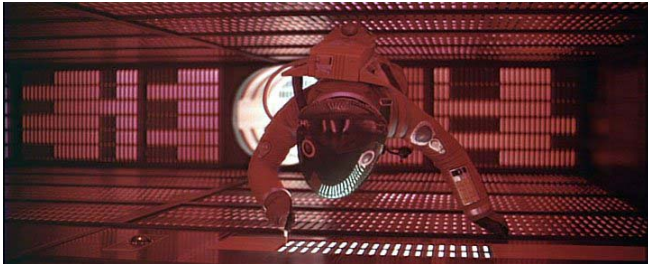


Da oggi parlare al computer è facile (Corriere Scienza, novembre 1991)

IBM ha sviluppato il primo computer con riconoscimento vocale - Utilizzabile un vocabolario di 20.000 parole - Applicazione nella refertazione radiologica - Accuratezza nel riconoscimento superiore al 97% - Costo del prodotto 30 milioni.

E' caduta un'altra barriera che divideva l'uomo dal computer. Il centro di ricerca IBM di Roma (oggi IBM SEMEA) in collaborazione con quello di Yorktown negli Stati Uniti, ha infatti sviluppato il primo sistema di riconoscimento vocale in grado di utilizzare un vocabolario fino a 20.000 parole interpretate in tempo reale e con un'**accuratezza di oltre il 97%**. L'uomo può così farsi capire dal computer con un'errore nel riconoscimento di sole 3 parole su cento. Si tratta senza dubbio di un evento importante perché per la prima volta ed in modo semplice e veloce l'uomo può comunicare con il computer attraverso il linguaggio naturale della parola.



Il sistema funziona su una piattaforma hardware composta da un Personal System/2 IBM modello 80 equipaggiato con microprocessore **Intel 80386** (o meglio 80486), per il riconoscimento vocale sono necessari circuiti elettronici aggiuntivi formati da chip specializzati nell'interpretazione dei segnali acustici, segnali che variando con continuità nel tempo caratterizzano tutti i parametri della voce umana. Naturalmente, per completare l'apparecchio sono indispensabili un microfono ed un altoparlante. Come software di base si ricorre all'AIX (un sistema operativo di classe Unix) attraverso il quale agisce il programma specifico sviluppato da IBM, il vero cervello dell'intero sistema.

Il riconoscimento della voce da parte delle macchine ha da sempre affascinato il mondo scientifico, in particolare quello dell'Intelligenza Artificiale che ha analizzato il problema attraverso l'uso dei sistemi esperti e della programmazione logica ricercando le regole sintattiche dei fonemi elementari. IBM ha invece percorso una strada diversa. Partendo dal presupposto che l'uomo pronuncia la stessa parola in modo diverso dipendente dal suo stato di salute, dal suo umore e da altri fattori sociali e culturali, ha studiato un'approccio statistico del problema.

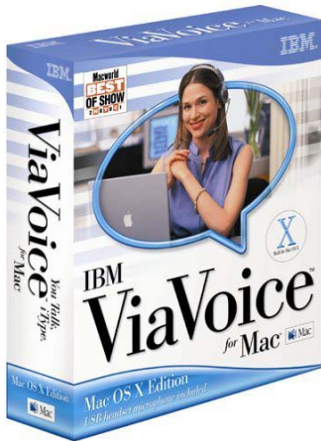
Si è fatto uso dei **modelli di Markov** nei quali il linguaggio viene considerato nella sua natura probabilistica non riferendosi a campioni fonetici della parola, ma a modelli matematici nei quali si percorrono i cammini che presentano la maggiore probabilità nel ricostruire il termine completo. In pratica ogni parola viene spezzata in tanti frammenti ognuno dei quali è confrontato attraverso l'analisi probabilistica con quello che più si avvicina nel vocabolario posto in memoria.

Bisogna osservare che questi risultati sono resi possibili grazie ai continui sviluppi della tecnologia microelettronica che ha messo a disposizione memorie di grande capacità nelle quali vengono memorizzate migliaia di parti sillabiche o vocaboli interi. Si supera così l'ostacolo di codificare complesse regole di pronuncia o elaborate costruzioni della frase incapaci di affrontare le eccezioni e le variazioni di pronuncia.



Si supera così l'ostacolo di codificare complesse regole di pronuncia o elaborate costruzioni della frase incapaci di affrontare le eccezioni e le variazioni di pronuncia.

La prima interessante applicazione del sistema vocale IBM riguarda la dettatura dei **referti medici** in particolare quelli del settore radiologico. Si è già conclusa con successo la sperimentazione attuata sotto l'egida del Ministero della Sanità in 10 ospedali italiani - dice l'Ing. Giorgio Sommi Direttore del centro di ricerca di Roma - i vantaggi riscontrati sono una notevole riduzione del tempo complessivo di refertazione ed una maggiore efficienza del sistema informativo ospedaliero. Dal punto di vista operativo il medico accede al computer e dopo essersi presentato con il proprio nome, analizza il materiale diagnostico e detta al microfono il referto che appare immediatamente sullo schermo in forma scritta quindi controllabile e modificabile nell'ortografia.



Terminata la dettatura il referto è disponibile in stampa e può essere archiviato in tempo reale nel sistema informativo dell'ospedale. Il radiologo può dettare fino ad 80 parole al minuto scelte tra le 13.000 del vocabolario disponibile e non deve sottostare ad alcuna regola se non quella di interporre una breve pausa tra una parola e la successiva.

Sostanzialmente - continua Sommi - sono possibili due modalita' di lavoro: la prima quella che garantisce un tasso di riconoscimento del 97% prevede una procedura iniziale di **apprendimento da parte del computer**, da compiersi una sola volta; il medico legge al microfono un testo campione di 20 minuti ed attraverso questo vengono identificati in modo univoco i parametri vocali. La seconda ipotesi non prevede questa fase di "calibrazione" ma l'accesso diretto alla refertazione che risulta comunque fattibile con un riconoscimento garantito del 90%. Il costo del sistema già disponibile sul mercato (computer escluso) è, chiavi in mano, attorno ai 30 milioni.

Per il futuro IBM ha intenzione di rendere disponibili, con l'inserimento di specifici vocabolari, prodotti applicativi per il settore assicurativo, bancario ed amministrativo pensando anche a sistemi dedicati ai portatori di handicap che non possono accedere all'uso della tastiera.

###