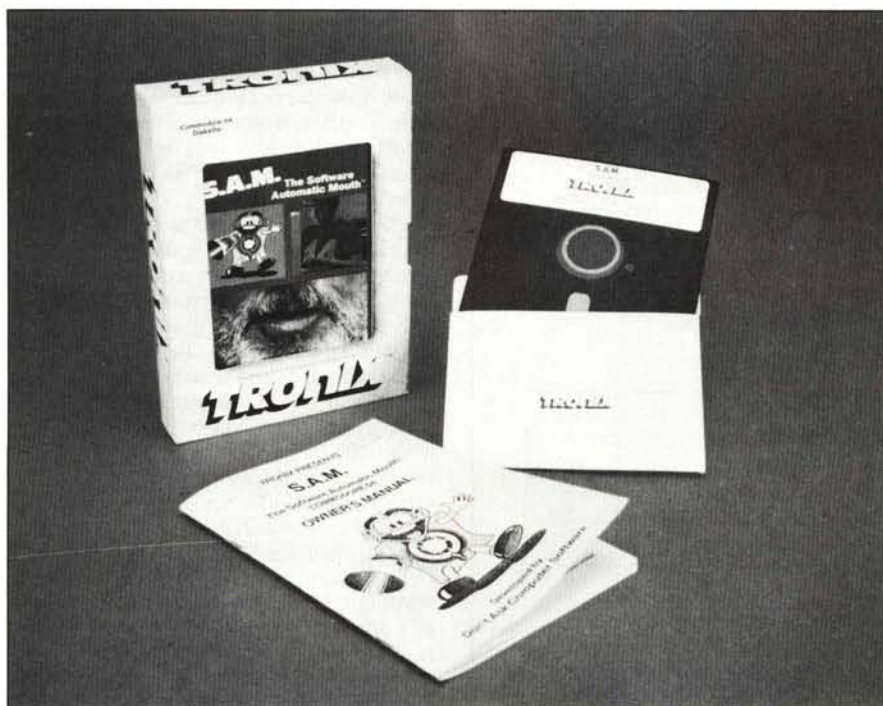


Sintesi vocale via software con il Commodore 64:



Software Automatic Mouth

di Leo Sorge



La tecnica per la sintesi vocale attualmente più sfruttata (vedi l'articolo nelle pagine precedenti) è quella per allofoni, consistente nel costruire ogni parola con successivi suoni da prendersi in un elenco fisso per ogni lingua e comunque assai limitato (40-50 elementi). Questo procedimento, basato su un'approssimazione (in frequenza ed ampiezza) delle forme d'onda che caratterizzano i suoni emessi dall'apparato fonatorio umano, è in pratica un programma molto complesso, per cui può essere realizzato non solo da hardware, con un chip specializzato ovvero con più chip, ma anche da software, realizzando un programma che sfrutti una tavola di dati relativi agli allofoni e li metta in uscita sull'audio del computer.

È questa la scelta fatta dalla californiana Don't Ask Computer, che produce questo dischetto per il Commodore 64. Vediamolo nei dettagli.

Costruttore:
Don't Ask Computer
2265 Westwood Br., Ste. B-150, Dept. C,
Los Angeles, 90064 California.



Una volta caricato il SAM, questo programma vi segnalerà — scrivendolo anche sullo schermo — la direzione che impostate con il joystick, più la pressione del tasto di fuoco. Tutte le REM e le righe con i soli due punti possono essere omesse.

Cosa fa

Il SAM, sigla di Software Automatic Mouth, bocca automatica da programma, permette di inserire parole nel computer in due modi: o come stringhe fonetiche (es: KAA4MAHDOHR per Commodore) oppure come semplici parole e numeri (Commodore 64). Per la seconda opzione bisogna caricare un programma aggiuntivo, Reciter, che fa l'analisi fonetica della parola. È evidente che questo secondo modo, anche se assai più semplice, fornisce un risultato meno comprensibile rispetto al primo, ma in questo modo non si deve essere in grado di eseguire di persona la scomposizione in allofoni.

Il modo più versatile, comunque, rimane proprio il primo, che ci permette di inserire pause ed intonazioni (vedi tab. 1) sia per i suoni vocalici che per quelli consonantici, alterare la velocità del parlato, la sua tonalità (che determina se la voce è maschile, femminile, di un adulto, ecc...).

Una volta caricato SAM (con la solita LOAD "SAM",8,1) bisogna dargli il RUN, con il quale lui si carica tutti i dati (circa 8K in linguaggio macchina) più le routine di un linguaggio di una decina di comandi che va ad inserirsi dentro il Basic del 64. Di questi comandi ci interessano in modo particolare PITCH, SPEED, LIGHT, SAM, RECITER e SAY.

PITCH frequenza accetta un parametro da 0 a 255 (0 dà la massima frequenza), e varia il timbro della voce (il valore di default, cioè quello che vale al caricamento è 64);

SPEED velocità, con velocità sempre da 0 a 255 (0 = max vel.), altera il numero di suoni al secondo. È molto importante per gli effetti del comando LIGHT.

LIGHT p si rivolge allo schermo video, e lo accende (p=1) oppure lo disabilita (p=0); la presenza dell'uscita video peggiora la comprensibilità e rallenta l'esecuzione, quindi (se possibile) tenetelo spento oppure accelerare la lettura;

SAM specifica che seguono ingressi di tipo fonetico;

RECITER specifica che seguono parole da codificare;

SAY stringa, ove la stringa è sia diretta ("GIORGIO") che per variabile (AS =

Tabella 1

SAM consente di alterare la pronuncia delle parole in funzione di eventuali inflessioni. Il metodo adottato consiste nell'uso del programma SAYIT — che abilita alla lettura di stringhe fonetiche — più alcuni codici che denotano l'emozione nella lettura del computer. Gli otto valori ammessi hanno i seguenti significati:

- 1) molto emotivo;
- 2) molto enfatico;
- 3) piuttosto forte;
- 4) normale;
- 5) leggero;
- 6) neutro (senza alterazioni di frequenza);
- 7) con diminuzione di frequenza;
- 8) con grande diminuzione di frequenza.

We will have S.A.M. say

"Why should I walk to the store?"

in a number of different ways.

1. WAY2 SHUH7D AY WAO5K TUX DHAH STOHR.
(You want a reason to do it.)
2. WAY7 SHUH2D AY WAO7K TUX DHAH STOHR.
(You are reluctant to go.)
3. WAY5 SHUH7D AY2 WAO7K TUX DHAH STOHR.
(You want someone else to do it.)
4. WAY5 SHUHD AY7 WAO2K TUX7 DHAH STOHR.
(You'd rather drive.)
5. WAY5 SHUHD AY WAO5K TUX DHAH STOHR2OH7R.
(You want to walk somewhere else.)

Cinque diverse intonazioni della frase "Why should I walk to the store?" (trad.: "Perché dovrei camminare fino al negozio?"). Le interpretazioni sono le seguenti:

- 1) vuoi una ragione per farlo;
- 2) non vorresti andarci;
- 3) vuoi che lo faccia qualcun altro;
- 4) preferiresti andare in auto;
- 5) vuoi passeggiare altrove.

burglar = BER4GULER
bus = BAH4S
business = BIH4ZNIXS
busy = BIH4ZIY
by = BAY4
byte = BAY4T

C

cabinet = KAE4BUNIXT
cable = KEY4BUL
calculate = KAE4LKYAXLEYT
calendar = KAE4LUNDER
call = KAO4L
calorie = KAE4LERIY
cancel = KAE4NSUL
candy = KAE4NDIY
can't = KAE4NT
capacity = KAXPAE4SIXTIY

coaxial = KOHAE4KSIYUL
coffee = KAO4FIY
coherent = KOW/HEH4RIXNT
cold = KOW4LD
college = KAA4LIXJ
color = KAH4LER
comfortable = KAH4MFTERBUL
Commodore = KAA4MAHDOHR
common = KAA4MUN
company = KAHM4PUNIY
complain = KUMPLEY4N
complex = KUMPLEH4KS
component = KAHMPOH4NUNT
computer = KUMPYUW4TER
condition = KUNDIH4SHUN
conscience = KAA4NSHUNTS
console = KAA4NSOHL
control = KUNTRH4L

Alcuni esempi, tratti dal dizionario in dotazione al SAM, di costruzione fonetica di una parola.

"GIORGIO" SAY A\$) che con tutti gli operatori su queste quantità (LEFT\$, RIGHT\$, MID\$, concatenazione) purché ovviamente con meno di 255 caratteri.

Da far notare che la massima durata continuata del parlato del SAM è di circa 2,5 secondi, per cui mettete qualche pausa o ce le metterà lui.

Tutti i comandi (questi e gli altri abilità) vanno fatti precedere dal simbolo "]" (SHIFT + punto e virgola), tranne il SAY che non lo vuole, e sono accettati anche nella forma accorciata, lunga almeno 2 caratteri dopo il simbolo].

Un po' di operazioni con questo programma ci hanno rivelato diversi aspetti da conoscere sui comandi aggiunti. Questi vengono accettati sia in inizio di linea che dopo i due punti, cosa interessante dato che spesso tool anche assai più complessi (come il Simon's Basic) non lo consentono; peraltro non possono esser inseriti in verifiche fatte con l'IF-THEN. Molto spesso, inoltre, come stringhe per il SAY non vengono accettate quelle con indice: questo problema può spesso esser risolto raggrup-

pando tutte le variabili letterali in un'unica senza indice, e poi tirar fuori le componenti con l'istruzione MID\$ e un'opportuna tabella (vedi altrove nell'articolo il programma Talkin'Joy).

Compatibilità

Il SAM ha bisogno di una sua zona di memoria: usualmente, nella massima configurazione, lascia comunque circa 26K di RAM, che però potrebbero anche non bastare. A tale scopo vengono fornite diverse indicazioni sul modo di limitare le possibilità del programma onde evitare il caricamento di parte del SAM. Alcune di queste economie hanno però degli effetti collaterali sulla gestione del sistema, per cui il nostro consiglio è di evitare sempre riduzioni del programma di sintesi, concentrandosi invece sulle ridondanze del programma dell'utente.

Un secondo problema è dato dagli effetti collaterali del metodo usato per generare l'uscita vocale: il SAM, mentre genera il segnale, intanto disabilita tutti gli sprite,

poi (ovviamente) arresta qualsiasi altra operazione in corso, quindi disabilita gli interrupt (con essi il tasto di Stop e quello del Restore).

Infine, anche se compatibile con il DOS 5.1 (quello sul dischetto in dotazione al lettore di dischi), il wedge del SAM può interferire con quello di base del 64 (ma questa eventualità, citata dal manuale del SAM, a noi non è mai capitata).

Conclusioni

Le seccature imposte dalla natura software di un generatore di segnale vocale sono essenzialmente due: la prima è che quando il computer "parla" non può far altro, mentre un modulo esterno ha bisogno dei soli dati in ingresso; inoltre, poiché l'uscita audio del 64 è sull'altoparlante del televisore, i vari passaggi del segnale audio (modulazione - demodulazione - amplificazione - riproduzione tramite l'altoparlante del TV) peggiorano di gran lunga la distorsione totale, contribuendo a rendere le frasi incomprensibili.

Peraltro far dei giochini parlanti è molto bello, specie sugli home computer, che costano poco: chi desiderasse collaborare con la rubrica di software, comunque, farebbe bene a prevedere il funzionamento del programma anche senza il parlato.

Il prezzo italiano di questo prodotto, non importato ufficialmente e quindi reperibile unicamente per l'interessamento di negozianti organizzati, è di circa 150.000 lire: non lo riteniamo fondamentale per il gioco, ma molto più per scopi pratici come ausilio e addestramento per disabili fisici e non vedenti, nel qual caso diventa un elemento praticamente indispensabile. Per quest'uso consigliamo l'impiego di un altoparlante amplificato esterno, il cui segnale di ingresso va preso dal piedino 3 del DIN pentapolare che va al video, e l'alimentazione derivata dai 9 V AC dei piedini 10 e 11 della user port. L'uscita di un tale circuito è disturbata dalla radiofrequenza presente sui rimanenti piedini del DIN delle connessioni audio-video.

```

1 REM =====
2 REM === TALKIN' JOY PROGRAM ===
3 REM =====
4 :
5 DATA "NORTH SOUTH WEST NORTHWEST EAST NORTHEAST SOUTHEAST"
6 DATA 0,0,1,5,7,5,13,0,13,4,18,9,28,9,38,0,38,4,43,9,53,9
7 READ Y#: X#="YOU FIRED!"
8 PRINT"D   S   SPEAKING JOY PLAYER   MCM"
9 FORT=0T010:READ F%(T),G%(T):NEXT
10 :
11 REM == SCHERMO & VELOCITA'
15 JLI1:1SP50
16 :
17 REM == FRASE DI APERTURA
20 SAY"CIAO, SONO IL --TAWL--KEYN --JOMEY"
30 :
170 REM == PROGRAMMA PRINCIPALE
180 GOSUB1000:IF JV=0THEN190
182 C#=MID$(Y#,F%(JV),G%(JV))
183 PRINTC#-SAY C#
190 IFFR=16THEN180:REM = FUOCO PREMUTO
192 PRINTX#-SAY X#:GOTO180
200 :
1000 REM == SUB LETTURA JOYSTICK   1010 :
1020 JV=PEEK(56320):FR=JVAND16
1030 JV=15-(JVAND15)
1040 RETURN

```