

INFORMATICA MUSICALE DAGLI ALBORI AL PROSSIMO FUTURO

L'applicazione di metodi e tecnologie dell'informatica in campo musicale ha progressivamente interessato tutti gli aspetti della fruizione, della performance, della produzione, dell'educazione e del supporto alle attività creative. L'attenzione delle ricerche, degli strumenti e dei media più avanzati è oggi posta sull'interazione e sulla personalizzazione delle modalità di accesso e trattamento dell'informazione musicale nelle sue varie tipologie (spartiti, audio, video, testi), soprattutto se tra loro integrate e correlate.

1. INFORMAZIONE MUSICALE E INFORMATICA MUSICALE

L'applicazione dei metodi e delle tecnologie informatiche in campo musicale consente la rappresentazione, la conservazione, la trasmissione, la riproduzione, l'elaborazione, l'analisi e la sintesi dell'informazione musicale. È opportuno ricordare che l'informazione musicale può essere considerata a vari livelli di rappresentazione che caratterizzano le attività musicali più comuni:

strutturale

composizione e analisi musicologica: strutture e relazioni strutturali, musica descritta in termini di oggetti musicali ad alto livello, forme, regole, procedimenti ecc.;

simbolico

esecuzione ed interpretazione, editoria musicale: successioni e sovrapposizioni ordinate e temporizzate di eventi;

esecutivo

orchestrazione, strumentazione, organologia, liuteria: generazione e controllo di vibrazioni udibili, flussi temporizzati di parametri per il controllo di modelli timbrici e interpretativi;

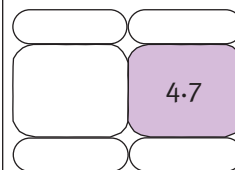
sonoro

registrazione, riproduzione, produzione, postproduzione: flussi temporizzati di segnali audio digitali.

È altresì opportuno ricordare che un'informazione può essere rappresentata tanto da diverse codifiche quanto da diversi procedimenti che possono produrla, secondo il vecchio slogan dell'informatica "algoritmi+dati=programmi" [1]. Per cui è stata di volta in volta una scelta, quella di rappresentare una certa informazione musicale come tabelle di dati piuttosto che come regole per generarle. Per esempio, si pensi alla tabella dei valori di una forma d'onda sinusoidale piuttosto che all'algoritmo per calcolarli, o anche si pensi alla scrittura su partitura di un certo contrappunto piuttosto che alle regole sintattiche per generarlo. Il viaggio nel tempo che si svolge nei prossimi paragrafi mostrerà come l'applicazione di metodi e tecnologie informatici all'informazione musicale si sia estesa progressivamente, arrivando oggi a interessare praticamente tutti gli aspetti della musica.



Goffredo Haus



2. GLI ALBORI (DAGLI ANNI CINQUANTA AGLI ANNI SESSANTA)

Ben prima della disponibilità dei computer si era pensato alla manipolazione dell'informazione musicale a livello di simboli: si pensi per esempio ai famosi "giochi" di composizione automatica di Haydn [2] e Mozart [3] della fine del settecento. Ma è proprio la comparsa del computer che ha permesso di avere lo strumento per l'esecuzione di procedimenti automatici.

Andando oltre i simboli e arrivando al suono, sin dalla presentazione al pubblico nel 1946 dell'ENIAC, il primo computer digitale, si pensò a generare suoni sfruttando i diversi stati dei componenti. Ma le prime applicazioni del computer che possano essere propriamente etichettate come musicali risalgono alla fine degli anni cinquanta. I primi esperimenti furono orientati alla composizione automatica, al controllo di componenti oscillatori per la sintesi del suono, alla trascrizione degli spartiti per l'esecuzione automatica al computer, al computer come esecutore diretto del pensiero compositivo.

2.1. Generazione di testi musicali

Al 1957 si può datare il primo esperimento di applicazione musicale del computer. Il compositore Lejaren A. Hiller, aiutato dal matematico Leonard Isaacson, programmò una composizione elettronica, "Illiac Suite" [4], sul computer ILLIAC all'Università dell'Illinois a Champaign; si tratta di un quartetto

d'archi composto sulla base di un gioco probabilistico (catene di Markov e metodo di Montecarlo tra gli strumenti formali utilizzati) in quattro sezioni derivate da quattro esperimenti diversi. Il musicista Hiller, frustrato dall'imprecisione e dall'inattendibilità dei sistemi analogici, nel computer digitale trovò il mezzo per creare suoni precisamente e completamente definiti.

2.2. Generazione di suoni digitali

Sempre nel 1957, negli USA, presso i laboratori della Bell Telephone, Max Mathews realizzò il primo programma per la sintesi del suono MUSIC I a cui seguono poi tutte le versioni successive fino a MUSIC V [5].

MUSIC V fu poi ripreso da altri programmatori-musicisti portando alla definizione di linguaggi più recenti come – in ordine temporale - il MUSIC 360 (Barry Vercoe), MUSIC1000 (Dean Wallraff), CMUSIC (Gareth Loy), CSOUND (autori vari) e i derivati, fino ai più recenti ad MPEG4 SASL e SAOL.

2.3. Trascrizione al computer e composizione pseudocasuale

Pietro Grossi, il pioniere dell'informatica musicale italiana, fu un precursore tanto dell'informatica musicale quanto della personalizzazione nella fruizione e nella generazione di arte multimediale (testo, musica, visuale).

Dapprima, ebbe l'intuizione di applicare questa possibilità alla musica. Nel 1967, ai laboratori ricerca della Olivetti General Electric, programmò un computer GE-115 per controllare in frequenza e in durata un banco di oscillatori ad onda quadra al fine di produrre suoni musicali. Grossi trascrisse per GE-115 il 5° Capriccio di Paganini e altri brani classici. Inoltre con la medesima tecnica di sintesi del suono, realizzò alcune composizioni elettroniche. Tutte queste opere sperimentali furono registrate e stampate su un disco 45 giri denominato *Computer Concerto* in occasione delle festività natalizie del 1967 (Figura 1).

Grossi proseguì nelle sue attività sperimentali sviluppando principalmente il filone delle trascrizioni e della personalizzazione dell'esecuzione di brani classici e serie di composizioni basate su metodi pseudocasuali.



FIGURA 1
Il lato A del 45 giri
"Computer
Concerto" del
M° Pietro Grossi

2.4. Computer: strumento per il compositore/interprete

Nel 1970, ancora Max Mathews realizza GROOVE (*Generated Real-time Output Operations on Voltage-controlled Equipment*) [6], il primo sistema ibrido sviluppato per la sintesi del suono, utilizzando un computer Honeywell DDP-224, dotato di tastiera musicale a 24 tasti, tastiera alfanumerica, 4 potenziometri e joystick: il tutto per controllare frequenze, ampiezze, durate, evoluzione temporale delle timbriche in tempo reale. Con questo prototipo Mathews intendeva creare un sistema per la programmazione e l'esecuzione in tempo reale. Con il sistema GROOVE non ci sono intermediari, il compositore produce il suono direttamente trattando l'informazione musicale nella sua completezza.

Ancora oggi si prosegue a progredire seguendo questa linea tracciata da Mathews che a buon titolo può essere considerato uno dei padri fondatori della *computer music*.

3. CRESCITA DELL'INFORMATICA MUSICALE (DAGLI ANNI SETTANTA AGLI ANNI NOVANTA)

Le tecnologie informatiche degli anni cinquanta e sessanta imponevano forti limitazioni alle applicazioni musicali possibili, soprattutto per la componente multimediale - audio e video - dell'informazione musicale: non si poteva elaborare il suono in tempo reale, non si poteva interagire con i contenuti musicali di un brano non essendo possibile riconoscere in tempo reale le principali proprietà musicali, non erano disponibili memorie digitali ad accesso diretto per registrare suoni e relativi video, non erano disponibili canali per la trasmissione dei dati audio e video di banda sufficiente ecc..

I metodi, i modelli, le tecniche e le tecnologie specifiche per l'informatica musicale hanno tratto dall'evoluzione generale stimolo per evolvere a loro volta e giungere per molti aspetti a piena maturazione.

È bene ricordare che il testo musicale - partiture, spartiti ecc. - è codificabile con poche centinaia di simboli per pagina e poteva quindi essere elaborato fin da allora da processori *general purpose*, essendo la complessità computazionale richiesta paragonabile a

quella del testo in linguaggio naturale. Il suono invece richiede decine di migliaia di numeri al secondo e necessita perciò di un costo computazionale superiore di alcuni ordini di grandezza. Per questo motivo, l'evoluzione dei processori *general purpose*, dei *digital signal processor*, delle memorie, dei dispositivi per l'interazione uomo-macchina, e in generale delle architetture, seguita dall'introduzione del web e dei dispositivi mobili, ha aperto la strada al trattamento del suono e più in generale dell'informazione multimediale.

Per comprendere come questa evoluzione abbia influito sull'informatica musicale, alcuni aspetti specifici meritano di essere brevemente messi in luce: la sintesi digitale dei suoni musicali, le architetture dedicate al trattamento dei suoni musicali, l'editoria elettronica del testo musicale; questi tre aspetti sono stati di particolare rilievo, anche perchè hanno caratterizzato i principali cambiamenti nella prassi di alcune professioni musicali tipiche.

3.1. Modelli e tecniche per la generazione digitale di segnali audio musicali

Quello che in principio non era possibile, poi è diventato gradualmente sempre più alla portata dei laboratori di ricerca, poi degli studi di produzione, poi degli esperti e infine degli appassionati: la programmazione timbrica, meglio nota come Computer Music o Musica Informatica [7, 8].

Si possono distinguere fondamentalmente tre approcci alla generazione sintetica delle timbriche musicali:

- modelli per campionamento;
- modelli matematici;
- modelli fisici.

Nel primo caso, suoni naturali di strumenti musicali od altre sorgenti sonore sono registrati digitalmente e poi riprodotti. Un approccio qualitativamente apprezzabile richiede però molta memoria, poiché molte varianti timbriche sono proprie del repertorio sonoro di uno strumento al variare dell'altezza, dell'intensità, della durata e della tecnica strumentale adoperata per produrre il suono. In ogni caso il costo computazionale è irrilevante. I limiti di questo metodo sono evidenti - principalmente staticità dei timbri registrati e scarse possibilità di controllo delle varianti - ed è utilizzato principalmente per abbattere i costi dei

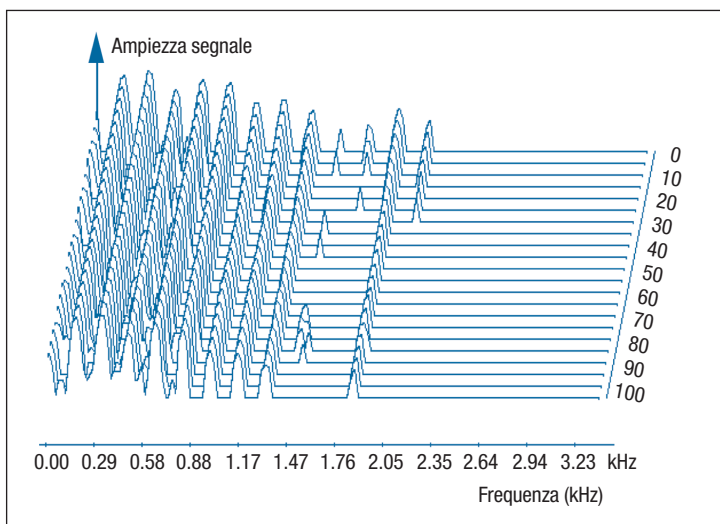


FIGURA 2

Modello tridimensionale di rappresentazione grafica del suono per la sintesi additiva: le tre variabili sono la frequenza, il tempo e l'ampiezza; ogni linea del grafico rappresenta la composizione del suono ad un certo istante di tempo come spettro di diverse frequenze, ciascuna dotata di una propria ampiezza relativa a quell'istante

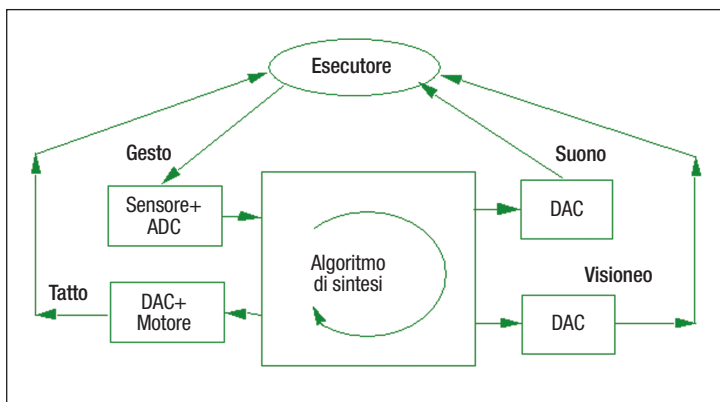


FIGURA 3

Schema del modello di interazione tra esecutore umano e strumento musicale virtuale basato su modelli fisici dei corpi vibranti: oltre alla retroazione costituita dai suoni prodotti suonando lo strumento virtuale e alla visualizzazione grafica di forma d'onda e composizione spettrale, è presente la componente tattile, retroazione indispensabile affinché l'esecutore possa suonare adattando i propri gesti, proprio come accade con gli strumenti tradizionali

suoni naturali prodotti “ad hoc” da musicisti. Quasi tutti i PC odierni dispongono di schede audio dotate di campionario audio.

I modelli matematici si basano sulla simulazione di una forma d'onda data mediante una formula che la approssimi, generalmente polinomiale. Questi modelli si distinguono per la quantità di termini presenti nella formula e per la semplicità o meno nella genera-

zione di timbriche predeterminate. Da citare almeno due modelli di questa famiglia:

□ la sintesi additiva: è basata sull'idea che una forma d'onda complessa può essere costruita sommando con opportuni “pesi” – eventualmente variabili nel tempo – forma d'onda semplici, tipicamente sinusoidi (Figura 2); permette di costruire in modo accurato una forma d'onda data che, una volta analizzata per identificare le sue componenti in frequenza, può poi essere ricostruita sommando tali componenti; peraltro, richiede molti termini per essere qualitativamente apprezzabile, e da questo deriva la sua scarsa semplicità d'uso – troppi parametri per controllare il suono – e quindi il suo successo commerciale pressoché nullo;

□ la sintesi per modulazione di frequenza o di fase [9]: è basata sull'idea che una forma d'onda semplice può essere molto arricchita spettralmente modulando con termini temporari la sua frequenza o la sua fase (i due metodi sono riconducibili uno all'altro); il fatto che con poche onde sinusoidali si possano costruire timbriche musicali dinamiche e ricche ha determinato il grande successo commerciale dei sintetizzatori basati su questa tecnica; si pensi al sintetizzatore Yamaha DX7; oggi, quasi tutti i PC dispongono di librerie standard di timbriche FM.

I modelli matematici sono stati tra gli anni settanta e novanta il punto di riferimento per i compositori di computer music, ma hanno mostrato significativi limiti rispetto alla variazione delle timbriche in funzione dei gesti interpretativi, proprio per il fatto di partire dalla simulazione di una forma d'onda che è già il risultato del processo di interazione tra chi suona e lo strumento che viene suonato.

Per questo motivo, l'approccio dei modelli fisici ha subito indicato di essere quello su cui puntare per la piena soddisfazione nella generazione e nel controllo delle timbriche musicali [10]. L'idea di fondo è di simulare il comportamento dei corpi vibranti (corde, ponticelli, casse armoniche ecc.) e delle loro interazioni, sia tra di essi stessi che con il musicista. In questo modo, il modello fisico può essere “suonato bene” o “suonato male” (Figura 3), come accade con gli strumenti tradizionali; tutte le soluzioni del liutaio diventano “programmabili”, potendo andare anche oltre ciò che è fisicamente possibile o ragionevole. Si

può - per esempio - programmare un tamburo di 200 km. di diametro o una chitarra con corde di un nanometro di spessore, perfino costruiti con un materiale qualsivoglia.

I motivi per cui questo approccio non ha ancora trovato piena applicazione negli strumenti musicali offerti dal mercato sono individuabili nella grossa richiesta di capacità computazionale richiesta, e ancor più per il fatto che non è ancora matura la letteratura scientifica sui modelli fisici degli strumenti musicali. Per ora, la maggior parte delle applicazioni concrete di questo approccio si è limitata a usare modelli fisici "parziali", che simulano solo una o comunque poche componenti della catena di generazione del suono; come - per esempio - gli amplificatori digitali a modelli fisici che simulano i diversi modelli di amplificatori valvolari, o le unità di effetti che applicate ai suoni degli strumenti tradizionali ne alterano le sonorità simulando banchi di effetti analogici, o i vari tipi di chitarre (acustiche, elettriche, basse) che producono le sonorità di decine di altre chitarre di cui lo strumento digitale implementa i modelli fisici per trasformare il suono della chitarra "fisica" nel suono di una delle chitarre "virtuali" di libreria.

3.2. Evoluzione delle architetture hardware

Gli avanzamenti nella modellistica e nelle tecniche per il trattamento dell'informazione musicale, specialmente considerando i livelli esecutivo e sonoro, sono andati di pari passo con la disponibilità di nuove tecnologie hardware. Al tempo dei mainframe si era arrivati a considerare come prima ipotesi operativa quella di usare il computer a livello simbolico e strutturale per controllare mediante parametri una serie di dispositivi analogici esterni per la sintesi e il trattamento del suono (oscillatori, sintetizzatori, mixer, banchi di filtri ecc.): sono i cosiddetti *sistemi ibridi*, dove le quantità di informazione più modeste (strutturale e simbolico) sono gestite digitalmente e quelle invece più considerevoli (modelli timbrici e segnali audio) sono controllate da computer ma gestite da periferiche analogiche.

Il motivo di questa soluzione è che non sussisteva la disponibilità di una potenza computazionale che permettesse di trattare il segnale audio senza interruzioni o ritardi. In questa fase i compositori di computer music potevano

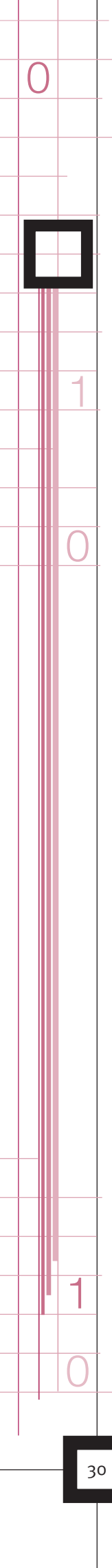
solo usare linguaggi operanti in tempo differito (tipo i linguaggi MUSIC-n); una volta date tutte le direttive al computer su come produrre i suoni corrispondenti a una composizione, veniva calcolato il segnale audio finale, in un tempo non predeterminabile e senza interazione con il compositore.

Negli anni settanta furono sviluppati i primi sintetizzatori digitali prototipali in alcuni centri di ricerca statunitensi (Stanford, Cambridge, Berkeley, Urbana), ed è nel 1978 che la Digital Music Systems, una piccola ditta spin-off del MIT, produce il primo sintetizzatore digitale - denominato DMX-1000 - programmabile da un qualunque minicomputer o mainframe [11]. Diventa realtà il controllo interamente digitale dell'informazione musicale ai vari livelli: un *computer general purpose* si occupa dell'informazione strutturale, simbolica e dei parametri che - inviati al DMX-1000 - permettono la generazione in tempo reale dei suoni. Si tratta di un'*architettura digitale mista*: tecnologia *general purpose* per controllare la tecnologia più spinta necessaria per il trattamento digitale dei segnali audio.

Solo tre anni dopo, il fisico Giuseppe Di Giugno progetta e realizza all'IRCAM di Parigi un sintetizzatore digitale - denominato 4X - capace di gestire una grande quantità di strumenti virtuali contemporaneamente e di sopportare un grande carico computazionale [12]. I modelli timbrici implementabili diventano così molto complessi, paragonabili a quelli richiesti per il controllo di un'orchestra digitale intera. Un grosso passo avanti, seguendo il tipo di architettura digitale mista introdotta con il DMX-1000.

Un cambiamento di rotta fu introdotto nel 1983 con la pubblicazione dello standard MIDI [13] e la comparsa degli strumenti musicali digitali che ne sfruttavano i modi di comunicazione: l'informatica musicale diventa repentinamente "individuale", non servono più i grandi centri di ricerca per avere a disposizione l'hardware necessario a produrre il suono digitale. Ai centri di ricerca resta da sviluppare nuove soluzioni più avanzate e da far progredire modelli e tecniche per poi tradurli in applicazioni software. L'accoppiata PC+MIDI si afferma decisamente sul mercato. Questa diffusione improvvisa porta però con sé tutti i limiti dello standard MIDI: tra-

0



smisione seriale, velocità di trasmissione bassa (poco più di 30.000 bit/s), pochi canali (16 al massimo), scarsa programmabilità dei modelli timbrici.

Infatti è l'hardware dei sintetizzatori MIDI a determinare intrinsecamente i modelli eseguibili: Yamaha con la modulazione di frequenza, Casio con la distorsione non lineare, AKAI con i campionatori, Roland con soluzioni miste. Peraltro i ricercatori capaci di costruire modelli timbrici erano e restavano pochi; la grande utenza si è sempre accontentata dei suoni di libreria forniti dalle case costruttrici, oltre a copiare le librerie di set timbrici degli amici.

Quindi diffusione sì, ma progresso scientifico ben poco grazie al MIDI. Si sviluppa moltissimo la possibilità di registrare i gesti dell'interprete su dispositivi che hanno l'aspetto di strumenti musicali ma che servono a produrre dati MIDI; dati che vengono registrati con applicazioni software chiamate *sequencer*, programmi che sono registratori multitraccia di dati MIDI: si esegue e si registra, poi si può intervenire per correggere, modificare, aggiungere, togliere ovvero fare editing dei dati MIDI organizzandoli in molte tracce, anche centinaia. Si può sempre scegliere in un secondo momento quali di queste tracce usare e come combinarle tra loro; un po' come quando dalla registrazione multitraccia audio analogica si fa il mixdown al master stereo. Tant'è che nel giro di qualche anno i sequencer MIDI sono stati arricchiti delle tracce audio digitali e la registrazione multitraccia combina tracce MIDI e tracce audio, in certi casi anche tracce video.

Ma quello che inizialmente era arricchire il PC della possibilità di produrre suoni digitalmente con periferiche MIDI esterne, diviene poi gradualmente una soluzione unica e integrata, per cui nel PC – dopo un primo esperimento con gli Atari – viene fornita di serie l'interfaccia MIDI, librerie di campioni sonori, librerie di timbri FM, software musicali di *editing* e *sequencing* musicale, e perfino processori *special purpose* per l'audio, poi confluiti nelle moderne motherboard con una o più CPU e diversi coprocessori integrati. Oggi, CPU standard, coprocessori per l'audio, per il video, per la comunicazione interna ed esterna, MIDI e quant'altro, fanno parte di un'unica soluzione integrata. Tanto che i PC di oggi sono in grado di eseguire applicazioni software musicali che

in tempo reale eseguono tutti i tipi di tecniche per la sintesi digitale del suono; perfino i tanto dispendiosi modelli fisici, purché con opportune limitazioni, ma senza bisogno di hardware aggiuntivo. È forse il controller musicale (tipo tastiera, chitarra, strumento a fiato, percussioni o anche microfono) la parte specifica più comune che serve ancora aggiungere.

Infine, va menzionato il fenomeno delle workstation musicali: iniziato alla fine degli anni settanta con il NEDCO Synclavier e poi il Fairlight CMI, consiste nella produzione di workstation in cui sono integrate:

- tecnologie *general purpose* di tipo PC;
- tecnologie *special purpose* tipo *Audio Digital Signal Processor* e MIDI;
- applicazioni software per l'edizione di partitura;
- applicazioni software per il *sequencing* e l'*editing* MIDI e in un secondo momento anche audio e video;
- librerie di suoni digitali di qualità professionale.

Questa tipologia di strumento è molto costosa, destinata generalmente a professionisti che vogliono trovare i problemi tecnici già risolti in una postazione fornita "chiavi in mano" per il musicista. È sempre stata da allora una nicchia ad alto valore aggiunto, come spesso accade con gli strumenti professionali. La loro diffusione è stata comunque limitata non solo dai costi alti, ma anche dal fatto che le stesse funzionalità di queste workstation integrate potevano essere implementate integrando le diverse componenti in modo anche più flessibile ed economico a patto di affrontare le problematiche tecniche del caso.

3.3. Il testo musicale: editoria e riconoscimento automatico

La scrittura di testi in linguaggio naturale (*text editor* e *word processor*) è stata senz'altro una delle prime e più diffuse applicazioni del computer. Il riconoscimento automatico dei testi stampati (OCR: *Optical Character Recognition*) è meno popolare, ma – nonostante l'intrinseca complessità – ha raggiunto da molti anni piena maturità, in altre parole è quasi infallibile.

Applicazioni di grande interesse per la minoranza di utenti del computer che si occupa di musica è l'equivalente capacità di scrivere spartiti (*score editor*) e di riconoscere auto-

maticamente i simboli musicali dei testi stampati (OMR: *Optical Music Recognition*). A metà degli anni ottanta sono comparsi i primi software per scrivere spartiti al computer; ben presto sia gli editori che i singoli musicisti hanno apprezzato questa tipologia di applicazioni al punto da considerarla come lo strumento prevalente per scrivere e stampare musica. Già agli inizi degli anni novanta la tecnologia (font, editor, formatter, publisher) era matura e il mercato soddisfatto. Anche i musicisti più ostili all'uso del computer se ne sono fatti una ragione, per l'affidabilità, la convenienza economica, l'autonomia acquisibile con il software per l'edizione musicale. Per quanto riguarda la lettura ottica della musica, nonostante siano comparsi negli stessi anni del software di edizione i primi applicativi, l'affidabilità nel riconoscimento automatico non è ancora così soddisfacente. Soprattutto, la percentuale di errori è fortemente dipendente dalle caratteristiche del singolo testo musicale considerato – per esempio, la densità di simboli musicali, la quantità di casi di sovrapposizione di simboli diversi nella medesima posizione - e in generale la complessità grafica del testo musicale, in combinazione con gli investimenti limitati tipici delle problematiche di nicchia, fan sì che si possa oscillare tra un 10% e un 50% di simboli riconosciuti in modo scorretto. Nei casi peggiori, è ancora più conveniente inserire il testo musicale nel computer a mano, ovvero con un music editor.

La maturazione delle tecnologie di OMR è attesa con grande interesse, per poter rendere disponibile una grande quantità di archivi musicali online, non più solo in forma grafica, ma anche in forma simbolica, editabile, disponibile per rielaborazioni, trasposizioni, estrazione delle parti strumentali, etc. Soprattutto per l'enorme patrimonio culturale di testi musicali che non sono più soggetti alla normativa sul diritto d'autore.

4. INFORMATICA MUSICALE OGGI

Nell'arco della sua ancora giovane vita, l'informatica musicale ha visto la realizzazione di numerosi progetti di ricerca i cui risultati di maggior interesse sono poi stati trasferiti al-

l'industria e all'editoria. Sono abbastanza delineate le linee evolutive che sono in corso in questi ultimi anni ed è altresì identificabile l'insieme dei problemi rimasti ancora aperti. Peraltro, non tutte le innovazioni fornite dalla ricerca sono state sfruttate a livello produttivo, generalmente per motivi economici.

Per concludere questa escursione nelle tematiche dell'informatica musicale ecco quindi lo stato dell'arte odierno e quello che è oggettivamente prevedibile nel medio periodo. Si ricordi che da almeno vent'anni un PC arricchito con opportuni dispositivi specializzati costituisce la base di quasi tutte le tipologie di stazioni di lavoro individuali, editoriali e in genere produttive. In questo contesto devono essere considerati alcuni aspetti particolarmente significativi per l'informatica musicale: l'interazione con l'informazione musicale, il riconoscimento automatico dei contenuti, internet, i dispositivi mobili. L'interazione con i contenuti multimediali è al centro dell'evoluzione informatica di questi anni, ma: cosa vuol dire interagire con l'informazione musicale?

Si consideri che l'informazione musicale può essere espressa ad uno dei livelli di rappresentazione sopra indicati in accordo con la forma più appropriata per l'interazione. Per esempio, se si vuole fornire all'utente un'applicazione capace di estrarre automaticamente le parti strumentali da una partitura orchestrale risulterà opportuno trattare l'informazione musicale mediante la forma grafica di partitura; oppure, se si vuole permettere all'utente di una libreria digitale multimediale di ascoltare un brano di musica di cui l'utente non conosce né titolo, né autore, né altri metadati, ma l'utente è in grado di canticchiarlo poiché lo ha già sentito in un'altra circostanza, allora bisognerà predisporre un'interfaccia per interagire tramite il canto dell'utente, che naturalmente funzionerà in modo efficace solo se l'utente è capace di intonare correttamente il frammento musicale che cerca. Detto questo, è necessario un piccolo approfondimento per capire cosa implica quanto ipotizzato.

Sia per il trattamento ad un certo livello, che per la trasformazione da un livello ad un altro dell'informazione musicale è necessaria la disponibilità di un insieme integrato di dispositivi software e hardware specifici capaci

di trattare (memorizzare, elaborare, trasmettere, rappresentare) i flussi informativi corrispondenti.

Per esempio, per il trattamento dell'informazione musicale a livello di suono ad alta fedeltà sono generalmente richiesti dispositivi hardware speciali per l'acquisizione e la produzione di suoni, insieme con memorie centrali di grande dimensione e memorie di massa molto capaci e con velocità di trasferimento dati molto alta. Ricordiamo infatti che un secondo di suono stereofonico di qualità Compact-Disc audio contiene una quantità di informazione pari a circa 90 pagine di testo in linguaggio naturale.

4.1. Il riconoscimento dei contenuti musicali

L'analisi dell'informazione musicale può portare al riconoscimento automatico di contenuti non espressi in modo esplicito: per esempio, a livello di simboli, si può riconoscere la tonalità in cui è scritta una partitura, quali sono i frammenti tematici più importanti, qual è la forma musicale, quali sono le regole rispettate e quali quelle infrante della sintassi musicale; oppure, a livello sonoro, si può riconoscere quali strumenti stanno suonando, dove inizia e dove finisce la parte di suono che corrisponde ad un simbolo di nota, una serie di parametri fisici che permettono di avere informazioni, melodiche, ritmiche, dinamiche e anche relative all'interpretazione operata da chi ha prodotto i suoni allo strumento.

Questo tipo di analisi è caratterizzata da una mole di calcoli molto elevata, da modelli analitici complessi e non è ancora giunta a piena maturazione, ovvero ci sono ancora informazioni che non si riesce a riconoscere in modo automatico.

Già nel 1975, nella sua tesi di dottorato, Moorer [14] impostò il problema dalla base, ovvero come porsi ad osservare il segnale audio per estrarne informazioni musicali. Da allora molti progressi sono stati compiuti, e tanto i modelli di interazione uomo-macchina più avanzati quanto i metodi di classificazione automatica dei contenuti musicali nelle librerie digitali accessibili tramite il web [15] si basano proprio sulla capacità dei computer odierni di riconoscere contenuti multimediali e, in particolare, musicali.

4.2. Interazione multimodale con i contenuti musicali

Già nel 1985, Roger Dannenberg aveva realizzato un prototipo di sistema capace di accompagnare automaticamente un musicista durante una performance [16]; il suo sistema conosceva tutte le parti strumentali del brano da suonare ed era in grado di riconoscere le note suonate dal musicista ed eseguire in automatico tutte le altre parti strumentali, in modo sincrono all'umano. Prototipi in questa direzione sono stati via via sviluppati da vari ricercatori ma ben poco di queste esperienze è arrivato a essere disponibile per il musicista o l'appassionato comune.

Si è così approdati alla situazione odierna, dove la possibilità più innovativa offerta dalle tecnologie informatiche per la musica combina le possibilità di riconoscimento automatico dei contenuti musicali con i formati per la rappresentazione integrata dei contenuti informativi delle diverse tipologie di materiali musicali: supporti audio, supporti video con tracce audio, pagine di musica scritta, altri materiali "fisici" (foto, video, scenografie ecc.), supporti magnetici e ottici contenenti codifiche simboliche immateriali (XML, MIDI, UNICODE, ASCII ecc.).

Infatti, l'informazione musicale può essere espressa – e di solito è prevalentemente questo che accade nel caso degli standard più diffusi come MP3, AAC, vari tipi di MPEG - ad uno specifico dei livelli di rappresentazione considerati sopra, in accordo con la forma di rappresentazione più appropriata per l'interazione. Per esempio, se si vuole fornire all'utente un'applicazione capace di estrarre automaticamente le parti strumentali da una partitura orchestrale risulterà opportuno trattare l'informazione musicale mediante la forma grafica di partitura.

Ma oggi è diventato possibile integrare i diversi livelli in un unico ambito di rappresentazione, integrando audio, timbriche, partiture, forme e strutture. Un gruppo di lavoro della IEEE Standards Association per la musica ha definito per questo uno specifico dialetto XML per la musica, denominato ufficialmente IEEE1599-2008, anche noto comunemente come "MX"; IEEE1599-2008 è uno standard internazionale approvato il 12 giugno 2008 [17], ottenuto grazie alla collaborazione tra LIM dell'Università

degli Studi di Milano (gruppo del Prof. Haus) e SUPSI di Lugano (gruppo del Prof. Baggi), con il supporto parziale del progetto CTI 8357.1 "e-Music XML Standard" e nell'ambito del progetto mondiale IMS "e-Music XML".

Il formato IEEE1599-2008 consente di navigare e interagire tra/con i contenuti musicali multimediali, come se questi materiali fossero parte di un unico ambiente di conoscenza integrata. In altre parole, tale tecnologia innovativa consente di:

□ trovare tutti i file di partiture musicali e tutti i file audio digitali che contengano un dato frammento musicale, interrogando un sito Internet, un database multimediale ecc.; il frammento può essere uno spezzone audio, vocale (cantato) o strumentale (suonato), oppure una sequenza di note scritta su pentagramma;

□ sincronizzare automaticamente varietà ampie a piacere di file audio, video e di partitura per una fruizione integrata delle varie forme di comunicazione musicale (CD, DVD, file su PC ecc.); ognuno dei materiali considerati per la sincronizzazione permette di "guidare" la navigazione nei vari materiali integrando di fatto i diversi player in un unico player multitraccia dove ogni traccia può essere usata per "navigare";

□ interagire con l'informazione musicale, aggiungendo ai controlli standard sul volume, sulla posizione stereo o sulle frequenze un ampio repertorio di possibilità che contemplano: scorrere un brano nota per nota, battuta per battuta o frase per frase, modificare la partitura e di conseguenza i suoni corrispondenti ecc.:

Gli strumenti scientifico-tecnologici per produrre i materiali nel formato previsto dallo standard sono prevalentemente italiani: sono stati sviluppati nell'arco di trent'anni presso il LIM - Laboratorio di Informatica Musicale del Dipartimento di Informatica e Comunicazione dell'Università degli Studi di Milano ed sono stati applicati a livello prototipale con successo:

□ nel sistema MAIS (*Music Archive Information System*) sviluppato dal LIM tra il 1997 e il 1999 per l'Archivio Musicale del Teatro alla Scala [18];

□ nei dimostrativi tecnologici sviluppati per il Ministero per i Beni e le Attività Culturali in occasione delle mostre su tematiche musicali te-

nute a Roma (dicembre 2005-gennaio 2006), Torino (febbraio-giugno 2006), Milano (dicembre 2006-gennaio 2007), Salisburgo (maggio-giugno 2007), Lucca (novembre 2007);

□ nei dimostrativi tecnologici sviluppati nell'ambito dei succitati progetti CTI e IMS congiuntamente alla SUPSI di Lugano.

IEEE1599-2008 è una codifica XML multistrato che integra le codifiche standard delle partiture (file Finale, MusicXML, NIFF, e altri formati), dei metadati catalogafici, dell'audio compresso (MP3, AAC, WMA ecc.) e non compresso (Compact-Disc Digital Audio, WAV, DVD Audio, etc.), delle performance (Csound, MIDI, SASL/SAOL ecc.), e dei dati di sincronizzazione. Nella figura 4 si può vedere il principio che consente di trattare in modo integrato e sincrono i diversi materiali musicali considerati (nella figura di esempio si vedono audio, MIDI e partitura): una struttura di dati chiamata *spine* integra e sincronizza l'informazione spaziale delle pagine di musica scritta con l'informazione temporizzata dei processi MIDI e audio. Nella figura 5 si può vedere l'architettura



FIGURA 4

Esempio di schema di corrispondenze tra la posizione spaziale dei simboli di uno spartito con gli istanti di tempo dei suoni correlati e dei gesti necessari per produrre tali suoni: è la struttura portante, chiamata "spine" del nuovo standard musicale multilayer IEEE1599-2008

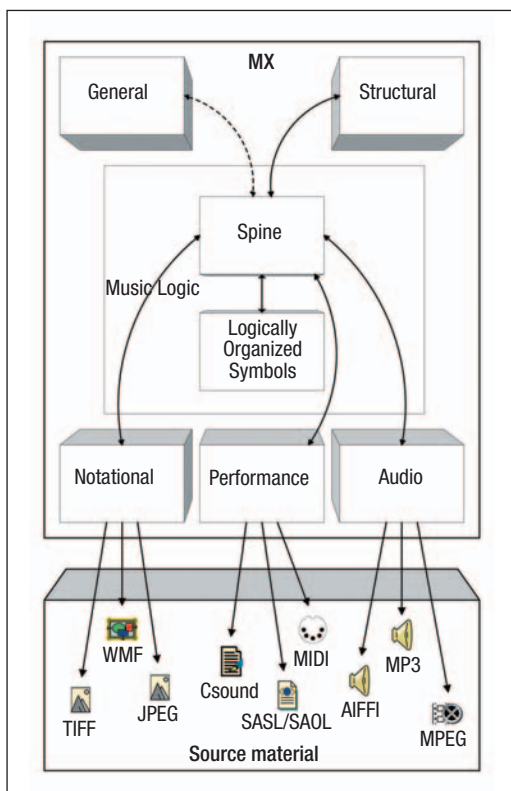


FIGURA 5

Schema delle relazioni tra i layer del nuovo standard musicale IEEE1599-2008: i file di notazione musicale, performance e audio con eventuale video sono tra loro integrati e correlati pur restando nelle forme di codifica degli standard propri delle varie tipologie di dati multimediali

tura che mette in relazione i diversi “strati” di informazione musicale, ciascuno in grado di recepire gli standard industriali di riferimento, nello standard IEEE1599-2008.

Molti elementi di innovazione vengono introdotti da IEEE1599-2008 nei modi di interagire tra umani e informazione musicale; tra i più evidenti e significativi, si rendono potenzialmente disponibili i seguenti:

- a.** fruizione sincrona di tutti i materiali musicali inerenti ad un brano (una o più partiture o parti strumentali, una o più interpretazioni audio ed eventualmente video, il multitraccia audio con uno o più *mixdown* disponibili, iconografia, elementi testuali);
- b.** interazione con i materiali musicali di un brano (definizione del *subset* di materiali per la fruizione, eliminazione di una o più “tracce audio” dalla fruizione);
- c.** navigazione multimodale dei contenuti: uso

di un qualunque materiale musicale per guidare la fruizione di tutti gli altri materiali sia nel testo (parole di una canzone e/o libretto d’opera), che nello spazio (materiali di partitura), che nel tempo (materiali audio e video);

d. “allineamento” di tracce audio non sincrone, sia rispetto ad altre tracce audio sia con i materiali di partitura;

e. dilatazione/contrazione di tracce audio senza modifiche timbriche e di intonazione (cambi di metronomo, sovrapposizione di tracce non sincrone);

f. indicizzazione dei materiali musicali (partiture, audio, video, testo letterario, iconografia) per contenuti (frammenti audio, frammenti di partitura, frammenti di testo).

Le ipotesi di nuovi media come quelle qui sopra formulate portano come effetto collaterale una potenziale forte valorizzazione dei patrimoni editoriali; i materiali degli archivi possono essere riutilizzati per le nuove iniziative editoriali con un’articolata varietà di modi e di potenziali ritorni economici.

Vale la pena sottolineare che un auspicabile effetto di valorizzazione dei patrimoni editoriali può essere ottenuto – prima ancora che pensando a nuovi media – presentando, elaborando e riorganizzando i materiali musicali degli archivi esistenti. Alcuni tipi di media, anche se già presenti sul mercato, possono essere concepiti in modo innovativo rispetto ai target del mercato di consumo, specie quando si considerino quelli professionale e formativo.

A titolo di esempio, si considerino le schermate nelle figure 6 e 7. Rappresentano casi esemplari del rappresentare e rendere fruibili i contenuti di brani musicali classici, con modalità multimodali e interattive.

Le applicazioni software – realizzate al LIM – da cui sono state prese le schermate consentono di interagire con i contenuti musicali dei rispettivi brani musicali considerati, in diversi “modi”: partitura, strumentazione, libretto, audio e video [19]. Ogni elemento informativo presente sullo schermo è attivo. La selezione con il mouse di una parola del libretto produce la riproduzione in audio o in video/audio a partire dall’istante in cui quella parola viene pronunciata, nonché il posizionamento sulla partitura di un cursore rosato che evidenzia la nota corrispondente al momento in cui la parola viene pro-

nunciata, nella parte che si è scelto di seguire per la fruizione.

Analogamente selezionare una nota sulla partitura con il mouse, posiziona automaticamente audio, video, libretto e quant'altro alla nota selezionata così come spostare il cursore del player multimediale sposta contemporaneamente e in modo sincrono i cursori sulla partitura, sul libretto e quant'altro.

In particolare nella figura 7, è mostrato un momento dell'interazione con le frasi melodiche del brano (un duetto tratto dalla Tosca di Puccini); i contenuti musicali sono percorsi mediante le schematizzazioni di frasi che possono essere selezionate nell'apposito riquadro in alto a destra nello schermo. La selezione di una frase posiziona il player multimediale, il cursore sulla partitura e il cursore sul libretto e avvia l'esecuzione della frase stessa evidenziando sulla partitura in colore giallo il segmento trasversale di partitura interessato.

La frontiera più avanzata dell'interazione consente di modificare i contenuti musicali dei brani, a partire dagli elementi informativi individuati e rappresentati. Ma questo discorso è per un futuro meno prossimo. Questa modalità di interagire con la struttura di un brano in tempo reale – l'audio e i frammenti di partitura vengono costruiti e riprodotti durante l'interazione senza ritardi – costituisce una soluzione per quando sarà disponibile una normativa sul diritto d'autore che disciplini questa forma di fruizione, oppure anche già subito, ma solo per le composizioni musicali che sono già uscite dal periodo disciplinato dalle attuali norme sulla proprietà intellettuale e industriale, o anche come strumento di produttività individuale. In questo caso la tecnologia è ben più avanti del diritto.

4.4. La formazione

Accanto a nuove possibilità, nuove problematiche, nuove economie e nuove modalità di accesso alla cultura musicale, si delinea sempre più la necessità di figure professionali interdisciplinari che abbiano competenze specifiche sulle tecnologie di interesse per la musica accompagnate da competenze propriamente musicali. Se fino agli anni novanta erano sufficienti pochi professionisti e ricercatori capaci di coniugare in virtù dei loro interessi e delle loro esperienze la doppia

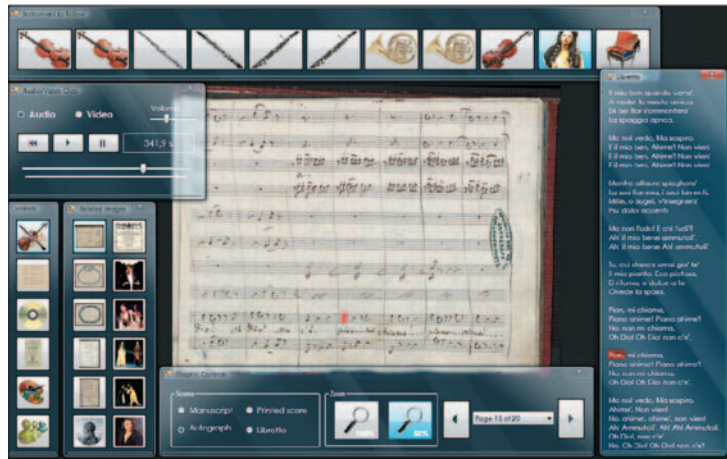


FIGURA 6

Esempio di schermata di musica interattiva; l'autografo di un'opera di Paisiello, il libretto, due performance (una in audio e l'altra in audio e video) sono fruibili in modo sincrono, tutti gli elementi informativi possono essere usati per muoversi "nel brano", ogni nota della partitura, ogni parola del libretto, ogni istante di tempo della barra del player



FIGURA 7

Esempio di schermata di musica interattiva: la riduzione canto e piano di un duetto della Tosca di Puccini, il libretto, tre diverse performance (due in audio e la terza in audio e in video) sono fruibili in modo sincrono e sono percorribili oltre che "per note", "per parole" e per "istanti", anche per "frammenti melodici" grazie alla finestra in alto a destra che evidenzia il materiale tematico alla base del brano, consentendo di visualizzare con evidenziazione gialla come il singolo frammento è "mappato" nella partitura del brano ogni volta che il frammento viene presentato

figura dell'informatico e del musicista, è ora necessario avere persone formate ad hoc e in quantità superiore. A questa esigenza cercano di rispondere alcune iniziative formative, sia a livello universitario che medio superiore. Da citare il corso di laurea triennale in "Scienze e tecnologie della comunicazione

musicale” attivo all’Università degli Studi di Milano dal 2002, che accetta 120 nuovi allievi ogni anno e riesce portarli al mondo del lavoro proprio in relazione alle specificità professionali acquisite in questo corso di studi.

5. PROBLEMI APERTI E PROSPETTIVE

A conclusione di questo percorso, è doveroso citare:

- cosa ancora non è possibile;
- verso dove si sta andando.

La ricerca di base deve ancora risolvere molti problemi, prima di passare le ipotesi di soluzione alla ricerca sperimentale e poi all’industria. I più rilevanti:

- maturazione delle tecniche basate su modelli fisici degli strumenti musicali e riconoscimento automatico delle relative timbriche;
- definizione di modelli di compressione multisensoriali, che sappiano sfruttare il modo in cui il nostro sistema fisiologico (cervello + sistema nervoso + organi di senso) funziona, al fine di ridurre la necessità di memoria per rappresentare l’informazione multimediale e musicale; in altre parole, integrazione degli attuali modelli percettivi monosensoriali (audio e video soprattutto) in modelli percettivi multisensoriali;
- definizione di modelli dell’interpretazione musicale capaci di eseguire automaticamente una partitura come un umano; alcune esperienze significative già sono state effettuate, ma non sono ancora pienamente soddisfacenti [20]; andando nel senso contrario, definizione di modelli interpretativi inversi, ovvero capaci di riconoscere gli interpreti e di trascrivere in partitura - fedele all’originale - le registrazioni audio;
- metodi affidabili per effettuare la separazione automatica delle tracce audio ovvero il “multitraccia” con l’audio separato per ogni voce o strumento a partire da originali mono o stereo. Alcune di queste problematiche sono più vicine a trovare soluzioni soddisfacenti, altre sono più lontane, e riguardano perciò il “domodani” dell’informatica musicale. Si può invece considerare che il domani dell’informatica musicale debba senz’altro prevedere che le prospettive di ulteriori avanzamenti, sia dal punto di vista degli arricchimenti del-

le modalità di interazione uomo-macchina che da quello della produttività e dell’economicità nella produzione di materiali musicali interattivi, passino attraverso la soluzione - o più verosimilmente la riduzione - di una serie di problemi aperti, verso cui orientare sforzi e risorse:

- incrementi di efficienza ed affidabilità nei procedimenti di riconoscimento automatico di caratteristiche dall’informazione multimediale;
- individuazione di metodiche efficaci per la separazione automatica di particolari “sorgenti di informazione”; per esempio: tracce audio di parlatori, cantanti, strumenti, personaggi statici o in movimento da immagini o video;
- ulteriori incrementi della larghezza di banda nelle telecomunicazioni e nella capacità di memoria dei supporti.

Sintetizzando al massimo, grazie al contributo dei progressi delle tecnologie informatiche specifiche, ci si può aspettare nel breve-medio periodo una sempre maggior tendenza alla personalizzazione nella fruizione musicale, sia a livello amatoriale che professionale, fino alla disponibilità globale del patrimonio di conoscenze musicali fin qui maturate, archiviate, pubblicate, suonate, scritte, discusse.

Bibliografia

- [1] Wirth N.: *Algorithms + Data Structures = Programs*. Prentice Hall PTR, Upper Saddle River, NJ, 1978.
- [2] Haydn J.: *Musikalisches Würfelspiel*. Manoscritto originale, 1783.
- [3] Mozart W. A.: *Musikalisches Würfelspiel*. K 516f, 1787.
- [4] Hiller L. A., Isaacson L. M.: *Experimental Music: Composition With an Electronic Computer*. McGraw-Hill, New York 1959.
- [5] Mathews M.V.: *The technology of computer music*. MIT Press, Cambridge, MA, 1969.
- [6] Mathews M.V., Moore R.: *Groove—A Program to Compose, Store, and Edit Functions of Time*. Communications of the ACM Vol. 12, 1970, p. 715.
- [7] Roads C.: *The Computer Music Tutorial*. MIT Press, Cambridge, MA, 1996.
- [8] Roads C., Strawn J., eds.: *Foundations of Computer Music*. MIT Press, Cambridge, MA, 1985.
- [9] Chowning J.: The Synthesis of Complex Audio Spectra by Means of Frequency Modulation. *Journal of the Audio Engineering Society*, Vol. 21, n. 7, 1973, p. 526-534.

- [10] Computer Music Journal: Two Special Issues on Physical Modeling. *MIT Press*, Vol. 16, n. 4 & Vol. 17, n. 1, Winter 92, Spring 93.
- [11] Wallraff D.: *The DMX-1000 Signal Processing Computer*. Proceedings of the 1978 International Computer Music Conference. San Francisco: International Computer Music Association, 1978.
- [12] Di Giugno G., Kott J., Gerzso A.: *Progress Report on the 4X Machine and Its Use*. Proceedings of the 1981 International Computer Music Conference. San Francisco: International Computer Music Association, 1981.
- [13] International MIDI Association: *MIDI Musical Instrument Digital Interface Specification 1.0*. North Hollywood: International MIDI Association, 1983.
- [14] Moorer J. A.: *On the Segmentation and Analysis of Continuous Musical Sound by Digital Computer*. Ph.D. thesis, Dept. of Computer Science, Stanford University, 1975.
- [15] Wold E., Blum T., Keislar D., Wheaton J.: Content-based classification, search and retrieval of audio. *IEEE Multimedia*, Vol. 3, 1996, p. 27-36.
- [16] Dannenberg R.: *An On-line Algorithm for Real-Time Accompaniment*. Proceedings of the 1984 International Computer Music Conference. San Francisco, CA: Computer Music Association, 1984.
- [17] AA.VV. IEEE Std: *1599-2008: Recommended Practice for Defining a Commonly Acceptable Musical Application Using XML*. IEEE Computer Society Press, New York, 2008.
- [18] Haus G.: Rescuing La Scala's Audio Archives. *IEEE Computer*, Vol. 31, n. 3, IEEE CS Press, 1988, p. 88-89.
- [19] Baggi D., Baratè A., Haus G., Ludovico L.: *NINA – Navigating and Interacting with Notation and Audio*. Proceedings of the 2-nd International Workshop on Semantic Media Adaptation and Personalization, London, IEEE Computer Society Press, 2007.
- [20] Sundberg J., Askenfelt A., Frydén L.: Musical performance: A synthesis-by-rule approach. *Computer Music Journal*, Vol. 7, 1983, p. 37-43.

GOFFREDO HAUS è professore ordinario di Informatica all'Università degli Studi di Milano, dove è anche direttore di dipartimento e consigliere di amministrazione. Ha svolto ricerche di informatica applicata alla musica, all'audio e al multimedia, producendo più di 250 pubblicazioni. Goffredo Haus ha diretto e coordinato numerosi progetti, gruppi di lavoro, congressi. Tra questi, da citare i progetti per la IEEE Computer Society, il Teatro alla Scala, il Teatro Bolshoi, il Ministero per i Beni e le Attività Culturali e il conseguimento del nuovo standard musicale IEEE1599-2008.
E-Mail: goffredo.haus@unimi.it