



Introduzione a SoX: il coltellino svizzero dei formati audio

SoX (il nome è una contrazione di SOUND eXchange) è una utility a riga di comando per il trattamento audio, la cui funzionalità principale è la conversione tra formati differenti. Essendo costruito secondo una architettura modulare, nel tempo, molti programmatori hanno potuto estenderne le funzionalità, e quindi si è arricchito di una vasta gamma di formati trattabili.

Per questo motivo, e per evitare di trasformare questo articolo in un elenco, citiamo solo alcuni dei formati supportati: oltre ai soliti mp3, ogg e wav, figurano cdr (cd musicali), voc (Sound Blaster), oss e als (dai rispettivi sottosistemi audio), txw (campionatori Yamaha) e gsm (quello dei telefoni cellulari). Per avere una panoramica completa sulla copertura dei vari formati si rimanda alla pagina di manuale ("man sox"), eventualmente coadiuvata dall'ottimo "Audio File Format FAQ"

<http://www.cnpbagwell.com/audio.html> scritto dallo stesso mantainer del programma ed alla "Analysis of sample rate conversion in SoX"

<http://leute.server.de/wilde/resample.html>

sempre linkata dalla home page del programma.

Come abbiamo già detto la conversione tra formati è solo una delle funzioni disponibili, infatti SoX lavora come una *pipe*, cioè interponendosi tra input e output ed applicando le trasformazioni volute.

Nel corso dell'articolo scopriremo come questa modalità di funzionamento permetta una grande elasticità d'uso, potendo essere sfruttato, ad esempio, per registrare e riprodurre in tutti i formati supportati. L'utility rende inoltre disponibili varie

funzionalità accessorie tra cui spicca una nutrita serie di effetti.

Da questa breve panoramica si può capire come SoX si sia guadagnato "sul campo" l'appellativo di "coltellino svizzero dei formati audio". Sottolineiamo comunque fin da subito che SoX non è un comando "facile", e sarà necessario un certo studio per poterne padroneggiare tutte le funzionalità. Richiede inoltre una certa conoscenza, hardware e software, del proprio sistema audio. D'altronde un solo comando per convertire tra formati, registrare, riprodurre, applicare effetti...

Download ed installazione

SoX è un software opensource, rilasciato sotto licenza GPL, e quindi in forma di sorgente. Sulla home-page sono inoltre disponibili binari precompilati per una nutrita serie di piattaforme tra le quali, ovviamente, Linux (anche come pacchetti rpm e deb), ma anche Windows, BeOS, BSD, Solaris e il buon vecchio Atari ST.

Nell'ipotesi che la vostra distribuzione non lo includa già è possibile scaricare l'ultima versione (al momento la 12.17.8) dal sito

<http://sox.sourceforge.net>

e dopo aver scompattato l'archivio in una directory temporanea digitare i soliti:

```
# ./configure && make
# make install
```

Naturalmente per piattaforme diverse da quella del pinguino la procedura può essere leggermente diversa e si troveranno istruzioni dettagliate nel file INSTALL, allegato ai

Introduzione a SoX, utility a riga di comando per il trattamento dei file audio. Un solo comando per convertire, mixare, registrare, riprodurre, applicare effetti...



Matteo Lucarelli

m.lucarelli@oltrelinux.com

Ingegnere aeronautico e programmatore dall'età di 11 anni, lavora come consulente nella progettazione e prototipazione di sistemi innovativi di monitoraggio ambientale e sicurezza. Mototurista e musicista dilettante (theorobians.it), non disdegna, a tempo perso, l'attività di web designer. Quando possibile preferisce lavorare con strumenti open-source (e rilasciare il codice sotto GPL).



sorgenti. Per la compilazione sono richieste, oltre ad una toolchain funzionante (cioè compilatore, linker e librerie di base), le librerie *ogg-vorbis* (per l'omonimo formato) e *liblame* (per il formato mp3). Nel caso in cui non si disponesse di queste librerie SoX verrà compilato privo del supporto per i rispettivi formati. Va notato che la dipendenza da librerie esterne per i due formati di compressione più diffusi è in realtà un vantaggio, perchè ci assicura, ad esempio, che l'mp3 prodotto da SoX abbia lo stesso livello qualitativo di quello prodotto da Lame, e quindi ci risparmia la difficile scelta tra due algoritmi di compressione differenti. Al termine dell'installazione, se è andato tutto bene, digitando da shell (anche da utente normale):

```
# sox -h
```

verrà presentato un utile elenco dei formati e degli effetti disponibili. Tale elenco rispecchia le reali possibilità della versione che avete appena compilato ed installato sul vostro sistema. Un'ultima nota: in rete esiste ancora la versione originale di Lance Norskog <http://www.spies.com/Sox> la cui pagina web è invariata dal 1995 (probabilmente un record). Quella attuale, ancora attivamente sviluppata, è invece mantenuta da Chris Bagwell e si trova su sox.sourceforge.net.

Sintassi di SoX

Tra i bug segnalati nella pagina man di SoX (per la precisione come unico bug) è elencata la sintassi (!). Secondo l'autore "la sintassi è orribile". Naturalmente prosegue specificando: "questo è quello che succede tentando di controllare tutto dalla riga di comando". Il problema è vero ed è peraltro condiviso da molti comandi Unix-style. SoX infatti appare troppo potente per essere "solo" un comando di shell.

Naturalmente questa caratteristica lo rende invece molto comodo per essere utilizzato da altri programmi o per essere inserito all'interno di script.

Ma andiamo con ordine: la *figura 1* illustra schematicamente la logica del comando tipo di SoX. La prima sezione comprende il comando `sox` e i flag opzionali generali, tra cui la verbosità

(-v), l'utile help dinamico (-h) e la variazione generale di volume (-v).

La seconda e la terza sezione descrivono rispettivamente l'input e l'output secondo uno schema analogo, cioè specificando prima il formato, con tutti i flag necessari, e successivamente il percorso o il nome del file. In particolare il formato può essere specificato con il desiderato dettaglio, lasciando al "buonsenso", ed alle ottime capacità di autoriconoscimento di SoX, i parametri non specificati.

Il nome del file può anche essere sostituito dal segno meno (-), intendendo in questo modo di utilizzare lo standard input oppure lo standard output. Questa possibilità è utile laddove si voglia usare SoX all'interno di comandi più articolati, ad esempio in una pipe (cioè si voglia leggere l'input o passare l'output direttamente ad un altro programma).

L'ultima sezione descrive gli effetti applicati, e può essere replicata nel caso di applicazione simultanea di più effetti.

Come si vede uno dei problemi, almeno iniziali, dell'uso di SoX risiede nel fatto che la sintassi utilizza una logica posizionale, cioè il significato dei vari parametri di chiamata varia in funzione della loro posizione nel comando digitato: lo stesso flag posizionato in punti differenti del comando agirà sull'argomento seguente, sia questo il file di output o quello di input.

Utilizzi "al volo"

Visto così sembrerebbe tutto terribilmente complicato (e probabilmente lo è), ma arrivati a questo punto della lettura merita un premio: evidentemente conscio della difficoltà, l'autore ha previsto una seconda pagina di manuale (oltre la canonica "man sox") denominata "soxexam" ("SoX examples").

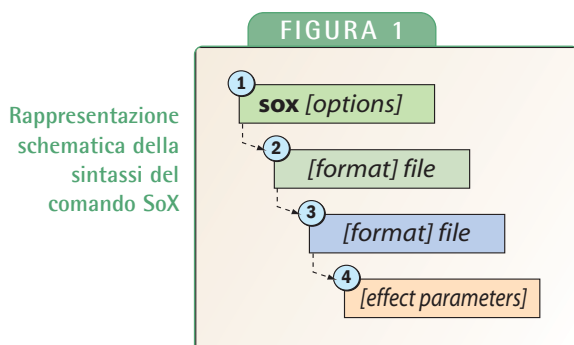
I più perspicaci avranno già capito che aprendo un terminale e digitando "man soxexam" ci verrà presentato un sintetico, ma utilissimo elenco ragionato di esempi di utilizzo. Dall'elenco, finalmente, apprendiamo che per convertire un file audio da un formato ad un altro (ad esempio per comprimere un file wav in formato mp3) è sufficiente digitare:

```
# sox file1.wav file2.mp3
```

senza specificare altro. SoX infatti ricava il formato desiderato dall'estensione, applicando una serie di ragionevoli valori di default per tutti i parametri non specificati. Naturalmente è possibile controllare con il desiderato livello di dettaglio tutti i parametri della conversione effettuata.

Ad esempio per comprimere un file mp3 rispettando uno specificato bitrate (ad esempio 16KHz):

```
# sox file1.wav -r 16000 file2.mp3
```





(si noti come il flag `-r 16000`, essendo posto prima del file di output, si intenda riferito a quest'ultimo, e quindi specifichi il risultato finale desiderato).

Se volessimo invece comprimere in `ogg-vorbis` mono un certo campione:

```
# sox filename.wav -c 1 filename.ogg
```

(il flag `-c` è seguito dal numero di canali).

Naturalmente questo utilizzo è sfruttabile per la conversione tra tutti i formati supportati, e considerato anche il grado di controllo possibile, basterebbe questo a considerare di sostituire con SoX tutto l'insieme di comandi tipo `lame`, `oggenc`, `oggdec`, ecc, anche solo per non doversi ricordare le diverse sintassi. Ma non finisce qui...

Hard-disk recording

Come abbiamo già visto tra i formati disponibili ci sono anche OSS e ALSA, cioè i formati nativi dei due sottosistemi audio utilizzati in Linux (per la precisione nel manuale di SoX vengono definiti pseudo-formati, infatti si occupano anche dei comandi necessari al settaggio del sistema audio).

Chi mastica un po' di Unix avrà già intuito che questo offre la possibilità di effettuare direttamente registrazioni e riproduzioni di file audio in qualsiasi formato, convertendo lo stream nel passaggio tra l'hard-disk e il device audio. Infatti, com'è noto, nel mondo Unix tutti i dispositivi sono accessibili nella directory `/dev` come file. Il che, tradotto in parole più semplici, significa che è sempre possibile registrare il suono campionato dalla nostra scheda audio digitando semplicemente:

```
# cat /dev/dsp > file.raw
```

(il nome esatto del device può variare in funzione dell'hardware e della distribuzione utilizzata, per una spiegazione più esauriente si veda il riquadro 1).

In questo modo l'input proveniente dal convertitore analogico/digitale della nostra scheda (tutte le schede audio ne hanno almeno uno) viene scritto direttamente nel file specificato, realizzando quindi una registrazione.

Naturalmente invertendo input e output con il comando:

```
# cat file.raw > /dev/dsp
```

è possibile riprodurre il suono registrato (anche in questo caso il device può essere leggermente diverso sul vostro sistema). Questo semplice sistema trova però un limite nel momento in cui si vogliono controllare i parametri della registrazione (frequenza di campionamento, numero di canali, ecc) o si vogliono registrare o riprodurre formati differenti o compressi. E qui entra in gioco il nostro SoX, interponendosi tra il device e l'hard disk ed effettuando per noi tutte le conversioni necessarie.

Il seguente comando, ad esempio, registra un file wav stereo ad 48KHz (qualità DAT):

```
# sox -r 48000 -c 2-t ossdsp /dev/dsp filename.wav
```

Per tutti quelli che si fossero spaventati leggendo il comando, il pacchetto sorgente di SoX include anche due comodi script (che si chiamano `guardacaso play` e `rec`) pronti a fare il "lavoro sporco" per noi. Il comando:

RIQUADRO 1



Il device DSP: di cosa si tratta?

Il device `dsp` (il cui nome deriva da "digital signal processor") è il punto di accesso al convertitore analogico/digitale cioè al circuito della scheda audio che si occupa di trasformare in bit il suono udibile, e viceversa. Normalmente scrivendo sul device si opera una trasformazione da dato a suono (D/A), e quindi si ottiene la riproduzione del file, mentre leggendo dal device si opera la trasformazione inversa (A/D) e quindi la registrazione (o, più precisamente, il campionamento). Come si può intuire qualsiasi scheda audio deve avere almeno un convertitore A/D (altrimenti non potrebbe produrre suono). Molte ne hanno più di uno,

quindi troveremo più di un `/dev/dsp`. In tal caso i device saranno nominati `dsp0`, `dsp1`, ecc, mentre `/dev/dsp` sarà probabilmente un link al primo di questi. Avere almeno due `dsp` assicura la possibilità di operare un full-duplex, cioè di poter registrare e riprodurre contemporaneamente. E' inoltre possibile (in realtà è molto probabile) che la nostra scheda audio abbia un solo `dsp` pur avendo più di un ingresso fisico (ad esempio CD e microfono). In tal caso il collegamento tra ingresso e convertitore, cioè la scelta di quale canale verrà registrato, si effettua tramite un altro device audio, che, guardacaso, si chiama `/dev/mixer`.

L'accesso a tale device è naturalmente possibile in modo diretto, anche se è consigliabile pilotarlo tramite uno dei tanti mixer grafici disponibili, con i quali è ovviamente possibile variare anche i livelli di registrazione e riproduzione, ed avere tutto sotto controllo con un colpo d'occhio.

NOTA: Il discorso fatto si applica propriamente solo al vecchio sottosistema audio, denominato OSS. Il nuovo sottosistema (ALSA), più completo e complesso, include comunque anche un livello di emulazione delle funzionalità dell'OSS, quindi tutto dovrebbe funzionare in entrambi i casi.

RIQUADRO 2



Massimizzare il volume di un file audio

Come si è visto nel corso dell'articolo, SoX permette di massimizzare il volume di un file audio, sfruttando così tutta la gamma dinamica a disposizione, ed evitando, ad esempio in una compilation, di avere pezzi a volumi diversi.

Il seguente comando permette di effettuare l'operazione in un colpo solo (naturalmente a finput e foutput vanno sostituiti rispettivamente il nome del file da massimizzare ed il nome del file dove porre il risultato):

```
# sox -v $(sox finput -e stat -v 2>&1) finput foutput
```

Ovviamente è possibile porre il comando in uno script: apriamo il nostro editor preferito e digitiamo il seguente testo:

```
#!/bin/bash
print_usage(){
    echo "utilizzo: somax fileinput fileoutput"
}
# controlla che ci siano due parametri
[ $# != 2 ] && print_usage && exit 0
# e che il primo sia un file
[ ! -f $1 ] && print_usage && exit 0
# ottiene il valore da utilizzare
VAL=`sox "$1" -e stat -v 2>&1`
```

```
# play file.wav
```

ad esempio è perfettamente equivalente a:

```
# sox file.wav -t ossdsp /dev/dsp
```

ed ha naturalmente l'effetto di riprodurre il file specificato. Mentre il comando:

```
# rec -r 16000 file.mp3
```

registra ad esempio un file mp3 campionato a 16KHz. Anche tramite `rec` e `play` è possibile applicare tutti gli effetti disponibili per registrare e riprodurre in tutti i formati supportati. Si fa notare come la registrazione diretta su hard-disk porti degli innegabili vantaggi, primo fra tutti il non avere alcun limite di lunghezza massima, se non quello imposto dalla capienza del vostro hard-disk (naturalmente poi dovrete fare i conti con un file di tali dimensioni!). Inoltre la possibilità di registrare direttamente in formati compressi (come ogg e mp3) non è normalmente disponibile neanche nei registratori inclusi negli ambienti Gnome e KDE.

Gli effetti disponibili

Oltre alle comode e complete funzionalità di conversione SoX include una nutrita serie di effetti audio. Oltre naturalmente a

```
# se il comando non ha dato errori
if [ $? == 0 ]; then
    # massimizza
    sox -v $VAL "$1" "$2"
# altrimenti stampa l'errore
else
    echo $VAL
fi
```

salviamo con il nome di "somax" e rendiamo eseguibile:

```
# chmod 755 somax
```

Copiamo lo script in un opportuno punto del PATH (questa operazione va eseguita da root):

```
# cp somax /usr/local/bin/
```

Fatto ciò avremo a disposizione nel nostro sistema un nuovo comando per massimizzare il volume dei file audio:

```
# somax finput foutput
```

NOTA: file di input e file di output, come un qualsiasi altro utilizzo di SoX, non possono coincidere.

tutti gli effetti necessari alla conversione, quindi resampling, mix dei canali, ecc. all'interno del repertorio troviamo:

- o **tutti i più comuni effetti** come echi, riverberi, chorus, phaser e flanger;
- o **filtri passa-alto e passa basso**, che permettono di tagliare dal segnale audio frequenze al di sopra (o al di sotto) di una certa soglia;
- o **fade-in e fade-out**, per effettuare sfumature in volume all'inizio o alla fine del file;
- o **variazione volume**, compressione ed espansione del range dinamico;
- o **pitch-shift e speed-shift**, per variare la velocità e/o l'intonazione del segnale;
- o **reverse**, che permette di ascoltare al contrario il suono registrato (e scoprire così eventuali messaggi satanici subliminali!).

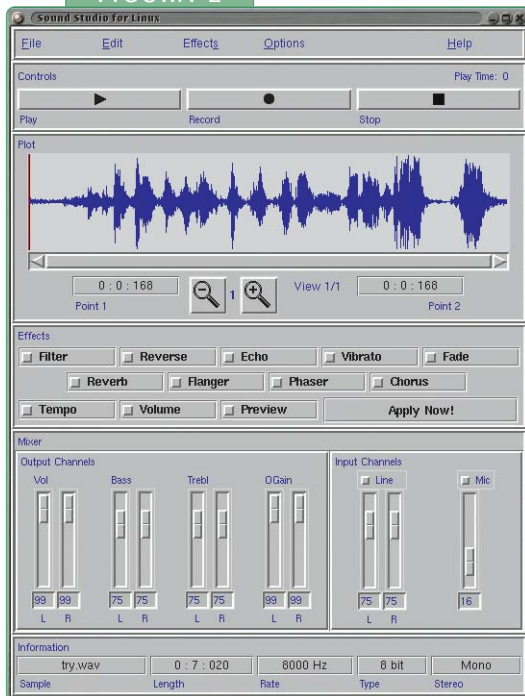
Per un elenco completo si rimanda naturalmente alla pagina di manuale. Il risultato di tali effetti può essere preascoltato prima dell'applicazione effettiva sfruttando la possibilità dello streaming direttamente sul device audio.

La seguente istruzione, ad esempio, ci farà ascoltare il campione `file.wav` a velocità dimezzata:

```
# sox file.wav -t ossdsp /dev/dsp stretch 2
```



FIGURA 2



Una schermata di Sound Studio per Linux

Mentre, se la vostra scheda audio supporta il full-duplex (cioè può registrare e riprodurre contemporaneamente), potrete, ad esempio, applicare effetti in tempo reale alla vostra voce (il seguente esempio applica un eco):

```
# sox -t ossdsp /dev/dsp0  
-t /dev/dsp1 echo 1 0.5 500 0.5
```

Se il preascolto ci avrà soddisfatto sarà sufficiente sostituire il device in output con il nome di un file per rendere definitive le modifiche apportate al campione sonoro.

Naturalmente l'utilizzo di effetti richiede una certa padronanza della materia. Infatti, visto il meccanismo di funzionamento, non è possibile variare i parametri dell'effetto durante la riproduzione, come si farebbe ad esempio con le "manopole" di un dispositivo analogico.

Un'ultima nota: giocando con gli effetti è facile ottenere risultati pericolosi per le vostre casse (e per le vostre orecchie), infatti SoX non pone particolari limiti ai parametri applicabili, quindi durante le prove è meglio tenere sempre una mano pronta sul CTRL-C!

Altre funzionalità

Se hard-disk recording, conversione ed effetti audio non vi bastano elenchiamo brevemente alcune delle altre possibilità del comando, il cui approfondimento lasceremo, per motivi di

spazio, al lettore. Il comando:

```
# sox foo.wav -e stat
```

stampa una schermata di utili informazioni audio sul file specificato. Tra queste risulta molto utile il parametro di "volume adjustment" che, passato a quest'altro comando:

```
# sox -v parametro foo.wav foymax.wav
```

ha l'effetto di massimizzare il volume del file specificato, e quindi di sfruttare tutto il range dinamico a disposizione senza incorrere in indesiderate distorsioni.

Questo utilizzo risulta particolarmente comodo quando si vogliono realizzare compilation di brani provenienti da fonti diverse, per uniformarne il volume, ed evitare così di avere sullo stesso CD brani a volumi differenti.

Tramite l'effetto "silence" è poi possibile rimuovere in modo automatico le parti di silenzio all'inizio o alla fine di un file. Oltre ai comandi già visti, lo stesso sox, e poi rec e play, il pacchetto include un ulteriore comando, soxmix, che può essere usato per miscelare due o più file. Infine è possibile tramite l'effetto synth utilizzare SoX come generatore di varie forme d'onda, anche complesse, e di rumore neutro (bianco, rosa, ecc).

Conclusione

Non sarebbe meglio una bella finestra per comandare tutta questa potenza? Evidentemente non siamo i soli a porci questa domanda, e infatti SoX è utilizzato come "lavoratore nell'ombra" da molte applicazioni, alcune delle quali sono dei veri e propri "frontend" (xsox ad esempio), mentre altre ne sfruttano le potenzialità all'interno di un più ampio ventaglio di funzionalità. In quest'ultimo gruppo segnaliamo l'ottimo Studio

<http://studio.sourceforge.net>

un completo editor non lineare scritto in Tcl/Tk (figura 2) che utilizza SoX per gli effetti e per la conversione.

SoX è inoltre utilizzato come plugin audio dai programmi di modellazione Octave e Matlab.

Concludendo si è visto come il comando offra un ventaglio di possibilità veramente ampio, del quale in questo articolo si è data solo una breve panoramica. Si lascia alla fantasia del lettore l'applicazione all'interno di scenari più complessi.

Infine è possibile affermare che SoX non è, per vari motivi, uno strumento audio tutt'altro, anche se si avvicina molto a questa definizione, ma che, vista la ricchezza di funzioni, soprattutto in relazione alle possibilità di conversione, è uno di quei tool che, chiunque lavori spesso con segnali audio, dovrebbe sempre "tenere nel cassetto", pronto all'uso.

