeva in lingua accadica, e sono quindi trattati in un secondo intervento dedicato allo scioglimento dell'ideogramma
stesso, mentre quelli usati per una migliore comprensione di
un valore sillabico raro sono distinti dal corpo della parola
da un simbolo specifico, in modo tale che l'analisi non sia disturbata dalla loro presenza, senza tuttavia che siano eliminati materialmente dal testo. Un altro simbolo, anteposto ad
ogni linea del testo, basta per la numerazione automatica delle linee. Una semplice indicazione permette infine di avere, in
fase di stampa, l'indicazione delle barre orizzontali divisorie e degli spazi bianchi presenti nella tavoletta, oltre che
il numero e la posizione dei sigilli.

I risultati della memorizzazione del testo cosi' codificato sono di varia natura. Anzitutto l'analisi grafemica. Il computer e' ora in grado di analizzare ogni tipo di testo cosi' codificato, di separare ogni segno, cui e stata data la numerazione tradizionale del Manuel d'Epigraphie Akkadienne di René Labat, e di riportarlo con la parte del contesto in cui e' inserito secondo una elencazione che tiene conto delle numerazioni dei segni immediatamente precedenti e/o seguenti. Quel che piu' importa, il segno viene riportato al giusto posto indipendentemente dal valore, sia sillabico che ideografico, espresso in quel momento, valore che viene riportato comunque tale e quale nella lista. E'evidente l'utilita' di un tale tipo di analisi, cui si puo giungere solo con l'uso del computer: confronti per stabilire l'esatto valore del segno in uno specifico contesto, constatazione rapida del valore del segno in contesti completi ai fini di una comprensione di contesti analoghi parzialmente abrasi o frammentari, studio statistico volto a stabilire l'uso di un segno piuttosto che quello di un altro avente il medesimo valore, studio sulla preferenza di scribi alcuni segni piuttosto che per altri, e di conseguenza attribuzione di testi redatti da scribi finora sconosciuti, e cosi' via. Un primo risultato di questo tipo, costituito di sei volumi provvisori e'gia' stato raggiunto a Pisa, a livello sperimentale, due anni fa (14). Particolarmente importante e' sottolineare che questo tipo di analisi, interessando l'epigrafia, puo' essere applicato anche subito a qualsiasi altro scritto cuneiforme di qualsiasi lingua (ittito, hurrico ecc.) previo qualche elementare intervento sul programma di analisi grafemica in nostro possesso.

Sul corpus, memorizzato una volta per sempre, intervengono anche i programmi di analisi grammaticale, oltre a quelli di
analisi sintattica ancora sotto studio. L'elaborazione elettronica riesce ad eseguire c. il 75% della analisi grammaticale,
isolando ogni parola ed analizzandola mediante il successivo
passaggio attraverso tabelle diverse, aventi ciascuna indicazioni di compatibilita o meno con le entrate delle altre, e capaci di isolare poi via via le forme invariabili, poi i vari
morfemi, poi le radicali. Successivi programmi sono in grado
naturalmente di categorizzare i risultati di queste analisi secondo i piu svariati criteri, che possono andare, ad esempio,
dall'isolamento di una forma verbale in tutte le sue attestazioni, a lavori lessicografici concernenti un intero corpus.

Ritornando alla codificazione, un discorso piu' lungo richiede quella delle schede di testa. Va premesso che le schede di testa sono suscettibili di un continuo aggiornamento (realizzabile con elementare facilita', come d'altronde quello delle traslitterazioni dei testi) ed anche ampliamento a categorie finora non considerate. Allo stato attuale le grandi categorie, identificabili ognuna con una sigla che automaticamente esclude le altre, sono cinque; una sesta, dedicata alla analisi storica, e' in fase di studio ed i risultati parziali sono stati discussi a Los Angeles anche con la parte italiana. La prima categoria di questi "dati secondari" riguarda le notizie di carattere bibliografico, tutte rapportate ad un dato "primario" che coincide con la pubblicazione principale del testo (di solito quella con copia autografa). Una serie di simboli

non da soltanto la sigla della pubblicazione, l'anno, il numero della pagina, la lingua usata; viene fornita infatti ogni notizia anche sulla natura del lavoro svolto a proposito testo in questione: ad esempio se si tratti di una copia soltanto, o di una trascrizione, o traduzione,o trattazione,o una qualsiasi delle loro combinazioni, con distinzione tra studi parziali o completi. Vengono inoltre indicate notizie di altro tipo: ad esempio se il dato bibliografico e uno studio su eventuali sigilli impressi nel documento, o se si tratta di uno studio limitato alla cronologia, o all'onomastica, o all'aspetto religioso. I simboli si riferiscono anche a eventuali collegamenti tra il documento analizzato ed altri documenti: se cioe' si tratti di un duplicato, ed in questo caso di quale testo, o se sia un joint, e di quale testo, o del contenuto dell'involucro, o dell'impressione del sigillo, e cosi'via. In questa categoria sono poi compresi i "dati personali" del testo, quali le sigle di scavo, di museo, o il numero della fotografia.

I dati geografici che costituiscono la seconda categoria, si riducono a tre, cioe' al luogo di ritrovamento del testo, a quello della sua provenienza (nome antico della localita', che puo' essere coincidente con il luogo di ritrovamento, oppure no, come nel caso delle lettere), ed infine nomi geografici nominati nel documento. Va da se' che questa semplice categorizzazione si allarga alle varianti: ad esempio vengono considerati tra i luoghi nominati anche i termini generici quali ana abulli nei testi di Nuzi, mentre tra i luoghi di provenienza sono compresi simboli che indicano un'area generica, quali i dintorni di una citta', o la parte nord-occidentale del paese (Ḥabūr), e cosi' via.

Piu' complessa e' stata la codificazione della terza categoria, relativa ai dati prosopografici. Trascurati i nomi di persona, tranne lo scriba e nei dati cronologici l'eponimo nei testi medio-assiri (e' in corso di stampa infatti un aggiornamento dell'opera-base), i dati prosopografici sono invece ap-

plicati ai testi nuziani, in vista di una elaborazione che possa fornire tutte le funzioni e le parentele dei numerosissimi personaggi attestati. Per chiarire meglio, l'uso del computer non mira certamente a creare un duplicato di Nuzi Personal Names (15) che, se porta titolature e parentele, e' basilare soprattutto per l'analisi linguistica dell'onomastica. L'intervento del computer mira invece a chiarire e unificare ogni dato relativo alla personalita del personaggio analizzato, quale appare dall'insieme dei testi. Cosi' viene indicata da una parte la posizione sociale che il personaggio aveva nell'ambito del suo entourage, dall'altra la posizione specifica che assumeva nell'ambito della situazione particolare descritta da ogni singolo testo. Per meglio chiarire, nella prima specificazione della sua personalita'si troveranno non solo le titolature classiche di professione mestiere o classe sociale, ma anche quelle che si evincono dal contesto e che non sono nominate: artigiano, generale, soldato, eponimo geografico, proprietario di immobili, o di bestiame, o di schiavi, e cosi'via; nella seconda specificazione invece una serie di simboli chiarisce se il personaggio funge nel testo da scriba, o da giudice, o nel caso di contraenti se e'un adottato, o un debitore, o un acquirente e cosi' via, fino alla specificazione di funzioni ancor piu particolari, quali quella di chi fornisce un avallo, o genericamente acquisisce o perde un bene (quando non si verificano le condizioni-tipo in cui avviene questo fenomeno, cioe' vendite, adozioni, matrimoni, eredita') o stipula contratti sostituendo uno dei contraenti, e cosi'via. Una terza indicazione e' dedicata alla parentela; evidenzio qui l'elaborazione di questo tipo di analisi non perche' sia piu' importante degli altri, ma perche'e'un contributo alla programmazione specificamente fornito dalla parte italiana. I nomi sono riportati in scrittura continua, secondo la lettura di NPN, quando naturalmente gli studi successivi non l'hanno modificata (ad esempio Síl-lí-Ku-bi e Zi-li-Ku-bi sono Silli-Kūbi). Una serie di simboli, di solito corrispondenti alle lettere iniziali dei termini accadici perche possano essere piu facilmente compresi dai
collaboratori stranieri, stabilisce il grado di parentela (o di
relazione, se si tratta ad esempio di schiavi e padroni; o di
riferimenti ad una generica genealogia) tra il soggetto ed altre persone nominate nel testo. Aggiunte, senza simboli, sono
anche le titolature ufficiali che nel testo sono attribuite al
personaggio. Un esempio ipotetico:

### NP1+TAMKA:RU M+NP2+M+NP3+DE:NU+B+NP4

(dove M =  $m\bar{a}ru$  sta per "figlio di", B =  $b\bar{e}lu$  sta per "padrone di" e tamkāru e dēnu, con la lunga indicata da due punti successivi alla vocale, sono due titolature presenti nel testo). Questo tipo di formulazione non solo permette di conoscere le parentele ed i titoli dei parenti che il testo fornisce per ciascun personaggio, e di confrontarli e integrarli automaticamente con i dati presenti negli altri testi, ma mediante un programma elaborato da Eugenio Picchi permette anche di intervenire su uno qualsiasi dei parenti stessi, in modo da poter trarre una caratterizzazione di ciascuno anche quando in tutti i documenti, o in parte di essi, e'attestato non in funzione attiva, ma solo come patronimico o capostipite di una generazione. Ad esempio se incontrassimo altrove il personaggio corrispondente a NP3, e volessimo cercarne altre attestazioni, il computer arriverebbe ad analizzare anche la scheda di testa in cui e' riportata la formulazione riferita sopra e, con una analisi inversa del ruolo delle persone, arriverebbe a presentare un NP3 dēnu, padrone di NP4, padre di NP2 e nonno di un tamkāru di nome NP1.

La quarta categoria (dati cronologici) ha costretto ad un cambiamento di codificazione a seconda del corpus considerato. Il perche e ovvio: se nei testi assiri abbiamo spesso indicazione di giorno, mese ed anno determinato dal nome del funzionario-eponimo, ma viceversa non sappiamo sempre sotto il regno

di quale sovrano debba essere collocato il testo, nelle lettere reali babilonesi il nome del sovrano e' piu' semplice da stabilire; nei testi di Nuzi abbiamo invece una conoscenza molto
nebulosa della cronologia, perche' le formule di datazione sono
rare e non puntualizzanti (tipo in arki šudūti ecc.) La diversita' di codificazione non significa di per se' una complicazione del lavoro del computer, perche' una specificazione del corpus introduce automaticamente la programmazione adatta al tipo
particolare di codifica, che si svolge secondo le esigenze (spesso complicate) e le conoscenze che abbiamo per ciascun corpus.

Infine, l'analisi dei tipi di testi: questa quinta categoria e'stata particolarmente sviluppata in Italia, data la varieta' dei testi, di cui ho dato sopra un breve accenno. In occasione della citata relazione al CNR i testi assiri memorizzati fino a quel momento sono stati presentati appunto secondo la tipologia, con una selezione avvenuta elettronicamente.

E'evidente che da tutta questa codificazione "secondaria", alla quale hanno collaborato più persone (Negri, Dosch, Giannotta, Aragona, Gonella) si può trarre automaticamente, con opportuni programmi, una serie quasi infinita di notizie per cosi' dire "incrociate", che evitano lunghe ricerche. A titolo di ipotetico esempio, una ricerca sull'attivita' e le parentele di un personaggio anche secondario, limitata al periodo di regno di un preciso re, ad una precisa area geografica e magari anche ad un particolare tipo di documenti, può essere svolta facilmente e rapidamente. E'poi altrettanto evidente che si adatta a molti altri tipi di testi dell'Oriente antico.

Avvenuta la codificazione, il testo viene passato alla memorizzazione al computer (16). L'elaborazione dei dati avviene
invece, ho detto, a Pisa, presso il "Laboratorio di Linguistica Computazionale" (diretto da Antonio Zampolli), ed a cura di
due collaboratori, Eugenio Picchi e Manuela Sassi. I primi risultati sono in relazione al numero dei testi memorizzati. Essi gia costituiscono una "banca dei dati" sia pure necessaria-

mente parziale. Gia' dal principio si e' cercato di dare una categorizzazione ai testi, in modo da costituire, nell'ambito di un medesimo corpus, dei gruppi coerenti. Per l'Assiro é stato seguito il criterio degli archivi, tranne che per i testi di tipo eccezionale, ad esempio le Leggi, raccolte in un gruppo omogeneo seppur provenienti da luoghi differenti. Per il Nuziano il criterio e'il medesimo, ma il computer può rapidamente ridisporre i testi e ripresentarli secondo la provenienza geografica, cioe' secondo una categorizzazione seguita a Heidelberg. Le pubblicazioni tengono conto di una tale ripartizione. La sede piu importante per l'edizione dei testi e degli studi e' "Cybernetica Mesopotamica" (Undena Pub. Malibu) fondata ediretta da Giorgio Buccellati. Questa collana ha loscopo di pubblicare i lavori elaborati al computer suddivisi in tre differenti settori. Il primo (Mechanism) e essenzialmente dedicato ai Manuali di codificazione; il secondo (Data) presenta invece i materiali memorizzati, con o senza un relativo commento; nel terzo (Results) vengono invece pubblicati i risultati dei vari tipi di analisi: la grafemica, la morfo-lessicale, la sintattica, la storica, ecc. I tre settori comprendono anche le pubblicazioni relative allo studio dei materiali archeologici. La parte italiana e' rappresentata nell'ambito di questa dalla Analisi Elettronica del Cuneiforme (AEC), componente italiana di Cybernetica Mesopotamica. E'nell'ambito di questa AEC che sono usciti i due primi lavori italiani (17), corrispondenti al settore Data Sets, Cuneiform Texts (DSC1 e DSC2). Per quanto limitate alla presentazione di testi (con le traduzioni ed il commento equivalgono in pratica alle monografie tradizionali) e con una utilizzazione del computer ridotta alla scelta automatica dei testi ed alla stampa, queste due pubblicazioni costituiscono un buon inizio ed un buon segno per il futuro. Esse sono, tra l'altro, le prime pubblicazioni in assoluto di "Cybernetica Mesopotamica"; altre sono gia'in preparazione,sia di tipo analogo a quelle gia edite (anche per il Nuziano) sia

di altro tipo che prevede una utilizzazione piu' radicale del computer (es. l'analisi secondaria dei testi economici, a cura di Saporetti-Aragona-Giannotta, e l'analisi grafemica dei segni e dei loro valori utilizzati dagli scribi che hanno redatto i documenti di Tell Billa, citta' assira a nord di Assur, a cura di Saporetti e Ghiroldi. E' in progetto un lavoro identico sulle cancellerie private della capitale). Contributi minori, studi per l'ampliamento ed il perfezionamento dei sistemi di elaborazione, notizie sui lavori di ogni équipe andranno poi raccolti in una rivista, "Bulletin for Computer Aided Research in Near Eastern Studies", inserita in un sistema di pubblicazioni periodiche concepite secondo una formula nuova e di cui fa parte anche "Assur", a cui ho gia' fatto cenno.

Studiosi affermati si aggiungono intanto ai collaboratori, Istituzioni di profonde tradizioni culturali prendono ormai in considerazione questi nuovi aspetti dello studio dell'Accadico, le nuove metodologie gia sembrano interessare gli studenti, almeno a Pisa; ed e evidente quali utilita essi possano trarre dallo studio dell'Informatica a fianco dello studio dell'Accadico, anche semplicemente per una formazione personale più completa. Personalmente credo che ormai il più sia stato fatto, che la fase più dura sia superata, e che sia tempo ormai di allargare i contatti per una esamina concreta dell'antico-Assiro e forse anche del medio-Babilonese cassita. In tal modo gli orizzonti saranno più vasti ma anche più definiti, ed il lavoro finora eseguito non sara stato inutile.

#### NOTE

- (\*) L'autore ringrazia vivamente il Centro Linceo Interdisciplinare di Scienze Matematiche e loro Applicazioni, in particolare il Direttore Professor Giorgio Salvini, per l'invito a tenere questa conferenza, ed il Professor Sabatino Moscati che ha voluto proporla.
- (1) WHITTAKER R., Concordance of the Ugaritic Literature (Cambridge, Mass. 1972).
- (2) in BUCCELLATI G., The Old Babylonian Linguistic Analysis Project, Biblioteca dell "Archivum Romanicum", Serie II, 36 (1977), 385 (p.392: Encoding: Programming; p.393: Graphemic Categorization; p.395: Morphological and Lexical Categorization; p.397: Text Processing and Photocomposition.
- (3) Ad esempio il segno ŠE puo essere presente sia con il significato sumerico di "orzo, grano" (e dunque usato con il suo originale valore ideografico, che puo ricordare il disegno primitivo o puo anche averlo ormai totalmente perduto) sia con il valore sillabico se in una qualsiasi parola accadica con la sillaba ŠE, tipo se-er-ru ("bambino"), a-se-ra-tu (un tipo di veste) e cosi via.
- (4) Cosi'per es. in a-li-du DU, dove il primo DU è l'ultima sillaba del termine "ālidu" ("genitore"), ed il secondo la lettura sumera dell'ideogramma con significato "andare". Il segno è lo stesso, derivante da un originale disegno di piede, ma l'uso è chiaramente diversissimo. Altro aspetto di questo fenomeno, da considerare ai fini di una memorizzazione al computer, e'l'uso diverso degli ideogrammi che raffigurano i numeri, perche'al loro impiego comune siaffianca talvolta l'abitudine di indicare, con uno di essi, una divinita' (ad es, il N. 30 corrisponde al dio Sîn, il dio-luna da cui deriva la ripartizione del mese in 30 giorni, ed il N.7 agli dèi-7 cioè i Sibittu).
- (5) Ad esempio l'ideogramma IGI puo essere letto in vari modi come  $\bar{i}nu$ , "occhio",  $am\bar{a}ru$  "vedere",  $\tilde{s}\bar{i}bu$  "testimone", mahar, "davanti a", e cosi'via; se dunque e' citato con il complemento fonetico -bi (IGI-bi), possiamo anzitutto stabilire che nel contesto va letto  $\tilde{s}\bar{i}bu$ , ed inoltre che la parola va interpretata al Genitivo singolare Ugualmente TI ha, tra le altre, la lettura  $bal\bar{a}tu$  con significato "vita" e "vivere", e dunque in un contesto in cui sia preceduto da  $\underline{u}$  ( $\underline{u}$ -TI) non puo avere che il significato "vivere", indicando la forma intensiva (D) del verbo (es. uballit).

- (6) Ad esempio il Preterito III m. s. di kšd e' ikšud, di solito scritto ik-šu-ud, ma talvolta anche ik-šud, cioe' con un segno di valore šud, poco usato. In tal caso, lo scriba ha preferito aggiungere a questo segno, che puo' avere anche altri valori, quali sir, su3, sar4, un altro segno di valore ud, per suggerire che la lettura giusta era šud (ik-šud-ud).
- (7) Ad esempio la parola naš lam tu, "contraccambio", normalmente scritta na-aš-la-am-tu, talvolta e' scritta na-aš-lam-tu, sicche' e successo che lo scriba, per eccesso di zelo e preoccupato che il segno con valore lam, raro, non fosse subito capito, ha premesso un segno con valore la, con il risultato che la parola risulta scritta na-aš-la-lam-tu.
- (8) Attualmente la Divisione Linguistica non e'più legata al CNUCE di Pisa, ma e'diventata Laboratorio di Linguistica Computazionale.
- (9) ZAMPOLLI A., CALZOLARI N. (Editors), Computational and Mathematical Linguistics. Proceedings of the International Conference on Computational Linguistics, Pisa 27/VIII-1/IX 1973, Biblioteca dell' "Archivum Romanicum" Serie II, 36 (cf. nota 2).
- (10) Questo concetto fondamentale e' stato poi evidenziato anche in una conferenza tenuta al Consiglio Nazionale delle Ricerche nel 1976 ed in un articolo esplicativo sulla rivista Egitto e Vicino Oriente I (Pisa 1978).
- (11) Il Progetto americano si svolge essenzialmente presso lo "Institute of Archaeology" e lo "Office of Academic Computing" dell'Universita di California, ma ha coinvolto gli interessi del "Research Committee" della stessa Universita, la "Ambassador-International Cultural Foundation" di Pasadena, il "National Endowment for the Umanities" di Washington D.C. Il Progetto, divenuto ora "Computer Aided Analysis of Mesopotamian Materials" (cui se ne e aggiunto un altro, dedicato all'analisi archeologica, "Mesopotamian Area Program") si sta allargando alla collaborazione di altre Istituzioni, americane ed europee; cito, per tutte, l'intensa collaborazione con Olivier Rouault, del "Cabinet d'Assyriologie" del Collège de France.
- (12) Ad esempio egli ha ripreso la Keilschriftbibliographie (iniziata da Padre Alfred Pohl una quarantina di anni fa per la rivista Orientalia, e proseguita poi a cura di vari autori) ristrutturandola, ampliandola e compilandola prima con la collaborazione di Horst Klengel, e poi anche con quella di Richard Caplice.
- (13) In occasione di una ricerca condotta in collaborazione con Helmut Freydank, dell'Accademia delle Scienze della Repubblica Democratica Tedesca.

- (14) Basato su questa analisi grafemica è anche un lavoro di Giorgio Buccellati, che considera l'uso statistico dei segni, con reciproci confronti, nei vari sotto-corpora dell'antico-Babilonese e dell'Accadico occidentale, in corso di stampa.
- (15) GELB I.J., PURVES P.M., MAC RAE A.A., Nuzi Personal Names, OIP 57 (Chicago 1943), con aggiornamento in E. Cassin e J.J. Glassner, Anthroponymie et Anthropologie de Nuzi, I (Malibu 1977).
- (16) Dopo tentativi di utilizzazione di ditte romane e pisane di limitato rilievo, il Progetto e'approdato ad un organismo assai considerevole, cioe' la Ditta Montiani di Firenze, dove ad una esecuzione tempestiva e molto corretta si e'aggiunto, inopinatamente, anche un interesse per l'Accadico che e'ulteriore garanzia di cura e quindi di perfezione.
- (17) SAPORETTI C., Assur 14446: la Famiglia A. Ascesa e declino di persone e famiglie all'inizio del medio-regno assiro, I, DSC 1 (Malibu 1979); IDEM, Le Leggi medioassire, DSC 2 (Malibu 1979). Electronic Data Processing by E. Picchi and M. Sassi.

