

LA SIMULAZIONE NUMERICA NELLA PROGETTAZIONE DI CIRCUITI INTEGRATI

La progettazione di circuiti elettronici, sia digitali, sia analogici, è sempre più condotta con l'ausilio di simulatori numerici. Oggi, chiunque sia coinvolto nello sviluppo di circuiti elettronici necessita di un simulatore circuitale. Questo strumento ha rivoluzionato l'industria elettronica, consentendo lo sviluppo degli attuali circuiti integrati ad alta densità (circuiti *VLSI*, *Very-Large-Scale Integration*), che possono contenere anche più di un miliardo di transistori. In questo contributo vengono analizzate le problematiche, gli strumenti e i metodi di progetto impiegati nello sviluppo dei moderni circuiti integrati, con particolare riguardo alla simulazione circuitale.

1. INTRODUZIONE

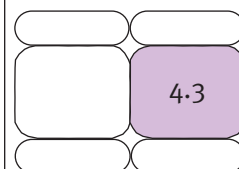
La progettazione di un circuito elettronico si sviluppa, essenzialmente, in tre diverse fasi. Il circuito viene innanzitutto concepito a livello di idea. In questa fase il progettista, sfruttando creatività ed esperienza, ottiene una prima definizione del sistema elettronico che, mediante modelli mentali di alto livello, risulta suddiviso in blocchi logico-comportamentali elementari opportunamente interconnessi tra loro al fine di svolgere le funzioni desiderate. A questa fase creativa iniziale seguono la definizione delle specifiche di progetto e una prima sintesi del circuito, che si basa su modelli reali, seppure estremamente semplificati, dei diversi elementi circuitali (quali transistori, amplificatori, resistori, condensatori, induttori e generatori di segnale) necessari alla sua realizzazione. In questa fase viene definita la topologia del sistema e viene progettata la rete elettrica del circuito elettronico. Nella terza fase, il circuito progettato viene analizzato al fine di verificarne la corretta funzionalità. In funzione dei risultati

ottenuti, il circuito viene opportunamente modificato al fine di ottimizzarne le prestazioni. Questo processo di analisi e sintesi viene ripetuto fino a quando non risultano soddisfatte tutte le specifiche di progetto. Il risultato finale sarà la definizione di tutti i componenti (incluse le loro dimensioni) e le interconnessioni del circuito, che consentirà la successiva fase di produzione.

La fase di analisi di un circuito elettronico è un processo che può risultare estremamente complesso e costoso, se il circuito stesso è complesso. Da questa fase dipendono la qualità del prodotto finale e, quindi, il successo del progetto. L'analisi può essere svolta in due diversi modi: realizzando e valutando sperimentalmente un prototipo su scheda elettronica (*breadboard*) oppure sviluppando un modello numerico in grado di riprodurre ed emulare, in modo sufficientemente accurato, il circuito reale e impiegandolo per simulare il circuito. La scelta fra i due diversi approcci dipende essenzialmente dal tempo richiesto e dal costo complessivo dell'operazione di analisi,



Alessandro Cabrini
Guido Torelli



nonché dalla confidenza sui risultati ottenibili. L'analisi che prevede la realizzazione di un prototipo è conveniente solo nel caso in cui si debbano progettare semplici circuiti a componenti discreti. Nel caso dello sviluppo di un circuito integrato, infatti, la realizzazione di un *breadboard* può risultare eccessivamente onerosa (si pensi al caso di un circuito contenente centinaia di milioni di transistori) e, soprattutto, non consente di emulare adeguatamente il comportamento del circuito reale, a causa dell'intrinseca differenza tra le prestazioni dei transistori integrati e di quelli discreti e tra i componenti parassiti presenti nelle due realizzazioni (questo vale in particolar modo per i transistori MOS, a causa della loro limitata capacità di erogare corrente di uscita). La simulazione numerica è pertanto la scelta unanimemente adottata per la progettazione di circuiti integrati. Una volta dimostrata, tramite simulazione numerica, la funzionalità del circuito progettato, questo sarà integrato mediante il processo di fabbricazione e subirà il vero e proprio collaudo elettrico. È evidente come, per seguire questo approccio, sia necessario assicurare la massima corrispondenza fra modello e realtà. L'accuratezza della simulazione numerica deve garantire la massima confidenza nei risultati al fine di rendere attendibile l'analisi svolta, in modo da limitare il numero di prototipi integrati intermedi (l'obiettivo è anzi ottenere silicio funzionante al primo tentativo) e, quindi, il *time-to-market* e i costi complessivi del progetto.

Dall'idea iniziale al progetto finale, il numero delle variabili in gioco e, quindi, la complessità dei modelli utilizzati nelle diverse fasi progettuali crescono esponenzialmente. Consideriamo per esempio un semplice resistore. Per il progettista, un resistore è un elemento che, concettualmente, mette in relazione lineare, tramite il suo valore di resistenza, la corrente che lo attraversa con la tensione che si sviluppa ai suoi capi. Non importa come il resistore verrà realizzato. Quello che il progettista utilizza è, al primo ordine, il valore nominale della sua resistenza (oltre alla legge di Ohm). Tuttavia, un resistore reale, realizzato per esempio in un circuito integrato in tecnologia CMOS, è molto più complesso. Per descrivere adeguatamente il componente reale, il simulatore circuitale utilizza un modello in

cui il valore di resistenza è solo uno dei tanti parametri. La resistenza di un resistore integrato può variare al variare della corrente da cui è attraversato, come pure può cambiare al variare della tensione che lo polarizza rispetto ad altri elementi del circuito (per esempio il substrato del circuito integrato) e al variare della temperatura. Il resistore può presentare correnti di perdita verso il substrato del circuito integrato. Possono osservarsi comportamenti diversi in funzione del materiale (silicio policristallino, silicio drogato ecc.) con cui il resistore è realizzato. Un resistore integrato presenta inoltre capacità parassite delle quali talvolta è necessario tenere conto. Questo ragionamento si può facilmente estendere agli altri componenti passivi e, a maggior ragione, ai componenti attivi (transistori).

In questo contributo, vengono analizzate le problematiche legate alla simulazione numerica dei circuiti integrati, con particolare riguardo alla simulazione circuitale, che rappresenta la tecnica di simulazione più accurata e attualmente più diffusa per l'analisi dei circuiti elettronici. Nel paragrafo 2 viene introdotto il simulatore circuitale certamente più accettato e utilizzato al mondo: SPICE. Nel paragrafo 3 sono evidenziate le differenti esigenze legate alla simulazione di circuiti analogici e a quella di circuiti digitali, e vengono passati in rassegna i diversi tipi di simulatore per questi ultimi. Il paragrafo 4, infine, prende in esame, con maggiore dettaglio, gli aspetti della simulazione circuitale di circuiti analogici.

2. IL SIMULATORE CIRCUITALE SPICE

I simulatori circuitali furono sviluppati a partire dai primi anni '70. Da allora, la diffusione dei simulatori circuitali è avanzata di pari passo con l'evoluzione delle tecnologie di integrazione. I primi circuiti integrati contavano al massimo qualche centinaio di transistori. Oggi, grazie alla continua riduzione delle dimensioni fisiche dei dispositivi MOSFET, la densità di integrazione è cresciuta fino a superare il miliardo di dispositivi in una singola piastrina di silicio, come previsto dalla legge di Moore [1] (enunciata negli anni '60 e seguita fino ai nostri giorni con sorprendente precisione), secondo la quale il numero di componenti in-

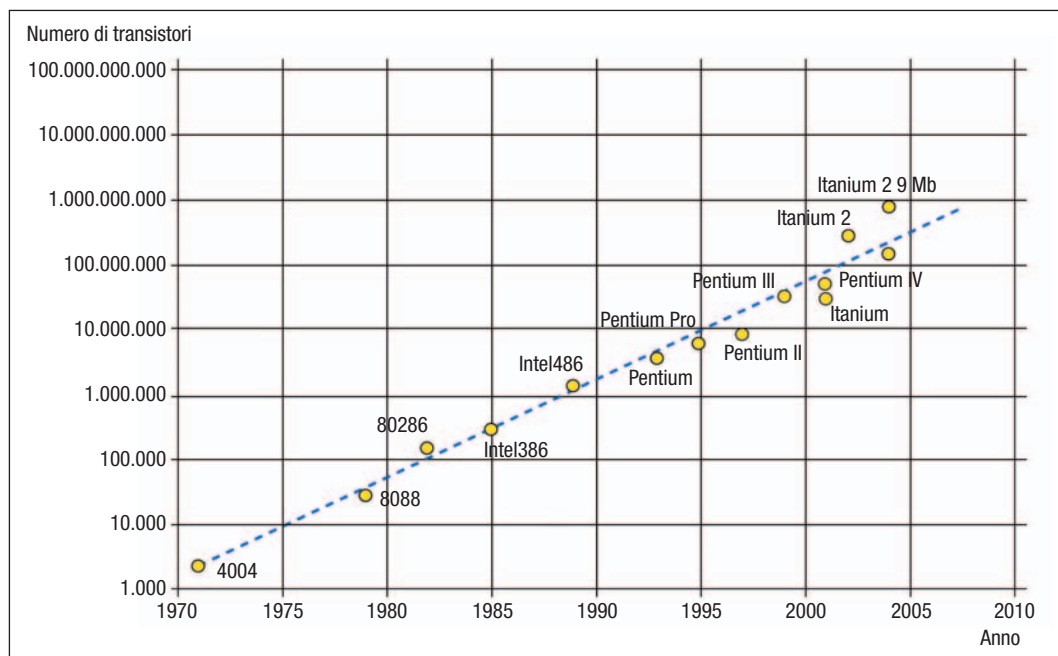


FIGURA 1
Legge di Moore per i microprocessori (raddoppio della densità di integrazione circa ogni due anni)

tegrabili in una singola piastrina di silicio cresce in modo esponenziale con il passare del tempo (Figura 1). Progettare un circuito elettronico con un miliardo di transistori sarebbe impossibile senza l'ausilio di un simulatore circuitale. Analogamente, simulare in un tempo ragionevole un circuito composto da un miliardo di transistori o, almeno, i blocchi circuitali ritenuti critici in esso contenuti sarebbe impossibile senza disporre di dispositivi integrati in grado di garantire ai computer potenze di calcolo estremamente elevate. L'uso combinato di simulatori circuitali sempre più potenti e tecnologie di integrazione sempre più spinte ha consentito lo sviluppo dirompente dell'elettronica moderna e dei settori a questa legati, primi tra tutti quelli dell'informatica e delle telecomunicazioni.

Sviluppato attorno al 1970 nei laboratori di ricerca dell'Università di Berkeley (California, U.S.A.) dal gruppo del Prof. Donald Pederson, SPICE (*Simulation Program with Integrated Circuit Emphasis*) fu il primo simulatore circuitale accettato estesamente nel mondo. Come lascia intendere il nome, l'interesse specifico di questo strumento era rivolto alla simulazione di circuiti integrati. La prima versione del simulatore (SPICE1) venne presentata durante una conferenza nel 1973 [2]. SPICE1 era scritto nel linguaggio di programmazione Fortran, e la descrizione del comportamento elettrico del

circuito si basava sull'analisi nodale, *node-voltage analysis*. La descrizione del circuito include i componenti elettrici (sia passivi, sia attivi) e i nodi (punti ai quali sono connessi almeno due componenti); a ogni nodo è associata una tensione, mentre a ogni ramo (un ramo rappresenta il collegamento tra due nodi) è associata una corrente. Ogni componente elettrico è rappresentato mediante un opportuno modello matematico, che descrive la relazione tra le tensioni presenti ai suoi terminali e le correnti che entrano in questi. In tal modo, sostanzialmente, il circuito risulta descritto utilizzando una matrice di ammettenza, che mette in relazione la corrente che scorre in ciascun ramo del circuito con le tensioni ai capi dei vari rami. La soluzione del sistema di equazioni (tipicamente non lineari) che descrive il circuito si basa su tecniche iterative (di tipo Newton-Raphson [3]): viene prima stimato un valore iniziale delle variabili elettriche (per esempio, nel caso di un'analisi di transitorio temporale, questo coinciderà con il valore assunto dalle variabili nell'intervallo temporale precedente a quello considerato), dopo di che il valore delle variabili viene variato in modo iterativo, fino a che non si raggiunge un livello considerato corretto, secondo un criterio prestabilito di errore massimo ammesso (quanto minore è questo errore, tanto migliore sarà l'accuratezza dei risultati, ma tanto maggiore sarà il tempo richie-

sto per completare la simulazione). Per quanto riguarda la variabile tempo, nelle analisi di transitorio temporale SPICE1 (come le successive versioni di SPICE) discretizza l'asse dei tempi in passi (*time step*) discreti (un passo più piccolo darà luogo a un'accuratezza migliore dei risultati e a un tempo di simulazione superiore).

Il tipo di analisi impiegato (cioè l'analisi nodale) rappresentava uno dei maggiori limiti di SPICE1 (componenti circuitali quali per esempio gli induttori e alcuni tipi generatori controllati introducono singolarità nella matrice), limiti che vennero risolti con la successiva versione del simulatore (SPICE2) introducendo la cosiddetta analisi nodale modificata (*modified nodal analysis*) [4]. SPICE1 utilizzava la *fixed step analysis* (simulazione a passo costante), mediante la quale l'intervallo temporale di simulazione viene suddiviso in una successione di istanti temporali equispaziati e lo stato del circuito viene risolto per ognuno degli istanti di questa successione. In SPICE2 venne inoltre introdotta la simulazione a passo variabile (*variable-step analysis*), nella quale il passo temporale di simulazione viene variato in funzione dell'effettiva attività del circuito (più precisamente, il passo viene reso più elevato quando le grandezze elettriche nel circuito risultano meno variabili), ottenendo una maggior velocità complessiva di simulazione. Nel 1989 venne introdotto SPICE3, scritto in linguaggio C, il quale presenta l'aggiunta di funzionalità per la visualizzazione di grafici e forma d'onda.

Il simulatore SPICE rappresenta, di fatto, la base dei diversi simulatori circuitali sviluppati negli anni e oggi in uso nell'industria elettronica. A partire da questo simulatore, molte aziende hanno sviluppato strumenti di progettazione commerciali, ottimizzandone le prestazioni. I simulatori commerciali oggi maggiormente diffusi e utilizzati includono infatti HSPICE (inizialmente commercializzato da Meta Software e oggi di proprietà di Synopsis) o PSPICE (di Cadence Design Systems). I modelli dei componenti utilizzati da un simulatore circuitali giocano ovviamente un ruolo fondamentale nell'accuratezza e nella robustezza dei risultati ottenuti. Normalmente, un'elevata accuratezza richiede lo sviluppo di modelli complessi, i quali possono portare a tempi di simulazione assai elevati e, in casi estremi, dare luogo

a problemi di convergenza. Occorre quindi un'adeguata scelta di compromesso tra accuratezza e complessità dei modelli (questo vale soprattutto per i modelli dei transistori). I modelli sviluppati vanno poi validati e calibrati mediante confronto con il comportamento sperimentale dei componenti. Nel corso degli anni sono stati sviluppati, anche da parte dei produttori di circuiti integrati, diversi modelli del transistor MOS, e sono stati proposti diversi livelli (*level*), orientati a garantire diverse prestazioni in termini di accuratezza e tempo di esecuzione. Il modello attualmente adottato come standard anche industriale, pure esso sviluppato presso l'Università di Berkeley, prende il nome di BSIM (*Berkeley Short-Channel IGFET Model*) [5].

3. CIRCUITI ANALOGICI E CIRCUITI DIGITALI: MONDI DIVERSI, ESIGENZE DIVERSE

Il simulatore SPICE nasce e si sviluppa come strumento per l'analisi di tipo analogico di circuiti elettrici nel dominio tempo-continuo. Amplificatori per audio, filtri, riferimenti di tensione o sensori di varia natura sono solo alcuni dei più noti blocchi elettronici analogici. La simulazione numerica di questi circuiti è concettualmente una simulazione di tipo tempo-continuo (anche se in realtà non può essere realizzata che in passi temporali discreti, come sopra evidenziato) con elevato grado di accuratezza "analogica". Il progetto di un circuito elettronico per la generazione di un riferimento di tensione stabile in temperatura, per esempio, richiede una simulazione che sia in grado di determinare con grande accuratezza il valore analogico della tensione di uscita del circuito al variare della temperatura e/o della tensione di alimentazione. Parametri fondamentali per un circuito analogico sono soesso il rapporto segnale/rumore, la linearità e la risposta in frequenza (per approfondimenti si veda il riquadro 1 a p. 36).

Lo sviluppo dei moderni sistemi elettronici digitali, quali per esempio i microprocessori o i sistemi di memorizzazione, ha determinato lo sviluppo di un diverso tipo di simulatori. I sistemi digitali, infatti, trattano il segnale non più come una grandezza continua di tipo tempo-continuo, bensì come un'informazione co-

dificata tramite bit (o *binary digit*, da cui il termine digitale) con evoluzione essenzialmente di tipo tempo discreto. Simulare un sistema digitale con un simulatore per circuiti analogici, quale SPICE, è un inutile spreco di risorse. Solitamente, infatti, di un circuito digitale si vuole verificare il corretto funzionamento logico, nonché la correttezza delle temporizzazioni e del sincronismo tra i diversi segnali. Non si richiede insomma, se non in casi specifici quali applicazioni ad altissima velocità, un'elevata accuratezza nella determinazione dell'evoluzione temporale delle tensioni ai nodi interni del circuito. Importanti nei circuiti digitali sono le transizioni dei segnali da un livello logico all'altro. In assenza di commutazioni, tutte le tensioni ai nodi interni del circuito sono stabili nel tempo. Una simulazione analogica di tipo tempo-continuo di un (intero) sistema digitale risulterebbe, ovviamente, estremamente costosa in termini di potenza di calcolo e i risultati ottenuti sarebbero caratterizzati da un'accuratezza non necessaria. Diverso è il caso in cui si voglia determinare nel dettaglio, per esempio, l'evoluzione temporale della tensione di uscita di una porta logica (quale un invertitore), come è richiesto durante lo sviluppo di celle elementari che verranno poi impiegate per realizzare un sistema digitale. In questo caso, la porta digitale deve essere considerata dal punto di vista analogico, e l'utilizzo di un simulatore circuitale quale SPICE risulta pienamente adeguato.

Un sistema elettronico digitale può essere modellizzato e descritto secondo quattro livelli di astrazione (Figura 2), cioè un primo livello di astrazione, tramite il quale si definisce la funzione che il sistema deve svolgere (indipendentemente dalla sua struttura fisica), un livello architetturale (o RTL, *Register Transfer Level*), nel quale il sistema è descritto come interconnessione di componenti di elaborazione standard (quali registri, sommatore e moltiplicatori), un livello di descrizione logico, nel quale ogni componente architetturale viene descritto tramite un'opportuna rete logica e, infine, il livello circuitale vero e proprio, nel quale ogni singola rete logica viene realizzata e simulata con uno specifico schema circuitale a livello transistore. A ognuno di questi diversi livelli di astrazione corrisponde uno specifico ambiente di simulazione numerica.

Da un progetto finanziato dal Dipartimento della Difesa americano, nel 1987 nacque il VHDL (*VHSIC Hardware Description Language*, [6] dove l'acronimo VHSIC indica *Very High Speed Integrated Circuit*), che è un linguaggio ad alto livello per la descrizione e la progettazione di sistemi digitali elettronici. Tramite VHDL è possibile descrivere il sistema sia a livello comportamentale, cioè mediante le relazioni tra le variabili di ingresso e le variabili di uscita, sia a livello strutturale, cioè mediante i moduli funzionali da cui esso è costituito e le relative interconnessioni (le stesse tipologie di descrizione sono possibili pure per i vari moduli); la descrizione del sistema permette di determinarne la risposta in funzione dei segnali presenti in ingresso. Un altro linguaggio ad alto livello per la progettazione di circuiti digitali è Verilog (di Synopsis). Una volta terminata con successo la fase di analisi, questi programmi consentono di realizzare la sintesi del circuito integrato in modo automatico, mediante l'impiego di celle digitali (completamente caratterizzate) disponibili in libreria.

Parametri critici per la progettazione e la conseguente simulazione digitale sono i tempi di ritardo, i tempi di commutazione e i tempi di propagazione dei segnali attraverso le diverse porte logiche. Un circuito digitale è completamente descritto una volta che siano noti la funzione logica che esso implementa e i tempi di elaborazione richiesti (i quali dipendono dalla

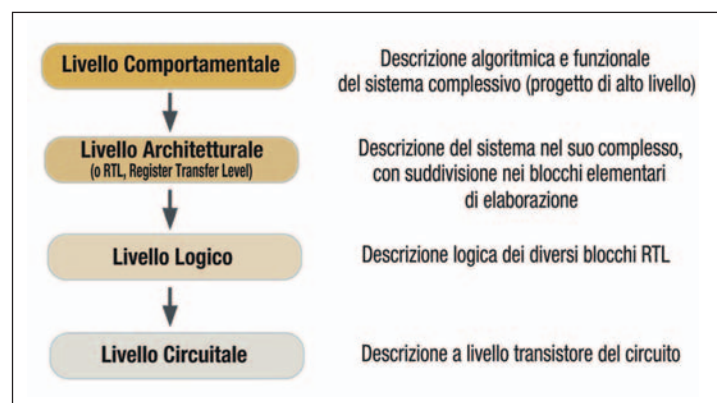


FIGURA 2

Livelli di astrazione nella descrizione di circuiti digitali. A questi livelli viene in genere aggiunto il livello fisico, che riguarda il layout, cioè il disegno delle figure che daranno origine alle maschere per la fabbricazione del circuito: questo livello specifica quindi come è costituito fisicamente il circuito. Il livello fisico, peraltro, non ha interesse specifico in questa sede



realizzazione fisica della porta logica). Sono pertanto stati sviluppati simulatori che descrivono e analizzano i circuiti digitali nel dominio del tempo. Esistono peraltro diversi livelli di dettaglio nella descrizione e, quindi, nell'analisi (che si riflettono, naturalmente, in diverse prestazioni in termini di accuratezza da una parte e di complessità computazionale dall'altra). Sono disponibili simulatori logici, per i quali gli elementi di base sono le porte logiche (queste non vengono descritte in base alla loro realizzazione circuitale, ma hanno associati i tempi di ritardo, che tipicamente dipendono dal carico capacitivo di uscita), e simulatori a livello interruttore (*switch-level*), che descrivono le singole porte logiche in base alla loro realizzazione circuitale e simulano ciascun transistor come un interruttore che risulterà spento (resistenza infinita) o acceso (resistenza pari alla resistenza serie media del corrispondente transistor nelle condizioni operative di accensione) in funzione dell'evoluzione dei segnali. I simulatori di tipo *switch-level* tengono conto delle capacità presenti ai nodi, in modo da fornire una stima sufficientemente accurata dei tempi di commutazione.

A livello ancora più basso possono essere classificati i simulatori temporali (*timing simulator*), il cui obiettivo fondamentale è valutare parametri temporali di una rete logica (tempi di propagazione, di salita e di discesa). Questi simulatori descrivono un circuito digitale attraverso la sua realizzazione fisica a livello transistor, ma impiegano per i componenti modelli assai semplificati (pur tenendo conto di numerosi effetti non ideali), in modo da permettere la simulazione anche di un circuito assai complesso i tempi ragionevoli. Il loro impiego può risultare particolarmente utile per la simulazione globale di un intero circuito nella fase finale della progettazione (per un approfondimento si consideri l'esempio relativo a una memoria digitale riportato nel riquadro 2 a p. 37).

4. LA SIMULAZIONE DI CIRCUITI ANALOGICI

A differenza di quanto accade nella simulazione dei circuiti digitali per i quali, fondamentalmente, vengono effettuate simulazioni aventi come obiettivo la verifica funzionale del circuito e la corretta temporizzazione delle diverse

parti del circuito stesso, l'analisi di un circuito analogico si articola in quattro diverse fasi, alle quali corrispondono quattro distinte tipologie di simulazione. La prima fase è l'analisi in continua (DC). In questo caso, il circuito è simulato in condizioni statiche, in assenza cioè di segnali tempo varianti. Obiettivo della simulazione in DC è quello di valutare il punto operativo (*Operating Point*, OP) del circuito, ovvero analizzare le condizioni di polarizzazione dei diversi componenti. Infatti, per ognuno dei diversi componenti (in particolare per i dispositivi attivi quali i transistor), è necessario analizzare accuratamente la regione di funzionamento. Un transistor MOS può operare in saturazione, in triodo, in debole inversione o in interdizione. In ognuna di queste regioni, il comportamento del transistor e, quindi, le sue caratteristiche elettriche di piccolo segnale sono differenti che nelle altre. In un circuito elettronico analogico, la polarizzazione e, quindi, la regione di funzionamento dei transistori devono essere quelle previste dal progettista. Un transistor progettato per uno stadio di guadagno, per esempio, può introdurre non linearità pesanti quando passa da una regione di funzionamento a un'altra: tipicamente, questo deve essere evitato. Lo strumento per verificare quanto sia robusto il progetto di un circuito dal punto di vista della polarizzazione è appunto l'analisi tramite simulazioni in DC.

Conclusa l'analisi in DC, il comportamento del circuito deve essere analizzato nei domini della frequenza e del tempo. A tal fine sono effettuate simulazioni in AC e simulazioni in transitorio. Nella simulazione in AC, viene analizzato il comportamento del circuito a piccolo segnale. In questo caso si analizza la risposta in frequenza del circuito nell'intorno dello stato di polarizzazione in DC. La caratteristica elettrica dei diversi componenti viene linearizzata nell'intorno del punto operativo del circuito e il circuito viene simulato valutando la risposta a uno stimolo sinusoidale con ampiezza unitaria e frequenza variabile in un intervallo predefinito. La simulazione in AC è fondamentale per analizzare la stabilità in frequenza di un circuito, per valutare il guadagno in continua e la banda passante di un amplificatore e per determinare la risposta in frequenza di un filtro. Il risultato di una simulazione in AC è solitamente il diagramma di Bode, tramite il quale è pos-

sibile valutare la posizione e il peso dei poli e degli zeri nella funzione di trasferimento di un circuito. Dato che è un'analisi di piccolo segnale, la simulazione in AC non è di per sé sufficiente per valutare in modo esaustivo la risposta del circuito a un generico stimolo esterno. I limiti dell'analisi in AC vengono superati effettuando un'analisi in transitorio. Con la simulazione in transitorio, il progettista valuta la risposta nel dominio del tempo del circuito progettato. In una simulazione in transitorio, lo stato di polarizzazione dei dispositivi varia nel tempo come avviene in un circuito reale e, quindi, possono essere analizzate le non linearità e le distorsioni introdotte nella risposta. Nell'analisi in transitorio, il progettista valuta la risposta a diversi stimoli di ingresso quali, ad esempio, gradini e rampe di tensione o corrente, e stima i tempi di recupero dello stato di polarizzazione per effetto di stimoli impulsivi (quali per esempio quelli causati da *sharing* capacitivi dovuti alla connessione, al nodo di uscita, di un condensatore inizialmente scarico). Oltre ai tipi di analisi indicati, il simulatore circuitale offre ulteriori possibilità, quali l'analisi di rumore e l'analisi di distorsione armonica (quest'ultima viene eseguita sul risultato di un'analisi in transitorio, al fine di tenere conto degli effetti non lineari dei componenti).

Le analisi in DC, in AC e in transitorio sono, quindi, necessarie per valutare il comportamento del circuito in modo esaustivo. Tuttavia, esse non sono sufficienti. Prima che il circuito progettato possa essere approvato per la successiva fase di layout, durante la quale esso viene tradotto dalla sua vista a livello schema alla vista fisica che verrà poi realizzata su silicio tramite il processo di integrazione (Figura 3), occorre completare le analisi considerando anche le variazioni di processo. A causa dell'accuratezza finita del processo di integrazione, infatti, tutti i parametri elettrici dei dispositivi sono ottenuti con una certa tolleranza (che varia a seconda del parametro considerato, ma si può stimare essere nell'ordine del 10%). Il circuito deve quindi essere analizzato considerando anche le dispersioni dei parametri di processo. Fra questi, i più critici sono la tensione di soglia e il fattore di transconduttanza dei transistori MOS, la resistività dei diversi materiali (silicio monocristallino drogato, silicio policristallino, metallizzazioni), le capacità per

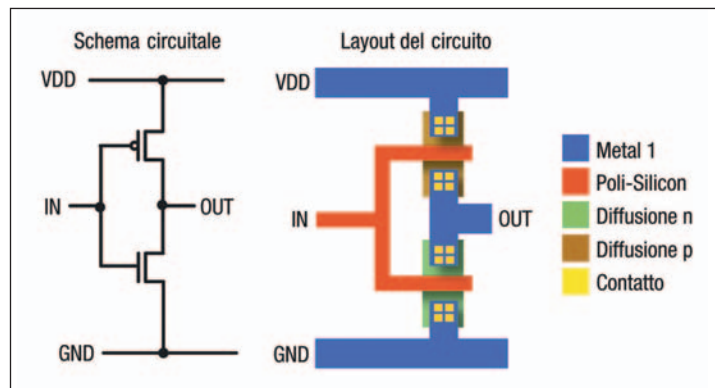


FIGURA 3

Vista a livello schema (sinistra) e vista layout (destra) di un invertitore CMOS

unità di superficie, le dimensioni geometriche dei componenti. Durante la fase di sviluppo del progetto, questa analisi viene in genere condotta imponendo ai diversi parametri i valori estremi delle tolleranze che la linea di produzione è in grado di garantire (analisi dei *corner* di processo). I simulatori circuitali rendono inoltre disponibile la cosiddetta simulazione di Monte Carlo, [7] che viene in genere eseguita nella fase conclusiva della progettazione di uno specifico blocco circuitale. Nella simulazione di Monte Carlo, il circuito progettato viene simulato (in DC, AC e/o transitorio, a seconda delle esigenze) un numero N di volte. In ognuna di queste N simulazioni, i parametri dei componenti vengono scelti in modo non correlato (quindi casuale) seguendo la distribuzione di probabilità propria, per ciascun parametro, del processo di integrazione considerato. La simulazione di Monte Carlo emula quindi un insieme casuale di realizzazioni possibili del circuito sotto analisi, dove per ciascuna realizzazione la probabilità di essere presa in esame risulta pari alla probabilità che essa si verifichi durante la fase di produzione. Su ciascuno dei campioni casuali così determinati, la simulazione esegue la valutazione delle grandezze di interesse (quali per esempio il guadagno in continua o la frequenza di taglio di un amplificatore, o il tempo di ritardo di un invertitore), permettendo in tal modo di avere una stima delle distribuzioni dei valori di tali grandezze cui il circuito progettato darà luogo in fase di produzione e, quindi, una stima della resa di produzione (fattore di fondamentale importanza a livello industriale).

Come mostrato nella figura 4, una volta ulti-

mata l'analisi circuitale, basata sullo schema elettrico, il flusso di progetto di un circuito integrato prosegue con la fase di *layout*. Come accennato in precedenza, durante questa fase vengono definite le figure geometriche che dovranno essere ottenute sui diversi strati di materiale (metalli, ossido di silicio, diffusione di tipo n e di tipo p, silicio policristallino) per realizzare i componenti e le interconnessioni che costituiranno, nel loro complesso, il circuito integrato voluto dal progettista; queste figure geometriche saranno utilizzate per la generazione delle maschere del processo di integrazione. Senza entrare nei dettagli operativi della fase di *layout*, in questa sede è opportuno puntualizzare che la realizzazione fisica di un circuito ha un impatto che può essere non trascurabile sulle prestazioni di questo, a causa della presenza inevitabile di elementi parassiti (quali le resistenze delle linee di interconnessione o gli accoppiamenti capacitivi fra elementi fisicamente adiacenti). A causa degli elementi parassiti introdotti, le prestazioni del circuito precedentemente simulate potrebbero risultare degradate. Una volta ultimato il *layout*, quindi, un buon progetto di un circuito integrato prevede l'uso di strumenti per l'estrazione degli elementi parassiti. Si rende a questo punto necessaria un'ulteriore fase di simulazione finale volta a verificare che, anche tenendo conto degli effetti degli elementi parassiti, il circuito risulti sempre in specifica di progetto, consentendo quindi il passaggio alla fase finale di produzione.

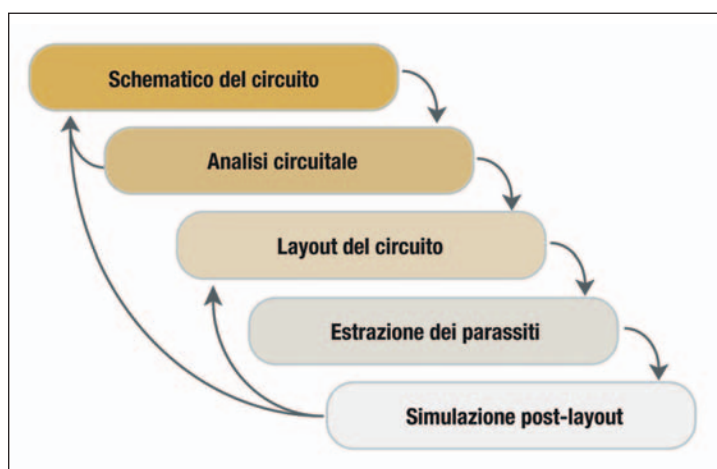


FIGURA 4

Flusso di progettazione di un circuito integrato

5. CONCLUSIONI

I moderni circuiti integrati sono sistemi realizzati con milioni (e talvolta anche miliardi) di transistori. Progettare un circuito elettronico di tale complessità sarebbe impossibile senza l'ausilio della simulazione numerica, mediante la quale le prestazioni di un sistema vengono stimate e affinate durante la fase di sviluppo. La simulazione può essere eseguita a diversi livelli di complessità, in modo da ottimizzare da una parte i tempi di calcolo e, dall'altra, l'accuratezza dei risultati. L'uso combinato di simulatori sempre più potenti e tecnologie di integrazione sempre più spinte ha consentito lo sviluppo dirompente dell'elettronica moderna e dei settori a questa legati, primi tra tutti quelli dell'informatica e delle telecomunicazioni. Questo lavoro ha analizzato le problematiche relative alle tecniche di simulazione, con specifico riguardo alla simulazione circuitale, che costituisce la tecnica più accurata per l'analisi dei circuiti integrati. Sviluppati a partire dai primi anni '70, i simulatori circuitali hanno subito un'evoluzione costante. Dal primo prototipo SPICE, si è giunti oggi a sistemi di simulazione in grado di affiancare il progettista durante tutto il percorso di sviluppo del circuito integrato. Dalla concezione dell'idea fino al rilascio del progetto finale per la fase di integrazione e realizzazione del dispositivo, il simulatore rappresenta uno strumento indispensabile, senza il quale la realizzazione dei moderni sistemi elettronici non sarebbe praticabile.

Acronimi

CMOS	Complementary Metal-Oxide-Semiconductor
IGFET	Insulated-Gate Field Effect Transistor
MOS	Metal-Oxide-Semiconductor
MOSFET	Metal-Oxide-Semiconductor Field Effect Transistor

Bibliografia

- [1] Moore G.E.: Cramming more components onto integrated circuits. *Electronics*, Vol. 38, n. 8, April 1965.
- [2] Nagel L.W., Pederson D.O.: *SPICE (Simulation Program with Integrated Circuit Emphasis)*. EECS Department, University of California, Berkeley, Tech. Rep. No. UCB/ERL ERL-M382, Apr. 1973.

- [3] Press W.H., Flannery B.P., Teukolsky S.A., Vetterling W.T.: "Newton-Raphson Method Using Derivatives" and "Newton-Raphson Methods for Nonlinear Systems of Equations". §9.4 and 9.6 in Numerical Recipes in FORTRAN: The Art of Scientific Computing, 2nd ed. Cambridge, England: Cambridge University Press, 1992, p. 355-362, p. 372-375.
- [4] Nagel L.W.: *SPICE2, a Computer Program to Simulate Semiconductor Circuits*. EECS Department, University of California, Berkeley, Tech. Rep. n. UCB/ERL-M520, 1975.
- [5] Liu W., Jin X., Chen J., Jeng M-C., Liu Z., Cheng Y., Chen K., Chan M., Hui K., Huang J., Tu R., Ko P.K., Chenming Hu: *BSIM 3v3.2 MOSFET Model Users' Manual*. EECS Department, University of California, Berkeley, Technical Report n. UCB/ERL M98/51, 1998.
- [6] AA. VV.: IEEE Standard VHDL Language Reference Manual. *IEEE Std*, 1988, p. 1076-1987.
- [7] Metropolis N., Ulam S.: The Monte Carlo Method. *Journal of the American Statistical Association (American Statistical Association)*, Vol. 44, n. 247, 1949, p. 335-341.

ALESSANDRO CABRINI si è laureato (con lode) in Ingegneria Elettronica nel 2000 presso l'Università di Pavia dove, successivamente, ha conseguito il Dottorato di ricerca in Elettronica ed Informatica con un lavoro su memorie non volatili di tipo flash multilivello. Dal 2004 svolge attività di ricerca presso il Dipartimento di Elettronica dell'Università di Pavia occupandosi di progettazione e caratterizzazione di memorie non volatili a cambiamento di fase basate su materiali calcogenuri. Fra le sue principali attività di ricerca risulta inoltre la progettazione di circuiti analogici per la conversione di tensioni DC-DC ad alta efficienza energetica basati su topologie a *charge pump* e *buck-boost*.

E-mail: alessandro.cabrini@unipv.it

GUIDO TORELLI, nato a Roma nel 1949, si è laureato (con lode) in Ingegneria Elettronica nel 1973 presso l'Università di Pavia dove, dopo la laurea, ha goduto di una borsa di studio annuale. Nel 1974, entrò in SGS-ATES (ora parte di STMicroelectronics), dove ha lavorato come progettista di circuiti integrati MOS, raggiungendo la qualifica di Dirigente (1987). Nel 1987 entrò nel Dipartimento di Elettronica dell'Università di Pavia, dove è attualmente Professore Ordinario. I suoi interessi di ricerca sono nell'area della progettazione, analisi e caratterizzazione di circuiti integrati MOS (attualmente l'enfasi è sulle memorie non volatili e sui circuiti misti analogico-digitali CMOS). Ha ricevuto (con colleghi) il Premio Ambrose Fleming della IEE (sessione 1994-95). È Senior Member della IEEE.

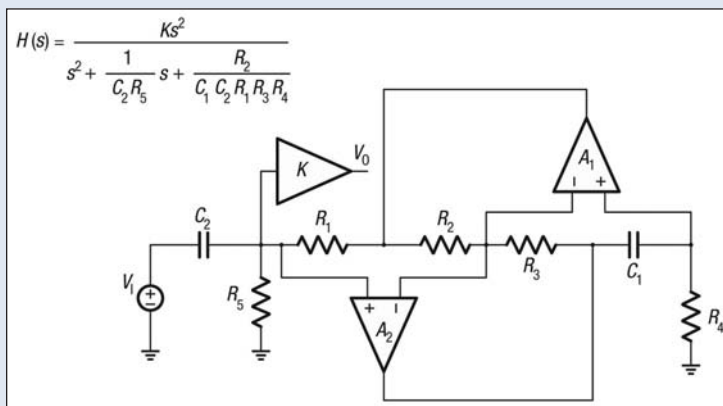
E-mail: guido.torelli@unipv.it

Riquadro 1 - Simulazione di un filtro elettronico

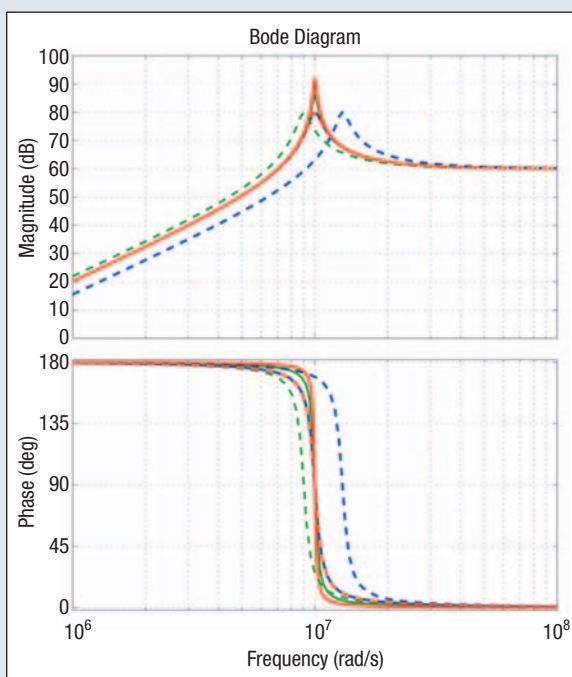
Un filtro elettronico è un circuito che introduce un'attenuazione e/o un'amplificazione selettiva per ciascuna delle frequenze del segnale applicato al suo terminale di ingresso. Esistono filtri di tipo passa basso (che attenuano le componenti ad alta frequenza del segnale), filtri passa alto (che, viceversa, attenuano le componenti a bassa frequenza del segnale), filtri passa banda (che attenuano le componenti del segnale al di fuori di un prefissato intervallo, o banda, di frequenze), filtri notch (che eliminano una determinata banda di frequenza, molto stretta, del segnale in ingresso) ecc.. Un filtro si caratterizza per la sua risposta in frequenza, che fornisce l'amplificazione del filtro in funzione della frequenza. Quando si progetta un filtro elettronico, la sua risposta in frequenza viene analizzata accuratamente mediante simulazione in AC. A questo proposito, è molto importante valutare la sensibilità della risposta in frequenza del filtro alle variazioni dei parametri dei componenti che lo costituiscono. Consideriamo, per esempio, il circuito della figura A, tramite il quale si realizza un filtro attivo passa alto del secondo ordine (cioè con due poli). La risposta in frequenza del circuito è data da:

$$H(s) = \frac{Ks^2}{s^2 + \frac{1}{C_2 R_5} s + \frac{R_2}{C_1 C_2 R_1 R_3 R_4}}$$

In particolare, i parametri K (guadagno a centro banda), ω_0 (frequenza caratteristica) e Q (fattore



A - Schema circuitale di un filtro attivo del secondo ordine (V_1 = segnale di ingresso; V_0 = segnale di uscita)

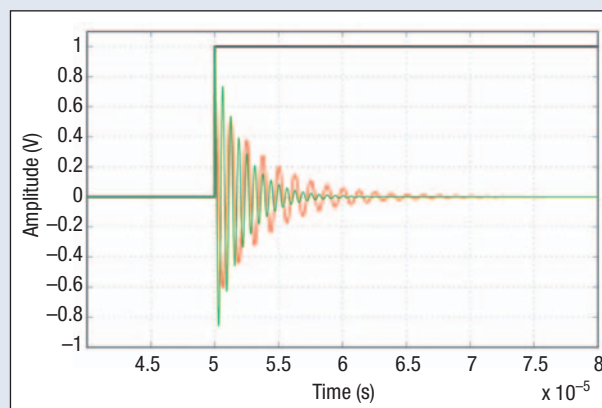


B - Risposta in frequenza del filtro di figura A per diversi valori dei componenti (e, quindi, di K , ω_0 e Q); la risposta è stata ottenuta mediante simulazione parametrica in AC

lizzare, per esempio, i transistori di funzionamento, la risposta del filtro al gradino, come quella mostrata nella figura C, che è essenziale, assieme a simulazioni in AC, ai fini della valutazione della sua stabilità in frequenza, e le prestazioni di distorsione armonica, che vengono valutate mediante la trasformata di Fourier della risposta temporale a uno stimolo sinusoidale) e, ove occorra, simulazioni particolari mirate a obiettivi specifici.

di merito) dipendono dall'accuratezza dei valori dei condensatori e dei resistori impiegati, oltre che dalle caratteristiche degli amplificatori operazionali utilizzati. Mediante simulazioni parametriche (che consistono nell'esecuzione di successive simulazioni, eseguite variando il valore di determinati componenti entro un dato intervallo con un passo prefissato) su resistori e condensatori, si ottengono i risultati riportati nella figura B: da questi è possibile stimare quali sono gli elementi maggiormente critici per la risposta in frequenza del sistema. Ciò consente di effettuare la scelta migliore tra le tipologie disponibili dei componenti, o di cambiare la topologia del filtro in modo che le specifiche richieste vengano rispettate anche in presenza delle inevitabili tolleranze di produzione. Un'analisi di Monte Carlo sulla topologia definitiva potrà fornire dati interessanti in tal senso.

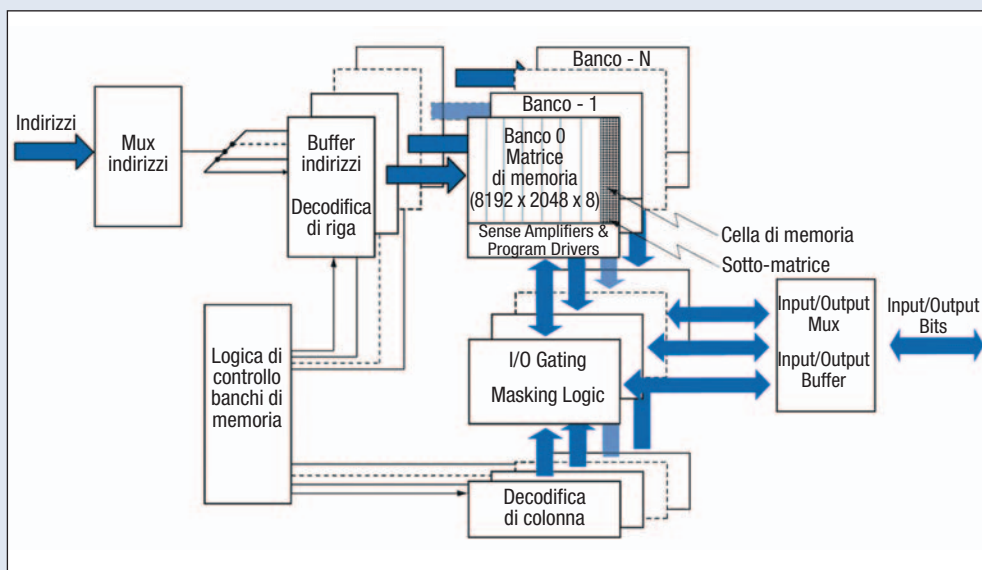
Naturalmente, una buona progettazione di un filtro come quello mostrato prevede pure simulazioni in DC (per analizzare le condizioni di polarizzazione dei componenti), simulazioni in transitorio (per ana-



C - Risposta al gradino del filtro di figura A, ottenuta mediante simulazione in transitorio, per due differenti combinazioni di (K , ω_0 e Q). Traccia nera: gradino di tensione applicato all'ingresso; tracce rossa e verde: uscita del filtro nei due diversi casi

Riquadro 2 - Simulazione di una memoria digitale

Una moderna memoria digitale include centinaia di milioni - nei casi più avanzati, miliardi - di transistori. La simulazione di sistema elettronico di tale complessità ha determinato lo sviluppo di sistemi di simulazione dedicati. In questo ambito troviamo la cosiddetta *hierarchical simulation* (simulazione gerarchica). Una caratteristica comune a tutti i sistemi di memoria è infatti la possibilità di suddividere il sistema stesso in un insieme gerarchico di blocchi elementari a complessità crescente via via che si sale dal basso verso l'alto (Figura A). Il più elementare di questi blocchi è la cella di memoria, mentre quello al livello più alto è il sistema completo. Durante il normale funzionamento di un dispositivo di memoria, solo una piccola parte dei transistori è effettivamente attiva in un certo istante di tempo. Infatti, le celle non indirizzate per una lettura o una programmazione (che solitamente rappresentano la grande maggioranza del sistema) sono inattive. Per questo, la simulazione del loro comportamento è, di fatto, non essenziale. I simulatori gerarchici sfruttano appunto questa caratteristica. In particolare, quando viene dato in ingresso uno stimolo, in una prima fase della simulazione tutte le celle di memoria sono rimosse dal sistema e lo stimolo applicato viene utilizzato per determinare le *bit-line* e le *word-line* che saranno attivate dallo stimolo stesso. In seguito, vengono inserite per la simulazione le celle ritenute necessarie (in quanto connesse o prossime alle *word-line* e alle *bit-line* attivate dallo stimolo applicato). La procedura prende il nome di *hierarchical array reduction*, e consente una notevole riduzione dei tempi di simulazione. Altri miglioramenti nei tempi di simulazione sono poi ottenuti sfruttando le tecniche note come *hierarchical storage* e *isomorphic matching*. Tramite la prima, elementi fondamentali (cella di memoria, flip-flop, porte logiche) che sono ripetuti più volte nel circuito, sono riconosciuti e memorizzati come singola istanza di un medesimo "modello" e vengono inseriti "on the fly" (al volo) nella simulazione solo quando ritenuti necessari. Tramite l'*isomorphic matching*, viceversa, vengono eliminati dalla simulazione cicli identici dedicati a blocchi uguali che lavorano nelle stesse condizioni operative. L'*isomorphic matching* elimina la ripetizione di simulazioni identiche, con evidenti vantaggi in termini di tempo complessivo di simulazione.



A - Schema a blocchi di un dispositivo di memoria digitale

AMBIENTI DI CALCOLO PER LA SIMULAZIONE NUMERICA

La simulazione numerica di sistemi continui, con lo sviluppo di modelli di calcolo sempre più raffinati, permette di indagare fenomeni fisici, chimici, biologici anche estremi ed è strumento fondamentale in molti campi della ricerca e dell'innovazione tecnologica. L'articolo presenta in breve lo stato attuale e le possibili evoluzioni future di infrastrutture e strumenti di calcolo che permettono di studiare la natura anche nei suoi aspetti più nascosti o elusivi e rendono possibile l'innovazione in molti settori avanzati della ricerca, della progettazione e della produzione.

1. INTRODUZIONE

Un sistema è un insieme di componenti che interagiscono ed operano nello spazio e nel tempo: in particolare nei sistemi continui possono essere continui sia lo spazio sia il tempo. Un modello è una rappresentazione semplificata di un sistema in un certo punto dello spazio o in un certo momento del tempo, ideato per permettere lo studio di un sistema reale e del suo funzionamento in un assegnato ambiente. Una simulazione è la manipolazione di un modello in modo da comprimere il tempo e/o lo spazio, permettendo così di cogliere e caratterizzare le interazioni tra le parti del sistema e con l'ambiente, interazioni che altrimenti non si potrebbero analizzare essendo troppo dilazionate nel tempo e/o distribuite nello spazio.

I modelli di sistemi dinamici continui sono modelli matematici alle equazioni differenziali, totali o parziali, la cui integrazione per mezzo di un calcolatore richiede appositi metodi numerici.

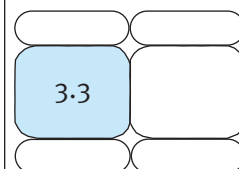
I sistemi e i processi citati hanno una carat-

teristica in comune: la *complessità*. Essa si presenta in svariate forme quali: l'esistenza di comportamenti altamente non lineari e/o di fenomeni aleatori, la coesistenza di scale multiple, l'elevato numero di componenti e la molteplicità delle loro interconnessioni, l'incompletezza della conoscenza dei meccanismi di base che li governano, la grande quantità di dati e le tecniche avanzate per la loro analisi.

In questo contesto è richiesto uno studio sistematico che includa: l'analisi del problema, la formulazione del modello, la soluzione analitica e/o numerica, la validazione del modello, l'interpretazione dei risultati e infine lo sviluppo di opportune strategie di controllo. Le competenze modellistiche e le possibilità di simulazione numerica giocano quindi un ruolo fondamentale in quanto permettono un'analisi accurata del fenomeno, la più completa possibile rispetto alle varie componenti, che possono essere atomi o batteri in campo fisico o biologico, oppure persone, macchine o aziende in un contesto socio-economico.



Ivo De Lotto
Giovanni Meloni
Giovanni Sacchi
Claudio Arlandini
Maurizio Cremonesi
Paride Dagna
Alice Invernizzi
Paolo Ramieri



I grandi progressi ottenuti nel campo dello sviluppo e della messa a punto di tecniche numeriche avanzate costituiscono un altro importante elemento per la simulazione, l'interpretazione e il controllo di fenomeni complessi. Si pensi, per esempio, all'impiego di metodi statistici e di ottimizzazione per la distribuzione dell'energia elettrica; alle diverse tecnologie (di tipo matematico e algoritmico) disponibili per la progettazione di strumenti per l'elaborazione dell'immagine in vista di una sua applicazione nei campi della biometria e dell'*imaging* medica; allo sviluppo dei cosiddetti metodi di *analisi isogeometrica* che rappresentano un nuovo strumento per la simulazione di problemi governati da equazioni a derivate parziali. Queste tecniche consentono di rappresentare il problema fisico di partenza utilizzando direttamente le approssimazioni CAD (*Computer Aided Design*) senza passare attraverso la discretizzazione FEM (*Finite Element Method*), riducendo notevolmente i tempi di progettazione e migliorando il rapporto accuratezza-tempo computazionale.

È infine necessario citare i *metodi di decomposizione di dominio*. Essi consentono di presentare la soluzione di un problema computazionale come successione di soluzioni di sottoproblemi, in generale più semplici da trattare e con un buon livello di indipendenza tra loro. Questa caratteristica rende i *metodi di decomposizione di dominio* particolarmente interessanti per applicazioni in ambienti di calcolo di tipo *High Performance Computing* (HPC) e *Grid*: infatti permettono di implementare, in modo naturale, sia un *parallelismo matematico* (a grana grossa) a livello della soluzione dei vari sottoproblemi sia un *parallelismo algoritmico* (a grana fine) a livello della risoluzione del singolo sottoproblema.

Un contributo importante per una descrizione e un controllo efficaci dell'evoluzione di sistemi e di processi di alta complessità è l'aumento costante della potenza di calcolo che le attuali architetture per l'HPC e il *Grid Computing* mettono a disposizione di una sempre più vasta e meno specializzata fascia di utenti. I sistemi HPC e *Grid* forniscono un supporto efficiente per lo sviluppo di applicazioni *compute-intensive* e *data-intensive*

in molteplici aree applicative scientifiche, industriali, economiche e commerciali. Tuttavia, per trarre beneficio dalla grande potenza di calcolo e dalle capacità offerte è necessario, da un lato, avere a disposizione strumenti (modelli, linguaggi ecc.) di programmazione in grado di permettere alle applicazioni di sfruttare al massimo le potenzialità di questi sistemi; dall'altro, è necessario disporre di tecniche di simulazione e di progettazione efficienti, affidabili e robuste.

Le previsioni meteorologiche per una data area geografica, la diffusione di un'infezione all'interno di una popolazione, la stabilità di una nave da crociera in funzione delle condizioni del mare, il disegno di nuovi materiali, l'analisi dei grandi *datasets* provenienti dai telescopi, l'analisi e l'interpretazione di immagini biomediche sono alcuni esempi di sistemi complessi il cui comportamento viene sempre più spesso studiato e descritto attraverso simulazioni al calcolatore.

Negli ultimi anni, l'approccio modellistico-simulativo si sta estendendo dalle scienze che tradizionalmente si basano sulla rappresentazione di problemi complessi tramite modelli (fisica, chimica, ingegneria), a molti altri ambiti scientifici, quali l'economia, la biologia, la medicina, la linguistica, la sociologia e, in generale, le scienze umanistiche. Inoltre, l'aumento della potenza di calcolo e il progresso nelle tecniche di simulazione permettono di affrontare problemi di crescente complessità, portando la simulazione a diventare in alcuni casi un complemento dell'esperimento, in altri casi un sostituto dell'esperimento tradizionale (per esempio perché troppo costoso o perché su una scala temporale troppo lunga per essere osservata), in altri casi ancora ad essere l'unica rappresentazione possibile di test non ripetibili. Si pensi, ad esempio, alla previsione degli effetti di un terremoto o alla simulazione della diffusione di sostanze inquinanti al fine di valutarne le conseguenze.

La continua evoluzione della potenza di calcolo disponibile permette ai ricercatori di affrontare problemi sempre più complessi, praticamente non pensabili nel passato perché richiedenti tempi per la soluzione assolutamente impossibili da gestire.

Assistiamo ad una continua rincorsa tra l'evo-

luzione delle architetture di calcolo e il progresso della ricerca nell'analisi e nella simulazione di fenomeni sempre più complessi.

Vengono quindi sviluppati e adattati metodi e algoritmi per la costruzione di potenti codici di simulazione in cui il tempo occorrente per la simulazione cresce linearmente con la complessità del modello da analizzare, per esempio con il numero di atomi componenti una molecola. Più potenza il ricercatore ha a disposizione, più è in grado di aggiungere atomi alla molecola di cui deve studiare le caratteristiche e le proprietà.

Gli stessi programmi di origine commerciale, nati dalla ricerca nel mondo accademico e poi ingegnerizzati per permetterne l'uso in modo efficiente in ambito industriale, hanno subito negli anni una profonda evoluzione. Si sono dovuti adattare ai cambiamenti delle architetture di calcolo: seriali, vettoriali, parallele distribuite e a memoria condivisa; tutto questo senza perdere di efficienza e di correttezza dei risultati numerici, introducendo nuove funzionalità e algoritmi, estendendo le possibilità di interazione con parti di codice sviluppato dall'utente, migliorando le prestazioni nei tempi di soluzione, elevando il grado di scalabilità in ambiente parallelo, inserendo interfacce utente sempre più duttili e facili da usare nella gestione dei dati di ingresso e di uscita.

Si sono sviluppati in questi anni anche linguaggi avanzati per la costruzione di codici di calcolo dotati di molte funzionalità per lo sviluppo di algoritmi, per la gestione dei dati, per l'esecuzione della simulazione, in grado di essere applicati ad un ampio spettro di applicazioni. Dotati di interfacce utente molto semplici da usare ma nello stesso tempo articolate e sofisticate, permettono un'efficiente gestione della visualizzazione dei dati di ingresso e di uscita e della costruzione del codice da eseguire. Nati per ambienti come PC e Workstation si sono via via evoluti in modo da sfruttare sia le nuove architetture parallele sia gli ambienti operativi di *grid* e *cloud computing*. Permettono quindi di trasformare le applicazioni in ambiente parallelo con minimi interventi di riadattamento mediante l'impiego di moduli e direttive specializzate.

Nei prossimi anni l'evoluzione degli strumenti

di calcolo verso architetture complesse dall'alto grado di parallelismo richiederà grossi sforzi per sviluppare nuove tecniche di gestione degli ambienti di calcolo in maniera efficiente, nella costruzione di programmi robusti in grado di sostenersi e di continuare l'esecuzione anche in caso di cedimento di alcuni dei componenti *hardware* sui quali sono in esecuzione (per esempio programmi che utilizzeranno centinaia di processori), nel pensare nuovi problemi di simulazione che chiederanno lo sforzo congiunto di gruppi di ricerca formati da numerosi esperti, geograficamente distribuiti, ma ben coordinati tra di loro.

2. ALCUNI ESEMPI DI APPLICAZIONI

I processi di simulazione nella maggior parte dei casi sono condotti al calcolatore numerico ed è a questi che fanno riferimento i contributi raccolti in questo volume di Mondo Digitale.

A livello propedeutico si è ritenuto opportuno presentare in un precedente numero di Mondo Digitale una valutazione filosofico - epistemologica della simulazione numerica [1] come strumento di indagine del comportamento di sistemi dinamici reali, visti la particolare natura del processo di simulazione e il ruolo dei modelli e della simulazione nello stesso processo conoscitivo e decisionale dell'uomo.

In questo numero, dato lo sviluppo inarrestabile delle tecnologie digitali, degli strumenti di calcolo e degli ambienti per il loro uso offerti all'utenza, si parlerà di simulazioni fatte attraverso l'utilizzo di sistemi di supercalcolo, per esempio quelli offerti dal consorzio CILEA.

In questa introduzione si presentano soprattutto gli sviluppi in corso per gli ambienti di calcolo, con riferimento alle tecnologie disponibili presso i grandi centri di supercalcolo e alle loro evoluzioni, quali quelle cosiddette *petascale* o addirittura *exascale*, del *green computing*, del *quantum computing* e dell'utilizzo delle tecnologie fotoniche, all'ecosistema dei centri di supercalcolo con riferimento al *cloud computing*, all'evoluzione delle tecnologie di rete, con particolare riferimento alla gestione di grandi quantità di dati e a internet 3D, e infine ai paradigmi di programmazione.

Come esempi applicativi si richiameranno in questo volume un'applicazione all'industria



automobilistica, un'applicazione nel progetto di circuiti integrati, la simulazione degli effetti di un terremoto in una valle alpina e sull'attuale città di Messina, l'applicazione alla fluidodinamica nella progettazione di barche a vela da regata e nello studio della circolazione sanguigna, la chimica computazionale, le nuove sfide della bioinformatica, le simulazioni meteorologiche per lo studio della qualità dell'aria e infine la finanza computazionale. Questi esempi non sono che un modesto campionamento dei molteplici settori dove la simulazione numerica è oggi proficuamente utilizzata, significativo tuttavia per farne comprendere l'importanza come strumento d'indagine e di progettazione.

3. L'HARDWARE COME MOTORE DELLA SIMULAZIONE

La qualità della simulazione cresce negli anni assieme alla potenza di calcolo disponibile. Quest'ultima ha seguito negli ultimi decenni un andamento conosciuto come prima legge di Moore. La legge è tratta da un'osservazione empirica di Gordon Moore, cofondatore di Intel con Robert Noyce: nel 1965 Gordon Moore, che all'epoca era a capo del settore R&D della Fairchild Semiconductor e tre anni dopo fondò la Intel, scrisse infatti un articolo su una rivista specializzata [2] nel quale illustrava come il numero di *transistors* che possono essere inseriti in maniera economicamente ragionevole su un circuito integrato raddoppia approssimativamente ogni due anni. Viene spesso enunciata in maniera non corretta come "raddoppia ogni 18 mesi", seguendo una citazione di David House dell'Intel. In realtà, attualmente il tempo di raddoppio è di circa 20 mesi. Poiché al diminuire delle dimensioni dei *transistor* la velocità alla quale possono operare aumenta e il numero di *transistor* è una misura approssimativa delle prestazioni computazionali, è prassi comune citare la legge di Moore per riferirsi all'evoluzione della potenza di calcolo per unità di costo.

La legge di Moore ha mantenuto la sua validità attraverso diversi cambiamenti tecnologici epocali, quali il passaggio alla tecnologia CMOS, tanto da aprire una discussione sul fatto che non sia in realtà una profezia autoavverante, perché ormai le industrie di semi-

conduttori usano la legge per stabilire la pianificazione e gli obiettivi a lungo termine di ricerca e sviluppo. L'ultimo cambiamento lo stiamo avvertendo ora, con l'avvento della tecnologia *multicore*. Per anni i produttori di microprocessori hanno ampliato la potenza di calcolo delle CPU, sfruttando l'aumento del numero di *transistor* per incrementare la frequenza di clock e il parallelismo a livello di istruzioni, con il vantaggio di avere codice seriale eseguito più velocemente su nuovi processori senza alcuna modifica. Poiché il calore dissipato è essenzialmente proporzionale al quadrato della frequenza di clock, mantenere immutato lo stesso andamento tecnologico avrebbe significato che nel giro di pochissimi anni la densità di calore sul chip, che ha ampiamente superato quella di un fornello elettrico, avrebbe raggiunto quella del *core* di un reattore nucleare. Non potendo ovviamente gestire in maniera semplice ed efficace il raffreddamento dell'oggetto, si è gioco-forza optato per un diverso approccio: mantenere la frequenza pressoché costante e approfittare dei progressi nella miniaturizzazione dei *gate* logici per collocare più *core* computazionali sullo stesso chip. In questo modo il mantenimento della legge di Moore è garantito per qualche anno in più, anche se è necessario riscrivere il software in maniera *multi-threaded* o multi-processo per consentire il pieno sfruttamento dell'hardware.

Il *laptop* che avete sul vostro tavolo ha già con tutta probabilità un processore *quad-core*. Intel ha appena lanciato sul mercato la nuova CPU denominata Westmere a tecnologia 32 nm a 6 *core* [3]. AMD, che aveva CPU a 6 *core* già dal 2009, lancerà tra poche settimane la nuova CPU denominata Magny-Cours a 8 e 12 *core* [4]. La guerra dei *core* impazza. Intel ha già inviato un prototipo di CPU a 48 *core* (progetto Barrelfish [5]) a università e centri di ricerca nel mondo per riceverne un *feedback* e trasformarlo in un prodotto. Allo stesso tempo fornisce alla ricerca una piattaforma sulla quale cominciare a lavorare per definire il software del futuro.

Le cose poi si complicano ulteriormente con l'introduzione di architetture eterogenee a livello di nodo, che sono diventate attrattive negli ultimi anni per diverse ragioni: rispetto alla sola CPU tradizionale offrono un'alta pre-

stazione di picco e hanno un'elevata efficienza energetica e/o un buon rapporto prezzo/prestazioni. Le tecnologie più comuni di questo tipo affiancano ad una CPU *general purpose* una o più unità di tipo *Cell Broadband Engine Architecture*, *General Purpose Graphics Processing Units* (GPGPU), o *Field Programmable Gate Arrays* (FPGA). Una panoramica aggiornata e approfondita sullo stato dell'arte ed una visione sul futuro del calcolo eterogeneo si può trovare nell'articolo di Brodtkorb et al [6].

La miniaturizzazione dei gate logici della tecnologia CMOS ha ancora margini di incremento per qualche anno, prima di arrivare al prossimo limite tecnologico. Se uno dei problemi, il *leakage* della corrente di gate, sembra essere ridotto usando ossido di afnio invece di ossido di silicio, molto più minacciosi appaiono quelli legati alla natura quantica della materia, con il *quantum-tunneling* degli elettroni tra circuiti adiacenti che diventa tanto più probabile quanto più la distanza tra due *transistor* si riduce. 11 nm potrebbero essere il limite invalicabile, facilmente raggiungibile in una decina d'anni [7].

Se le attuali tecnologie arriveranno ai propri limiti nei prossimi 10 anni, quali ne prenderanno verosimilmente il posto?

Nel 1982 Richard Feynman suggerì l'idea di calcolatori quantici [8]. L'idea di base è di utilizzare al posto dei *bits* dei calcolatori odierni dei *quantum bits*, detti *qubits*, implementati usando sistemi quantomeccanici a due stati, ovvero sistemi non confinati ai due stati base 0 e 1, ma che possono assumere anche sovrapposizioni degli stessi, cioè essere contemporaneamente in entrambi. L'idea si può espandere: un registro composto da n *qubits* può contenere contemporaneamente 2^n valori e su di esso si possono eseguire 2^n operazioni contemporanee.[9] I problemi di costruire calcolatori quantici sono squisitamente ingegneristici. Più *qubits* interagenti sono coinvolti, più difficile è ingegnerizzare l'interazione che mostri il risultato dell'operazione, perché aumenta la probabilità che l'informazione quantica contenuta sul componente di calcolo interferisca con l'ambiente in cui è immerso. Questo processo è detto "decoerenza". È necessario quindi costruire sistemi sub-microscopici in cui *qubits* interagiscano tra loro, ma non con

l'ambiente. Ottime introduzioni si possono trovare su questi articoli precedentemente pubblicati dalla rivista [10, 11].

L'entusiasmo per il quantum computing esploso qualche anno fa sembra essere largamente scemato. Le grandi case, HP e IBM in testa, che avevano investito notevoli cifre nel settore pochi anni fa, paiono avere chiuso i rispettivi laboratori con poco clamore, segno che i tempi per la maturità di questa tecnologia sono ancora lontani, una ventina d'anni almeno. I progressi comunque ci sono: un gruppo di ricerca americano [12] ha appena pubblicato di aver calcolato lo spettro di energia della molecola più semplice, quella di idrogeno, a 20 *bits* di precisione usando un sistema quanto-fotonico a 2 *qubits*. Questo rappresenta, di fatto, la nascita della chimica computazionale con questa tecnologia di calcolatori.

C'è molto clamore anche sul cosiddetto *DNA-computing*. Leonard Adleman nel 1994 infatti dimostrò come si poteva risolvere il problema matematico del percorso Hamiltoniano tra 7 punti usando del DNA.[13] Da allora varie macchine di Turing sono state realizzate con questa tecnologia. Per alcuni problemi specializzati calcolatori a DNA saranno certamente più piccoli e più veloci di sistemi tradizionali. Tuttavia, sussistono serie limitazioni alla possibilità di costruire calcolatori "general purpose" in questo modo. Rimandiamo il lettore all'ottima panoramica di Giancarlo Mauri sullo scorso numero della rivista [14].

Se quindi ragionevolmente tra dieci anni non vedremo *qubits* o fiale di enzimi a bordo delle *workstations*, ci sono comunque un paio di nuove tecnologie che daranno un forte impulso alla simulazione.

La prima è quella del "memristore". È dagli anni '70 che in linea teorica si sa che oltre ai tre elementi passivi fondamentali dell'elettronica, resistori, condensatori e induttori, dovrebbe esistere un quarto, un resistore di memoria, o *memristore* [15]. Ma solo nel 2008 i ricercatori degli HP Labs sono riusciti a realizzarne uno [16]. Le proprietà dell'oggetto sono di mantenere il proprio ultimo livello di resistività quando la tensione viene ridotta a zero e conservarlo indefinitamente fino a quando la tensione è riattivata. Richiedono quindi molta meno energia per funzionare delle memorie attuali, e secondo le previsioni entro 3-

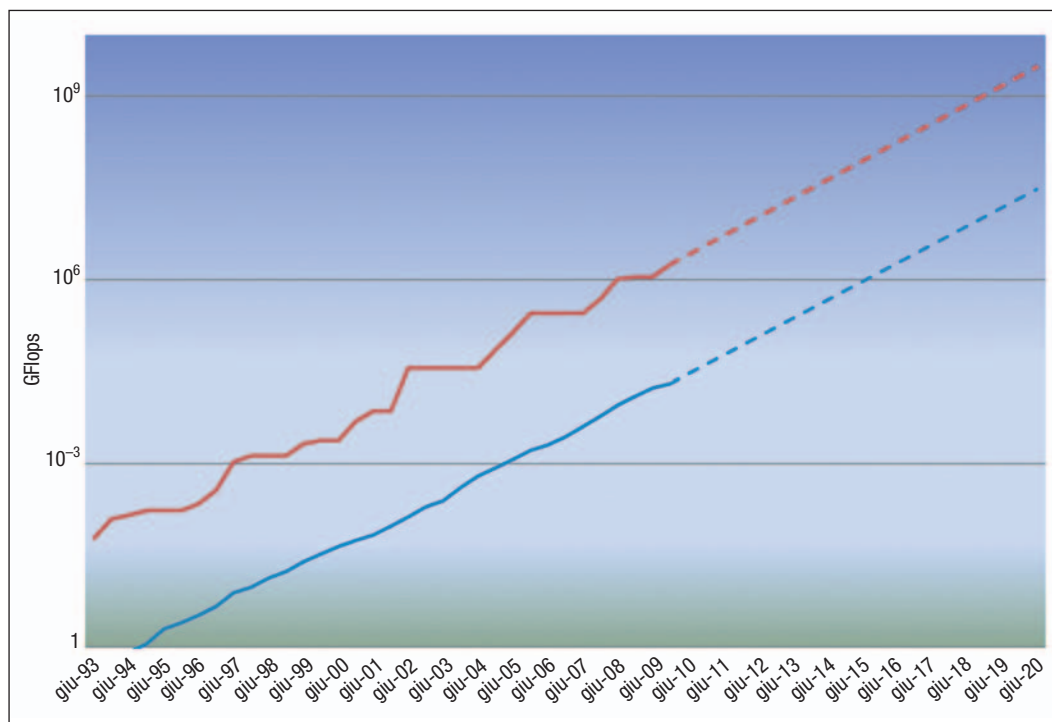


FIGURA 1

L'evoluzione della potenza sostenuta di calcolo nella classifica TOP500 a partire dalla sua prima edizione (giugno 1993) fino a novembre 2009, con un'estrapolazione degli andamenti fino al 2020. In rosso i valori corrispondenti alla posizione n. 1 e in blu quelli relativi alla n. 500

5 anni saranno disponibili sul mercato. Non solo ci permetteranno di avere a disposizione RAM di un ordine di grandezza superiore a quelle attualmente disponibili, con densità di 20 GB/cm² già entro il 2013, ma saranno sufficientemente veloci da consentire di creare processori che vedano accoppiate unità computazionali e memorie di grandi capacità. E questo potrebbe risultare veramente rivoluzionario per quanto riguarda le prestazioni dei codici applicativi.

La seconda è quella della "fotonica basata sul silicio", un settore nel quale IBM e Intel stanno investendo moltissimo. Lo scopo è quello di sviluppare un insieme di tecniche e tecnologie ottiche capaci di andare progressivamente a sostituire le normali connessioni in rame all'interno di un calcolatore e, ancora più oltre, all'interno dei chip e dei microprocessori. L'obiettivo è duplice: minori consumi perché usando fotoni invece di elettroni per trasportare informazione si riduce l'energia dissipata, e, di conseguenza, una maggior integrazione. Il grande vantaggio di questa tecnologia è che richiede un semplice adattamento delle

procedure usate per costruire circuiti CMOS [17]. Il primo laser a germanio, capace di produrre luce a lunghezze d'onda utilizzabili per la comunicazione ottica è stato annunciato quest'anno dal MIT [18].

In sostanza, la previsione che la prima legge di Moore rimanga valida per i prossimi 10 anni appare più che fondata. Come si traduce questo in termini di potenza di calcolo a disposizione di un ricercatore? Per saperlo non ci resta che estrapolare i trend della classifica TOP500 [19].

Quello che ci aspettiamo (Figura 1) è quindi che nel 2020 il calcolatore più potente al mondo abbia una potenza sostenuta di 3 *exaflop/s* (10¹⁸ operazioni in virgola mobile al secondo) e che esistano almeno 500 sistemi con una potenza superiore a 30 *petaflop/s* (10¹⁵ operazioni in virgola mobile al secondo). Giusto per dare un confronto, l'attuale numero 1, *Jaguar dell'Oak Ridge National Laboratory*, supera di poco il *petaflop*. Questo apre notevoli possibilità, ma l'hardware è naturalmente solo uno degli aspetti del problema della simulazione del futuro.

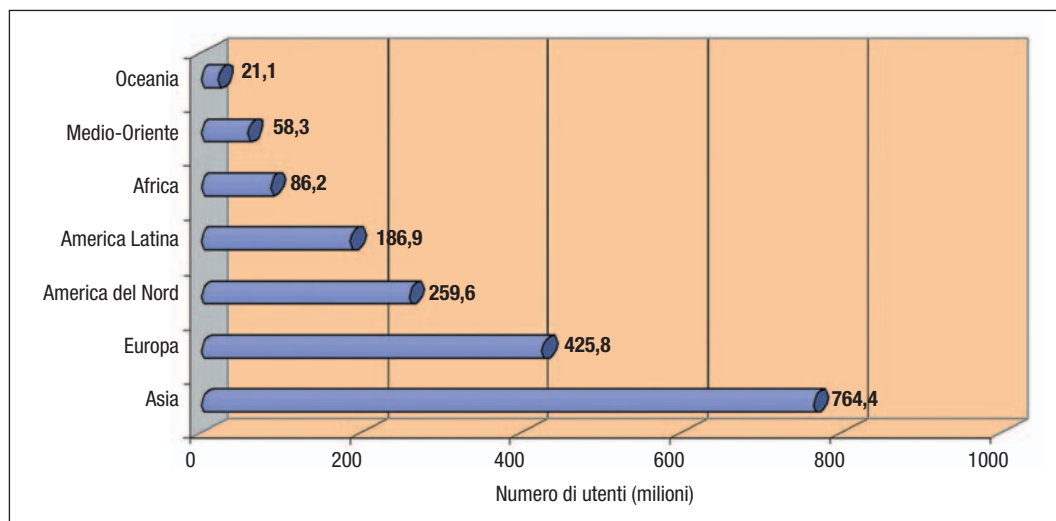


FIGURA 2
Distribuzione
utenti Internet
nel mondo (2009)

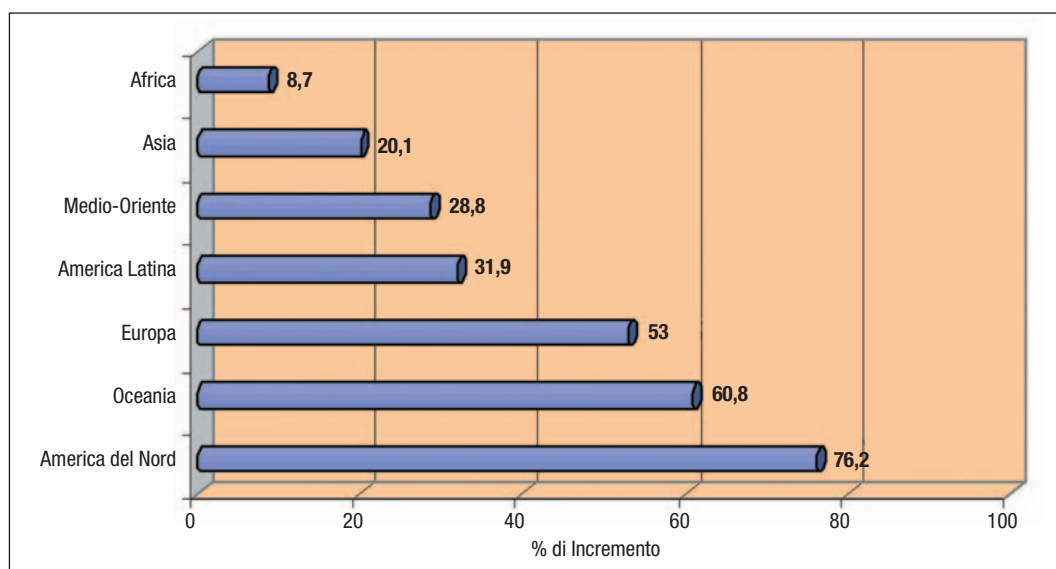


FIGURA 3
Tasso di incremento
del numero di utenti
tra il 2008 e il 2009

4. L'INFRASTRUTTURA DI RETE PER LA SIMULAZIONE DEL FUTURO

Si deve al Web tradizionale l'innescare della diffusione e del successo di Internet, intesa come rete di comunicazione basata su protocollo IP, come nuova frontiera della comunicazione e dello scambio di dati. Osservando le statistiche di utilizzo ci si accorge che Internet ha conosciuto negli anni una crescita di tipo esponenziale. I fattori che ne hanno favorito l'incremento sono molteplici. In primo luogo il numero di utenti, che ad oggi, ha quasi raggiunto i due miliardi di individui [20] come mostrato nella figura 2 e nella figura 3. In secondo luogo il traffico del singolo utilizzatore medio, che è cresciuto grazie anche alla moltitudine di ap-

plicazioni attualmente disponibili, quali l'evoluzione dei *social network*, dell'*e-commerce*, del *p2p file sharing*, e del *semantic web*. Per sostenere la crescente domanda di utilizzo, crescono di conseguenza le infrastrutture di rete sottostanti in termini di banda passante mentre la latenza *point-to-point* si abbassa. La ricerca scientifica non può che beneficiare dell'evoluzione crescente dell'infrastruttura di rete. Scienza ed evoluzione tecnologica vanno avanti di pari passo: sono spesso le esigenze della ricerca scientifica di base a fare da traino e da lancio per importanti innovazioni tecnologiche. Ricordiamo infatti che è proprio al CERN che agli inizi degli anni '90 videro la luce il primo server e il primo browser di pagine html della storia.

Il progresso del calcolo su larga scala è strettamente legato all'evoluzione delle tecnologie di rete. Sarebbe, infatti, impensabile la distribuzione di una grande mole di dati da un punto all'altro del Pianeta senza un'infrastruttura veloce e affidabile. Pensiamo per esempio agli esperimenti del Large Hadron Collider del CERN che producono più di 15 *petabyte* di dati all'anno, che devono essere distribuiti ed elaborati in centri ed università sparse in più parti del mondo.

L'infrastruttura di rete italiana ha ottenuto sviluppi notevoli grazie al Consortium GARR che gestisce e sviluppa la rete informatica della ricerca italiana. L'attuale rete GARR-X, basata su fibra ottica, è dotata di capacità 40 volte maggiori della precedente GARR-G e mette in condizioni di poter sfruttare il protocollo IPv6 *multicast*.

A fronte della crescita dell'infrastruttura di rete che si ha a disposizione si aprono nuove prospettive di utilizzo che offrono potenzialità inusitate per lo scienziato che si occupa di simulazione: ad esempio il passaggio dall'internet delle persone all'internet degli oggetti, e il passaggio dall'internet 2D all'internet 3D.

Internet sta penetrando sempre di più nella vita quotidiana, entrando negli oggetti di uso comune. Una delle sfide tecnologiche alla base dell'internet del futuro è rappresentata proprio dallo sviluppo di una "internet degli oggetti", che permetterà ad oggetti a distanza di comunicare tra loro. L'introduzione della tecnologia RFID (*Radio Frequency Identification*), che garantisce un meccanismo di identificazione univoco, viene considerato il punto di partenza per lo sviluppo dell'internet delle cose. Questa tecnologia è già adottata in diversi settori industriali e permette agli oggetti di dialogare con l'utilizzatore e tra loro: nella grande distribuzione e nei punti vendita, nell'identificazione ospedaliera dei pazienti, nella fruizione di servizi nel trasporto pubblico, per citare qualche esempio. L'internet del futuro sta cercando di vincere una nuova sfida: mettere in rete milioni di oggetti in grado di comunicare tra loro a distanza, connessi ad una Rete Globale. La tecnologia RFID combinata con un indirizzo IP, assegnato a ciascun oggetto, consentirà a oggetti distanti di poter parlare. Con l'internet degli oggetti viene aggiunta una nuova dimensione al mondo della comunica-

zione e dell'informazione: dalla connessione in ogni luogo e ad ogni momento, si aggiunge la connessione ad ogni cosa.

Il paradigma del *Pervasive Computing* è strettamente associato al concetto di Internet delle cose. Il termine sta ad indicare la diffusione e il coordinamento di piccoli dispositivi di calcolo e di comunicazione all'interno dell'ambiente fisico ed è contrapposto al concetto di realtà virtuale. La realtà virtuale inserisce idealmente le persone in un mondo virtuale, il *pervasive computing* all'opposto è integrato nel mondo fisico. Una panoramica più dettagliata sull'argomento si trova in questo articolo della rivista [21]. Con la prospettiva di un ambiente totalmente connesso e di una capacità di calcolo distribuita ovunque, molti sono i pre-requisiti tecnologici affinché questo paradigma diventi realtà, grazie all'evoluzione delle tecnologie RFID in veri e propri *smart object*, oggetti intelligenti in quanto dotati di microprocessori, memoria e *device* per la comunicazione Wireless.

L'evoluzione del modo di lavorare dello scienziato che si occupa di simulazione è evidente: scomparirà, o tenderà a scomparire la differenza tra il ricercatore che lavora in vitro e quello che lavora in silico. Strumenti di misura o di laboratorio comunicheranno in maniera diretta con il calcolatore consentendo indagini al momento impensabili. Allo stesso modo nasceranno nuove figure trasversali e interdisciplinari di scienziato.

Scaturiscono, tuttavia, numerose implicazioni, non solo a livello tecnologico, dal tentativo di applicare questo nuovo paradigma. La standardizzazione delle tecnologie, che in alcuni settori risulta ancora troppo frammentata, è il primo obiettivo da raggiungere per consentire alla più vasta quantità di oggetti di avere accesso alla rete. Sul fronte sicurezza e privacy, lo sviluppo dell'internet degli oggetti deve garantire il rispetto per la vita privata e la protezione dei dati personali, nonché la libertà del singolo di accedere alla rete con implicazioni evidenti sulle normative inerenti la sicurezza dell'informazione e sulla gestione dell'identità digitale di cose e persone.

L'internet degli oggetti non è l'unica frontiera che si prospetta sul lato dell'evoluzione delle tecnologie internet. Con 3D Internet si intendono le tecnologie 3D utilizzate per la visualiz-

zazione delle informazioni su Web. Secondo Sean Koehl, esperto dell'Intel Lab, la tecnologia sta evolvendo al punto tale che le applicazioni 3D diventeranno parte integrante del Web nei prossimi 5-10 anni. Il 3D Internet consente una navigazione interattiva nel Web garantendo la gestione di un mondo virtuale più complesso della semplice consultazione di riviste, audio, video e animazioni disponibili con il 2D Internet.

Il 3D Internet unisce a informazione e interazione, già presenti nel Web tradizionale, un ambiente 3D immersivo. La struttura tecnologica del web tradizionale non risulta però adeguata a coprire le esigenze del 3D Internet. Gli oggetti virtuali non sono più semplici documenti statici, ma sono automi in grado di rispondere a determinati input, programmi che eseguono specifiche operazioni: con il 3D Internet appare evidente la necessità di un ambiente intelligente. La latenza delle reti, inoltre, è un problema cruciale in un mondo virtuale dove l'interattività è un requisito essenziale. Le reti vanno pertanto riprogettate per garantire tempi di risposta ridotti tra server e client.

Quali sono le implicazioni che seguono l'evoluzione delle tecnologie Internet nell'ambito della simulazione scientifica? Second Life ha trovato applicazioni anche nell'ambito della chimica computazionale [22]. La tecnologia alla base di Second Life garantisce la possibilità di visualizzazione di dati 3D, di simulazioni di reazioni chimiche e la facilità di *scripting* rendendola uno strumento appetibile anche per applicazioni in ambito scientifico. A questo si aggiungono le funzionalità del *Social Network* che permettono di creare una vera e propria comunità scientifica, garantendo la collaborazione su progetti comuni.

Questo non è l'unico esempio di applicazioni del 3D Internet in ambito scientifico. ScienceSim [23], presentato al SCo9, è una piattaforma per la visualizzazione di dati 3D via Web basata su tecnologia *opensource* OpenSim pensato per creare una comunità scientifica virtuale.

La visualizzazione scientifica (e non solo) può trarre beneficio dall'introduzione di una nuova piattaforma per la visualizzazione via Web. La piattaforma Nvidia RealityServer coniuga la potenza dell'hardware Nvidia Tesla con software per servizi 3D via Web. Reality Ser-

ver si appoggia ad un ambiente GPGPU-based accessibile via Web tramite PC, *netbook*, *smartphone* garantendo interattività e visualizzazione di dati 3D. Con questa soluzione le capacità elaborative su lato client sono ridotte al minimo. La manipolazione e il processamento dei dati avvengono su lato server, al client vengono inviati unicamente dati video compressi. La larghezza di banda minima richiesta rimane pressoché costante indipendentemente dalla complessità del dato 3D. In questo prodotto le tecnologie di visualizzazione 3D convergono con la necessità di un ambiente HPC in grado di garantire scalabilità, parallelismo e performance per la manipolazione di dati 3D via Web.

5. VERSO UN ECOSISTEMA DEL SUPERCALCOLO EUROPEO

Nel 2001 il Consiglio dei Ministri Europeo ha invitato la Commissione Europea, in collaborazione con gli Stati Membri, a esplorare nuovi approcci per affrontare i problemi relativi alle infrastrutture per la ricerca scientifica.

Questa iniziativa ha portato alla creazione di ESFRI (*European Strategy Forum on Research Infrastructures*), costituito nel 2002, con il mandato di supportare un piano strategico per lo sviluppo coerente delle infrastrutture per la ricerca europea e incoraggiare iniziative per un utilizzo più efficiente delle stesse. ESFRI è costituito da rappresentanti degli stati membri della Comunità Europea, dei Paesi associati e della Commissione Europea. Soltanto il sesto programma quadro, attivo negli anni 2002-2006 aveva un finanziamento di 18 miliardi di Euro, mentre per gli anni 2007-2013 il settimo programma quadro mette a disposizione della ricerca scientifica in Europa più di 50 miliardi di Euro. Queste risorse sono destinate alla realizzazione di diverse infrastrutture scientificamente rilevanti, che assumono in questi anni di crisi economica un importante ruolo di sostegno della capacità di innovazione tecnologica dei Paesi europei. Lo scopo non è solo di sostenere l'economia in settori tecnologicamente avanzati, ma anche di preparare un ambiente di ricerca che sia un polo attrattivo per le nuove generazioni di scienziati [24].

Limitandoci al contesto del calcolo ad alte prestazioni per la ricerca scientifica è particolar-

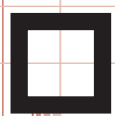


TABELLA 1
L'elenco dei Paesi
aderenti all'iniziativa
PRACE

Paesi aderenti all'iniziativa PRACE
Austria
Bulgaria
Cipro
Finlandia
Francia - partner principale
Germania - partner principale
Grecia
Irlanda
Italia - partner principale
Norvegia
Paesi Bassi - partner principale
Polonia
Portogallo
Regno Unito - partner principale
Repubblica Ceca
Serbia
Spagna - partner principale
Svezia
Svizzera
Turchia

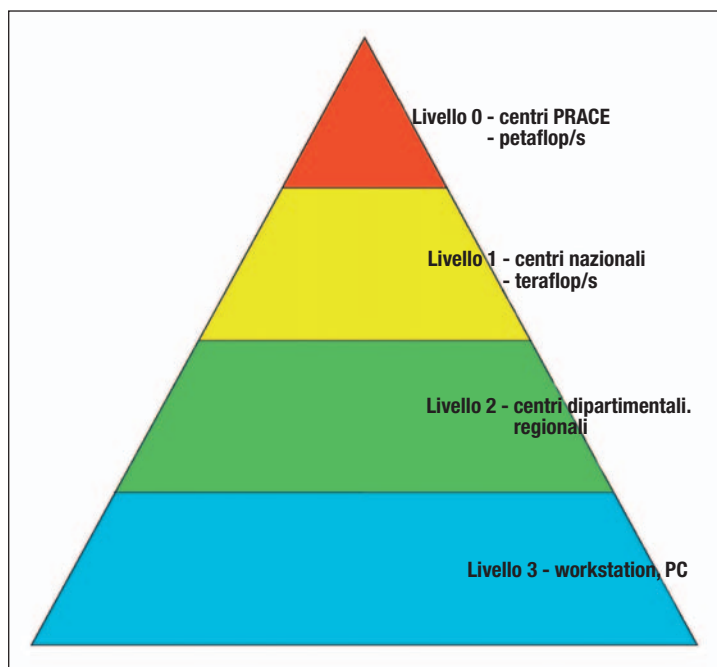


FIGURA 4
Schematizzazione della piramide del supercalcolo europeo

mente rilevante l'iniziativa PRACE (*Partnership for Advanced Computing in Europe*), nata per volontà di 14 Paesi europei tra cui l'Italia. Questa iniziativa è sostenuta anche da un progetto, finanziato da ESFRI per gli anni 2008-2009 con 10 milioni di Euro. L'obiettivo è la costruzione di un'infrastruttura di calcolo ad alte prestazioni permanente con capacità di calcolo allo stato dell'arte dei massimi livelli mondiali, ma soprattutto si vuole creare un'organizzazione che coinvolga i centri e le comunità di ricerca nazionali, per garantire la sostenibilità dell'infrastruttura e lo sviluppo delle conoscenze necessarie per poter usufruire di questa importante realizzazione nei modi più efficaci. Altri Paesi si sono poi aggregati all'iniziativa PRACE; nella tabella 1 sono elencati i 20 Paesi aderenti, compresi quelli che hanno espresso l'intenzione di ospitare e gestire gli elaboratori di potenza più elevata (partner principali).

Il concetto di ecosistema del supercalcolo è stato introdotto nel 2007 dall'*High Performance Computing in Europe Taskforce* (HET), che ha avuto il compito, come attività iniziale del progetto PRACE, di individuare una strategia per sviluppare infrastruttura e servizi di calcolo per la scienza in Europa, con un'enfasi particolare nella creazione e nel mantenimento di centri di calcolo con potenze dell'ordine del *petaflop*. Questa impresa più che rappresentare una sfida tecnologica e organizzativa è una necessità per mantenere la ricerca scientifica europea ai massimi livelli mondiali. Il rapporto della *taskforce* HET presenta infatti una lista piuttosto nutrita di grandi sfide [25], ovvero problemi scientifici che necessitano di risorse di calcolo a partire dal *petaflop* per poter essere affrontati con successo.

Dal lavoro della *taskforce* europea per il supercalcolo nasce l'idea di organizzare l'ecosistema del supercalcolo europeo come "piramide del calcolo". Interpretando il concetto di ecosistema in senso ampio come l'insieme delle figure professionali coinvolte, delle conoscenze scientifiche e tecnologiche oltre che degli strumenti tecnologici utilizzati, l'infrastruttura del supercalcolo europeo può essere schematizzata come una piramide a 4 livelli, legati da una rete di professioni, conoscenze, tecnologia e comunicazioni (Figura 4).

Il livello 3, quello più basso, è il più popolato e

vicino alla maggioranza dei ricercatori. Esso è costituito soprattutto dagli utenti finali, ricercatori e studenti e dagli strumenti di calcolo più comuni ovvero terminali, *workstation* e personal computer. Il livello appena successivo è popolato da ricercatori e studenti che usano le risorse di calcolo dei centri locali, come quelli di ateneo o i centri regionali minori.

Il livello 1 è costituito da ricercatori e studenti che per le loro indagini necessitano di strumenti di calcolo con elevate capacità computazionali e di grandi capacità di memorizzazione dei dati e quindi usano i calcolatori disponibili presso i centri di calcolo regionali o nazionali. Afferiscono a questo livello personale altamente specializzato, produttori di hardware e sviluppatori di codici di calcolo parallelo. I centri di calcolo di livello 1 hanno, infatti, nell'ambito dell'ecosistema, l'importante compito di preparare l'utenza all'uso delle risorse di calcolo più prestigiose afferenti al livello 0, l'ultimo e il più esclusivo della piramide del supercalcolo europeo.

Gran parte dell'attività svolta dall'iniziativa PRACE è rivolta alla definizione e integrazione dei centri di supercalcolo di livello 0, che ospiteranno i calcolatori con le prestazioni più elevate, dell'ordine dei *petaflops*. A questo livello della piramide afferisce personale fortemente specializzato nella gestione di elaboratori dalle caratteristiche molto particolari, che dovrebbero essere in grado di fornire risposte a diverse tra le sfide più ardue che l'umanità abbia mai affrontato. L'accesso a questi elaboratori sarà inevitabilmente molto selettivo e ristretto ai gruppi di ricerca che dimostreranno di saper sviluppare e utilizzare programmi di calcolo adatti e poter gestire le ingenti quantità di dati prodotti. Le preziose risorse di calcolo di livello 0 saranno gestite in modo da formare un'unica risorsa disponibile per la ricerca da qualunque Paese europeo. Questo progetto è particolarmente ambizioso in quanto le risorse di calcolo di livello 0 devono essere rinnovate ogni 2-3 anni per mantenere le prestazioni ai massimi livelli mondiali. Il costo per la loro realizzazione è stimato nell'ordine dei 200-400 milioni di Euro, da rinnovare ogni triennio circa, mentre il costo di mantenimento annuale si prevede intorno ai 100-200 milioni.

È facile capire perché l'infrastruttura del calcolo europeo è organizzata a livelli se si tiene

conto che diversi problemi scientifici non hanno bisogno delle risorse di calcolo con le prestazioni più elevate, che un singolo centro di calcolo non può fornire servizi in modo efficiente a tutti i possibili utenti europei e che le risorse di maggior pregio non possono essere sprecate con programmi inadatti o per lo sviluppo dei codici.

Tuttavia la piramide del supercalcolo prevede una rete di interconnessione tra i livelli che permetta la condivisione delle risorse a livello pan-europeo. Questa rete deve intendersi sia come tecnologia capace di connettere centri di calcolo di tutta Europa, sia come insieme di accordi e iniziative che garantiscano l'accesso alle risorse, anche le più prestigiose, da parte dei ricercatori di qualunque Paese europeo.

Esempi importanti di collaborazioni simili, cui PRACE fa riferimento per la definizione dell'infrastruttura europea del supercalcolo, sono il consorzio DEISA (*Distributed European Infrastructure for Supercomputing Applications*) e GEANT2.

DEISA ha lo scopo di organizzare i centri di calcolo nazionali, con elaboratori della classe dei *teraflops* (10^{12} operazioni in virgola mobile al secondo), in un unico ambiente globale di calcolo utilizzando tecnologie *Grid*.

GEANT2 è la rete che unisce i centri per la ricerca in Europa e collega 34 nazioni e 30 reti nazionali, con tratte fino a 10 Gbit/s [26].

In prospettiva quindi i centri di calcolo che faranno parte dell'ecosistema europeo saranno federati in una griglia di calcolo globale basata sulla coordinazione dei *Grid* nazionali. Per realizzare questo risultato sarà preziosa l'esperienza del progetto EGEE (*Enabling Grid for E-science in Europe*), che attualmente comprende 250 istituzioni di 51 Paesi diversi.

La difficoltà sia economica che tecnologica di realizzare e gestire calcolatori con prestazioni dell'ordine del *petaflop* limita il numero dei centri di calcolo in grado di ospitarli a una decina circa, distribuiti tra i Paesi più sviluppati della Comunità Europea. Proprio quest'anno si conclude il lavoro organizzativo e preparatorio del progetto PRACE con l'individuazione di nove prototipi architetture che dovrebbero dimostrare come realizzare al meglio i prossimi super calcolatori europei. Può essere interessante notare che alcune di queste macchine sono costruite accoppiando i più diffusi pro-



cessori *general-purpose* con i cosiddetti acceleratori, di tipo GPGPU o FPGA [27].

Se i centri di livello 0 sono pochi e hanno l'ambizione di servire i progetti di ricerca più impegnativi di tutti gli stati europei, i centri di calcolo di livello regionale sono meglio distribuiti sul territorio, più accessibili alla maggioranza dei ricercatori. Inoltre, mentre l'hardware è fortemente localizzato, il software, l'addestramento e l'istruzione sono diffusi più capillarmente e possono essere efficacemente forniti anche dai centri di calcolo più piccoli, che assumono così un ruolo importante per formare gli esperti in grado di lavorare con i calcolatori di potenza elevata. È anche auspicabile che il software destinato ai calcolatori di punta sia sviluppato sulle macchine di più facile accessibilità, quelle disponibili presso i centri di tipo regionale o locale.

Lo sviluppo dell'ecosistema di calcolo per la ricerca in Europa dipenderà da come i protagonisti coinvolti interagiranno tra loro e da come gli organismi di coordinamento riusciranno a indirizzare le scelte tecnologiche verso soluzioni capaci di coniugare efficienza ed economie di scala. Prevedere come si evolverà questo sistema è ancora più problematico a causa del rapido sviluppo della tecnica. In particolare proprio in questi anni nuove tecnologie stanno sostituendo quelle vecchie nelle connessioni di rete, nel calcolo ad alte prestazioni, nella fornitura dei servizi di calcolo [28].

Per diversi aspetti è auspicabile che l'evoluzione dell'infrastruttura del supercalcolo si sviluppi verso soluzioni di tipo IaaS (*Infrastructure as a Service*), con servizi simili al cosiddetto *cloud computing*, un paradigma di calcolo basato su internet, sviluppato per permettere agli utilizzatori di usare risorse di calcolo, di memoria e di software a richiesta. Anziché investire in hardware e software, l'utente acquista i servizi già pronti dai fornitori, al pari di altri servizi di pubblica utilità. Questo concetto non deve essere confuso con le griglie di calcolo o *Grid*, che sono caratterizzate invece da organizzazioni virtuali in cui enti diversi mettono in condivisione le proprie risorse realizzando così una singola piattaforma con lo scopo di raggiungere un fine comune. Mentre il *Grid* si sviluppa in ambito accademico per superare i limiti dei singoli centri di calcolo, il *cloud computing* nasce in ambito privato per razionaliz-

zare l'uso delle risorse [29]. Il termine *cloud* è anche una metafora per tutto ciò che di tecnologico e organizzativo e non meglio specificato è interposto tra gli utenti e i fornitori del servizio: il paradigma concentra invece l'attenzione sulla qualità e l'affidabilità dei servizi.

La realizzazione dell'infrastruttura del calcolo europeo secondo i concetti di *cloud computing* promette di poter sfruttare nel migliore dei modi le risorse a disposizione della ricerca: ricercatori e sviluppatori potrebbero contare sulla disponibilità delle risorse informatiche come e quando ne hanno bisogno; i centri di calcolo potrebbero gestire le risorse senza preoccuparsi di riservarne per eventuali picchi di carico.

La disponibilità di servizi di calcolo ad alte prestazioni, utilizzabili come beni fruibili da chiunque, aprirà opportunità di crescita professionale a molti studenti che troveranno impiego nelle piccole e medie imprese, che potranno permettersi di accedere a risorse e servizi attualmente non ancora alla loro portata.

6. NUOVE APPLICAZIONI, NUOVI PARADIGMI

Abbattuto il muro del *petaflop*, la nuova sfida che attende il mondo del calcolo ad alte prestazioni è il raggiungimento dell'*exaflop*, limite superabile solamente attraverso innovazioni tecnologiche e concettuali assolutamente rivoluzionarie. Questo nuovo traguardo, oltre all'investimento legato al semplice incremento di prestazioni hardware, richiederà l'esplorazione di soluzioni software innovative, che permettano lo sviluppo di paradigmi di programmazione in grado di sfruttare in modo efficiente tale potenza di calcolo e caratterizzati da un livello di astrazione tale da rendere quasi invisibile al programmatore la complessità sottostante.

Come risposta concreta a tale sfida, USA, UE e governo giapponese hanno finanziato il lancio dell'*International Exascale Software Project* (IESP) [30], un progetto ambizioso la cui missione è quella di sviluppare un piano volto a determinare entro il 2020 un ambiente di calcolo standard e di alta qualità per il raggiungimento e il mantenimento di sistemi *peta/exascale*, oltre al coordinamento dello sviluppo di infrastrutture software comuni e pos-

sibilmente *open source* che permettano la massima cooperazione da parte della comunità scientifica internazionale. Al momento, gli esperti afferenti a IESP si sono occupati di mettere a punto una *roadmap* dove sono state inserite le sfide che dovranno essere affrontate nel prossimo futuro, con particolare attenzione a tre diversi fattori critici:

- raggiungere in tempi rapidi un miglioramento dell'accesso alla crescente potenza informatica a disposizione;
- superare la chiara inadeguatezza delle attuali infrastrutture software al fine di supportare il miglioramento delle performance;
- ovviare alla quasi completa mancanza di coordinamento nel mondo scientifico nell'affrontare gli enormi ostacoli tecnologici che andranno superati.

La *roadmap* è stata realizzata dopo circa un anno di studio in base ai consigli e alle raccomandazioni della comunità mondiale che si occupa di software per il calcolo scientifico ad alte prestazioni. Al suo interno sono stati delineati un insieme di obiettivi e una serie di problematiche che dovranno essere affrontate in merito ai diversi aspetti che concorreranno alla creazione di un software adatto alla gestione di un calcolatore di tipologia *exascale*, tenendo anche conto dei possibili mutamenti delle esigenze scientifiche e delle disponibilità economiche e tecnologiche, prestando particolare attenzione alla rivoluzione architeturale dei prossimi sistemi di calcolo.

L'idea è quella di raggiungere una piattaforma integrata di software scientifici, denominata *X-stack*, utilizzabile entro il 2018-2020, che abbia le seguenti specifiche:

- permettere di progettare applicativi capaci di sfruttare al massimo le risorse dei più potenti sistemi di calcolo. Basti pensare che le applicazioni con la migliore efficienza parallela attualmente disponibile sono in grado di scalare fino ad alcune decine di migliaia di *threads*, cifra del tutto inadeguata per i supercomputer progettati per l'*exascale*, dove le applicazioni dovranno essere in grado di scalare fino ad un milione di *threads*, ovvero con un salto di circa 2 ordini di grandezza; risulta pertanto evidente che la sfida tecnologico-scientifica e concettuale è davvero impressionante;
- gestire codici caratterizzati da differenti livelli di parallelismo e necessità computazio-

nali, venendo incontro alle esigenze di tutti i gruppi di ricerca che li utilizzeranno;

- essere altamente modulare, al fine di facilitare l'utilizzo di componenti alternativi (co-processor) che potrebbero caratterizzare le future architetture *multicore* quali GPGPU e schede FPGA, componenti altamente e intrinsecamente parallele capaci di una notevole potenza di calcolo;

- essere *open source*, favorendo in tal modo lo scambio e l'integrazione di parti di codice sviluppati da gruppi di lavoro indipendenti.

Le diverse fasi di sviluppo di *X-stack* sono state dettagliate e delineate secondo precise scadenze temporali. Ogni aspetto implementativo è stato analizzato a fondo e, per far meglio comprendere quanto sia complessa e articolata questa impresa, è sufficiente scorrere l'elenco dei componenti da sviluppare:

- un nuovo sistema operativo che permetta un miglior *scheduling* delle risorse a disposizione, comprensivo di strategie per la gestione energetica e delle *fault tolerance*;
- un nuovo sistema per la gestione e l'ottimizzazione delle risorse hardware in funzione delle richieste delle applicazioni;
- un accesso ai dati che permetta l'utilizzo ottimale ed efficiente di *file system* paralleli;
- nuovi modelli di programmazione;
- piattaforme orientate al calcolo scientifico;
- compilatori in grado di recepire le nuove funzionalità software e hardware;
- adeguate librerie e algoritmi numerici;
- nuovi strumenti per il *debug* delle applicazioni;
- nuovi strumenti per la visualizzazione e l'analisi dei dati.

Parallelamente all'*International Exascale Software Project*, la comunità scientifica internazionale e alcuni organismi come quelli preposti alla difesa militare e del territorio, stanno già portando avanti diversi progetti che prevedono lo sviluppo di linguaggi di programmazione innovativi basati sul paradigma di programmazione parallela PGAS (*Partitioned Global Address Space*) [31] fra cui X10 [32], Chapel [33], Fortress [34], Co-Array Fortran [35], UPC [36] e Titanium [37] rappresentano una realtà in notevole fermento.

Il modello PGAS nasce da una fusione e un'evoluzione delle note direttive per la comunicazione interprocessuale MPI (*Message Passing Interface*) e OpenMP (*Open Multi-Proces-*

sing), assumendo i pregi di entrambe. Essendo basato su uno spazio di indirizzamento globale della memoria, il modello previsto per la comunicazione fra i *threads* è molto più efficiente, completamente asincrono e risolto direttamente dal compilatore, permettendo un notevole incremento di performance e scalabilità di ogni applicazione parallela. Questa soluzione, unitamente alla creazione di nuovi paradigmi di programmazione dotati di un elevato livello di astrazione, facilita il programmatore nella gestione dei *threads* per l'accesso di strutture dati complesse con una sintassi molto più snella.

X10, Chapel e Fortress, sviluppati rispettivamente da IBM, Cray e Sun Microsystems come parte del programma governativo USA *High Productivity Computing System* appartenente all'organo DARPA (*Defense Advanced Research Projects Agency*), sono linguaggi di programmazione distribuita orientati agli oggetti, caratterizzati da un ambiente di sviluppo altamente modulare e moderno, che saranno in grado di apportare un incremento di performance di un ordine di grandezza per le applicazioni massicciamente parallele. Co-Array Fortran, UPC e Titanium costituiscono piuttosto un'estensione dei linguaggi Fortran, C e Java pensati con l'intento di incrementare drasticamente le potenzialità di tali linguaggi sulle future architetture *multicore*.

Gli algoritmi che fino ad ora si basavano su librerie matematiche standard, quali Blacs, ScaLapack, o FFTW, dovranno essere ripensati e il progetto open-source PLASMA (*Parallel Linear Algebra for Scalable Multi-core Architectures*) [38] ne costituisce un esempio concreto.

Inoltre, sarà fondamentale costruire il software tenendo in conto l'affidabilità delle architetture *exascale*, incorporando soluzioni per ovviare al problema del *fault tolerance*. È evidente, infatti, che il "*mean time to failure*" decrescerà da centinaia di giorni a poche ore su sistemi composti da centinaia di migliaia di *cores* computazionali; di conseguenza, l'unica tecnica attualmente disponibile ai programmatori per ovviare a tale problema tramite il metodo "*checkpoint/restart*" diventerà del tutto inefficace e poco praticabile. Il sistema sperimentale HARNESS (*Heterogeneous Adaptive Reconfigurable Networked SyStem*) [39], e in particolare l'interfaccia FT-MPI (*Fault Tolerant*

MPI) [40] presente al suo interno, costituisce un primo contributo e una soluzione concreta al problema. FT-MPI fornisce una libreria di comunicazione per l'utente finale che implementa completamente le specifiche MPI-2 e ne estende la semantica per il *recovery* automatico dei processi falliti.

Alla luce di quanto detto è evidente che la strada da percorrere è ancora molto lunga ma è altrettanto evidente che la comunità scientifica sa che non si può permettere in alcun modo di perdere tale sfida.

Bibliografia

- [1] Schiaffonati V.: Computer simulations ed esperimenti in virtuo. *Mondo Digitale*, n. 2, 2010, p. 3-12
- [2] Moore, G.E.: Cramming more components onto integrated circuits. *Electronics Magazine*, 1965, p. 4.
- [3] AA.VV.: *Intel server processors*. http://www.intel.com/p/en_US/products/server/processor
- [4] AA.VV.: *AMD Sets the New Standard for Price, Performance, and Power for the Datacenter*. <http://www.amd.com/us/press-releases/Pages/-amd-sets-the-new-standard-29mar2010.aspx>
- [5] AA.VV.: *Futuristic Intel Chip Could Reshape How Computers are Built, Consumers Interact with Their PCs and Personal Devices*. http://www.intel.com/pressroom/archive/releases/2009/20091202comp_sm.htm
- [6] Brodtkorb A.R., Dyken C., Hagen T.R., Hjelmervik J.M., Storaasli O.O.: State-of-the-Art in Heterogeneous Computing. *Scientific Programming*, Vol. 18 n. 1, 2010, p. 1-33.
- [7] AA.VV.: *Advanced Integrated Circuit Fabrication Processes Course*. <http://www.stanford.edu/class/ee311/NOTES/TrendsSlides.pdf>
- [8] Feynman R.P.: Simulating physics with computers. *Int. J. Theor. Phys.* Vol. 21, 1982, p. 467-488.
- [9] AA.VV.: *What is Quantum Computation*. http://www.quantiki.org/wiki/index.php/-What_is_Quantum_Computation
- [10] Angeleri E.: Quantum computing: sogno teorico o realtà imminente? *Mondo Digitale*, n. 1, 2003, p. 36-50.
- [11] Di Zitti E., Ricci D., Caviglia D.D.: Nanotecnologie e ICT: potenzialità e prospettive. *Mondo Digitale*, n. 3, 2004, p. 3-20.
- [12] Lanyon B. P. et al.: Towards quantum chemistry on a quantum computer. *Nature Chemistry*, Vol. 2, 2010, p. 106-111.



- [13] Adleman L.M.: Molecular Computation Of Solutions To Combinatorial Problems. *Science*, Vol. 266, n. 11, 1994, p. 1021-1024.
- [14] Mauri G.: Il computer vivente: calcolo molecolare e cellulare. *Mondo Digitale*, n. 1, 2010, p. 3-15.
- [15] AA.VV.: *The Mysterious Memristor*. <http://spectrum.ieee.org/semiconductors/-design/the-mysterious-memristor>
- [16] Strukov D.B.: The missing memristor found. *Nature*, Vol. 453, 2008, p. 80-83.
- [17] AA.VV.: *Silicon Photonics*. <http://techresearch.intel.com/articles/Tera-Scale/1419.htm>
- [18] AA.VV.: *First germanium laser*. <http://web.mit.edu/-newsoffice/2010/first-germanium-laser.html>
- [19] AA.VV.: *TOP500*. <http://www.top500.org>
- [20] AA.VV.: *World Internet Users and Population Stats*. <http://www.internetworldstats.com/-stats.htm>
- [21] Saracco R.: Ubiquitous Computing. *Mondo Digitale*, n. 3, 2003, p. 19-30.
- [22] Lang A., Bradley J.: Chemistry in Second Life. *Chemistry Central Journal*, Vol. 3, n. 14, 2009.
- [23] AA.VV.: *ScienceSim*. <http://www.sciencesim.com/-wiki/doku.php>
- [24] AA.VV.: *European Roadmap for Research Infrastructures, Implementation Report 2009*. Luxembourg Office for Official Publications of the European Communities, 2009.
- [25] AA.VV.: *Towards a Sustainable High-Performance Computing Ecosystem through Enabling Petaflop Computing in Europe*. <http://www.hpcineurope-taskforce.eu>
- [26] Llurba R., Dias A.B.S.: *Academic Supercomputing in Europe*. Facilities & Policies, fall 2009 - Eleventh edition, <http://www.nwo.nl/ncf>
- [27] AA.VV.: *PRACE*. NEWSletter, Vol. 8, 2009.
- [28] AA.VV.: *e-IRG Roadmap 2010*. <http://www.e-irg.eu/publications/roadmap.html>
- [29] *Cloud computing*. Wikipedia, http://en.wikipedia.org/wiki/Cloud_computing
- [30] AA.VV.: *International Exascale Software Project*. <http://www.exascale.org>
- [31] AA.VV.: *Partitioned Global Access Space*. <http://www.pgas.org/>
- [32] AA.VV.: *X10 Language*. <http://x10-lang.org/>
- [33] AA.VV.: *Chapel Language*. <http://chapel.cray.com/>
- [34] AA.VV.: *Fortress Language*. <http://projectfortress.sun.com/Projects/Community>
- [35] AA.VV.: *Co-array Fortran Language*. <http://www.co-array.org/>
- [36] AA.VV.: *UPC Language*. <http://upc.gwu.edu/>
- [37] AA.VV.: *Titanium Language*. <http://titanium.cs.berkeley.edu/>
- [38] AA.VV.: *Parallel Linear Algebra for Scalable Multi-core Architectures*. <http://icl.cs.utk.edu/-plasma/index.html>
- [39] AA.VV.: *Heterogeneous Adaptive Reconfigurable Networked SyStem*. <http://www.csm.ornl.gov/-harness/>
- [40] AA.VV.: *Fault Tolerant MPI*. <http://icl.cs.utk.edu/-ftmpi/index.html>

Ivo DE LOTTO è professore ordinario di Calcolatori Elettronici dal 1971 presso la Facoltà d'Ingegneria di Pavia. Ricercatore presso Nuclit, Euratom, CNR, CISE, Università di Bologna e di Pavia ha pubblicato oltre 250 lavori. Ha insegnato Elettronica Nucleare, Strumentazione Elettronica, Elettrotecnica, Calcolatori Elettronici, Reti Logiche, Intelligenza Artificiale, Grafica 3D, Sistemi informativi direzionali. Ha ricoperto varie funzioni organizzative tra le quali Direttore del CILEA (1975-89), Preside della Facoltà d'Ingegneria di Pavia (1996-02), Presidente di AICA (2003-06). Nel 1988 è medaglia d'oro del Ministero quale benemerito della scuola, arte e cultura.
E-mail: ivo.delotto@unipv.it

GIOVANNI MELONI ha conseguito la Laurea Magistrale in Ingegneria Elettrica. Dal 1970 si occupa di applicazioni ingegneristiche. Dal 1988 è Dirigente della Sezione Calcolo ad alte prestazioni ed Applicazioni Scientifiche e Ingegneristiche del CILEA. È stato responsabile CILEA in alcuni progetti europei nel calcolo ad alte prestazioni e in ambito beni culturali. È responsabile da oltre 10 anni del progetto CILEA Digital Library che mette a disposizione della ricerca più di 7.000 riviste scientifiche in formato digitale.
E-mail: meloni@cilea.it

CLAUDIO ARLANDINI si è laureato in Fisica all'Università di Torino nel 1996 e ha conseguito il Dottorato di Ricerca in Astrofisica Nucleare presso l'Università di Heidelberg. Svolge la sua attività presso il Consorzio CILEA, dove è Tecnologo coordinatore del gruppo di competenza per il calcolo ad alte prestazioni. È coautore di circa 50 pubblicazioni internazionali.
E-mail: arlandini@cilea.it

MAURIZIO CREMONESI, laureato in matematica all'“Università degli Studi di Milano” nel 1984, dallo stesso anno lavora al CILEA, dove si è occupato dello studio e diffusione delle problematiche relative alla ottimizzazione del software su calcolatori vettoriali e paralleli. È stato coinvolto nel progetto EU CAPRI-PIERS, ambiente di calcolo parallelo per l'affidabilità di impianti industriali e nel progetto FIRB LitBio. I suoi interessi attuali riguardano lo studio delle prestazioni di programmi di calcolo parallelo e l'organizzazione di corsi di programmazione su elaboratori ad alte prestazioni.
E-mail: cremonesi@cilea.it

PARIDE DAGNA si è laureato in Fisica nel 2001 presso l'università del Piemonte Orientale sede di Alessandria. Dal 2002 lavora nel gruppo di High Performance Computing del CILEA dove si occupa della gestione, dell'ottimizzazione e dello sviluppo di applicazioni tecnico-scientifiche su architetture massicciamente parallele. In tale ambito i suoi interessi riguardano la parallelizzazione ed il porting degli applicativi su tali architetture e su piattaforme GPGPU based. È inoltre docente dei corsi di programmazione per l'HPC tenuti all'interno del CILEA riguardanti i linguaggi Fortran, C/C++, Cuda.
E-mail: dagna@cilea.it

ALICE INVERNIZZI ha conseguito la Laurea Magistrale in Ingegneria Matematica presso il Politecnico di Milano nel 2007, con specializzazione in metodi numerici per la scienza e l'ingegneria. Dal 2007 è parte dello staff di Supercalcolo presso il CILEA dove si occupa di management e assistenza di codici e software per l'analisi strutturale. È inoltre coinvolta nell'insegnamento di linguaggi e tecniche per la programmazione in ambito scientifico, con particolare interesse per i linguaggi C/C++ e metodologie OO, programmazione scientifica in Python, GPGPU Computing e GUI programming.
E-mail: invernizzi@cilea.it

PAOLO RAMIERI ha conseguito la Laurea Magistrale in Fisica (1998) presso l'Università di Torino, sede di

Alessandria. Dal settembre 1999 è nel Team di Supercalcolo del CILEA, dove è responsabile della gestione dei software per la CFD e si occupa di corsi di formazione relativi a linguaggi e tecniche di programmazione per l'HPC, di infrastrutture per l'utilizzo dei cluster di calcolo via web, dell'assistenza agli utenti per l'ottimizzazione di codici e l'uso dei sistemi HPC. Ha collaborato a progetti europei (BioInfoGrid) e FIRB (LitBio) ed è stato per due anni docente di Calcolo Parallelo presso l'Università di Bergamo.

E-mail: ramieri@cilea.it

GIOVANNI SACCHI è Dirigente di Ricerca presso l'Istituto di Matematica Applicata e Tecnologie Informatiche CNR, Pavia e responsabile del Progetto Modellistica e Simulazione di Sistemi Complessi del Dipartimento Tecnologie dell'Informazione e delle Comunicazioni del CNR. Interessi di ricerca: modellistica e simulazione numerica di problemi provenienti dal calcolo strutturale; implementazione di algoritmi e codici di simulazione numerica; sviluppo, implementazione e porting di codici di calcolo su architetture ad alte prestazioni. Attività Didattica: titolare del corso di Calcolo Scientifico e Modellistica Numerica presso il Dottorato in Matematica e Statistica dell'Università degli Studi di Pavia.

E-mail: gianni.sacchi@imati.cnr.it

BIOINFORMATICA E CALCOLO AD ALTE PRESTAZIONI

Con l'introduzione delle più moderne tecnologie di biologia molecolare si è cominciato a conoscere molto più in dettaglio la struttura del DNA, del RNA e delle proteine. Questa enorme mole di informazioni saranno necessarie per comprendere le malattie genetiche partendo dal Genoma Umano. La Bioinformatica e la *System Biology* contribuiranno alla creazione di nuove metodologie utili per la comprensione dei genomi dei diversi organismi. Su questa base verranno illustrati recenti progressi di alcune applicazioni di Bioinformatica basate sul calcolo ad alte prestazioni e il Grid computing.

1. INTRODUZIONE ALLE APPLICAZIONI DELLA BIOINFORMATICA

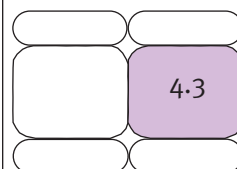
La biologia e la medicina sono la base per studiare l'organismo umano a tutti i livelli, sia morfologico che funzionale [1, 2]. Con i recenti progressi tecnologici nell'ambito delle biotecnologie e della biologia molecolare, è possibile conoscere in pochi giorni il genoma del singolo individuo. Quest'enorme mole di informazioni si sta accumulando in banche dati e, mediante l'utilizzo di appositi programmi di analisi, consentirà la comprensione dell'espressione dei geni, la loro regolazione e le mutazioni nel DNA che sono alla base di malattie genetiche [4]. Per gestire le sequenze genomiche e proteiche e le informazioni che ne descrivono la complessità funzionale dei geni, individuati nel corso dei progetti per il sequenziamento dei genomi di diversi organismi e specie [12, 21, 15], è stata creata una nuova disciplina, denominata "Bioinformatica", che si è recentemente estesa fino ad occuparsi del trattamento dell'informazione biologica e medica a tutti i livelli [5, 6].

La Bioinformatica si occupa quindi dell'acquisizione, memorizzazione, distribuzione, analisi e interpretazione dei dati prevalentemente nell'ambito della biologia molecolare con collegamenti sempre più importanti con la medicina. Questa nuova disciplina scientifica utilizza metodi di matematica, informatica, biologia, medicina e fisica allo scopo di migliorare la comprensione dei fenomeni biologici. Di seguito elenchiamo, a titolo esemplificativo, alcuni obiettivi della bioinformatica:

- sviluppo di strumenti per la generazione e il mantenimento dell'informazione proveniente dalle varie fonti: mappa fisica, mappa genetica, mappa cromosomica, mappa citogenetica, polimorfismi, e le informazioni relative alle sequenze genomiche e proteiche ecc.;
- raccolta e organizzazioni delle informazioni genetiche associate alle patologie mediche;
- sviluppo di programmi di calcolo per l'analisi delle sequenze;
- sviluppo di Interfacce grafiche in grado di visualizzare in maniera efficace l'informazione richiesta;



Luciano Milanesi



- ❑ sviluppo di metodi software che consentano di agevolare tutte le fasi dei progetti;
- ❑ sviluppo di strutture per database specializzate e integrate;
- ❑ sviluppo di strumenti informatici, includendo software, hardware e algoritmi per l'organizzazione e l'analisi dei dati;
- ❑ realizzazione di standard per lo scambio e la descrizione dei dati;
- ❑ realizzazione di una rete dati per la raccolta, la distribuzione e l'aggiornamento costante di tutta l'informazione prodotta;
- ❑ raccolta della bibliografia, dei brevetti e di altri database di supporto all'informazione specifica;
- ❑ predizione dei geni nelle sequenze di DNA;
- ❑ predizione delle strutture tridimensionali delle proteine partendo dalle sequenze primarie;
- ❑ predizione delle funzioni biologiche e biofisiche sia dalle sequenze e sia dalle strutture;
- ❑ simulazione dei processi metabolici e cellulari basati su queste funzioni.

Per far fronte a queste sfide, molti metodi sviluppati storicamente in altre discipline (matematica, scienze dell'informazione, fisica ecc.) sono stati applicati per risolvere queste problematiche. Fra le più significative metodologie impiegate per scoprire i principi biologici di base, partendo dai dati contenuti nelle diverse banche dati dalle sequenze genomiche fino alle strutture proteiche, si possono citare: teoria dell'informazione, statistica, intelligenza artificiale, logica, linguistica computazionale, metodi stocastici, teoria dei grafi, sistemi d'apprendimento basati su reti neurali, sistemi esperti ecc..

2. APPLICAZIONI DELLA BIOINFORMATICA PER LO STUDIO DI MALATTIE COMPLESSE

Le malattie complesse, quali le malattie cardiovascolari e le malattie neurodegenerative, sono dovute all'azione combinata di fattori genetici e ambientali. Il numero di geni implicati in queste malattie, il loro effetto e le loro possibili interazioni con i fattori ambientali non sono stati ancora determinati.

Negli ultimi anni la comunità scientifica ha rivolto una particolare attenzione allo studio

delle popolazioni isolate per identificare geni responsabili di malattie complesse. La scelta delle popolazioni isolate si basa sulla possibilità di semplificare il modello genetico alla base di tali malattie. Infatti, le caratteristiche di queste popolazioni quali il ridotto numero di fondatori e l'elevata consanguineità, causerebbero una riduzione del numero di geni implicati nello sviluppo di queste malattie. D'altra parte, l'uniformità dello stile di vita e dell'ambiente, permetterebbero di ridurre la variabilità dei fattori di rischio non genetici.

Recentemente sono stati sviluppati nuovi strumenti bioinformatici per l'identificazione e la ricerca nuovi geni e varianti geniche trovate nella popolazione e ritenute responsabili delle malattie complesse del sistema cardiovascolare e nervoso [14]. Il potere genetico del modello "isolati genetici" deriva dalla possibilità di costruire pedigree estesi che rapportano la popolazione vivente a pochi fondatori in un numero relativamente basso di generazioni. In aggiunta, le caratteristiche di queste popolazioni quali il ridotto numero di fondatori e l'elevata consanguineità, causerebbero una riduzione del numero di geni implicati nello sviluppo di queste malattie. D'altra parte, l'uniformità dello stile di vita e dell'ambiente, consente di ridurre l'interferenza dei fattori di rischio non genetici permettendo la dissezione dei tratti ereditari. Il CNR è particolarmente attivo in questo ambito, al fine di far progredire la ricerca nel settore dell'informatica applicata alla genetica di popolazioni mediante l'utilizzo di metodologie e algoritmi di *data mining* per il supporto alle decisioni e la predizione del rischio. In particolare, sono stati sviluppati nuovi software applicativi e nuovi algoritmi per la gestione, l'analisi e l'implementazione di banche dati di popolazioni, al fine di facilitare l'identificazione dei geni coinvolti nei tratti quantitativi e nelle malattie complesse.

L'analisi dei dati inerenti a questa tipologia di studi si basa su software di analisi quali (SOLAR, SimWalk2, MERLIN ecc.). La piattaforma sviluppata presso il CNR [16] ha consentito l'ottimizzazione degli algoritmi attualmente disponibili, riducendo così notevolmente il tempo di calcolo mediante l'uso dell'infrastruttura ad alte prestazioni HPC (*High Performance Computing*). Queste soluzioni hanno consentito un miglioramento significativo rispetto alle

tecnologie esistenti in termini di velocità di analisi, predisposizione all'automazione, affidabilità e riproducibilità, visualizzazione grafica di ampie genealogie e infine hanno permesso anche di ottenere una migliore accuratezza nel calcolo statistico-matematico.

3. SVILUPPI DELLE APPLICAZIONI PER IL CALCOLO INTENSIVO E IL GRID COMPUTING

Lo studio del DNA genomico abbinato allo sviluppo di strumentazione di sequenziamento massivo di interi genomi in pochi giorni, stanno rivoluzionando la conoscenza dei processi che sono alla base dei profili strutturali e funzionali di molte malattie genetiche e di alcune tipologie di tumori [17]. Infatti, la conoscenza dei processi di replicazione, trascrizione, traduzione e di localizzazione mediano e controllano l'espressione dell'informazione genetica. Lo scopo principale della biologia molecolare è quello di capire i meccanismi che stanno alla base delle capacità di regolazione di questi processi. Diversi metodi sono impiegati per lo studio e la predizione delle funzioni, basandosi sull'analisi dalle sequenze o delle strutture proteiche [18, 19]. Fra questi ricordiamo a titolo d'esempio: la predizione delle regioni codificanti per le proteine; la predizione delle strutture secondarie e terziarie partendo dalla sequenza genomica e/o amminoacidica; la dinamica molecolare; la produzione di modelli molecolari; la simulazione metabolica partendo dalle funzioni biochimiche dell'organismo umano ecc..

Dal punto di vista prettamente bioinformatico ognuno di questi processi è una possibile sorgente di studio e di applicazione di nuove metodologie d'indagine. Uno dei principali obiettivi della Bioinformatica basata sulle tecnologie messe a punto dalle tecniche Informatiche di Grid computing e ICT, è di dare in tempi rapidi le informazioni e le metodologie d'indagine, che consentano ad esempio di fornire alla scienza medica le necessarie informazioni per comprendere i meccanismi alla base di tutte le possibili disfunzioni d'origine genetica. Dalle mutazioni in regioni funzionali del DNA, alla mancata produzione di una certa proteina a causa di un anomalo funzionamento dei fatto-

ri di trascrizione, fino a comprendere il funzionamento d'ogni singolo gene in relazione con gli altri geni nei diversi processi metabolici che continuamente avvengono nel corso dell'intera vita [20, 22].

4. BIOINFORMATICA IN SUPPORTO ALL'INNOVAZIONE E ALL'INDUSTRIA FARMACEUTICA IN AMBITO EUROPEO

L'industria farmaceutica, sempre di più utilizza i dati, le metodologie e le conoscenze prodotte nei progetti di ricerca nazionali e internazionali al fine di accelerare lo sviluppo di nuove medicine, proteine specifiche e terapie genetiche in grado di intervenire in maniera selettiva per risolvere le possibili cause di malattie.

Da questa intensificazione nella produzione di nuove sequenze, nasce l'esigenza di riconoscere nuove strutture e funzioni dei geni coinvolti nei vari processi metabolici. Per questo è necessario identificare i modelli quantitativi che stanno alla base di questi processi. È quindi diventato essenziale poter dedurre le informazioni utili partendo dall'analisi delle strutture e dalle funzioni a loro associate, utilizzando al meglio le informazioni contenute nelle varie banche dati genetiche, e sia dai metodi e programmi messi a disposizione dalla Bioinformatica.

Nell'ambito del forum europeo ESFRI "European Strategy Forum on Research Infrastructures" sono state promosse iniziative quali ELIXIR per la Bioinformatica e BMMRI per le Biobanche [23]. Questi progetti, in congiunzione con altre iniziative europee, si pongono come obiettivo la possibilità di sviluppare nuove applicazioni, per esempio nei seguenti settori:

- gestione di laboratori biomedici distribuiti;
- calcolo di espressione dei geni su larga scala;
- studio di malattie genetiche basate su dati genotipici e fenotipici;
- analisi filogenetica di diverse specie;
- accesso a database biologici distribuiti;
- analisi funzionale delle proteine e del proteoma di diverse specie;
- simulazioni di dinamica molecolare basate su tecnologie di HPC e Grid computing.

Mediante iniziative come EGEE, DEISA, BioInfoGRID, EDGE, BMMRI [3, 4, 9, 11] e il progetto eu-

ropeo EGI (*European Grid Infrastructure*), per esempio, è stato possibile utilizzare con estrema efficacia i più avanzati centri di calcolo europei basandosi sull'avanzatissima infrastruttura per il calcolo HPC distribuito (DEISA) e su Grid Computing in collaborazione con i maggiori centri di supercalcolo.

In particolare sono state messe a punto nuove applicazioni di "Drug Discovery" per la ricerca di nuovi farmaci nel campo delle Malaria e dell'Influenza Aviaria [10], mediante il calcolo automatico per la ricerca basata sul *Virtual Screening* utilizzando migliaia di computer distribuiti in Europa. In questo progetto è stato totalmente automatizzato il calcolo di milioni di possibili molecole utilizzabili per la creazione di nuove medicine.

5. PROGETTI INTERNAZIONALI

Nel seguito verranno brevemente illustrate alcune iniziative a cui si è attivamente contribuito. La descrizione di altri progetti in ambito europeo è reperibile direttamente dal sito della Comunità Europea.

5.1. Enabling Grids per E-science (EGEE-III) - European Grid Infrastructure (EGI)

Il Grid computing è diventato indispensabile per una scienza basata su dati intensivi. Il concetto è stato dimostrato attraverso pro-

getti sostenibili come EGEE - *Enabling Grids per E-science* [11] il cui logo è rappresentato nella figura 1. EGEE si propone di sviluppare un *middleware* per una vasta gamma di comunità scientifiche, come l'astrofisica, la chimica computazionale, la terra e le scienze della vita, la fusione e la fisica delle particelle e la Bioinformatica, costruito sulla rete paneuropea GÉANT2. A seguito del progetto EGEE, la Comunità Europea ha promosso recentemente il progetto EGI (*European Grid Infrastructure*) il cui logo è rappresentato nella figura 2 con lo scopo di consentire ai ricercatori presenti in diversi Paesi di collaborare in ricerche comuni basate su infrastrutture Grid per perseguire i seguenti obiettivi:

- espandere e ottimizzare la più grande produzione di Grid in Europa con un continuo funzionamento delle infrastrutture, un sostegno a più comunità di utenti e l'aggiunta di ulteriori risorse di calcolo e dati;

- preparare la migrazione degli attuali Grid da un progetto base ad un modello sostenibile di infrastrutture federato basato su Iniziative di Grid Nazionale (NGI).

EGI fornirà una rete europea coerente e affidabile, assicurando che l'Europa sia in prima linea per la possibilità di utilizzare il calcolo scientifico.

5.2. Biobanking e Infrastrutture di Ricerca di Risorse Biomolecolari (BBMRI)

La conservazione e il trattamento di campioni e di dati biomedici sono essenziali per il mondo accademico e l'industria di ricerca al fine di curare e prevenire le malattie umane. La fase preparatoria si basa sul progetto BBMRI - *Biobanking and Biomolecular Resources Research Infrastructure* [9] con i seguenti scopi principali:

- preparare l'integrazione delle biobanche esistenti, mediante l'adozione di componenti innovativi e in linea con gli standard etici, giuridici e sociali a livello europeo secondo lo schema rappresentato nella figura 3;

- aumentare l'efficacia nelle ricerche delle scienze biomediche e nello sviluppo delle risorse biologiche nella sanità pubblica;

- ampliare e garantire la competitività della ricerca europea e dell'industria in un contesto globale;

- sviluppare un quadro finanziario sostenibile.

FIGURA 1
Logo del progetto
Enabling Grids per
E-science (EGEE)



FIGURA 2
Logo del progetto
European Grid
Infrastructure (EGI)



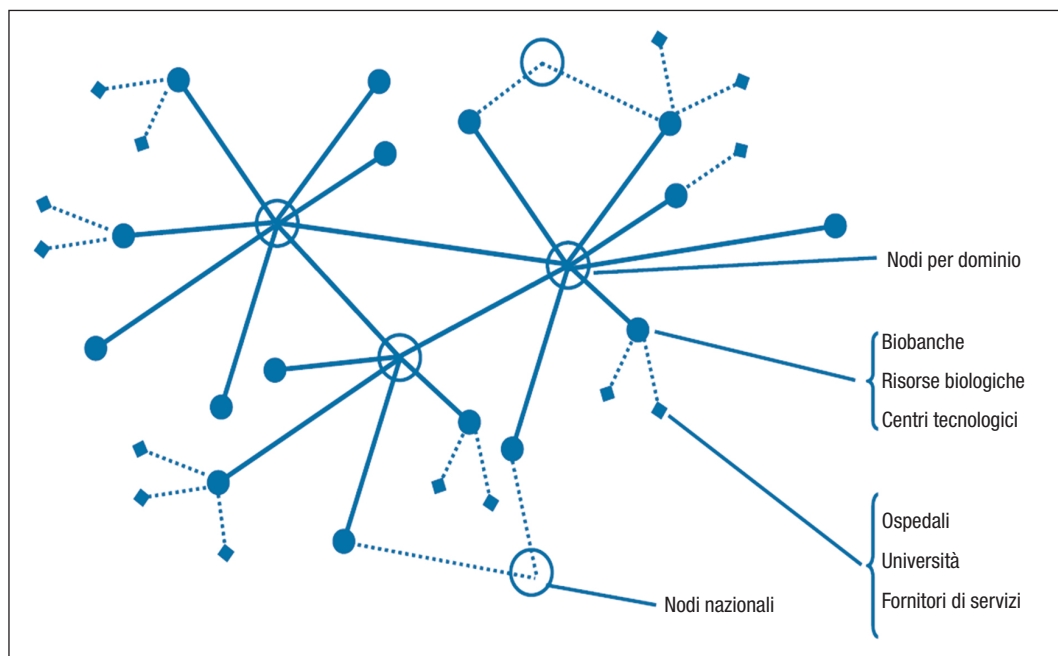


FIGURA 3
Schema di coordinamento e di collegamento del progetto BBMRI fra le Biobanche federate in Europa

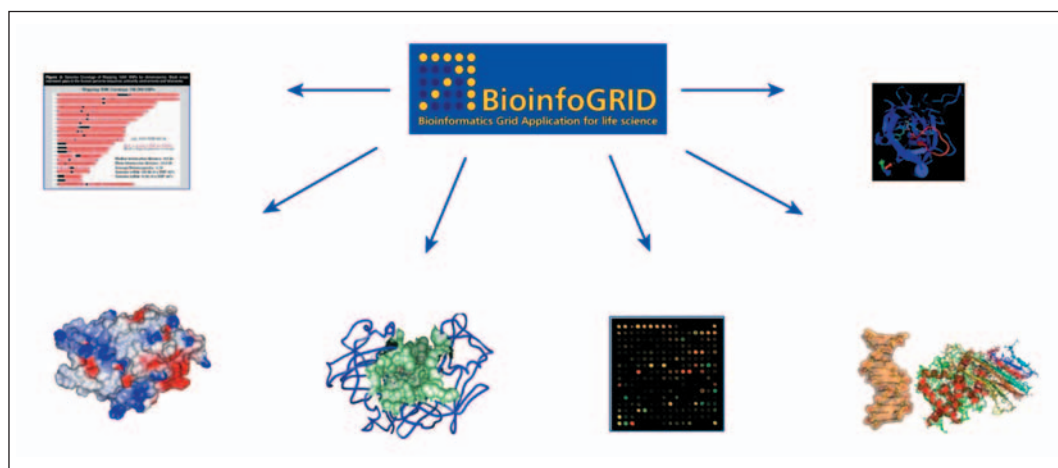


FIGURA 4
Ambiti applicativi del progetto europeo BioinfoGRID per applicazioni in ambito bioinformatico basate su Grid Computing

5.3. Applicazioni GRID di Bioinformatica per vita scientifica (BioinfoGRID)

Il progetto europeo BioinfoGRID (<http://www.bioinfoGRID.eu>) [4] ha contribuito a migliorare i servizi e le applicazioni bioinformatiche per gli utenti di biologia molecolare mediante l'utilizzo delle infrastrutture Grid, sviluppando applicazioni di genomica, proteomica, trascrittomica e dinamica molecolare. Nella figura 4 sono mostrati schematicamente i vari ambiti applicativi della Bioinformatica basata sulla tecnologia Grid.

6. PROGETTI NAZIONALI

L'Italia ha già promosso nell'ambito della Bioinformatica, progetti nazionali finanziati

dal MIUR come LITBIO [13], LIBI e ITALBIO-NET (*Italian Bioinformatics Network*). Queste iniziative hanno consentito di raggiungere risultati in diversi settori tecnologici e metodologici della Bioinformatica e della *System Biology* per lo studio dei sistemi biologici complessi [7, 17].

È importante sottolineare il forte legame internazionale delle iniziative dei progetti nazionali in Bioinformatica.

Come esempio di forte internazionalizzazione, il progetto interdipartimentale "CNR-Bioinformatics" [8] del CNR il cui logo è rappresentato nella figura 5, in congiunzione con il progetto europeo EGEE-III finanziato dalla Commissione Europea e coordinato dal CERN

di Ginevra, ha promosso un'iniziativa di carattere internazionale per realizzare analisi su larga scala finalizzata all'individuazione di nuove molecole a scopo terapeutico.

Mediante questi progetti e in collegamento con l'Accademia Sinica in Taiwan, CNRS-IN2P3 in Francia e il CNR-ITB in Italia, è stata intrapresa un'iniziativa per l'analisi di 300.000 possibili composti farmacologici per combattere il virus H5N1 per l'Influenza Aviaria. Lo scopo principale di questa iniziativa è consistito nell'identificare dei potenziali composti in grado di inibire le attività di un enzima sulla superficie del virus capaci di causare l'influenza, denominato Neuraminidasi del sottotipo N1. La figura 6 mostra il posizionamento di un inibitore relativamente all'enzima neuraminidasi del virus dell'aviaria H5N1. Mediante l'uso dell'infrastruttura GRID composta da 50.000 computer messa a punto dal progetto europeo EGEE-III è stato possibile identificare le più promettenti molecole da sottomettere a test biologici al fine di accelerare il processo di sviluppo di farmaci contro il virus influenzale.

L'infrastruttura di calcolo messa a punto da questo progetto ha consentito la riduzione del numero dei possibili farmaci in modo da passare alla seconda fase di laboratorio, per con-

trollarne la tossicità e le altre caratteristiche biochimiche necessarie prima dello sviluppo del farmaco finale. Con questa tecnologia sarà possibile utilizzare con estrema efficacia i più avanzati centri di calcolo europei per la ricerca di nuovi potenziali molecole per la produzione di farmaci per combattere nuovi tipi di virus influenzali e per la ricerca di nuovi farmaci per malattie quali la malaria e l'influenza suina.

7. PROSPETTIVE FUTURE

Solo una piccola frazione del genoma umano è utilizzata per la costruzione delle proteine: si può stimare che il DNA totale codificante sia circa 10^8 basi vale a dire corrisponda al 3% circa del genoma totale. Per la ricerca e la predizione dei geni, sono stati utilizzati algoritmi basati su reti neurali e metodi statistici. Il metodo sperimentale per l'identificazione delle regioni codificanti di una regione del DNA, consiste nel confrontare la sequenza di cDNA e la corrispondente sequenza genomica, mediante un multiallineamento fra sequenze. Nella maggioranza dei casi la sequenza di cDNA non è nota, risulta quindi necessario sviluppare una serie di metodi per il riconoscimento delle regioni del DNA in cui risiede il gene. Una parte importante di questo processo consiste nel riconoscere i siti di *splicing*, delle regioni codificanti e delle regioni promotrici. È stato recentemente finanziato a livello internazionale un progetto per il sequenziamento di 1000 genomi di persone appartenenti a popolazioni diverse. Questo processo in congiunzione con molti altri studi, metterà in luce le differenze genetiche e le mutazioni puntiformi (SNP *Single Nucleotide Polymorphisms*). Proprio la correlazione delle modifiche funzionali derivate da queste mutazioni caratterizzano il futuro della medicina personalizzata, cioè la possibilità di costruire farmaci specifici e cure per ogni persona che tengano conto delle modifiche del suo genoma. Questo processo appena iniziato vedrà importanti ricadute nel prossimo decennio con un notevole impatto sulla salute dell'uomo e sulla produzione e dosaggio di nuovi farmaci mirati. Attualmente le fasi per la ricerca di nuovi farmaci è estremamente costosa e richiede tempi lunghi. Quando il genoma sarà completamente disponibile,

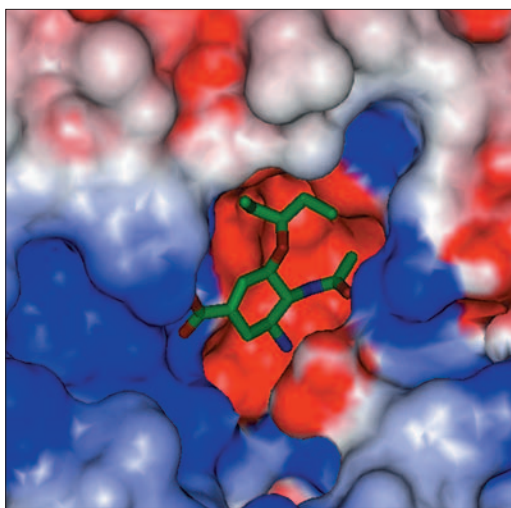
FIGURA 5

Logo del progetto
interdipartimentale
del CNR
di Bioinformatica



FIGURA 6

Interazione tra
l'inibitore e l'enzima
neuraminidasi del
virus dell'aviaria
H5N1



sarà possibile studiare l'influenza sulla salute in base all'interazione dei geni fra loro e agli altri fattori ambientali e alimentari. Inoltre, con il progresso dell'ingegneria genetica e della medicina molecolare sarà possibile prevenire e curare alcune malattie ancora prima del loro manifestarsi. Tutto questo porterà lo sviluppo di nuovi strumenti Informatici per l'analisi di una enorme mole di dati, e lo sviluppo di nuove metodologie di indagine. La Bioinformatica e la *System Biology* forniranno la base metodologica alla biologie e alla medicina molecolare per il raggiungimento di tali obiettivi basandosi su piattaforme avanzate per il calcolo ad alte prestazioni HPC e il Grid Computing.

Riconoscimenti

Questo lavoro è basato sui seguenti progetti Europei: INBIOMEDvision, ELIXIR, EGEE-III, BioinfoGRID, BBMRI, EDGE; progetti nazionali: CNR-BIOINFORMATICS, LITBIO (RBLA0332RH), ITALBIONET (RB-PR05ZK2Z_001), GENPOP (RBIN064YAT_003).

Bibliografia

- [1] Alfieri R., Merelli I., Mosca E., Milanesi L. *The Cell Cycle DB: a systems biology approach to cell cycle analysis*. Nucleic Acids Res. 2008 Jan;36 (Database issue):D641-5.
- [2] Alfieri R., Barberis M., Gaglio D., Chiaradonna F., Milanesi L., Vanoni M., Klipp E., Alberghina L.: *Towards a systems biology approach to mammalian cell cycle: modeling the entrance into S phase of quiescent fibroblasts after serum stimulation*. BMC Bioinformatics. 2009 Oct 15;10 Suppl 12:S8.
- [3] Barbera R., Donvito G., Falzone A., La Rocca G., Milanesi L., Maggi G.P., Vicario S.: *The GENIUS Grid Portal and robot certificates: a new tool for e-Science*. BMC Bioinformatics. 2009 Jun 16;10 Suppl 6:S21.
- [4] BIOINFOGRID: *Bioinformatics Grid Application for Life Science*. <http://www.bioinfoGRID.eu>
- [5] Castellani G., Verondini E., Giampieri E., Milanesi L., Bersani F., Zironi I., Remondini D.: Large scale modelling of neuronal systems. *Il Nuovo Cimento*. Vol. 32 C, n. 2, March - April, 2009, p. 9-12.
- [6] Calabria A., Mosca E., Viti F., Merelli I., Milanesi L.: SNPranker: a tool for identification and scoring of SNPs associated to target genes. *J Integr Bioinform*. 2010 Mar 25, Vol. 7, n. 3, doi: 10.2390/biecoll-jib-2010-138.PMID: 20375450.
- [7] Chiappori F., D'Ursi P., Merelli I., Milanesi L., Rovida E.: *In silico saturation mutagenesis and docking screening for the analysis of protein ligand interaction: the Endothelial Protein C Receptor case study*. BMC Bioinformatics. 2009 Oct 15;10 Suppl 12:S8.
- [8] CNR-BIOINFORMATICS: *Progetto interdipartimentale CNR*. <http://www.cnr-bioinformatics.it>
- [9] BBMRI - Biobanking and Biomolecular Resources Research Infrastructure. <http://www.bbMRI.eu>
- [10] D'Ursi P., Chiappori F., Merelli I., Cozzi P., Rovida E., Milanesi L.: Virtual screening pipeline and ligand modelling for H5N1 neuraminidase. *Biochem Biophys Res Commun.*, Jun 12, Vol. 383, n. 4, 2009, p. 445-9. Epub 2009 Apr 14.
- [11] EGEE-III - Enabling Grids for E-science-III (INF-SO-RI-222667). <http://www.eu-egEE.org/> Francesconi M., Remondini D., Neretti N., Sedivy J.M., Cooper L.N., Verondini E., Milanesi L., Castellani G.: *Reconstructing networks of pathways via significance analysis of their intersections*. BMC Bioinformatics, 2008 Apr 25;9 Suppl 4:S9. 8.
- [12] Lazzari B., Caprera A., Cestaro A., Merelli I., Del Corvo M., Fontana P., Milanesi L., Velasco R., Stella A.: *Ontology-oriented retrieval of putative microRNAs in Vitis vinifera via GrapeMiRNA: a web database of de novo predicted grape microRNAs*. BMC Plant Biol. 2009 Jun 29;9:82.
- [13] LITBIO - *Laboratory of Interdisciplinary Technologies in Bioinformatics*. <http://www.litbio.it>
- [14] ITALBIONET - Italian Bioinformatics Network.
- [15] Merelli I., Guffanti A., Fabbri M., Cocito A., Furia L., Grazini U., Bonnal R.J., Milanesi L., McBlane F.: *RSSsite: a reference database and prediction tool for the identification of cryptic Recombination Signal Sequences in human and murine genomes*. Nucleic Acids Res. 2010 May 16. doi:10.1093/nar/gkq391.
- [16] Milanesi L., Calabria A., Di Pasquale D., Gnocchi M., Orro A., Trombetti G.: A HPC and Grid enabling framework for genetic linkage analysis of SNPs. *Il Nuovo Cimento*, Vol. 32 C, n. 2, March - April 2009, p. 249-253.
- [17] Milanesi L., Romano P., Castellani G., Remondini D., Lio' P.: *Trends in Biomedical Complex Systems*. BMC Bioinformatics. 2009 Oct 15;10 Suppl 12:S8.
- [18] Mosca E., Bertoli G., Piscitelli E., Reinbold R., Zucchi I., Milanesi L.: *A strategy for the identification of functionally related genes based on data mining and data integration*. BMC Bioinformatics. 2009 Oct 15;10 Suppl 12:S8.
- [19] Mosca E., Bertoli G., Piscitelli E., Reinbold R., Zucchi I., Milanesi L.: *Identification of functionally related genes using data mining and data integration: a breast cancer case study*. BMC Bioinformatics 2009, 10(Suppl 12):S8.

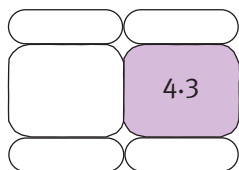
- [20] Orro A., Guffanti G., Salvi E., Macciardi F., Milanesi L.: *SNPLims: A Data Management System for Genome Wide Association Studies*. BMC Bioinformatics 2008 Mar 26;9 Suppl 2:S13.
- [21] Panzitta F., Caprera A., Merelli I., Luciano Milanesi, Williams J.L., Lazzari B., Stella A.: Mining the Bovine Genome with the "Bovine SNP Retriever". *Journal of Heredity*, Nov-Dec, Vol. 99, n. 6, 2008, p. 696-8.
- [22] Torri F., Akelai A., Lupoli S., Sironi M., Amann-Zalcenstein D., Fumagalli M., Dal Fiume C., Ben-Asher E., Kanyas K., Cagliani R., Cozzi P., Trombetti G., Lievers L.S., Salvi E., Orro A., Beckmann J. S., Lancet D., Kohn Y., Milanesi L., Ebstein R. B., Lerer B., Macciardi F.: *Fine mapping of AHI1 as a schizophrenia susceptibility gene: from association to evolutionary evidence*. FASEB J. 2010 Apr 6. PMID: 20371615.
- [23] Viti F., Merelli I., Caprera A., Lazzari B., Stella A., Milanesi L.: *Ontology-based, Tissue MicroArray oriented, image centred tissue bank*. BMC Bioinformatics. 2008 Apr 25;9 Suppl 4:S4.

LUCIANO MILANESI è ricercatore dal 1983 presso il CNR - Istituto di Tecnologie Biomediche. Dal 1988 ha partecipato a diversi progetti nazionali ed internazionali volti allo sviluppo di nuovi metodi Bioinformatici per l'analisi del Genoma Umano. Ha partecipato dal 2000 ad oggi a più di 20 progetti di cui 10 europei, fra cui ha coordinato il progetto LITBIO "Laboratorio Interdisciplinare di Bioinformatica" e BIOINFOGRID "Bioinformatics Grid Application for Life Science". È autore di più di 260 pubblicazioni ed è membro dell'editorial board di diverse riviste scientifiche nel campo della bioinformatica, System Biology ed Informatica Medica.
E-mail: luciano.milanesi@itb.cnr.it



STUDIO COMPUTAZIONALE DELLA FLUIDODINAMICA VASCOLARE

Raffaele Ponzini



Negli ultimi decenni l'uso delle risorse computazionali per lo studio della fluidodinamica del sangue ha avuto uno sviluppo eccezionale. Una descrizione accurata e realistica di grandezze emodinamiche è oggi possibile grazie all'integrazione di tecniche numeriche, di *imaging* e di visualizzazione scientifica. Attraverso l'uso di risorse di calcolo ad alte prestazioni, i ricercatori possono affrontare problemi che non sarebbero stati risolti con tecniche di misurazione tradizionali. Questo articolo introduce i metodi, le problematiche e gli sviluppi più promettenti dell'uso di modelli numerici per lo studio dell'emodinamica.

1. INTRODUZIONE

Da circa trent'anni, una grande parte della ricerca in ambito cardiovascolare ha concentrato la propria attenzione sulle relazioni che intercorrono tra emodinamica locale e l'insorgenza di fenomeni fisiopatologici legati poi all'evolvere di patologie cardiovascolari. Come è noto, infatti, le malattie cardiovascolari colpiscono purtroppo una quantità di persone nel solo mondo occidentale tale da essere stimate come la principale causa di morte in Europa. Ogni anno, infatti, queste patologie sono causa circa del 48% dei decessi. L'enormità di questi dati ha evidentemente una controparte nelle voci dell'economia dei Paesi coinvolti; nel 2006 il trattamento di pazienti colpiti da patologia cardiovascolare solo in Europa ha comportato una spesa di circa 110 miliardi di Euro, pari a quasi il 10% del totale delle spese sanitarie [1].

Oggi è considerato plausibile che l'insorgenza di patologie sia correlata anche con la presenza di strutture fluidodinamiche e di *pattern* emodinamici alterati o altresì disturbati [2]. Fenomeni di questo tipo sono stati osservati

qualitativamente a vari livelli e con svariate tecniche di indagine in distretti vascolari caratterizzati dalla presenza di curvature, biforcazioni e diramazioni, o più in generale dove siano presenti variazioni nella direzione e nella portata del flusso ematico [3]. Più in dettaglio, laddove, per vari motivi, siano presenti basse velocità in prossimità della parete vascolare, tali cioè da causare il ristagno del sangue, è risultata più elevata l'incidenza dell'insorgenza e dello sviluppo di processi infiammatori [4, 5]. Tuttavia, al fine di studiare e analizzare quantitativamente le relazioni che intercorrono tra il sangue e l'insorgere di questi fenomeni degenerativi localizzati principalmente alla parete vascolare, è necessario un livello di controllo sulle variabili fluidodinamiche con un dettaglio spaziale e temporale estremamente elevato. Purtroppo i convenzionali metodi di indagine, basati cioè su studi teorici o su modelli in vitro e/o animale, si sono rivelati insufficienti o solo parzialmente efficaci per sostenere questo tipo di analisi. Le motivazioni sono da ricer-

carsi nel fatto che se da un lato i modelli teorici, in cui tutte le variabili sono controllabili o calcolabili, si dimostrano essere troppo semplicistici per descrivere la complessità del fluire del sangue nei distretti vascolari, dall'altro i modelli in vitro e su animale non permettono di controllare i valori dei parametri fisici in gioco con il livello di accuratezza richiesta dalla complessità dei fenomeni biologici studiati o si rivelano di difficile realizzazione pratica. Tutti questi ambienti di indagine "classici" risultano pertanto in buona sostanza inefficienti al fine di comprendere i meccanismi biomeccanici che intercorrono nell'interazione tra flusso ematico e stato di sforzo dei tessuti da esso irrorati che stanno alla base biologica di tutti processi infiammatori e degenerativi.

In questo contesto un approccio alternativo è fornito dai modelli di fluidodinamica computazionale (CFD, *Computational Fluid Dynamics*). Infatti, una volta fornite le condizioni al contorno, la geometria del modello e le caratteristiche della reologia del fluido, lo studio tramite modelli CFD consente di calcolare ed emulare in maniera del tutto realistica i campi di velocità e pressione all'interno di geometrie tridimensionali complesse quali sono i vasi sanguigni, ben rappresentandone le variazioni dinamiche sia nello spazio che nel tempo.

L'uso dei metodi computazionali fornisce, infatti, vantaggi unici rispetto agli studi sperimentali:

- permette di modificare facilmente i parametri del modello;
- permette di ottenere analisi ripetibili;
- è utilizzabile su sistemi non analizzabili sperimentalmente;
- permette di introdurre condizioni al contorno da misure in vivo;
- garantisce dati privi di rumore;
- permette di simulare condizioni fluidodinamiche che vanno dal sano al patologico.

Infine, dalle soluzioni della CFD risulta semplice estrarre le distribuzioni di grandezze di primario interesse come ad esempio la velocità del flusso sanguigno e di quantificare importanti coefficienti e indici emodinamici con un livello di accuratezza spaziale e temporale definibile dal ricercatore.

Alla luce sia della rilevanza socio-economica legata alla comprensione della fluidodinamica vascolare sia della difficoltà di ottenere infor-

mazioni quantitative utili a studiare e quantificare l'interazione biomeccanica tra il sangue e la parete vascolare con tecniche "classiche", appare evidente come la disponibilità di nuovi strumenti di indagine per l'analisi e lo studio di queste grandezze all'interno dell'albero arterioso umano, risulti di fondamentale interesse per cercare di colmare un limite conoscitivo rilevante e ad oggi ancora largamente esistente.

Il presente articolo si pone come obiettivo quello di descrivere attraverso alcuni studi rappresentativi svolti all'interno di una collaborazione scientifica tra vari atenei nazionali, internazionali e il CILEA, come oggi lo studio in silico dell'emodinamica, tramite cioè modelli computazionali, rappresenti un ambiente di indagine privilegiato per calcolare specifici descrittori dei *pattern* fluidodinamici e della parete vascolare, oltre che uno strumento innovativo e strategico in questo ambito di ricerca. Infine, un'applicazione specifica che sfrutta la CFD per ottimizzare il calcolo della portata ematica da misurazioni cliniche con strumentazione Doppler viene descritta nel riquadro a p. 70.

2. MODELLI CFD IN EMODINAMICA

Volendo schematizzare in maniera esemplificativa il processo della modellazione CFD in emodinamica (si veda la Figura 1) possiamo riconoscere tre macro-fasi:

- i. (box blu) fase di *pre-processing* in cui da dati relativi all'anatomia del distretto vascolare e alla fluidodinamica di alcune sezioni del vaso (i cosiddetti *boundary*) è possibile implementare un modello CFD del distretto di interesse in termini di modello geometrico 3D e di condizioni al contorno;
- ii. (box verde) fase di calcolo vero e proprio dove grazie a tecniche di calcolo parallelo e a strutture hardware e di interconnessione ad alta efficienza è possibile risolvere le equazioni costitutive del sangue (equazioni di Navier-Stokes) all'interno del modello ottenuto in tempi ragionevolmente brevi;
- iii. (box arancio) fase di *post-processing* in cui a partire dai campi di moto calcolati è possibile estrarre con tecniche di visualizzazione scientifica e tramite l'applicazione di algoritmi numerici i valori di interesse biologico che so-

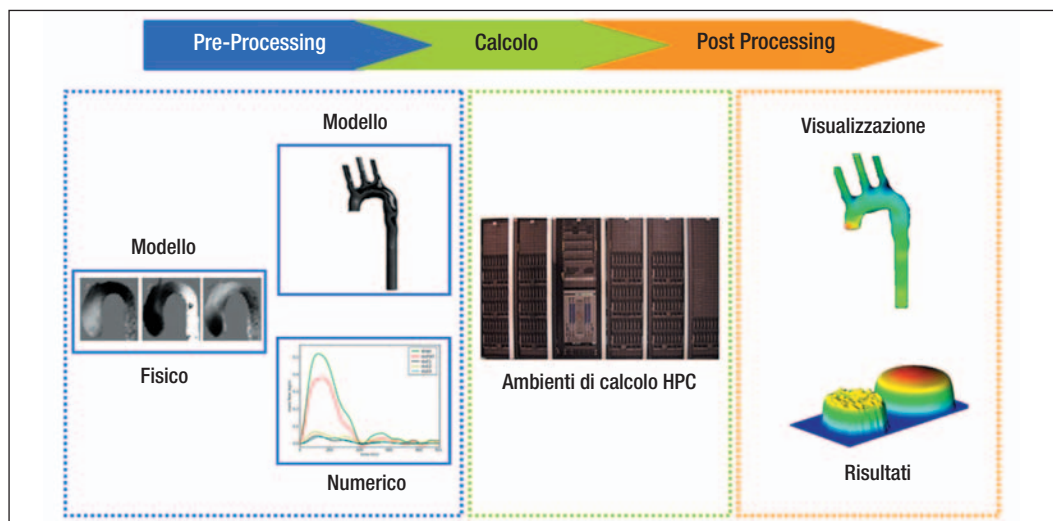


FIGURA 1
Schema di
modellazione CFD
in emodinamica.

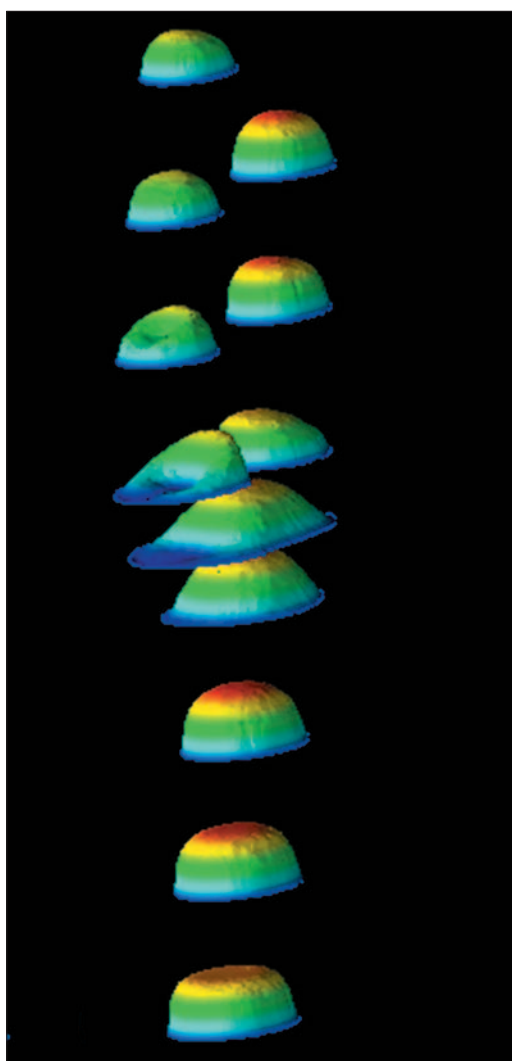


FIGURA 2
Profili spaziali di velocità all'interno di diverse sezioni di
una biforcazione carotidea

no necessari per caratterizzare la fenomenologia dell'interazione sangue-parete vascolare. Questo tipo di diagramma è accumulabile a molteplici ambiti di ricerca basati su queste tecniche, tuttavia nel caso dell'emodinamica è importante ricordare quali siano le ipotesi lavorative, i limiti e le semplificazioni che necessariamente devono essere applicate per poter concretamente formulare una modellazione numerica di questi problemi biologici. A margine si ricorda che ogni aspetto, sia esso tecnico piuttosto che prettamente scientifico, coinvolto in queste fasi costituisce un elemento rilevante per la corretta modellazione di questi problemi biologici e che al fine di ottenere modelli affidabili in ogni fase si devono cercare di limitare le semplificazioni indotte nel naturale processo di schematizzazione al calcolatore di un problema evidentemente complesso e articolato.

Infatti, nella natura biologica del problema abbiamo a che fare con un fluido, il sangue, che è di fatto un fluido corpuscolato e non Newtoniano, le pareti dei vasi, che hanno strutture anatomiche complesse e ramificate, non sono affatto lisce né è nota in generale la legge costitutiva della parete stessa. Il profilo spaziale e temporale delle velocità che si possono imporre ai bordi numerici dei modelli variano secondo leggi che coinvolgono funzioni di ordine superiore (teoria di Womersley) come mostrato nella figura 2.

Il problema è poi evidentemente accoppiato per quanto riguarda l'interazione fluido-parete vascolare: l'azione del moto dell'uno si

riflette nella reazione dell'altro sul fluido stesso. Inoltre si deve considerare come - nonostante l'albero vascolare sia di fatto un continuo di ramificazioni interconnesse in cui lo stato fluidodinamico di una singola regione ha degli effetti rilevabili in tutte le altre - per una ragionevole modellazione di questi problemi, sia necessario sviluppare delle metodologie che permettano di considerare il singolo vaso sanguigno in analisi, identificandolo nella sua anatomia e studiandolo senza dover modellare con la stessa accuratezza tutto il resto del sistema cardiovascolare di cui pure fa parte e di cui evidentemente risente. Detto ciò, come vedremo, esistono delle strategie implementative che, supportate da dati teorici, permettono di semplificare ragionevolmente le caratteristiche del problema rendendolo risolvibile con moderni strumenti di calcolo.

Il sangue viene comunemente modellato come un fluido incompressibile, isotermico, Newtoniano, con densità pari a 1060 kg/m^3 e viscosità costante pari a 3.5 cP (valori di riferimento per il sangue in condizioni fisiologiche, rappresentativi di un ematocrito pari al 43%). La parete esterna del vaso è modellata come rigida assegnando all'interfaccia fluido-parete la condizione di non scivolamento.

Il profilo spaziale di velocità viene tipicamente imposto come parabolico oppure piatto estendendo la dimensione del dominio al fine di permettere al fluido di sviluppare il suo vero andamento nello spazio. Alternativamente, approcci più raffinati sono stati proposti da ricercatori italiani e applicati con

successo allo studio di problemi di pertinenza prettamente clinica [6, 7].

Poiché il sistema circolatorio è una rete chiusa in cui l'emodinamica in un singolo distretto è correlata alla dinamica globale dell'intero sistema, per tenere conto di tale effetto sull'emodinamica locale si ricorre ad un approccio detto multi-scala. Il distretto di interesse viene studiato mediante un modello tridimensionale (3D) che permette uno studio locale delle caratteristiche del flusso ematico e delle sollecitazioni che esso esercita sulla parete. Il restante albero ematico viene modellato tramite modelli semplificati ottenuti integrando le equazioni di Navier-Stokes per ottenere un "equivalente-elettrico" in termini di effetti resistivi, di impedenza e capacitivi dovuti alla presenza del tutto sistemico rispetto a quando modellato in 3D solo a livello locale.

L'approccio multi-scala consente di assegnare sulle sezioni di outlet del modello 3D locale, condizioni al contorno fisiologicamente realistiche, coerenti e consistenti con la variazione del flusso e della pressione sistemica durante il ciclo cardiaco. La soluzione del problema accoppiato (3D/0D) avviene mediante la separazione in due sotto-sistemi, all'interno dei quali vengono risolte le equazioni di Navier-Stokes: il sistema 3D, risolto all'interno del modello anatomico 3D, e il sistema 0D, che riduce le equazioni che governano il moto ad un'equazione differenziale ordinaria equivalente appunto a quella di una rete elettrica, vengono risolti separatamente durante la simulazione. Nella figura 3 viene mostrato lo

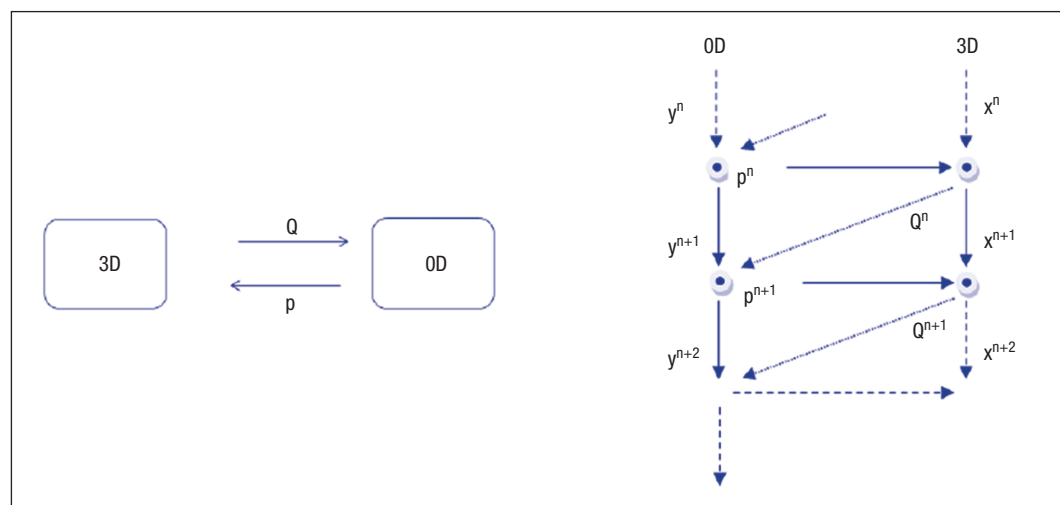


FIGURA 3
Schema di avanzamento in tempo dell'approccio multi-scala 3D-0D

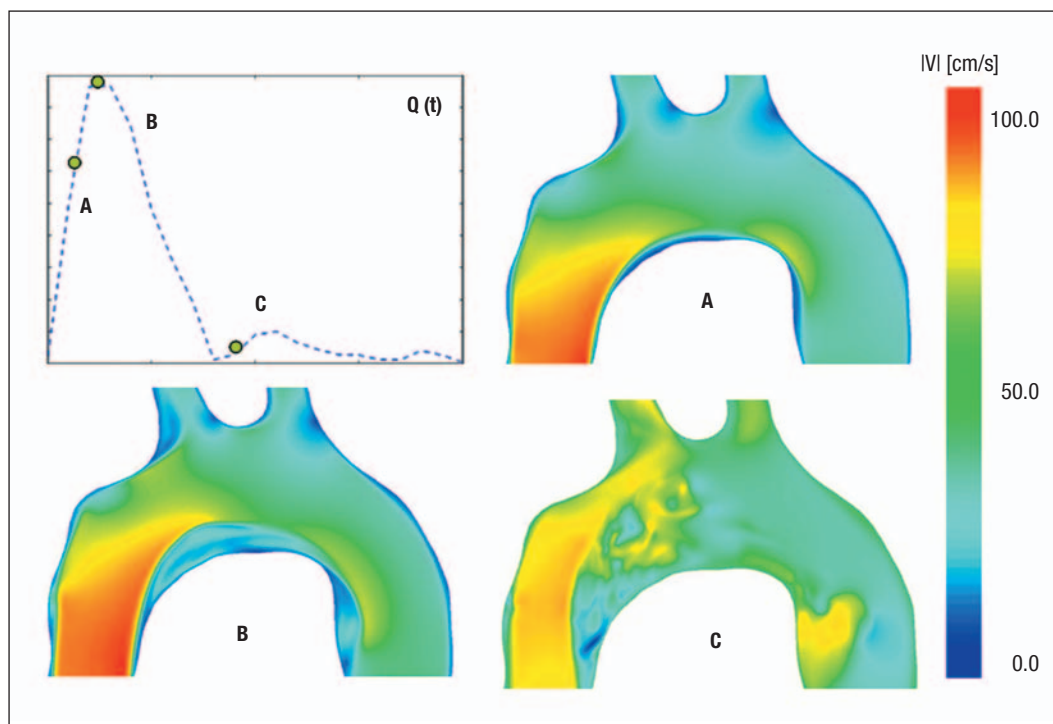


FIGURA 4

Campi di moto all'interno di un modello di arco aortico in diversi istanti del ciclo

schema di avanzamento adottato nell'evoluzione di questo tipo di modelli accoppiati. Anche se in teoria, ogni tipo di accoppiamento è possibile tra i due modelli, generalmente perché a livello clinico solo certe quantità ematiche sono misurabili, il modello 3D fornisce in ingresso al modello a parametri concentrati un dato di portata, che viene assunto come forzante per risolvere l'equazione differenziale associata alla rete e calcolare così il valore di pressione imposto sulle sezioni di uscita del modello 3D.

Attraverso l'accoppiamento, il valore di pressione sulla sezione di uscita nel modello 3D è assegnato in funzione dell'impedenza della rete a valle (modello 0D che simula la circolazione sistemica).

Grazie a questo tipo di approcci metodologici, ben identificati dal punto di vista matematico da vari gruppi di ricerca, è possibile implementare modelli CFD affidabili, in grado cioè di fornire campi di moto realistici che evolvono nel tempo secondo quanto osservato anche in analisi cliniche [8].

Riguardo all'accoppiamento fluido strutturale, necessaria per inserire nella modellazione l'effetto di deformazione delle pareti dei

vasi e peraltro di elevata complessità matematica e computazionale, si vuole in questa sede solo osservare che la strategia comunemente più utilizzata consiste nel risolvere in maniera consecutiva i due sottoproblemi (fluidodinamico e strutturale) accoppiando solutori specifici già validati. In particolare all'interno dei progetti sviluppati si annovera lo studio di interazione fluido-struttura per il problema di una valvola cardiaca impiantabile [9] ed oggi attualmente in commercio.

Dal punto di vista pratico è interessante osservare come software commerciali e *open source* sono oggi già a disposizione dei ricercatori come strumenti maturi per implementare in modelli CFD tutte queste "ragionevoli semplificazioni" che pur permettendo una valutazione realistica dei campi di moto calcolati rendono affrontabili dal punto di vista computazionale questi problemi di natura biologica che contengono un'elevata complessità intrinseca. In questo senso si vuole qui mostrare con la figura 4 l'evoluzione spazio-temporale della distribuzione di velocità all'interno di un arco aortico [10].

3. DISCUSSIONE

La necessità di sviluppare un modello computazionale dell'emodinamica di modelli vascolari assurge dall'esigenza di superare i notevoli limiti legati alla misura diretta, in vivo così come in vitro, o tramite modelli teorici del moto del sangue all'interno dei distretti vascolari. Infatti, da un lato la valutazione in vitro risente delle limitazioni intrinseche di quelle tecniche sperimentali storicamente usate in emodinamica, quali sono le tecniche anemometriche e quelle basate su velocimetria Doppler, a causa dell'elevata incertezza che le caratterizza.

Oggi, grazie a più moderne tecniche di misura basate sull'imaging medicale, è possibile ottenere delle misure in vivo di grandezze legate alla fluidodinamica e allo stato di sforzo della parete vascolare. Su tutte, uno strumento potenzialmente molto valido è rappresentato dalla risonanza magnetica (MRI, *Magnetic Resonance Imaging*). Tecniche innovative quali ad esempio la risonanza magnetica a contrasto di fase (PC MRI, *Phase Contrast Magnetic Resonance Imaging*) permettono di ottenere campi di moto 4D (3D nel tempo) del tutto analoghi a quelli ottenuti con la CFD come discusso in [10]. Tuttavia, al momento, un grave limite per quanto riguarda l'accuratezza spaziale e temporale sembra non permettere di effettuare con la dovuta precisione i calcoli di quegli indicatori, sia di vena fluida che di parete, che oggi vengono comunemente utilizzati nella valutazione dello stato emodinamico dei vasi arteriosi studiati tramite modelli numerici.

I metodi computazionali assumono quindi una notevole importanza poiché consentono di simulare realisticamente e con la dovuta accuratezza spazio-temporale lo stato fluidodinamico all'interno di vasi sanguigni anche in geometrie complesse e con pareti deformabili.

Poiché una ricostruzione accurata della geometria è basilare in quanto quest'ultima influenza fortemente i campi di moto, il modello computazionale in emodinamica viene oggi comunemente implementato a partire da immagini tomografiche soggetto-specifiche, tali cioè da consentire una valutazione fluidodinamica specifica per quel soggetto. Questo aspetto risulta di particolare interesse non solo per fornire un maggior livello di realismo nel calcolo dei campi di moto ma anche e soprattutto nello studio di patologie degenerative o in presenza di specifiche

configurazioni morfologiche del vaso tali da rendere inefficienti l'utilizzo di modelli generici e conoscenze teoriche pregresse.

Inoltre, dato che in vivo è noto che l'emodinamica in un singolo distretto è fortemente dipendente dal comportamento sistemico, qualora siano disponibili informazioni sufficienti (dati clinici sistemici in corrispondenza dei *boundary* del distretto vascolare) si cerca di tenere conto degli effetti della circolazione a monte e a valle dei distretti studiati. Per modellare tale interazione tra fluidodinamica locale e sistemica, si fa uso di modelli multi-scala accoppiando un modello 3D specifico del distretto con un modello 0D a parametri concentrati (equivalente elettrico della fluidodinamica sistemica dedotto dai dati clinici in corrispondenza dei *boundary* del distretto vascolare).

4. CONCLUSIONI E SVILUPPI FUTURI

Per quanto sin qui descritto e discusso, è possibile concludere che i modelli CFD risultano oggi un ambiente privilegiato per lo studio di problemi di emodinamica coinvolti nella comprensione di fenomeni infiammatori e degenerativi del sistema cardiovascolare. Queste problematiche sono di primissimo interesse per tutti i paesi occidentali sia da un punto di vista etico che socio-economico.

In questo contesto dunque l'uso della CFD si pone al di là delle sue più comuni applicazioni in altri ambienti di ricerca come ad esempio la validazione di procedure di misura e/o l'ottimizzazione di dispositivi e processi. Tuttavia anche in ambito biomedicale sono numerosi gli esempi di sviluppo per la progettazione di device medicali (impiantabili ed esterni) quali ad esempio valvole cardiache [9, 11] (come mostrato nella Figura 5), filtri arteriosi [12] e posizionatori magnetici per *bypass* [13]. In questo tipo di applicazioni ingegneristiche, la disponibilità di dati "a basso costo" e facilmente visualizzabili oltre che la facilità nel ripetere "esperimenti computazionali" per ottenere numerosi insiemi di dati ottimali in termini di accuratezza spazio-temporale per analisi statistiche e quantificazioni di indici sintetici, rappresenta senza dubbio un punto di forza a favore del sempre maggior utilizzo di queste tecniche.

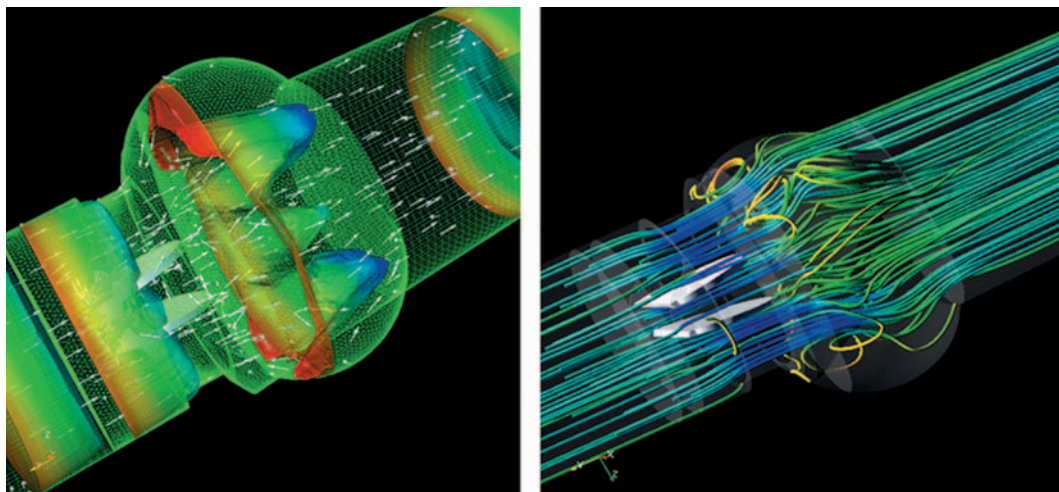


FIGURA 5

Visualizzazione delle linee di flusso, dei profili spaziali di velocità e dei campi di moto all'interno di un modello di valvola cardiaca impiantabile. Il modello CFD usato implementa un'interazione fluido-struttura

Esempio applicativo Ottimizzazione della stima della portata ematica per analisi di velocimetria Doppler in clinica

La stima della portata ematica all'interno dei vasi sanguigni è un problema di primario interesse in clinica al fine di valutare lo stato di perfusione dell'albero ematico e dei tessuti posti a valle di un punto di osservazione. Questo tipo di valutazione risulta cruciale ad esempio per valutare l'efficacia di interventi atti a ripristinare un livello sufficiente di perfusione dei tessuti come avviene comunemente durante procedure chirurgiche quali per esempio quelle di angioplastica o dell'effettuazione di *bypass* arteriosi. Grazie alla strumentazione clinica oggi disponibile è possibile ottenere con elevata precisione solo valori di velocità e non di portata del sangue. In particolare con tecniche di velocimetria Doppler basati sul cateterismo vascolare è possibile avere informazioni sul valore della velocità in un punto posto al centro del vaso sanguigno di interesse e, ponendo il catetere lontano da zone curve e da biforcazioni, è possibile ipotizzare che il profilo spaziale della velocità presenti una sostanziale simmetria assiale. Tuttavia a partire da questa sola informazione non è possibile calcolare il valore della portata. Ulteriori ipotesi sono pertanto necessarie. Una formula validata per stimare la portata in piccoli vasi (coronarie) e generalmente utilizzata per stimare il valore dell'integrale lungo un ciclo cardiaco della portata a partire dal solo valore di velocità di centro vaso è la seguente:

$$Q = 0.5 \times V_m \times A \quad (1)$$

dove V_m e A sono rispettivamente il valore di velocità massima in centro vaso e il valore dell'area della sezione dello stesso. Affinché i valori forniti da questa formula siano corretti è necessario che in media, nel tempo, la forma del profilo spaziale della velocità sia parabolico. Qualora questa formula venisse utilizzata in presenza di profili non noti a priori, risulterebbe inattendibile. Purtroppo nella pratica clinica molto spesso questa formulazione semplificata viene utilizzata anche per stimare il valore della portata ematica non solo come valore medio sul periodo in piccoli vasi ma anche per stimarne il valore all'istante di picco (sia esso sistolico o diastolico) in vasi di medio calibro. In queste condizioni tuttavia l'ipotesi di parabolicità del profilo di velocità necessaria per applicare con successo l'equazione 1 viene meno e le stime ottenute possono essere inesatte. Per questo motivo a partire dal 2006 sono stati effettuati degli studi basati su modelli e tecniche CFD al fine di costruire una formula più generale valida per effettuare delle stime della portata di picco in vasi di qualsiasi calibro utilizzabile senza modificare l'attuale pratica clinica nelle analisi di velocimetria Doppler [6, 7, 8]. La formula ottenuta è formalmente simile a quella di partenza fatto salvo per la presenza di un valore parametrico basato su una grandezza adimensionale (il numero di Womersley) e costruito grazie ad un'analisi statistica di centinaia di simulazioni numeriche effettuate in varie condizioni emodinamiche, reologiche e per vasi di differente forma e calibro. Da ogni simulazione il valore della velocità massima di centro vaso è stato utilizzato per effettuare stime secondo l'equazione 1 per poi calcolare il valore dell'errore rispetto al valore effettivo di

segue

portata (noto nelle simulazioni numeriche) e dalla curva degli errori ottenuti dalle varie simulazioni tramite un'ottimizzazione ed è stata costruita la formula che minimizza gli errori di stima al variare del numero di Womersley. La formulazione ottenuta è la seguente:

$$Q = Kw \times Vm \times A \quad (2)$$

dove Kw è funzione del valore del numero di Womersley definito come segue:

$$W = R \sqrt{2\pi f / \nu} \quad (3)$$

dove R è il raggio del vaso, f è la frequenza cardiaca e ν è la viscosità del sangue.

Pertanto, in questa nuova formulazione, al posto di utilizzare in maniera aprioristica un valore costante come fattore di moltiplicazione per stimare il valore dell'integrale della velocità sulla sezione del vaso, viene più opportunamente effettuata una stima che varia al variare del valore del numero di Womersley tenendo pertanto in considerazione sia l'effetto della frequenza cardiaca che il valore della sezione del vaso. La formula ottenuta, che è stata soprannominata "*Womersley number based*", è stata poi validata con successo in vivo utilizzando un protocollo di PC MRI su 10 soggetti sani [14] mostrando un significativo vantaggio nell'uso di questa formula (a parità di complessità di applicazione) con valori dell'errore di stima fino a 5 volte inferiori rispetto a quelli ottenuti con l'equazione 1 nelle stesse condizioni.

Questo è un esempio completo di applicazione delle tecniche CFD per l'ottimizzazione di procedure cliniche in cui la disponibilità di grandi quantità di dati controllabili, ripetibili e affidabili ma ottenibili a basso costo ha reso fattibile la determinazione di una formulazione più corretta attraverso una parametrizzazione su un vasto campione statistico di dati. Tuttavia è importante sottolineare come sia sempre necessario effettuare delle validazioni in vivo dei risultati ottenuti con tecniche numeriche al fine di valutare quanto le schematizzazioni indotte dal modello siano pertinenti con la complessità biologica dei fenomeni studiati.

Bibliografia

- [1] Allender S., Scarborough P., Peto V., Rayner M., Leal J., Fernandez R.L., Gray A.: *European cardiovascular disease statistics 2008*. European Heart Network, 2009.
- [2] Karino T.: Microscopic structure of disturbed flows in the arterial and venous systems, and its implication in the localization of vascular diseases. *International Angiology*, Vol. 5, n. 4, 1986, p. 297-313.
- [3] Karino T., Goldsmith H.L.: Role of blood cell-wall interactions in thrombogenesis and atherogenesis: a microrheological study. *Biorheology*, Vol. 21, n. 4, 1984, p. 587-601.
- [4] Caro C.G., Fitz-Gerald J.M., Schroter R.C.: Atheroma and arterial wall shear: observation, correlation and proposal of a shear dependent mass transfer mechanism for atherogenesis. *Proc R Soc Lond B Biol Sci*, Vol. 177, 1971, p. 109-159.
- [5] Gnasso A., Irace C., Carallo C., De Franceschi M.S., Motti C., Mattioli P.L., Pujia A.: In vivo association between low wall shear stress and plaque in subjects with asymmetrical carotid atherosclerosis. *Stroke*, Vol. 28, n. 5, 1997, p. 993-998.
- [6] Ponzini R., Vergara C., Redaelli A., Veneziani A.: Reliable CFD-based estimation of flow rate in haemodynamics measures. *Ultrasound Med Biol.*, Oct; Vol. 32, n. 10, 2006, p. 1545-55.
- [7] Vergara C., Ponzini R., Veneziani A., Redaelli A., Neglia D., Parodi O.: Womersley number-based estimation of flow rate with Doppler ultrasound: sensitivity analysis and first clinical application. *Comput Methods Programs Biomed.*, May, Vol. 98, n. 2, 2010, p. 151-60.
- [8] Ponzini R., Lemma M., Morbiducci U., Montevecchi F.M., Redaelli A.: Doppler derived quantitative flow estimate in coronary artery bypass graft: a computational multiscale model for the evaluation of the current clinical procedure. *Med Eng Phys.*, Sep, Vol. 30, n. 7, 2008, p. 809-16.
- [9] Nobili M., Morbiducci U., Ponzini R., Del Gaudio C., Balducci A., Grigioni M., Maria Montevecchi F., Redaelli A.: Numerical simulation of the dynamics of a bileaflet prosthetic heart valve using a fluid-structure interaction approach. *J Biomech.*, Aug 7, Vol. 41, n. 11, 2008, p. 2539-50.
- [10] Morbiducci U., Ponzini R., Rizzo G., Cadioli M., Esposito A., De Cobelli F., Del Maschio A., Montevecchi F.M., Redaelli A.: In vivo quantification of helical blood flow in human aorta by time-resolved three-dimensional cine phase contrast magnetic resonance imaging. *Ann Biomed Eng.*, Mar, Vol. 37, n. 3, 2009, p. 516-31.
- [11] Morbiducci U., Ponzini R., Nobili M., Massai D., Montevecchi F.M., Bluestein D., Redaelli A.: Blood damage safety of prosthetic heart valves. Shear-induced platelet activation and local flow dynamics: a fluid-structure interaction approach. *J Biomech.*, Aug 25, Vol. 42, n. 12, 2009, p. 1952-60.
- [12] Fiore G.B., Morbiducci U., Ponzini R., Redaelli A.: Bubble tracking through computational fluid dynamics in arterial line filters for cardiopulmonary bypass. *ASAIO J.*, Sep-Oct, Vol. 55, n. 5, 2009, p. 438-44.

[13] Morbiducci U., Lemma M., Ponzini R., Boi A., Bondavalli L., Antona C., Montevicchi F.M., Redaelli A.: Does the Ventrica magnetic vascular positioner (MVP) for coronary artery bypass grafting significantly alter local fluid dynamics? A numeric study. *Int J Artif Organs.*, Jul, Vol. 30, n. 7, 2007, p. 628-39.

[14] Ponzini R., Vergara C., Rizzo G., Veneziani A., Roghi A., Vanzulli A., Parodi O., Redaelli A.: Womersley Number-Based Estimates of Blood Flow Rate in Doppler Analysis: In Vivo Validation by Means of Phase-Contrast MRI. *IEEE Trans. on Biomed. Eng.*, Jul, Vol. 57, n. 7, 2007 p. 1807-1815.

RAFFAELE PONZINI ha una Laurea in Bioingegneria conseguita al Politecnico di Milano (2002) con una tesi di modellistica numerica applicata al campo dell'emodinamica (Analisi dei profili spaziali di velocità in un bypass coronarico da Y tramite un modello multiscala). Dal 2003 lavora all'interno del gruppo di calcolo ad alte prestazioni del CILEA dove è responsabile della gestione dei codici di fluidodinamica computazionale (CFD). Il suo campo di applicazione lavorativo include anche l'insegnamento di diversi linguaggi di programmazione (Python, C, C++) e di tecniche di programmazione ad oggetti. Nel 2007 ha conseguito un dottorato di ricerca in Bioingegneria presso il Politecnico di Milano con una tesi dal titolo "*Computational modelling of local haemodynamics phenomena: methods, tools and clinical applications*". I suoi interessi di ricerca includono la modellistica multiscala, il processamento di immagini e la visualizzazione scientifica. È attualmente autore di 13 articoli (*peer-reviewed*) internazionali nel campo della biomeccanica computazionale e di numerose pubblicazioni a congresso nel campo biomedicale e della modellazione CFD in ambito navale.
E-mail: ponzini@cilea.it

ICT E INNOVAZIONE D'IMPRESA

Casi di successo

Rubrica a cura di

Roberto Bellini, Chiara Francalanci

La rubrica *ICT e Innovazione d'Impresa* vuole promuovere la diffusione di una maggiore sensibilità sul contributo che le tecnologie ICT possono fornire a livello di innovazione di prodotto, di innovazione di processo e di innovazione di management. La rubrica è dedicata all'analisi e all'approfondimento sistematico di singoli casi in cui l'innovazione ICT ha avuto un ruolo critico rispetto al successo nel business, se si tratta di un'impresa, o al miglioramento radicale del livello di servizio e di diffusione di servizi, se si tratta di una organizzazione pubblica.



Ecotoys: il Web 2.0 al servizio della sostenibilità

Cinzia Cappiello

1. INTRODUZIONE

Gli ultimi anni sono stati caratterizzati da un'evoluzione del Web che ha fortemente cambiato l'interazione sito-utente. Infatti, dal Web inteso come strumento semplicemente orientato alla visualizzazione di contenuti, si è passati al Web 2.0 che raccoglie applicazioni progettate in linea alle esigenze degli utenti, caratterizzate da un'elevata interoperabilità e basate principalmente sulla condivisione di informazioni. Il Web 2.0 attribuisce un ruolo sempre più attivo all'utente dandogli la possibilità di pubblicare propri contenuti e accedere, modificare e commentare informazioni pubblicate da altri utenti. Per quanto riguarda le imprese, questa nuova forma di comunicazione dà, in primo luogo, la possibilità di aumentare la propria presenza sul Web tramite la pubblicazione di informazioni anche al di fuori del tradizionale sito Web aziendale. Inoltre, l'enorme mole di opinioni che gli utenti condividono su idee e prodotti costituiscono una notevole base di conoscenza che si può rivelare preziosa per un'efficace pianificazione delle azioni di marketing.

Nel presente articolo è descritto il caso di Ecotoys, un'azienda che ha deciso di adottare gli strumenti messi a disposizione dal Web 2.0 per

acquisire un maggior vantaggio competitivo e per diffondere la propria filosofia aziendale basata sullo sviluppo sostenibile e l'ecologia. In particolare, si vuole sottolineare come il Web 2.0 dia la possibilità alle piccole e medie imprese in generale di intraprendere campagne di marketing con un bassissimo costo di investimento.

2. ECOTOYS

Sempre di più si parla della “*green economy*” come una leva da sfruttare per dare una svolta alla crisi. Lo sviluppo sostenibile permette alle piccole e medie imprese di rivoluzionare il proprio business offrendo servizi nuovi e attenti alle moderne esigenze del mercato. La sostenibilità mira alla riconciliazione di tre pilastri fondamentali quali l'ambiente, l'economia e la società [1]. Questi tre aspetti, infatti, non sono mutuamente esclusivi e il loro sfruttamento congiunto può portare a notevoli risultati. Soprattutto in questi ultimi anni in cui la gravità dei problemi ambientali, dall'inquinamento ai cambiamenti climatici, dalla crisi energetica alla riduzione della biodiversità, accresce di giorno in giorno la rilevanza della “questione ecologica” e sempre più iniziative sono indirizzate a questi temi [2].

Un esempio di piccola media impresa che ha ba-

sato la propria attività secondo tutte le linee guida dettate dalla *green economy* è Ecotoys, produttrice e importatrice di prodotti e giocattoli ecologici con sede vicino a Como. Ecotoys nasce nel 2004 con due grandi obiettivi: diffondere prodotti ecosostenibili e contribuire alla creazione di un'etica della sostenibilità ambientale. L'azienda ha iniziato la propria attività producendo e commercializzando HappyMais, mattoncini realizzati in Mater-Bi che si possono attaccare l'uno all'altro semplicemente inumidendoli. Il Mater-Bi è un materiale alternativo alla plastica derivato dall'amido di mais e brevettato da un'azienda italiana. Ecotoys ha progettato una macchina per realizzare mattoncini colorati, valida versione ecologica ai classici in plastica. Con il tempo l'azienda ha ampliato il proprio catalogo con numerosi prodotti ecologici quali i primi soldatini biodegradabili al mondo, giocattoli solari, giocattoli in legno FSC, giocattoli equosolidali, giochi a manovella da usare senza pile e molto altro ancora.

Questo orientamento "*green*" ha avuto un impatto non solo sulla definizione dei prodotti da commercializzare ma anche sulla concezione del sistema impresa. Ecotoys è una vera bio-azienda che mira a migliorare anche il territorio in cui opera sia da un punto di vista ambientale che culturale. Infatti, l'azienda è immersa in un ampio spazio verde che ospita diverse coltivazioni e animali e arricchisce la propria attività facendosi spesso promotrice di incontri culturali per la sensibi-

lizzazione al tema della sostenibilità e rispetto per l'ambiente. L'azienda in questa visione mira ad essere un posto accogliente sia per i lavoratori sia per gli acquirenti che si recano allo spaccio aziendale.

Il marchio Ecotoys è stato diffuso dal 2004 al 2009 principalmente attraverso canali di vendita pressoché diretti. Infatti, la strategia di marketing ha sempre mirato prevalentemente al posizionamento di stand dimostrativi e di commercializzazione dei prodotti presso centri commerciali, grossisti, *garden center*, punti di vendita di giocattoli e articoli da regalo. Ecotoys ha approcciato l'utilizzo del Web tramite la realizzazione di un sito informativo corredato dalle classiche funzionalità di commercio elettronico. Tuttavia le potenzialità dell'ICT sono sempre state sottoutilizzate soprattutto per un'azienda che ha la volontà di diffondere una filosofia aziendale innovativa e un forte messaggio per uno stile di vita più consapevole e rispettoso della natura.

Nel 2009 Guido Gerletti, il proprietario dall'azienda, vista soprattutto la continua crescita dell'interesse nella *green economy* e l'aumento del numero di persone che è interessato all'argomento e privilegia prodotti ecologici, decide di investire maggiormente nell'ICT e, in particolare, inizia a valutare l'adeguatezza degli strumenti offerti dal Web 2.0 per il miglioramento del posizionamento aziendale.

3. IL WEB 2.0 COME STRUMENTO DI WEB MARKETING PER LE PICCOLE E MEDIE IMPRESE

Il Web 2.0 si pone al centro degli strumenti più efficaci per la realizzazione di strategie di marketing. Infatti, alla domanda: "Quali strategie di marketing *digitali* enfatizzeresti nel 2009?", uno studio ha evidenziato che le prime dieci strategie considerate più efficaci sono quelle riportate in tabella 1.

La tabella 1 conferma che gli strumenti tipici del Web 2.0 (esempio, *blogging* o *social network*) occupano le prime posizioni della classifica stilata nel sondaggio oggetto di studio. Dette soluzioni consentono, inoltre, di massimizzare l'efficacia del marketing tagliando nettamente i costi. *Bloggare* o partecipare a *social network* sono attività a costo ridotto rispetto alle tradizionali strategie di marketing (esempio, pubblicità tradizionale o digitale). Il Web 2.0 diventa quindi

Strategia	Preferenze
Blogging	34%
Microblogging (Twitter)	29%
Search engine optimization	28%
Social network participation (Facebook, LinkedIn)	26%
Email marketing	17%
Social media monitoring & outreach	17%
Pay per click	14%
Blogger relations	12%
Video marketing	10%
Social media advertising	7%

TABELLA 1

Valutazione delle strategie di marketing digitali [3]

una risorsa preziosa che, soprattutto, le piccole e medie imprese possono sfruttare per ottenere dei vantaggi competitivi. Oltre al palese vantaggio economico, è necessario aggiungere che l'uso di queste soluzioni è alla portata di tutti dal momento che non richiedono particolari competenze informatiche.

Per esempio, per quanto riguarda il *blogging*, la tecnologia per realizzare un blog è disponibile a un costo minimo. Vi sono numerose piattaforme *open source* (esempio, WordPress, Blogger o Splinder) che consentono a chiunque di aprire un blog e di iniziare a comunicare con clienti potenziali o fidelizzati. Dette piattaforme sono dei sistemi molto semplici in cui l'inserimento dei contenuti (testi, immagini, video ecc.) richiede una competenza assimilabile a quella necessaria per l'utilizzo di un normale applicativo di videoscrittura. La questione più importante, quando si progetta l'apertura di un blog è la scelta dei contenuti e lo stile da utilizzare. Infatti, il blog viene spesso usato, soprattutto dai marchi più conosciuti, per avvicinare e fidelizzare i clienti e per avere opinioni e *feedback* dai propri clienti o utilizzatori nella fase di promozione di un nuovo prodotto/servizio. Gli stili usati dalle aziende sono svariati: ci sono grandi aziende che hanno scelto di aprire blog focalizzati su singoli prodotti di loro produzione, aziende che raccontano segreti delle loro lavorazioni o curiosità del settore oppure aziende che comunicano con i propri clienti attraverso personaggi di fantasia.

Le piccole e medie imprese, e soprattutto le realtà artigianali in cui le competenze e l'esperienza sono la risorsa principale dell'azienda, hanno una grossa opportunità con il blog. Infatti, oltre ad essere uno strumento eccellente per intraprendere azioni di marketing, è un valido mezzo per condividere tecniche di lavorazione, trucchi del mestiere, o per semplicemente enfatizzare la cura per il design che costituiscono il vero punto di forza di tante aziende italiane. Altro vantaggio fondamentale del blog è che è anche uno strumento di *search engine optimization* in quanto consente di migliorare il posizionamento dell'azienda sui motori di ricerca e di conseguenza aumentare il traffico del sito. Il blog, infatti, ha una struttura che permette ai motori di ricerca di indicizzare i contenuti con grande facilità.

Di fianco al *blogging*, altra tecnica fundamenta-

le per le piccole e medie imprese vi è l'*E-mail marketing* che prevede l'uso delle e-mail per mantenere le relazioni con i clienti. Si può far ricadere sotto questo nome tante diverse forme di comunicazione tra cui le classiche newsletter aziendali. Anche in questo caso l'azienda deve definire molto attentamente e a priori qual è l'obiettivo principale a cui la newsletter deve puntare. Si può scegliere di mandare degli aggiornamenti, quali le informazioni sui nuovi prodotti e sulle ultime offerte oppure decidere di porre l'accento su un'approfondita conoscenza dell'azienda, inserendo articoli riguardanti lo staff e i processi interni.

Attraverso tutti i mezzi che mette a disposizione, il Web marketing dà anche la possibilità di conoscere e analizzare le preferenze degli utenti. Infatti, attraverso il monitoraggio dei siti aziendali (blog inclusi), molti strumenti (esempio, Google Analytics, ShinyStat) offrono tecniche di *data mining* capaci di studiare l'attività dei visitatori dei siti web analizzando: i modelli del traffico web, la navigazione del sito, le viste delle pagine e le transazioni al fine di comprendere gli interessi dei clienti, migliorare il design e la navigazione del sito Web e aumentare la redditività dei processi di vendita *on-line*.

4. IL CASO DI STUDIO: VERSO UN UTILIZZO STRATEGICO DEL WEB

L'analisi della situazione di Ecotoys ha messo in luce la potenzialità di raggiungere un forte vantaggio competitivo con lo sfruttamento del Web. L'obiettivo primario dell'azienda è quello di potenziare il canale Web per aumentare le vendite e diffondere il codice etico aziendale che risiede nel fatto di "lavorare in armonia con sé stessi e con la natura, dare informazioni trasparenti al consumatore sulla provenienza e sulla qualità dei prodotti, lavorare in una rete di relazioni armoniose e basate sul confronto e sul rispetto reciproco" (www.ecotoys.it). È stato definito per l'azienda un piano operativo di attività pensate per il rafforzamento del sito Web da una parte e per lo sviluppo di iniziative orientate al Web marketing dall'altra. In primo luogo, una riprogettazione del sito Web dell'impresa si è resa necessaria per veicolare il giusto messaggio alla clientela e per avere una buona base informativa per tutte le altre iniziative. In secondo luogo, l'azienda aveva la duplice esigenza di pensare alla

propria attività commerciale e alla diffusione della filosofia aziendale con iniziative di sensibilizzazione su diversi temi specifici. La valutazione di strumenti Web 2.0 ha portato alla decisione di appoggiarsi al *blogging*, alla *search engine optimization* e all'*e-mail marketing*. Il blog è, infatti, lo strumento più semplice ed efficace per diffon-

dere il progetto della bio-azienda e per divulgare notizie sulla crisi economica e ambientale. Vista l'eterogeneità dei contenuti da pubblicare si è pensato all'apertura di più blog:

- blog aziendale di stile informativo con notizie e dettagli sui prodotti che ha il fine primario di aggiornare i clienti e di informarli sulle dinamiche aziendali oppure sugli eventi eco-solidali in cui l'azienda interviene;

- blog informativi mirati alla divulgazione di contenuti incentrati su diverse tematiche.

La figura 1 riassume i principali siti che EcoToys ha deciso di aprire per potenziare la propria presenza sul Web e sensibilizzare gli utenti alle problematiche che sono alla base del progetto aziendale. In particolare, i blog sono progettati per la diffusione del messaggio aziendale e per avere riscontri sui prodotti commercializzati e le iniziative intraprese. L'attività commerciale invece è supportata dal sito principale veicolo di tutte le informazioni aziendali, da un sito di commercio elettronico (Eco-Spaccio) e da un sito dedicato a HappyMais, prodotto di punta dell'azienda. Tutti i siti sono tra di loro collegati attraverso link reciproci per ottimizzare il posizionamento dell'azienda sui motori di ricerca.

La figura 2 mostra il nuovo sito realizzato da EcoToys. La grafica è stata pensata come primo vei-

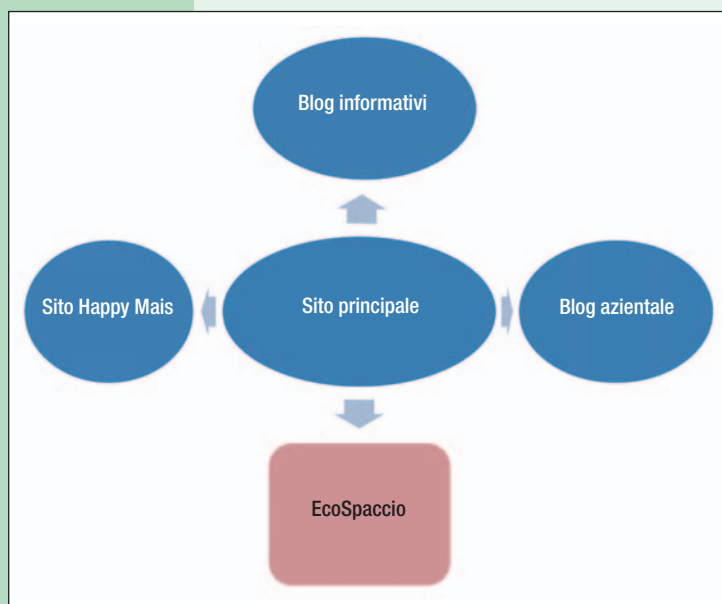


FIGURA 1

La definizione dei siti sui cui poggia la strategia di Web Marketing di EcoToys



FIGURA 2

Homepage del sito aziendale

colo per la trasmissione del messaggio aziendale: una fabbrica immersa nel verde e il ritorno a giochi tradizionali fabbricati con materiali naturali. L'homepage è la porta d'accesso ai diversi siti di Ecotoys: i diversi link portano ai diversi siti divulgativi e alla vendita dei prodotti aziendali. Tutti i siti finora sviluppati sono rappresentati nella figura 3. L'azienda sta ancora progettando i siti satelliti aggiuntivi. Detti siti saranno rivolti a utenti specifici (esempio, mamme, scuole) o saranno focalizzati su prodotti specifici (esempio, ecobomboniere) o indirizzati verso tematiche specifiche (esempio, diversi problemi ecologici). L'azienda ha anche iniziato la propria attività di *e-mail marketing* preparando una newsletter. La newsletter dà, infatti, la possibilità di iniziare a sfruttare il canale Web per promuovere iniziative, prodotti o offerte commerciali. Avendo diverse tipologie di utenza, si ritiene altresì opportuno distinguere due tipi di newsletter: una pensata per i clienti "retail" e un'altra pensata per i clienti "corporate" che nel contesto aziendale possono essere, per esempio, scuole e negozi. Inserendo notizie e suggerimenti nonché informazioni rilevanti sull'azienda, la newsletter è uno strumento molto efficace per stabilire e mantenere relazioni con i clienti evidenziando notizie utili per le diverse categorie di utenza.

La raccolta delle caratteristiche e delle preferenze dei clienti diventa fondamentale per implementare in futuro funzionalità di *data mining* che permettano di analizzare l'attività dei visitatori dei siti Web studiando i modelli del traffico Web, la navigazione del sito e le viste delle pagine e le transazioni al fine di comprendere gli interessi dei clienti e intraprendere campagne personalizzate di marketing.

5. RISULTATI DEL PROGETTO

I risultati raggiunti tramite l'utilizzo del Web 2.0 come veicolo di comunicazione e marketing dimostrano nettamente come la presenza e le attività sul Web siano strettamente correlate con l'attività commerciale. Il sito ufficiale, i blog correlati e la newsletter sono stati tutti messi *on line* e avviati nel mese di novembre 2009. La newsletter è stata inviata a un target fidelizzato di sole 400 persone di area *retail*, con immediato riscontro di circa 100 visite al giorno e con una media di tempo per visita superiore ai 2 min. In dicembre e gennaio le visite sono leggermente calate facendo registrare una media intorno al 75% dei picchi di lancio, ma con importanti puntate verso l'alto in coincidenza con azioni sul territorio o azioni di *cross web-marketing* (per esempio un calen-



FIGURA 3
Siti realizzati da Ecotoys

dario di beneficenza vicino a Natale). In breve si può dire che la vivacità commerciale delle feste ha avviato un trend di crescita, poi consolidatosi in gennaio e durato costantemente fino alla fine di marzo 2010, in controtendenza rispetto ai consueti trend del mercato di settore, che dopo il Natale vedono una lunga pausa pluristagionale.

Un trend simile contraddistingue le visite al sito www.ecotoys.it e ai blog correlati. Infatti, dopo il lancio nel mese di novembre 2009 si sono registrate 2500 visite mensili durante l'inverno, mantenendosi intorno alle 2000 visite in primavera e calando solo in estate. Questo elevato numero di accessi ha avuto un enorme beneficio sui siti correlati a partire da <http://shop.ecotoys.it>, che ha registrato un incremento del 20% delle visite nel periodo di lancio del nuovo sito istituzionale (arrivando intorno alle 6000 visite mensili intorno alle feste di Natale). Stesso andamento è stato riscontrato sul sito di prodotto www.happymais.it.

In sostanza, la nuova strategia di Web marketing, la comunicazione attraverso i blog e le iniziative di sensibilizzazione a favore di uno stile di vita sostenibile hanno portato un netto miglioramento delle vendite rispetto all'anno precedente. L'incremento dei valori assoluti di acquisto online segnala un aumento della fidelizzazione di alcuni clienti, mediata anche dal miglior impatto comunicativo del nuovo sito istituzionale (da cui arriva buona parte del traffico di acquisto).

Altro dato interessante, evidente soprattutto nella fase successiva al lancio e nonostante la targettizzazione della promozione a contatti già noti, è la percentuale di nuovi visitatori sempre superiore all'80%, indice di un efficace passaparola digitale.

6. CONCLUSIONI

Molti imprenditori, specialmente delle piccole e medie imprese, non considerano le grandi opportunità che Internet offre. In passato, l'utilizzo dell'IT per diffondere il proprio marchio non è stato considerato per gli elevati costi che le diverse iniziative richiedevano o per la mancanza di competenze. Il Web 2.0 abbatte queste barriere eliminando totalmente o quasi i costi di investimento e non richiedendo particolari conoscenze tecniche. Infatti, realizzare un blog per esempio ha un costo nullo e può essere ag-

giornato facilmente. I vantaggi risiedono nell'immediata presenza sul Web grazie al fatto che i blog sono uno strumento efficace per avere migliori *ranking* nei motori di ricerca; inoltre, i blog e in genere tutti gli strumenti offerti dal Web 2.0 (esempio, *microblogging* o partecipazione a *social network*) sono un efficace strumento per entrare in contatto con i propri clienti. L'importante è fornire contenuti sempre aggiornati, possibilmente interessanti, e consoni al settore merceologico in cui l'impresa si trova. Solo in questo modo gli articoli possono suscitare interesse e incentivare gli utenti a inserire commenti e opinioni sul prodotto o servizio in questione. Può inoltre essere un buon mezzo per pubblicizzare le novità o le offerte dell'attività aziendale. Attraverso questi strumenti l'azienda può aumentare la propria reputazione e credibilità, la fidelizzazione dei propri clienti e quindi la probabilità che le vendite aumentino nel breve periodo [4].

BIBLIOGRAFIA

- [1] Adams W.M.: *The Future of Sustainability: Rethinking Environment and Development in the Twenty-first Century*. Report of the IUCN Renowned Thinkers Meeting, 29–31 January 2006. Disponibile on line: http://cmsdata.iucn.org/downloads/iucn_future_of_sustainability.pdf.
- [2] La Vergata A., Ferrari G.: *Ecologia e sostenibilità. Aspetti filosofici di un dibattito*. Franco Angeli Editore, 2008.
- [3] Odden L.: *Top Digital Marketing Tactic for 2009: Social Media Marketing*. 2009 disponibile on line in Top Rank – On line Marketing Blog: <http://www.toprankblog.com/2009/03/social-media-marketing-tops-digital-marketing-tactics-for-2009/>
- [4] Gilroy K.: *Web 2.0 for Business Advantage, report from The Otter Group*. 2006 disponibile on line: <http://www.ottergroup.com/wp-content/themes/ottergroup/papers/Web%202.0%20for%20Business%20Advantage.pdf>

CINZIA CAPIELLO, Ricercatrice presso il Dipartimento di Elettronica e Informazione del Politecnico di Milano, dove svolge attività di ricerca nell'area dei sistemi informativi prevalentemente su temi che inerenti la gestione e la valutazione della qualità dei dati.
E-mail: capiell@elet.polimi.it

ICT E DIRITTO

Rubrica a cura di

Antonio Piva, David D'Agostini

Scopo di questa rubrica è di illustrare al lettore, in brevi articoli, le tematiche giuridiche più significative del settore ICT: dalla tutela del *domain name* al *copyright* nella rete, dalle licenze software alla *privacy* nell'era digitale. Ogni numero tratterà un argomento, inquadrandolo nel contesto normativo e focalizzandone gli aspetti di informatica giuridica.



Diritto d'autore e *fair use* nelle nuove tecnologie

David D'Agostini, Antonio Piva, Maurizio Sala

1. INTRODUZIONE

Tramite internet si accede ad una quantità di risorse e di fonti di informazione pressoché inesauribile, che vengono quotidianamente e costantemente incrementate.

L'avvento di internet e, soprattutto, la sua diffusione (conseguenza di una serie di fattori quali, tra molti altri, la diminuzione dei costi dei computer e delle connessioni, il miglioramento, anche nella facilità d'uso, dei *browsers*, il perfezionamento dei motori di ricerca, l'aumento dei servizi disponibili sul web) hanno cambiato il modo di vivere il quotidiano, le relazioni interpersonali e anche i profili di diritto ad esse applicabili.

Giusto per fare un esempio, nel mondo analogico, tempo e luogo sono concetti giuridicamente determinabili, di modo che un'azione, compiuta in un momento e in un luogo preciso, è sottoposta a norme e giurisdizione facilmente individuabili. Internet ha relativizzato concetti chiari e metabolizzati dalla scienza giuridica introducendone di nuovi, come quello di delocalizzazione. Un articolo pubblicato in internet è diffuso contemporaneamente in tutto il globo e poco importa che il suo autore si trovi in Alaska o in Nuova Zelanda, l'internauta italiano ne potrà disporre contemporaneamente a qualunque altro abitante della terra che, come lui, acceda alla "rete delle reti".

Le informazioni reperibili in internet sono, nella stragrande maggioranza dei casi, gratuite.

Alcuni assumono, anzi, che l'informazione in generale (indipendentemente dal modo di diffusione) dovrebbe essere esclusivamente gratuita, consentendo al genere umano di beneficiarne e con ciò progredire senza più alcuna barriera economica, sociale e di stato delle persone.

La facilità di accesso e di diffusione delle informazioni non può essere però confusa con l'anarchia e, entrando nel tema oggetto del nostro articolo, con l'uso indiscriminato delle fonti, quasi che qualsiasi cosa, frutto del lavoro e dell'intelligenza altrui, possa essere liberamente utilizzato per il semplice fatto di essere stato reperito ovvero diffuso in internet. Strumenti tecnologicamente avanzati, che oggi sono alla portata di tutti ci danno la possibilità di condividere, con un numero sempre maggiore di persone, un'enorme mole di dati, spesso non considerando gli sforzi con cui questi sono stati prodotti.

Di seguito affronteremo, evidentemente senza pretendere di esaurire la materia, il tema del diritto d'autore e delle modalità e dei limiti con cui l'opera altrui può esser utilizzata da terzi, con particolare riguardo al mondo internet, senza violare i diritti del soggetto cui essi competono.

2. IL DIRITTO D'AUTORE

Parlando di diritto d'autore in un contesto delocalizzato quale è quello disegnato dal web, va subito chiarita la "diversa visione" che di esso si ha nei Paesi di *civil law* tra cui l'Italia,



da quelli di diritto anglosassone (nei quali vi-
ge il diritto denominato *common law*).

Un'idea, in quanto tale, non è tutelata. Oggetto di
tutela è la concretizzazione dell'idea, vale a dire
la sua esteriorizzazione in una forma (che per le
arti figurative potrebbe essere, tra le altre, la pit-
tura, il disegno ovvero la fotografia e per le lettere
un testo oppure un trattato scientifico)¹. Ciò si
evincesse chiaramente dall'art. 2576 del Codice Civi-
le italiano ove il diritto d'autore viene individuato
nell'espressione, vale a dire nel prodotto, del la-
voro individuale. Una volta realizzata l'opera,
l'autore gode per tale (senza necessità di attività
ulteriore come per esempio la registrazione per il
caso del brevetto) di una serie di diritti riconducibi-
li a due grandi categorie: il diritto morale a tute-
la della personalità dell'autore e il diritto patrimoniale
di sfruttamento economico dell'opera.

La norma italiana di riferimento è la legge 22 apri-
le 1941 n. 633, modificatasi nel corso del tempo
anche a seguito del recepimento dei trattati inter-
nazionali e di alcune direttive europee.

Il diritto morale sull'opera si traduce nei diritti di
rivendicazione della paternità dell'opera, dell'in-
tegrità della stessa e di ritiro dal mercato.

Il diritto patrimoniale, a propria volta, si esprime
in vari modi, alcuni dei quali codificati nei diritti di
pubblicazione (e, quindi, anche di non pubblica-
zione, ovvero di inedito), di riproduzione, di di-
stribuzione, di comunicazione al pubblico (pre-
sente e non presente), di trasformazione, elabo-
razione e modificazione.

Nel diritto anglosassone il diritto d'autore è ri-
conducibile al *Copyright* (diritto di copia) che non
conosce il diritto morale dell'autore, ancorché la

giurisprudenza ne abbia tenuto conto nelle pro-
prie decisioni.

Il *Copyright* storicamente nasce in Inghilterra in-
torno al XVI secolo per assolvere alla funzione di
esercitare la censura sulle opere pubblicate².

Ai giorni nostri la diffusione di internet ha fatto
venire meno uno dei presupposti del *Copyright*,
vale a dire il costo per la produzione e la diffusi-
one delle opere³.

Questa è la ragione del recente braccio di ferro
tra l'agente letterario newyorkese Andrew Wyle
che ha stipulato con Amazon.com (la più grande
libreria *on-line* al mondo) un accordo per la di-
stribuzione di classici della letteratura moderna
in versione esclusivamente digitale (*e-book*),
aggirando le case editrici e la più importante ca-
sa editrice americana Randohom House che af-
ferma di detenere i diritti di distribuzione anche
digitale di quelle opere.

Andrew Wyle assume che la sua decisione sia
stata conseguente al rifiuto degli editori di rive-
dere l'ammontare dei loro compensi a favore
degli autori, posto che nell'era digitale i costi di
stampa, distribuzione, stoccaggio e invenduto
sono ben minori di quelli che un tempo hanno
consentito agli editori di trattenere la maggior
parte dei proventi della vendita dei libri; se poi
si trasferisce il ragionamento agli *e-book* i costi
appena enunciati si riducono a una cifra prossi-
ma allo zero.

La Society of Authors britannica ha, per parte
sua, denunciato come inique le *royalties* che gli
editori offrono agli autori di *e-book* (dal 10% al
25%): infatti, amazon.com riconosce agli autori
royalties in misura del 70%.

¹ Art 2 Convenzione di Berna (ratificata in Italia con legge 20 giugno 1978 n 399) "L'espressione «opere letterarie ed artistiche» comprende tutte le produzioni nel campo letterario, scientifico e artistico, qualunque ne sia il modo o la forma di espressione"; art. 2576 cod. civ. Acquisto del diritto. "Il titolo originario dell'acquisto del diritto di autore è costituito dalla creazione dell'opera, quale particolare espressione del lavoro intellettuale".

² Infatti, la presenza dei mezzi meccanici di riproduzione delle opere letterarie consentiva la diffusione tra la popolazione di scritti a vario oggetto. Per evitare che venissero messi in circolazione scritti contrari alla monarchia ovvero non graditi alla stessa, la corona britannica demandò alla corporazione dei librai (editori) di Londra il compito esercitare l'attività di controllo. Agli editori vennero quindi concessi i diritti di copia su ogni stampa prodotta: la concessione prevedeva non solo il diritto esclusivo di stampa, ma anche quello di confisca delle stampe e dei libri non autorizzati nonché di bruciare i testi non autorizzati. Gli autori non avevano, quindi, alcun diritto di proprietà sulla loro opera.

³ L'avvento di nuovi strumenti, come l'ipad o kindle, ha stravolto anche l'utilizzo di documenti elettronici che finora erano stati poco fruibili: si pensi alla difficoltà incontrata nel leggere un romanzo a monitor CRT ovvero all'impossibilità di sfogliare, con gesto naturale, le pagine di un quotidiano in un 32" lucido. Importanti giornali periodici hanno già implementato la loro versione digitale appositamente creata per questi *e-book reader*; le loro pagine scaricabili da chiunque, in qualsiasi parte del mondo si trovi, senza la necessità di passare in edicola, sono visualizzate in versione *old style* e su un dispositivo che rispetta alla carta stampata, lasciateci la battuta, ha l'unico difetto di non sporcare le mani.

L'era dell'impossibilità di utilizzo del portatile a letto per difficoltà di alimentazione, luce del monitor troppo intensa, difficoltà di approvvigionamento dei documenti è finita, è però palese la necessità di tutelare questi file, o meglio i loro autori.

In linea di principio, i paesi di *common law* offrono una tutela diversa rispetto alla legislazione italiana anche sotto il profilo dei diritti patrimoniali: in tali stati il diritto patrimoniale d'autore non è un diritto unitario con diverse manifestazioni, al contrario vi sono tanti *Copyright* quanti sono i media tramite i quali può essere diffusa l'opera (si pensi allo stesso film trasmesso su canale analogico o satellitare o via cavo ecc.).

Quanto al profilo regolamentare del diritto d'autore il primo atto sovranazionale è stata la Convenzione di Berna sulla protezione delle opere artistiche e letterarie (firmata a Parigi nel 1886 e recepita nel nostro ordinamento nella sua ultima revisione con legge 20 giugno 1978 n. 399).

A livello europeo ricordiamo, tra le altre, la direttiva 92/100 del 19 novembre 1992 (recepita in Italia con DLgs. 16 novembre 1994 n. 685) sul diritto di noleggio e di prestito degli originali e delle copie delle opere protette⁴ e la direttiva 93/82 del 27 novembre 1993 (recepita con DM 14 ottobre 1996 sui diritti d'autore applicabili alla radiodiffusione via satellite e via cavo).

In epoca più recente citiamo ancora:

□ il D.Lgs. 9-4-2003 n. 68 che ha dato attuazione della direttiva 2001/29/CE sull'armonizzazione di taluni aspetti del diritto d'autore e dei diritti connessi nella società dell'informazione, modificando la legge 22 aprile 1941, n. 633 sul diritto d'autore.

□ Il DLgs 16-3-2006 n. 140 che ha dato, poi, attuazione alla direttiva 2004/48/CE⁵ sul rispetto dei diritti di proprietà intellettuale apportando ulteriori modifiche alla legge 22 aprile 1941, n. 633 sul diritto d'autore.

Non si citano – perché al di fuori del tema in esame – le norme e direttive attinenti al software e ai database, anche se ci pare doveroso richiamare l'attenzione del lettore alla tutela giuridica dei programmi per elaboratore, in quanto anch'essi vengono tutelati dal legislatore median-

te il diritto d'autore. In particolare vengono tutelati i programmi espressi in qualsiasi forma, purché originali e/o creativi, unitamente al materiale preparatorio per la progettazione dello stesso; non rientrano nella previsione normativa i principi che stanno alla base delle interfacce uomo-macchina del software (si pensi alla contrapposizione Apple e Microsoft riguardante l'interfaccia a icone dei propri sistemi operativi). Viene pertanto estesa al software la tutela riconosciuta dal diritto d'autore alle opere letterarie, determinando i diritti esclusivi (moralì e patrimoniali) dei quali i soggetti titolari possono avvalersi e la relativa durata⁶.

Il Diritto d'autore nelle nuove tecnologie ha rilevanza anche nei siti web, nelle e-mail, e nelle moderne applicazioni quali blog o *social network* dove permangono diritti patrimoniali e morali in capo all'autore.

Infatti, a tutte le opere dell'ingegno presenti su Internet (testi, fotografie, musica ecc.), nonché ai siti web medesimi (qualora siano creazioni intellettuali nuove e originali) si estende la legislazione sul diritto d'autore, secondo la quale chi realizza l'opera ha *“il diritto esclusivo di utilizzare economicamente l'opera in ogni forma e modo, originale e derivata”*.

Per quanto riguarda i siti, come noto, le pagine web sono realizzate tramite un linguaggio denominato HTML (*Hypertext Markup Language*) che il *web browser* è in grado di interpretare leggendone le istruzioni, cioè gli appositi *tag* inseriti nelle pagine. Un particolare tag è costituito dal link che consente di collegarsi direttamente a un'altra pagina web, anche di un sito terzo.

La legittimità del link interno (il *deep link*, che collega a una pagina interna di un altro sito) è assai dibattuta in quanto alcuni sostengono che l'“aggiramento” della *home page* possa costituire una violazione della proprietà intellettuale provocando un danno al titolare del sito (per

⁴ La direttiva è stata abrogata dall'articolo 14 della direttiva 2006/115/CE, concernente il diritto di noleggio, il diritto di prestito e taluni diritti connessi al diritto di autore in materia di proprietà intellettuale.

⁵ La direttiva è denominata, in inglese, *“Directive on the enforcement of intellectual property rights”*, è nota anche con l'acronimo IPRED o anche come *“Direttiva enforcement”*. In data 25 aprile 2007 il Parlamento Europeo ha adottato la Risoluzione legislativa sulla proposta modificata di direttiva del Parlamento europeo e del Consiglio relativa alle misure penali finalizzate ad assicurare il rispetto dei diritti di proprietà intellettuale (COM(2006)0168 – C6-0233/2005 – 2005/0127(COD)) cd. IPRED2

⁶ La materia è disciplinata dalla Direttiva CEE del 14 maggio 1991 n. 250; il legislatore italiano ha recepito tale normativa mediante il Decreto Legislativo 29 dicembre 1992 n. 518 che ha aggiornato la già citata Legge sul diritto d'autore 22 aprile 1943, n. 633. Con queste disposizioni vengono aggiunte, inoltre, norme speciali inerenti la copia di *backup*, la decompilazione e il *debugging*. Per una completa trattazione sulla Tutela del Software si rimanda all'articolo *“La tutela giuridica dei programmi per elaboratore”* pubblicato in Mondo Digitale n. 1 - marzo 2003.



esempio, la diminuzione dei proventi pubblicitari). Viene considerato illecito il *framing* (all'interno della cornice del sito di partenza viene caricata ed inclusa la pagina di un altro sito) senza che l'avente diritto abbia previamente espresso il proprio consenso, in quanto normalmente una pagina web è costituita da opere protette dal diritto d'autore quali testi, fotografie digitali, immagini e grafica.

Oggetto di tutela del diritto d'autore sono inoltre le opere musicali e video anche in formato digitale. Il formato di codifica digitale della musica è l'MP3, abbreviazione di MPEG1-Layer3 (*Moving Picture Expert Group*), ossia un algoritmo di compressione audio che codifica file audio e video in formato digitale compresso. Tali opere, facilmente reperibili in rete attraverso siti specializzati o portali quali iTunes Store, vengono scaricate dagli utenti e trovano ampia diffusione nei supporti tecnologici di uso comune, come PC, palmari lettori mp3, cellulari. Anche in questo caso l'acquisizione è soggetta al riconoscimento dei diritti in capo all'autore; il brano musicale è, infatti, soggetto a termini e condizioni di utilizzo che normalmente non consentono di modificare, affittare, noleggiare, prestare, vendere o distribuire l'opera, e pertanto può essere ritenuto illecito passare ad altri o scambiare file musicali.

I diritti d'autore possono trovare protezione con l'utilizzo di sistemi di gestione dei diritti digitali, i DRM (*Digital Rights Management*), ovvero misure di "autotutela tecnologica" che permettono di identificare e descrivere i diritti di proprietà intellettuale, comunque validi e di tracciare le licenze d'uso e dell'utilizzo effettivo del contenuto. A tal proposito meritano una citazione le nuove modalità per concedere in uso le proprie opere, come le licenze Creative Commons, distinte per il livello di concessione dei diritti (non commercializzabile, non alterabile o trasformabile) spesso utilizzate per immagini, musica, testi, e-book⁷. Infine, la normativa vigente, con il d.lgs 9 aprile 2003 n.68, riconosce all'autore anche il diritto di proteggere tecnologicamente le proprie opere mediante tecniche di crittografia, misure anticopia, note informative indicanti l'utilizzatore e i diritti dell'autore⁸.

3. DIRITTO D'AUTORE E FAIR USE

Nel quadro normativo succintamente (e non esaustivamente) richiamato, vediamo ora di comprendere la misura in cui è tutelabile il diritto dell'autore in rapporto alla libera utilizzazione della sua opera da parte di terzi.

Si precisa subito che il diritto dell'autore sulla propria opera non si estende sino al punto di impedirne qualsiasi utilizzo da parte di altri soggetti: infatti, il legislatore, oltre alla tutela del lavoro creativo dell'autore, vuole altresì stimolare la circolazione delle idee e delle informazioni contenute nelle opere.

La citata legge italiana n. 633/41 individua agli artt. da 65 a 71 *decies* le condizioni alle quali l'opera altrui può essere liberamente utilizzata.

Per quanto qui ci interessa, anche in riferimento alle moderne tecnologie, l'art. 70 prevede che *"il riassunto, la citazione o la riproduzione di brani o di parti di opera e la loro comunicazione al pubblico sono liberi"*

- a. se effettuati per uso di critica o di discussione;
- b. nei limiti giustificati da tali fini;
- c. purché non costituiscano concorrenza all'utilizzazione economica dell'opera.

Lo stesso articolo prevede anche che se il fine non è di *"critica o di discussione"* bensì di *"insegnamento o di ricerca scientifica"*, le attività sopra indicate sono lecite se l'utilizzo dell'altrui opera avviene altresì *"per finalità illustrative e per fini non commerciali"*.

Inoltre tali attività devono essere *"sempre accompagnate dalla menzione del titolo dell'opera, dei nomi dell'autore, dell'editore e, se si tratti di traduzione, del traduttore, qualora tali indicazioni figurino sull'opera riprodotta"* (art. 70 terzo comma LDA).

Un accenno va poi fatto al comma 1 bis dell'art 70 (aggiunto dalla legge 9 gennaio 2008 n. 2) secondo il quale: *"è consentita la libera pubblicazione attraverso la rete internet, a titolo gratuito, di immagini e musiche a bassa risoluzione o degradate, per uso didattico o scientifico e solo nel caso in cui tale utilizzo non sia a scopo di lucro"*.

Con riferimento a tale disposizione, è stato le-

⁷ Per una completa trattazione sul DRM e sulle licenze Creative Commons si rimanda all'articolo *"Diritto d'autore tra Digital Right Management e Creative Commons"* pubblicato in Mondo Digitale n. 1 - marzo 2007.

⁸ Per una completa trattazione sul diritto d'autore nella rete e nelle nuove tecnologie si rimanda all'articolo *"Il diritto d'autore nella rete"* pubblicato in Mondo Digitale n. 3 - settembre 2004.

gittimamente osservato come appaia inspiegabile il trattamento (peggiorativo) riservato alla libera utilizzazione via internet dell'opera visiva o musicale rispetto a ciò che è consentito per la loro pubblicazione con mezzo diverso posto che, fermo il rispetto delle condizioni di cui al primo comma dell'art. 70 LDA sopra esposte, nel primo caso l'immagine e la musica possono essere riprodotte solo a "bassa definizione o degradate" mentre nel secondo caso possono essere liberamente comunicate al pubblico con qualsiasi mezzo e senza limitazioni di qualità. Queste condizioni di libero utilizzo sono rispettate ad esempio dai portali di vendita musicale online (esempio, iTunes Store, Beatport, Juno) nei quali all'utente, al fine di critica o discussione del prodotto, viene resa disponibile un'anteprima degradata di un'opera musicale in catalogo oppure l'intera opera interrotta ad intervalli regolari da un segnale acustico che ne impedisce la completa riproducibilità.

Ora ci si chiede se il "libero uso" di diritto italiano coincida o meno con il "fair use" di diritto statunitense.

Per poter rispondere è necessario precisare quali siano i contenuti del *fair use*.

Per *fair use* letteralmente s'intende l'*uso corretto (ovvero leale)* di un'opera altrui protetta da diritto d'autore.

Perché l'uso sia "fair" è necessario che chi utilizza l'opera altrui:

- a. lo faccia per determinati scopi quali: ricerca o studio, insegnamento, critica;
- b. senza scopo commerciale,
- c. in modo limitato (senza, quindi, appropriazione del lavoro altrui);
- d. senza incidere sul valore dell'opera.

Il *fair use* statunitense è, però, cosa diversa dal *fair dealing* (trattamento corretto o leale) previsto nelle nazioni del *Commonwealth*.

Il *fair dealing* è, per così dire, una serie di eccezioni al diritto esclusivo di *Copyright* che un soggetto può opporre in occasione di una controversia nella quale gli venga imputata la vio-

lazione del diritto d'autore. Le categorie di *fair dealing* sono un numero chiuso tipizzato, quindi non si può invocare il *fair dealing* al di fuori di esse⁹.

La differenza sostanziale tra *fair use* statunitense e *fair dealing* è che il primo è una categoria generale applicabile al caso concreto verificando di volta in volta se è riconducibile o meno a tale categoria generale, il secondo l'esatto opposto: esso coincide in una raccolta codificata di casi tra i quali, quello concreto, può o meno rientrare.

Tra i casi più celebri, nella giurisprudenza nordamericana, sull'applicazione del *fair use*, si ricorda la controversia tra Universal e Disney contro Sony relativa al video registratore Beta-max (che permetteva all'utente di videoregistrare liberamente un programma radiotelevisivo o un film).

La Corte suprema, riformando la sentenza d'appello, con una storica pronuncia del 1984 affermò che il criterio del *fair use* dovesse essere interpretato come possibilità concessa ad un autore di fare un "uso ragionevole dei risultati realizzati da altri autori al fine di creare opere nuove"; sulla stregua di tale massima, concluse che l'utente che videoregistra deve essere "considerato come un consumatore del materiale depositato e non come un utilizzatore (nel senso del *fair use*)".

Sedici anni dopo, il concetto di *fair use* venne ripreso nell'altrettanto noto caso "Napster": la difesa eccepì che la copia dei *files* musicali in formato "mp3" per uso non commerciale (ovvero "personale") fosse ammessa dalla dottrina di *common law* americana, in quanto rientrante appunto nel "fair use".

Così rappresentato il quadro generale della materia possiamo dire che il "libero uso" dell'opera altrui di diritto italiano è più determinato (essendo anch'esso codificato) del *fair use* di diritto statunitense che, forse solo apparentemente, appare più esteso¹⁰.

La mancanza, nel diritto statunitense, di codifi-

⁹ In Australia, per esempio, si rientra nel *fair dealing* nei casi di ricerca e studio, rassegna e critica, rassegna stampa. Secondo una norma australiana si può fotocopiare al massimo un capitolo di un libro e comunque non più del 10% del suo totale.

¹⁰ Riteniamo che tale differenza discenda anche dalla diversa origine culturale del diritto d'autore: l'approccio europeo (e italiano) ha quale scopo principale di tutelare il lavoro intellettuale dell'autore sia sotto il profilo morale che patrimoniale, mentre la visione anglosassone subordina il diritto d'autore al più ampio interesse della collettività alla diffusione dell'opera per il progresso scientifico e culturale. Vi è, naturalmente, chi è di contrario avviso e ritiene che il vigente art. 70 LDA abbia introdotto in Italia una disciplina analoga al *fair use* statunitense.

Riquadro 1

Convenzione di Berna - articolo 9

1. Gli autori di opere letterarie e artistiche protette dalla presente Convenzione hanno il diritto esclusivo di autorizzare la riproduzione delle loro opere in qualsiasi maniera e forma.
2. È riservata alle legislazioni dei Paesi dell'Unione la facoltà di permettere la riproduzione delle predette opere in taluni casi speciali, purché una tale riproduzione non rechi danno allo sfruttamento normale dell'opera e non causi un pregiudizio ingiustificato ai legittimi interessi dell'autore.
3. Qualsiasi registrazione sonora o visiva è considerata riproduzione ai sensi della presente Convenzione.

Riquadro 2

Trattato OMPI sul diritto d'autore (WCT) - articolo 10

1. Le Parti contraenti hanno la facoltà di prevedere, nella propria legislazione, limitazioni o eccezioni ai diritti esclusivi degli autori di opere letterarie e artistiche contemplati dal presente trattato, in taluni casi speciali che non siano in conflitto con la normale utilizzazione economica dell'opera e non comportino un ingiustificato pregiudizio ai legittimi interessi del titolare.
2. Le Parti contraenti, nell'applicare la Convenzione di Berna, circoscrivono le limitazioni o le eccezioni di cui sopra a taluni casi speciali che non contrastino con la normale utilizzazione economica dell'opera e non comportino un ingiustificato pregiudizio ai legittimi interessi del titolare.

Trattato OMPI sulle interpretazioni ed esecuzioni e sui fonogrammi (WPPT) - articolo 16

1. Le Parti contraenti hanno la facoltà di prevedere nella propria legislazione, per quanto riguarda la protezione degli artisti interpreti o esecutori e dei produttori di fonogrammi, limitazioni o eccezioni della stessa natura di quelle previste nella predetta legislazione per quanto riguarda la protezione del diritto d'autore sulle opere letterarie e artistiche.
2. Le Parti contraenti impongono le limitazioni o le eccezioni ai diritti contemplati dal presente trattato soltanto in taluni casi speciali che non siano in conflitto con la normale utilizzazione economica delle esecuzioni o dei fonogrammi e non comportino un ingiustificato pregiudizio ai legittimi interessi dell'artista interprete o esecutore e del produttore di fonogrammi.

Riquadro 3

Considerando 44 della Direttiva 2001/23/UE

La facoltà di applicare le eccezioni e le limitazioni previste nella presente direttiva deve essere esercitata nel rispetto degli obblighi internazionali. Le eccezioni e le limitazioni non possono essere applicate in modo da arrecare pregiudizio agli interessi legittimi dei titolari dei diritti o da essere in contrasto con la normale utilizzazione economica delle loro opere o materiali protetti. L'introduzione di tali eccezioni o limitazioni da parte degli Stati membri deve in particolare tenere debitamente conto dell'accresciuto impatto economico che esse possono avere nel contesto del nuovo ambiente elettronico. È pertanto possibile che la portata di alcune eccezioni o limitazioni debba essere ulteriormente limitata nel caso di taluni nuovi utilizzi di opere e materiali protetti.

ca dei casi riconducibili a quello che in Italia è il "libero uso" dell'opera altrui comporta, comunque la si voglia guardare, maggiore incertezza con conseguente necessità (teorica) di dover introdurre una causa ogni volta che si presenta un'ipotesi di *fair use*. In Italia, al contrario, la presenza di casi codificati consente una più rapida verifica della liceità o meno dell'uso dell'opera altrui.

Ciò è tanto più vero se si considera che a livello internazionale ed europeo, in particolare, l'eccezione alla violazione del diritto d'autore ovvero - detta in altro modo - la limitazione che la legge nazionale può apporre a tale diritto è sottoposta alle regole delle convenzioni internazionali (il cosiddetto *three step test*) previsto dalla Convenzione di Berna (riquadro 1) e ripreso dai trattati dell'Organizzazione Mondiale per la Proprietà Intellettuale (riquadro 2) e dalle norme comunitarie (riquadro 3).

Dall'interpretazione di tali norme, si può ricavare la seguente sintesi:

1. la limitazione è ammessa nei soli casi speciali espressamente previsti dalla legge;
2. devono essere evitati contrasti con lo sfruttamento normale dell'opera;
3. non deve essere arrecato pregiudizio ingiustificato agli interessi legittimi del titolare del diritto.

Il diritto di libera riproduzione e diffusione dell'opera altrui è, nell'ordinamento giuridico italiano, riconosciuto anche in altri casi.

Così per esempio:

- gli articoli di attualità, economici, politici e religiosi, se non riservati, possono essere riprodotti e diffusi in altri giornali riviste e radiodiffusi (art 65 LDA);
- i discorsi pubblici su argomenti di interessi politico ed amministrativo e gli estratti di conferenze pubbliche possono essere riprodotti e comunicati al pubblico a scopo meramente informativo in riviste o giornali anche radiotelevisivi e telematici purché ne vengano indicati la fonte, l'autore, la data ed il luogo (art. 66 LDA);
- la riproduzione di singole opere o brani di opere è libera se effettuata per uso personale dei lettori a mano o con mezzi di riproduzione non idonei allo spaccio ed alla diffusione dell'opera in pubblico (art 68, primo comma LDA);
- è libera la fotocopiatura di opere esistenti nelle biblioteche pubbliche o scolastiche, nei musei e negli archivi pubblici se effettuata dai pre-

detti organismi senza vantaggio economico diretto o indiretto (art 68 secondo comma LDA);
□ la riproduzione per uso personale tramite fotocopia xerocopia o sistema analogo è libera se attuata nei limiti del 15% di ciascun volume o fascicolo di periodico, escluse le pagine di pubblicità (art 68, terzo comma LDA).

Volendo fare degli esempi pensiamo al reperimento in rete di immagini scattate da soggetti che poi le abbiano trasferite in siti accessibili da terzi. In tal caso (senza entrare nel merito del diritto all'immagine del soggetto eventualmente ritratto) una cosa è prelevare e utilizzare tale fotografia per un commento estetico ovvero tecnico nell'ambito di un forum di fotografia ed altra è impiegare il frutto dell'opera altrui per inserirla come illustrazione in una nostra presentazione, sito o altro. Mentre nel primo caso ben si può parlare di uso lecito, nel secondo bisogna distinguere se l'utilizzo dell'immagine fosse stato dichiarato libero dall'autore oppure no. In tale ultima evenienza l'utilizzo non sarebbe riconducibile al *fair use*.

Altro caso può essere quello delle banche dati. È noto che, per esempio, accedendo a banche dati giuridiche si ottiene non solo il testo di un provvedimento o di una sentenza ma alcune volte anche un commento o dei richiami in nota. Se utilizzassimo tali note all'interno di un nostro lavoro facendole proprie e non rendendo nota la fonte e l'autore è evidente che violeremmo il diritto d'autore altrui.

Non diverso il caso delle presentazioni che si trovano on-line. Capita, infatti, che alcuni utenti diffondano in rete slide illustrative su vari argomenti, oppure che tali documenti redatti da terzi siano confezionati in modo sufficientemente esaustivo da poter essere impiegati così e semplicemente.

Anche in tale ipotesi il prelievo e l'uso come opera propria del prodotto intellettuale altrui (al di fuori dei casi che ne venga citato l'autore e la fonte e che l'utilizzo sia limitato ai casi sopra già trattati) è illecito.

Per concludere – senza pretesa di esaurimento degli esempi che sono ben più numerosi – pensiamo ai brani musicali o *jingle* reperibili in internet non tanto come pezzi singoli musicali ma come sottofondo a prodotti multimediali. Non è consentito estrarre la traccia sonora per utilizzarla autonomamente.

Al contrario se lo spezzone di opera multimediale (previa citazione della fonte e dell'autore) ve-

nisse prelevato, separato nelle sue componenti audio video e testo, per scopo di critica o in un contesto educativo si rientrerebbe, senso lato, in un caso di *fair use*.

4. CONCLUSIONI E PROSPETTIVE FUTURE

In definitiva si osserva che la disciplina regolatrice, nei vari Paesi e per quanto ci interessa in Italia, del possibile uso dell'opera altrui, pur non essendo uniforme, consente comunque di conoscere i limiti a tale utilizzo.

La tutela degli utilizzatori e degli autori di opere, che oggi sono sempre più scambiate elettronicamente, non deve essere quindi solo un concetto normativo, ma necessita anche di un riscontro formativo attivo che miri ad informare attivamente e rendere consapevoli gli utenti dei nuovi mezzi comunicativi¹¹.

Certamente ancora va fatto molto, soprattutto a livello di armonizzazione delle singole norme statali e per una migliore (e forse anche più coraggiosa) visione e loro applicazione al mondo digitale.

A tale proposito si osserva che il 16 luglio 2008, la Commissione europea ha adottato un libro verde dal titolo "*Il diritto d'autore nell'economia della conoscenza*", al fine di promuovere un dibattito sui migliori mezzi per assicurare la diffusione on line delle conoscenze per la ricerca, la scienza e l'istruzione.

Il libro verde si articola in due parti, entrambe concernenti il tema delle eccezioni ai diritti esclusivi previste dalla citata direttiva 2001/29/CE. In particolare, nella seconda parte vengono affrontate le tematiche che, secondo la Commissione europea, assumono maggior rilievo per la diffusione della conoscenza nell'era della comunicazione digitale: le eccezioni per le biblioteche e per gli archivi, l'eccezione in favore dei soggetti disabili, la diffusione delle opere creative per scopi didattici e di ricerca, i contenuti generati dagli utenti (*User Generated Contents* – UGC).

Le applicazioni web 2.0 (quali blog, *podcast*, *wiki*, *video sharing* ecc.), infatti, hanno la capacità di coinvolgere gli utenti nella creazione e nella distribuzione dei contenuti, condividendo facilmente testi, video o immagini e svolgendo un

¹¹ L'utilizzo improprio di contenuti digitali, di *social network* o blog, o di appositi software spesso deriva da una mancanza di conoscenza o consapevolezza degli utenti.

0

ruolo più attivo e collaborativo nella diffusione delle conoscenze.

Preso atto di un tanto, da un punto di vista *de iure condendo*, il 23 giugno 2009 alla Camera dei Deputati è stata presentata la proposta di legge C.2525 "Modifiche alla legge 22 aprile 1941, n. 633, in materia di libere utilizzazioni di contenuti pro-

tetti da diritto d'autore" che prevede importanti novelle alla predetta normativa, nonché l'introduzione di disposizioni sulle utilizzazioni libere.

Il testo è stato assegnato alla commissione Cultura, scienza e istruzione in sede referente il 24 settembre 2009, pertanto il suo iter legislativo appare ancora lungo. Speriamo non troppo.

DAVID D'AGOSTINI, avvocato, master in informatica giuridica e diritto delle nuove tecnologie, docente all'Università degli studi di Udine. Presiede la Commissione informatica dell'Ordine degli avvocati di Udine, è responsabile dell'area "Diritto& informatica" della rivista "Il foro friulano". Presiede l'Organismo di vigilanza di Autovie Venete SpA.

E-mail: studio@avvocatodagostini.it

ANTONIO PIVA, laureato in Scienze dell'Informazione, *Vice Presidente dell'ALSI* (Associazione Nazionale Laureati in Scienze dell'Informazione ed Informatica) e Presidente della commissione di informatica giuridica. Docente a contratto di *diritto dell'ICT e qualità* all'Università di Udine. Consulente sistemi informatici e Governo Elettronico nella PA locale, valutatore di sistemi di qualità ISO9000 ed ispettore AICA.

E-mail: antonio@piva.mobi

MAURIZIO SALA, avvocato in Milano e Bucarest, Mac user da sempre, si occupa di informatica applicata al mondo del diritto dal 1987, componente della commissione informatica dell'Ordine degli avvocati di Milano e dell'omologa commissione presso il Tribunale di Milano, collabora con il Ministero della Giustizia per il processo telematico. Coordinatore per l'informatizzazione dell'Unione Nazionale delle Camere Civili, Docente ai seminari e ai master informatici dell'Università di Milano Bicocca.

E-mail: m.sala@sala.it



PROFESSIONE ICT

Competenze e professionalità per l'innovazione digitale

Rubrica a cura di

Roberto Bellini, Federico Butera, Alfonso Fuggetta

Il tema dell'innovazione e della competitività del sistema Italia è all'ordine del giorno della discussione economica e di quella sulle politiche industriali; sono promosse iniziative istituzionali a supporto dell'innovazione e si auspica un maggiore contributo della ricerca a livello universitario e privato. Anche l'Unione Europea spinge sul tema dell'innovazione, in particolare sul ruolo che le tecnologie ICT possono svolgere sia nei sistemi industriali che nei sistemi di governo e sull'importanza che può avere la definizione di un *framework* comune delle competenze ICT, compatibile con quanto previsto dall'EQF - *European Qualification Framework* - recentemente approvato dall'Unione Europea (2006).

Mondo Digitale vuole sostenere la diffusione di una maggiore sensibilità sul contributo che le competenze e le professionalità relative alle tecnologie digitali possono fornire in termini di innovazione dei servizi e del business dell'Impresa e di servizi per la cittadinanza erogati dagli enti della Pubblica Amministrazione. Questa nuova rubrica è dedicata appunto all'approfondimento sistematico di tutti gli aspetti che riguardano i progetti di analisi e di miglioramento delle competenze per l'innovazione digitale, il monitoraggio dei bisogni di competenza richiesti dal mercato e la valutazione delle offerte di qualificazione e aggiornamento delle competenze proposte dalle istituzioni educative di base e dagli operatori della formazione professionale e permanente.

La rubrica analizzerà l'andamento del mercato del lavoro delle professionalità ICT, i casi di successo nella crescita di competenze del personale dei fornitori di tecnologie e servizi e degli specialisti ICT, sia delle imprese manifatturiere e di servizio che degli enti della Pubblica Amministrazione, nonché l'andamento delle retribuzioni a livello nazionale e internazionale, usando come riferimento i profili e le competenze dello Standard EUCIP che AICA promuove in Italia.

I bisogni di formazione nel settore ICT

Roberto Bellini, Pierfranco Ravotto

1. INTRODUZIONE

Dopo avere mostrato negli articoli precedenti¹ come attraverso l'analisi del gap delle competenze rispetto ai profili di prossimità e a quelli obiettivo sia possibile disegnare percorsi formativi personalizzati, in questo articolo e nel prossimo cercheremo, utilizzando le indicazioni dello Standard delle competenze e dei profili EUCIP, di rispondere ad altre due domande:

□ Come possiamo costruire uno scenario dei

“Bisogni di Formazione” richiesti da un robusto campione di specialisti che operano nel mercato del lavoro del settore ICT?

□ Qual è l'offerta formativa disponibile e quali indicazioni possono essere utilizzate da una società o ente specializzati in formazione ICT per soddisfare bisogni di formazione sia latenti che esplicitati?

Ci appoggiamo per questa analisi al risultato di un'indagine condotta dal Centro di Competenza Praxi² sull'insieme degli specialisti ICT che negli

¹ Caso Vittoria Assicurazioni, Mondo Digitale n. 4, 2008; Pianificare la Formazione nelle imprese a rete, Mondo Digitale n. 2, 2010.

² PRAXI SpA – Società di *management consulting* attiva dal 1966 – ha 250 dipendenti su 10 sedi in Italia e un fatturato pari a circa 30 milioni di euro. Praxi è articolata su 4 divisioni professionali: Organizzazione, Risorse Umane, Informatica, Valutazioni e Perizie.

ultimi 3 anni hanno risposto con la propria candidatura a decine di campagne di reclutamento e selezione condotte da Praxi per i propri clienti in tutta Italia. I risultati sono ovviamente validi solo per il campione analizzato.

2. IL CAMPIONE INDAGATO E IL METODO DI ANALISI DELLE COMPETENZE

La base dati su cui è stata condotta l'analisi è costituita da 1002 questionari completamente compilati, raccolti attraverso una rilevazione su circa 12.000 candidati; il questionario, molto lungo da compilare, richiede mediamente da 2 a 3 h di lavoro per essere completato essendo costituito da due tipologie di domande:

□ **variabili socio-demografiche** comprendenti: età, livello di istruzione, indirizzo degli studi, area geografica in cui il rispondente lavora, inquadramento contrattuale;

□ **variabili relative alle competenze possedute:** il rispondente valuta in autonomia le competenze possedute e il proprio livello di approfondimento.

L'elaborazione delle competenze rilevate ha permesso di ottenere per ciascun candidato il profilo professionale corrispondente nello Standard EUCIP. Nella figura 1 sono indicati i passi della metodologia con cui è stato prodotto l'Indice di Competenza medio per profilo, su cui si è poi sviluppata l'analisi delle carenze da coprire, per esempio, con interventi formativi.

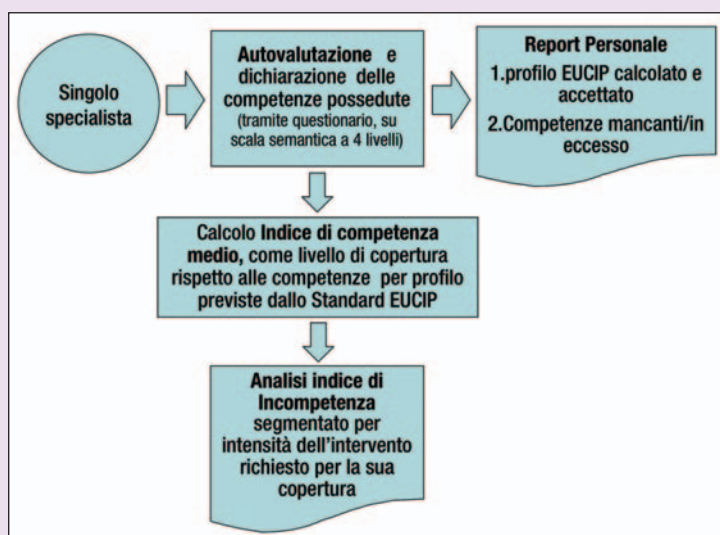


FIGURA1

Metodo di produzione dell'Indice di Incompetenza

Lo specialista ICT, a titolo individuale, risponde alle domande di un questionario in cui gli si chiede:

□ di completare i dati socio-demografici che nell'analisi vengono utilizzati per segmentare le carenze di copertura delle sue competenze;

□ di valutare e dichiarare il proprio livello di conoscenza e abilità sulle 156 categorie di competenza dello Standard EUCIP.

La risposta per ciascuna delle 156 categorie di competenza è articolata su una scala semantica a 4 livelli:

□ **nessuna:** quasi nessuna conoscenza dell'argomento, idee confuse;

□ **base:** conoscenza dei concetti principali, priva di approfondimenti;

□ **buona:** conoscenza dei concetti e capacità applicativa rafforzata da qualche esperienza diretta (anche nell'ambito di un corso con laboratorio);

□ **profonda:** padronanza completa e approfondita dei concetti e competenza applicativa derivante da ampia esperienza in un contesto complesso.

Ad ogni livello di conoscenza viene attribuito un punteggio che contribuisce a calcolare la prossimità a ciascuno dei 21 + 1 profili dello Standard EUCIP. Al termine del questionario lo specialista può leggere sul proprio Report personale quanto si avvicina a ciascuno dei profili: il profilo di prossimità è quello a cui si avvicina di più, su cui cioè risulta avere il punteggio più alto. Per tutti i profili, l'algoritmo incorporato nel questionario fornisce anche il dettaglio delle competenze mancanti e di quelle in eccesso rispetto al profilo standard. Il Report personale costituisce una preziosa risposta per chi lo compila e permette anche di considerare affidabile la risposta stessa, considerato che chi compila risponde prima di tutto a se stesso.

3. IL CAMPIONE ANALIZZATO

L'insieme dei dati rilevati sul campione di 920 rispondenti all'indagine Praxi ha portato a calcolare l'indice di competenza a livello aggregato per le 18 aree in cui sono raggruppate le categorie di competenza; l'indice viene calcolato come media quadratica dei valori attribuiti alle categorie che la compongono.

In tabella 1 è riportata la numerosità dei profili EUCIP che si ottengono dall'elaborazione delle competenze auto-valutate e dichiarate. Come abbiamo già anticipato, i profili emergono dall'elaborazione delle competenze che

ogni singolo rispondente ritiene di possedere. Risulta molto importante la distanza fra il profilo che il rispondente ritiene di avere all'inizio della compilazione del questionario e il profilo che viene effettivamente elaborato sulla base delle competenze possedute: solo il 23% dei rispondenti identifica correttamente il proprio profilo di prossimità, ossia quello che risulta avere il punteggio più alto rispetto ai profili dello Standard EUCIP. Due possono essere le cause principali di questa distanza:

1. l'oggettiva difficoltà a riconoscersi in uno dei profili dello Standard, data l'enorme dispersione delle definizioni relative ai mestieri che le imprese, i singoli specialisti e le istituzioni locali e centrali alimentano nel mercato del lavoro;

2. la ricostruzione del profilo di prossimità, ottenuta a partire dalle competenze dichiarate, tiene sia conto del ruolo e della specializzazione che il rispondente **dichiara di avere al momento** della risposta sia delle competenze acquisite attraverso i percorsi di studio, di lavoro e aggiornamento tecnico lungo tutta la sua vita professionale. Infatti, bisogna tenere presente che l'indagine è stata sviluppata su un sottoinsieme di specialisti che già lavorano, ma che si sono anche candidati a cambiare lavoro: Praxi è nota per avere un'importante attività di *recruitment* e selezione di personale tecnico e manageriale per medie e grandi imprese ed enti pubblici.

Osservando la tabella 1, emerge con evidenza che i profili più numerosi sono 3, con oltre 100 risposte: *Software Developer*, *X-Systems Engineer* e *IT Administrator*. Alcuni profili professionali mostrano invece, nel campione analizzato, una numerosità sorprendentemente bassa, presumibilmente per ragioni abbastanza diverse. Da una parte, infatti, la bassa numerosità di profili come quelli del *Client Manager*, del *Sales&Application Consultant*, dell'*IS Auditor* e dell'*IT Trainer* è da imputare al fatto che il campione è costituito da figure con prevalenti competenze tecniche; dall'altra invece la bassa numerosità di specialisti della sicurezza, dell'analisi di sistemi e di business, delle telecomunicazioni e di rete, può dipendere veramente dal fatto che questo tipo di specializzazioni sia effettivamente più scarso di quanto dovrebbe essere.

Una nota particolare va fatta per il profilo dell'*IT Administrator*: l'amministratore di sistema è il profilo che in assoluto dovrebbe essere il più frequente, dato che interessa sia le piccole unità periferiche delle grandi imprese ed enti della PA (pensiamo

a banche e assicurazioni, alle grandi reti della distribuzione organizzata, ai Commissariati e alle Stazioni dei Carabinieri, al sistema educativo italiano costituito da migliaia di scuole di un unico grande Ministero come quello della Pubblica Istruzione, agli sportelli postali ecc.), sia i circa 4 milioni di microimprese italiane con meno di 10 addetti. Il Profilo dell'*IT Administrator* identifica una specializzazione di gestione e supporto tecnico dei piccoli sistemi informatici che probabilmente vie-

	Profilo Eucip	Casi
P01	IS Manager	37
P02	IS Auditor	19
P03	Client Manager	25
P04	Sales & Application Consultant	27
P05	Enterprise Solutions Consultant	10
P06	Logistics & Automation Consultant	19
P07	Business Analyst	23
P08	IS Project Manager	68
P09	IS Analyst	18
P10	Software Developer	228
P11	System Integration & Testing Engineering	41
P12	Web & Multimedia Master	22
P13	IT Systems Architect	20
P14	Telecommunications Architect	19
P15	Security Adviser	6
P16	Database Manager	24
P17	Network Manager	65
P18	Data Center & Configuration Manager	24
P19	X-Systems Engineer	113
P20	Help Desk Supervisor	27
P21	IT Trainer	19
P22	IT Administrator	148
		1002

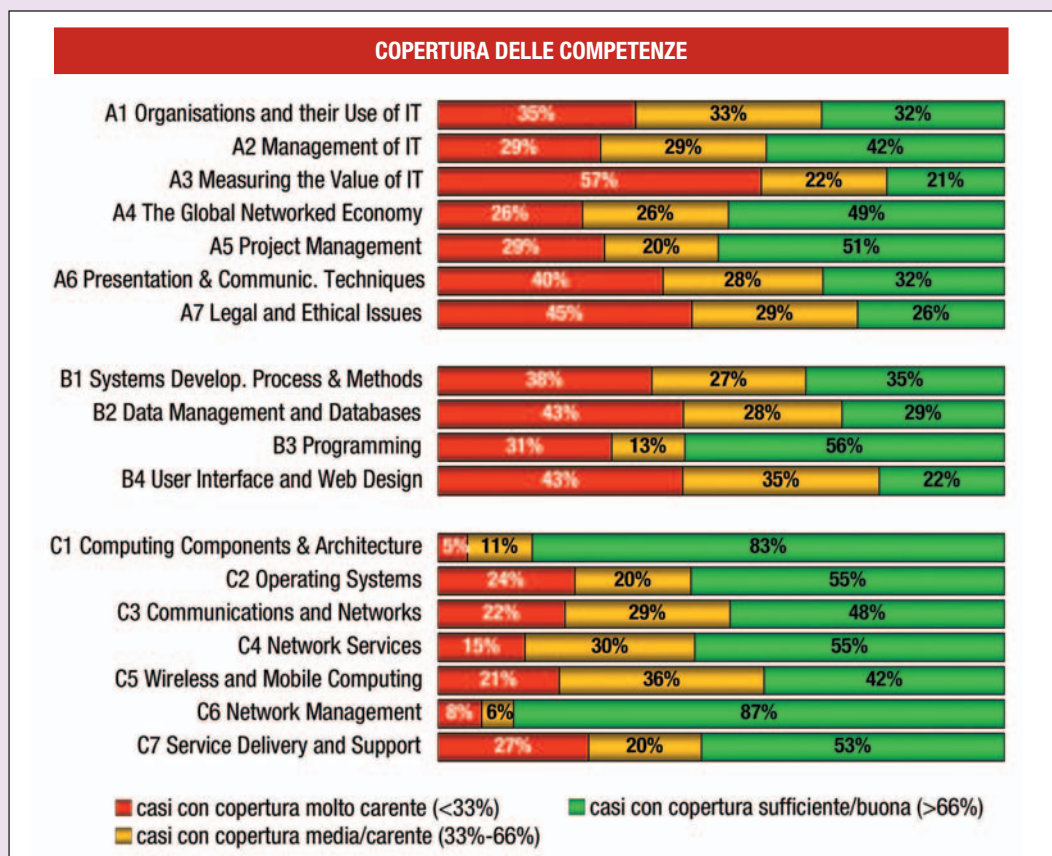
Il totale dei profili è superiore a quello del campione di 920 rispondenti; per i rispondenti che hanno generato fino a 3 profili equivalenti, si è scelto di considerarli tutti.

TABELLA 1

Profili EUCIP ottenuti da 920 rispondenti (Fonte: elaborazione AICA-EUCIP su dati Praxi)

FIGURA 2

*Analisi del grado di copertura delle competenze (base: tutti i rispondenti),
Fonte: elaborazione AICA-EUCIP su dati Praxi*



ne misconosciuta dal mercato del lavoro dei mestieri ad alta intensità di conoscenza. In tutti i casi, comunque, va sottolineato che il campione analizzato rappresenta solo se stesso.

4. ANALISI DELLA COPERTURA DELLE COMPETENZE RICHIESTE DAL MERCATO DEL LAVORO

Da quando si è cominciato a parlare di una gestione delle persone che lavorano basata sulle competenze, è la prima volta, a nostra conoscenza, che - pur limitatamente ad uno specifico segmento del mercato del lavoro (quello delle specializzazioni nel settore ICT) - è possibile analizzare il grado di copertura delle competenze richieste dalle imprese e dalle organizzazioni pubbliche: la granularità dello Standard delle Competenze e dei Profili EUCIP fornisce la base per l'analisi proposta.

Nella figura 2 è riportato il risultato dell'analisi del grado di copertura delle competenze aggregate nelle 18 aree dello Standard EUCIP, che raggruppano le 156 aree di competenza, a loro volta articolate in circa 3.000 unità di conoscenze elementari.

Per ogni area di competenza è indicata l'intensità della carenza di competenza segnalata dal campione, segmentata in 3 fasce:

- competenze molto carenti, con copertura inferiore al 33% (in rosso);
- competenze basse/medie, con copertura compresa fra il 33 e il 66% (in giallo);
- competenze sufficienti/buone, con copertura superiore al 66% (in verde).

Nella prima sezione A (area *Plan*) della figura 2, l'area di competenza "Measuring the Value of IT" (A3) è quella che mostra la carenza più significativa: il 57% dei rispondenti dichiara di avere un'importante carenza che si accompagna ad un altro 22% dei rispondenti con un grado di copertura medio/bassa (fra il 33 e il 66%): in totale il 79% dei rispondenti mostra di avere bisogno di interventi formativi su questa competenza; sempre considerando la somma dei rispondenti per aree con media e alta carenza di competenze, troviamo a ruota l'area "Legal and Ethical issues" (A7) con un livello di carenza pari al 74% dei rispondenti; altre aree nelle quali più del 50% dei rispondenti dichiara di avere carenze importanti e medie sono rispettivamente quelle del "Management of IT"

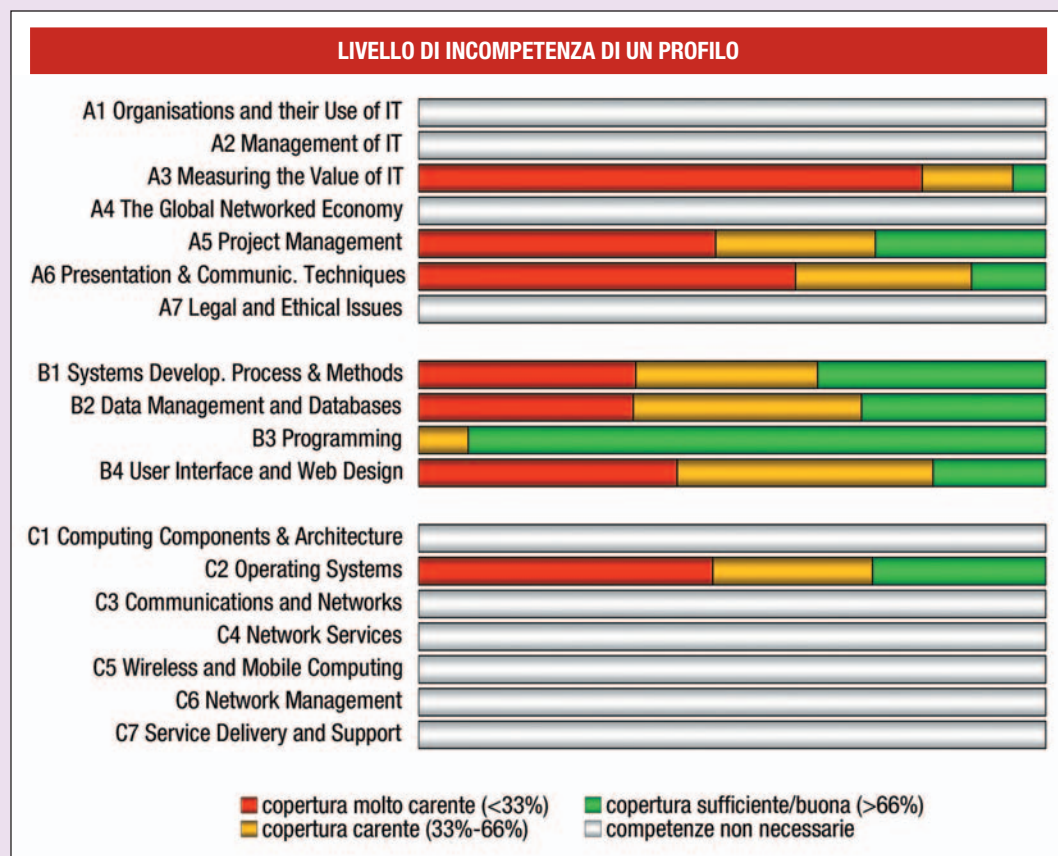


FIGURA 3

Analisi del livello di "incompetenza" del Profilo del Software Developer, Fonte: elaborazione AICA-EUCIP su dati Praxi.

Per "livello di incompetenza" si intende la mancanza di conoscenze specifiche rispetto all'ideale bagaglio di competenze richieste dallo standard EUCIP per ciascun profilo professionale

(A2) e quella relativa a "Presentation & Communication Technique" (A6).

Passando all'area B (Build) le carenze più importanti si trovano rispettivamente nelle aree B4 (User Interface & Web Design) con il 78% di rispondenti che dichiarano carenze, B1 (Systems Development, Process & Methods) e B2 (Data Management and Data Bases) in cui rispettivamente il 65% e il 71% dei rispondenti dichiarano carenza. I risultati dell'analisi permettono quindi di individuare quali siano le aree prioritarie su cui il mercato della formazione specialistica dovrebbe concentrare i suoi sforzi: tutte le aree sopra indicate richiedono evidentemente interventi formativi molto significativi.

Lo stesso tipo di analisi può essere condotto a livello di un singolo profilo EUCIP. Prendiamo ad esempio il profilo del Software Developer, che risulta essere quello maggiormente presente nel campione Praxi.

Il profilo nello Standard EUCIP è definito come se-

gue: "An EUCIP Software Developer is expected to play a considerable technical role in information systems design and to be very effective in carrying out the creation and maintenance of complex software modules that typically need to be integrated into a wider information system. Different specializations are possible, either in the field of applications and web services or in system-level software". La denominazione adottata in Borsa Lavoro è **Analista Programmatore**.

Nell'Osservatorio ASSINTEL 2009 la finalità di questo profilo è indicata sinteticamente come "ruolo chiave nella fase di realizzazione del sistema informatico in quanto utilizza linguaggi, tools, tecniche e metodi per trasformare le specifiche del sistema in codice eseguibile dai sistemi di elaborazione".

Nella figura 3 viene riportato il risultato dell'analisi delle competenze condotta sui 228 rispondenti interessati a cambiare lavoro con Praxi. Per facilitare l'analisi del risultato, abbiamo costruito un

indice di incompetenza per profilo, intendendo misurare con questo parametro la “carenza di conoscenze specifiche rispetto all’ideale bagaglio di competenze richieste dal Profilo dello Standard EUCIP”. In altre parole, abbiamo considerato le competenze pesate per ciascun profilo EUCIP come livello obiettivo rispetto a cui confrontare le competenze dichiarate da ciascuno specialista: l’indice di incompetenza è il complemento a 100 delle competenze possedute (in termini di autodichiarazione) dal rispondente al questionario Praxi, costruito sullo Standard EUCIP.

L’indice di incompetenza risulta molto elevato (sopra il 50% dei rispondenti) in 7 delle 8 aree di competenza necessarie per essere un buon *Software Developer*; nel campione di 228 rispondenti:

- oltre il 75% dichiara carenze profonde o medie nelle aree A3 (*Measuring the value of IT*), A6 (*Presentation & Communication Techniques*) e B4 (*User Interface & web design*);
- oltre il 50% dei rispondenti dichiara carenze profonde o medie nelle aree A5 (*Project Management*), B1 (*Systems Development Process &*

Methods), B2 (*Data Management and Databases*) e C2 (*Operating Systems*).

Per concludere, diamo una vista di insieme dei livelli di incompetenza su tutti i profili, come quella riportata nella figura 4: per ogni profilo singolarmente, rispetto al campione dei rispondenti, è stato indicato il relativo livello di incompetenza nelle 3 fasce già considerate (medio-alta, corrispondente ad una copertura < 33%, media compresa fra il 33 e il 66%, medio-bassa, corrispondente ad una copertura superiore al 66%).

Se proviamo a classificare i profili in funzione della profondità e dell’estensione del livello di incompetenza sulle 18 aree che contribuiscono alle competenze dei 22 profili dello Standard EUCIP, abbiamo il seguente risultato legato alla numerosità dei rispondenti:

- **4 Profili con alta priorità**, con almeno 3 aree rosse (xR) e le varie aree gialle (yG), che richiedono interventi di formazione profondi ed estesi: nell’ordine, *Web & Multimedia Master* (4R, 4G), *Enterprise Solution Consultant* (3R, 5G), *Data Base Manager* (3R, 3G), *X-Systems Engineer* (3R, 5G);

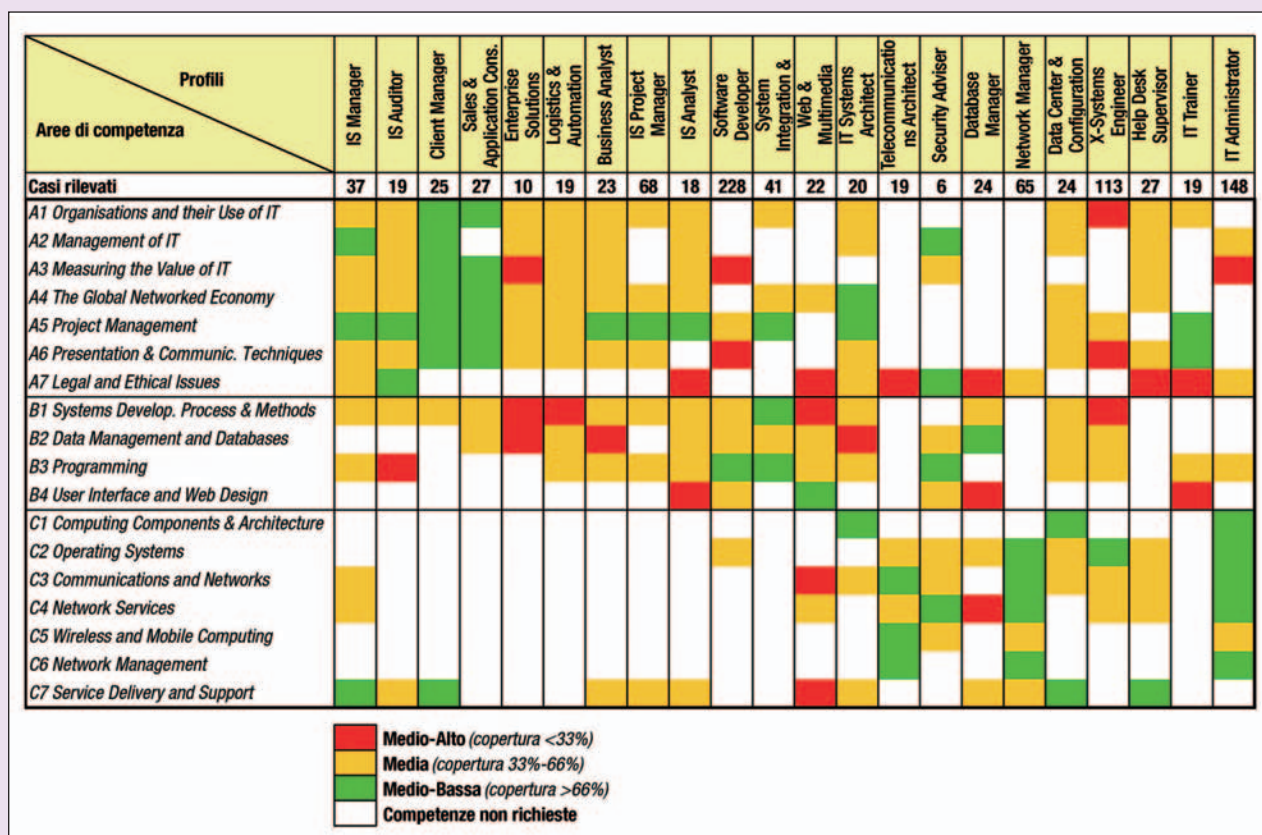


FIGURA 4

Livello di “incompetenza” per profilo (Fonte: elaborazione AICA-EUCIP su dati Praxi)

□ **10 Profili con media priorità**, con almeno 1 area rossa (xR) e le varie aree gialle (yG): nell'ordine, *IS Analyst* (2R, 8G), *Software Developer* (2R, 5G), *IT Trainer* (2R, 2G), *IS Auditor* (1R, 8G), *Logistic & Automation Consultant* (1R, 8G), *Business Analyst* (1R, 8G), *IT System Architect* (1R, 8G), *Help Desk Supervisor* (1R, 8G), *IT Administrator* (1R, 4G), *Telecommunication Architect* (1R, 2G);

□ **tutti gli altri Profili (8) con bassa priorità**: *IS Manager*, *Client Manager*, *Sales and Application Consultant*, *IS Project Manager*, *System Integration & Testing Engineer*, *Security Adviser*, *Network Manager*, *Data Center & Configuration Manager*.

Possiamo completare l'analisi con quella trasversale: le aree di competenza che richiedono interventi formativi prioritari sono quelle comuni a più profili: per esempio se prendiamo le aree di competenza che interessano almeno 3 profili troviamo, nell'ordine: *Legal & Ethical issues* (6), *Systems Development* (4), *Process & Methods* (4), *Measuring the Value of IT* (3), *Data Management & Databases* (3), *User Interface & Web Design* (3).

5. CONCLUSIONI

Riteniamo che i risultati ottenuti attraverso l'analisi dei livelli di incompetenza dei profili attuali costituiscano un passo avanti molto importante per quanto riguarda la costruzione di scenari per la formazione: sia dal punto di vista delle singole imprese che da quello dello scenario di mercato, poter rilevare sistematica-

mente e dettagliatamente quali siano le aree di competenza più carenti costituisce un input prezioso per quanto riguarda l'impostazione dei piani formativi. Naturalmente, a condizione che ogni area di competenza considerata corrisponda ad una numerosità di potenziali utenti sufficientemente significativa, cosa che in questa prima analisi è vera solo per alcuni profili; ad esempio, per quanto riguarda le competenze del profilo del *Security Adviser*, non sono sufficienti le risposte di soli 6 rispondenti.

Sappiamo fin da ora che sarà poi possibile arricchire gli scenari dei bisogni di formazione relativi ai profili attuali con il portafoglio delle competenze obiettivo: la pianificazione formativa si potrà appoggiare allora anche alla prevista evoluzione dei bisogni formativi.

I risultati indicati sono possibili solo alle seguenti condizioni:

□ che sia disponibile uno Standard di Competenze e Profili come quello EUCIP, condiviso sia dall'offerta che dalla domanda di formazione;

□ che sia applicabile un'adeguata strumentazione di rilevazione delle competenze disponibili e di quelle obiettivo.

Una volta evidenziati i livelli di incompetenza rilevati presso il mercato del lavoro e rilevata la disponibilità di offerta formativa, riconducibili entrambi ad uno standard condiviso, sarà anche possibile avere indicazioni per aggiornare con sistematicità la differenziazione fra i percorsi formativi di base sviluppati nella scuola e nell'università e quelli mirati a migliorare l'efficacia dei piani della formazione continua.

PIERFRANCO RAVOTTO, laureato in Ingegneria Elettronica al Politecnico di Milano nel 1974, è stato docente di Elettrotecnica, Elettronica, Automazione e Sistemi organizzativi in istituti tecnici occupandosi di corsi post-diploma, alternanza scuola-lavoro, scambi di giovani con alternanza all'estero ed e-learning. Ha coordinato diversi progetti europei nel programma Leonardo da Vinci ed ha partecipato alla progettazione e realizzazione di corsi di aggiornamento per docenti e dirigenti scolastici. Dal 2008 collabora con AICA per il framework EUCIP. È Project Manager del progetto Ensemble per il Dipartimento di Scienze dell'Educazione dell'Università di Firenze. È membro del direttivo nazionale della Sie-L, Società Italiana e-Learning, e del direttivo milanese di AICA. È certificato EUCIP IT Trainer.

E-mail: p.ravotto@aicanet.it

MODELLAZIONE DEGLI EFFETTI SISMICI IN UNA VALLE ALPINA

In tempi recenti, il terremoto del Cile, quello di Haiti e l'evento aquilano del 6 aprile 2009, hanno catalizzato l'attenzione dell'opinione pubblica. Nonostante la sismologia e l'ingegneria sismica siano discipline mature, è ancora oggi estremamente difficile prevedere quale potrebbe essere l'effetto di un sisma laddove tali eventi siano molto rari e/o caratterizzati da particolari condizioni geologiche e morfologiche. I moderni metodi numerici di calcolo possono aiutarci a simulare questi scenari generando dei terremoti simili a quello dell'Aquila ma localizzati lungo altre faglie considerate attive.

1. INTRODUZIONE

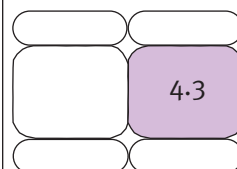
Lo sviluppo della civiltà umana ha da sempre dovuto fronteggiare diversi tipi di catastrofi naturali, tra le quali i terremoti. Tuttavia, solo all'inizio del XX secolo è stata proposta una teoria, nota come "tettonica delle placche" o "tettonica a zolle", in grado di fornire un'interpretazione fisica dell'origine dei terremoti. Semplificando notevolmente, essa può essere riassunta nel seguente modo: la litosfera terrestre, ovvero la parte più superficiale del nostro pianeta, risulta essere suddivisa in blocchi di dimensioni più o meno grandi (ordine di grandezza di un continente o sue porzioni) ed essi sono in continuo movimento a causa del materiale sottostante. Le velocità che caratterizza tale movimento è oggi nota grazie alle misurazioni attraverso il sistema GPS e risulta dell'ordine dei mm/anno. Ogni blocco, nel corso dei millenni, scivola addosso ad un altro e questa lentissima ma continua collisione genera un accumulo di energia in punti precisi, ad esempio lungo i bordi delle "placche". L'impossibilità di sostenere un certo livello di sforzo, innesca un terremoto, il quale dunque

non rappresenta altro che il rapido (alcune decine di secondi) rilascio dell'energia accumulata nel corso di centinaia o migliaia di anni; tutto questo con le conseguenze catastrofiche che ben conosciamo.

La nascita della moderna sismologia, basata su strumenti capaci di misurare lo scuotimento del terreno e l'affinarsi di diverse competenze scientifiche, hanno permesso di produrre delle mappe del rischio sismico ragionevolmente attendibili. Oggi siamo dunque in grado di sapere quale scuotimento massimo dobbiamo attendere in una particolare regione del pianeta e dopo quanto tempo in media un analogo evento si verificherà nuovamente. D'altra parte questo tipo di informazione è disponibile con attendibilità diversa a seconda della regione oggetto di indagine. Per fare un esempio concreto, il territorio italiano è colpito circa una volta ogni 10-20 anni da sismi di media magnitudo (Friuli 1976 Mw6.4, Irpinia 1980 Mw6.9, Umbria-Marche 1996 Mw6.0, Aquila 2009 Mw6.3). La magnitudo momento (Mw) è una misura quantitativa dell'energia rilasciata durante un evento sismico; una Mw pari a 5, è paragonabile all'energia rilascia-



Marco Stupazzini





ta dalla bomba atomica di Hiroshima, i sismi di Haiti (12/01/2010) e del Cile (27/02/2010) avevano una magnitudo rispettivamente pari a 7.0 e 8.8. Bisogna tenere presente che tale scala è logaritmica: all'incremento di un punto corrisponde un sisma 32 volte più intenso (il sisma del Cile è stato circa 1000 volte più forte di quello di Haiti). Le informazioni sugli effetti che questi eventi hanno prodotto sul costruito (scala macrosismica) sono opportunamente tabulati; per esempio in Italia il catalogo parte dal 79 d.C. e arriva fino ai nostri giorni [1], ovviamente con accuratezza e completezza diversa a seconda del periodo storico. Questi "cataloghi dei terremoti" sono il frutto del lavoro congiunto di équipes multidisciplinari formate da storici, geologi, geofisici e ingegneri sismici.

Non sempre però le informazioni storiche sono sufficienti: nella regione di Grenoble (Alpi Francesi), non si ha memoria scritta di terremoti catastrofici e questo potrebbe erroneamente portare a considerare la regione esente da rischio sismico. In nostro aiuto vengono i geologi strutturali e sismotettonici, i quali, attraverso un lavoro investigativo degno dei migliori detective, hanno ricostruito la storia geologica dell'area e hanno tracciato la posizione delle faglie attive, cioè delle "rotture" all'interno della crosta dove viene accumulata e successivamente rilasciata l'energia sotto forma di scuotimento.

Il nuovo interrogativo al quale si deve rispondere è come produrre una mappa dello scuotimento atteso per una zona di cui si conosce la posizione delle faglie anche se non è mai stato osservato fino ad ora un evento di intensità tale da produrre danni ($M_w > 5.5$) in quella regione.

Attorno alla seconda metà degli anni '60 due brillanti studiosi A. Cornell e L. Esteva gettarono le basi per uno studio probabilistico del rischio sismico [2] che venne successivamente formalizzato da Cornell in una famosa pubblicazione [3]. L'approccio probabilistico, in inglese *Probabilistic Seismic Hazard Analysis* (PSHA), è estremamente popolare oggi e è stato usato in modo soddisfacente a scala mondiale. Ad ogni sorgente viene associata una certa probabilità di generare un terremoto e attraverso opportune relazioni, funzioni della M_w , della distanza tra sorgente e sito di interesse e delle condizioni geologiche locali, è possibile calcolare lo scuotimento atteso per

una determinata posizione e un certo tempo di ritorno (intervallo medio di tempo tra due eventi sismici). Bisogna tenere presente che queste funzioni sono di solito calibrate con dati provenienti da sismi avvenuti in diverse zone del pianeta e, sebbene in tempi recenti vi sia stato un notevole sforzo per migliorarne la qualità (si veda il progetto "NGA" [4]), esse possono risultare piuttosto imprecise a causa di una serie di ineludibili incertezze epistemiche¹ o meno.

Da sottolineare che, proprio per quanto riguarda l'approccio probabilistico, è stata recentemente promossa un'iniziativa a livello mondiale [5] con l'obiettivo di raccogliere tutte le informazioni necessarie per definire i modelli attraverso i quali saranno successivamente prodotte le mappe di scuotimento per i diversi Paesi. Il metodo probabilistico, sebbene abbia offerto un'ottima risposta al problema concreto di definire il rischio sismico, non è tuttavia esente da limiti: questi si incontrano per esempio qualora sia importante studiare nel dettaglio la fisica del processo di propagazione delle onde e la loro interazione con strutture geologiche complesse, come nel caso delle valli alpine o appenniniche.

In effetti, la propagazione delle onde sismiche che vengono originate dalla sorgente (la faglia) è guidata da precise leggi fisiche ben note sin dal XVIII secolo. La soluzione di queste equazioni può risultare in alcuni casi molto difficile in "forma chiusa", cioè attraverso una semplice funzione, per esempio quando la geometria del problema è complessa. Solo l'avvento del calcolatore ha reso possibile lo studio del terremoto secondo il processo fisico da cui è governato. Abbandoniamo per un momento la sismica e immaginiamo una situazione decisamente più familiare: un bambino che getta un sasso in uno stagno e fa incresparsi l'acqua. I fronti d'onda, all'inizio circolari, si propagano nello stagno ma se il fronte incontra un ostacolo (quale un isolotto o un cigno) esso si incurva e genera ulteriori riflessioni. Il sisma nell'analogia è il sasso, l'acqua è il terreno e l'ostacolo è una delle complesse strutture geologiche che si possono incontrare.

¹ Per "incertezza epistematica" si intende una mancanza di conoscenza strettamente legata al fenomeno fisico studiato.

re decine o centinaia di metri sotto la superficie terrestre.

Nei prossimi paragrafi descriveremo brevemente i metodi numerici che vengono impiegati per questo tipo di studi, presenteremo alcuni risultati dell'analisi di risposta sismica condotta per la valle alpina di Grenoble, e infine trarremo alcune conclusioni sul futuro della moderna sismologia numerica.

2. IL METODO AD ELEMENTI SPETTRALI PER LO STUDIO DELLA PROPAGAZIONE DI ONDE SISMICHE

Durante e soprattutto dopo la seconda guerra mondiale, i calcolatori si sono sempre maggiormente diffusi ed evoluti, fino ad arrivare ai giorni nostri, in cui il "personal computer" è oramai omni presente.

Congiuntamente, il calcolo numerico, disciplina che rientra nella classe della matematica applicata, ha subito un notevole sviluppo nel progettare metodi per la risoluzione "pratica" di problemi matematici tramite opportuni algoritmi. Si noti che la soluzione calcolata da un algoritmo (cioè la soluzione numerica) è sempre un'approssimazione di quella esatta. Da questo si evince facilmente che la bontà di un metodo numerico può essere valutata attraverso due fattori:

i. l'accuratezza della soluzione, ovvero la differenza tra la soluzione esatta e quella numerica,
ii. la velocità con cui la soluzione viene ottenuta. Spesso è necessario un compromesso poiché una soluzione molto accurata può richiedere tempi di calcolo inaccettabili e viceversa una soluzione numerica ottenuta in tempi rapidi può discostarsi troppo dalla realtà e non essere utile. Una serie di fenomeni fisici possono oggi essere studiati attraverso dei metodi numerici opportunamente implementati al calcolatore (per esempio il calcolo statico per la progettazione di un edificio, le sollecitazioni a cui è soggetto un aeroplano durante il volo, la simulazione di eventi di piena ecc.).

Tornando dunque al tema della nostra discussione, il terremoto può essere visto come un caso particolare della ben più generale propagazione di onde in un mezzo elastico (onde elastodinamiche per l'appunto). I metodi numerici che consentono lo studio di questo tipo di problema fisico sono molteplici e ci limiteremo a ci-

tare il metodo ad elementi spettrali (*Spectral Element Method*, SEM [6, 7, 8]), sviluppato a partire dalla prima metà dagli anni '90 del secolo scorso, poiché tutte le simulazioni che verranno presentate in seguito sono state ottenute attraverso il codice di calcolo GeoELSE [9, 10] che implementa questo particolare metodo numerico. Senza scendere nei dettagli, è importante sottolineare che i SEM possono essere visti come un'evoluzione del ben più noto e diffuso metodo agli elementi finiti (*Finite Element Method*, FEM, [11]). I SEM risultano particolarmente adatti per lo studio di problemi dinamici e, in tale ambito, garantiscono una soluzione molto accurata in tempi di calcolo accettabili.

3. LA VALLE DI GRENOBLE: DALLA REALTÀ AL MODELLO NUMERICO

Le valli alpine sono il risultato dell'incessante lavoro di modellazione compiuto prima dai ghiacciai e poi, una volta ritirati o scomparsi, dai corsi d'acqua che ne hanno preso il posto. Il risultato sono delle valli che, semplificando, presentano una caratteristica sezione trasversale a forma di "V", nelle quali la punta risulta essere riempita di materiali decisamente più soffici (detrito di morena, ghiaie e sabbie compatte) rispetto a quello delle montagne circostanti (prevalentemente marmi, graniti o calcari); vista la configurazione geometrica la "V" viene chiamata "letto roccioso", perché proprio come un letto raccoglie i materiali più soffici depositati in epoche successive. La conoscenza della geometria dell'interfaccia tra materiali soffici e rigidi è una delle condizioni necessarie per poter analizzare la risposta sismica della valle. Si noti che questo tipo di studio richiede un notevole impegno sia finanziario che temporale, visto che questa interfaccia è sepolta a diverse centinaia di metri di profondità e per delimitarla è necessario compiere uno studio di sismica a riflessione o rifrazione lungo diversi allineamenti. Possiamo chiarire il concetto facendo un'analogia con quello che accade quando una persona si sottopone ad una TAC (Tomografia Assiale Computerizzata) in ospedale: durante l'esame il medico riesce ad ottenere un'immagine di quello che è presente all'interno del paziente in modo "non invasivo" cioè senza effettuare incisioni, grazie alla propagazione

e attenuazione di particolari onde che attraversano i diversi tessuti di cui il corpo umano è composto e raggiungono nuovamente la superficie, cioè la pelle.

Si tenga presente che la tomografia sismica ricopre oggi un ruolo fondamentale in diversi settori, per citarne uno su tutti, l'industria petrolifera che riesce ad individuare la presenza di giacimenti anche grazie a questa tecnica; inoltre, la comunità europea ha finanziato un progetto di ricerca incentrato proprio su queste tematiche [12].

Lo studio della risposta sismica della valle alpina di Grenoble si presenta come un problema particolarmente interessante, vista la precisione con cui è nota la forma tridimensionale del contatto tra materiali soffici e letto roccioso. Una ricostruzione così dettagliata è stata pos-

sibile grazie ai dati raccolti nel corso degli anni dal *Laboratoire de géophysique interne et tectonophysique* (LGIT) di Grenoble [13]. Nella figura 1 viene mostrata la zona di Grenoble e la forma tridimensionale del bacino alluvionale (contatto letto roccioso - materiali soffici).

Una volta acquisiti i dati che descrivono la geometria reale, tipicamente attraverso *Computer Aided Design* (CAD) o *Geographic Information System* (GIS), essi devono essere tradotti in un modello numerico (la griglia di calcolo) che il calcolatore sia poi in grado di leggere e interpretare. L'operazione può essere piuttosto complessa e sono stati creati appositi programmi per aiutare l'utente in questa delicata fase: il programma CUBIT [14] riesce a gestire la creazione di griglie di calcolo molto sofisticate. La figura 2 mostra quella realizza-

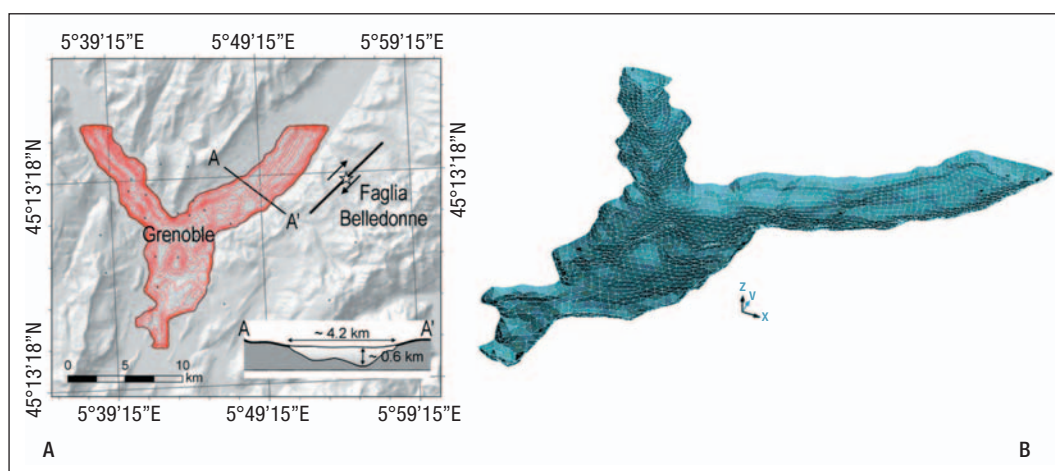


FIGURA 1

A - Il bacino alluvionale di Grenoble, la posizione della faglia Belledonne e le isobate del contatto letto roccioso – materiali soffici (linee in rosso). In basso è riportato lungo la sezione AA' l'andamento di un profilo bidimensionale. B - La complessa forma dell'interfaccia letto roccioso-materiale soffice in una sua ricostruzione tridimensionale

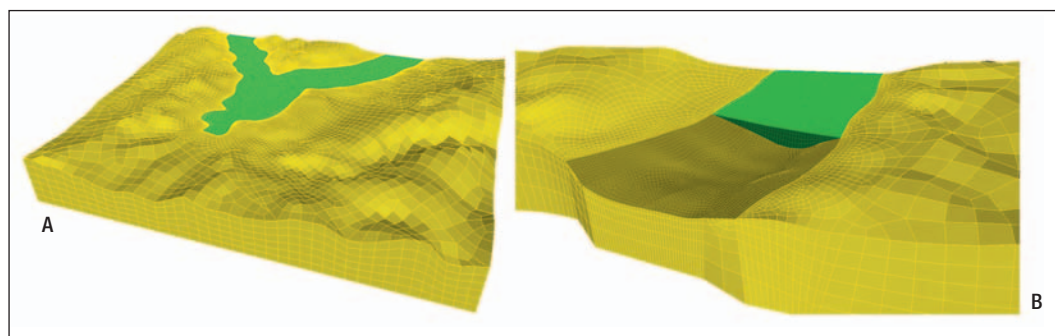


FIGURA 2

A - La griglia di calcolo realizzata per lo studio di risposta sismica della valle alpina di Grenoble (Alpi francesi). B - Dettaglio della griglia di calcolo nella zona di contatto tra il letto roccioso (giallo scuro) e il deposito alluvionale (colore verde)

ta per la valle di Grenoble. I materiali rigidi (colore giallo) sono discretizzati da elementi più grandi rispetto a quelli del deposito alluvionale (colore verde), questo si rende necessario per riuscire a campionare correttamente l'onda sismica durante il suo percorso. L'ottimizzare questa transizione attraverso griglie di calcolo non strutturate, passando cioè da una griglia fine ad una lasca senza sprecare troppi elementi, è uno dei punti cruciali per ottenere soluzioni numeriche accurate in tempi di calcolo ragionevoli.

4. LA RISPOSTA SISMICA DELLA VALLE DI GRENoble

Una volta raccolti tutti gli "ingredienti" necessari per la nostra simulazione, cioè:

- i. la posizione delle faglie attive e le loro caratteristiche (per esempio, la magnitudo massima ed il meccanismo di rottura),
- ii. le proprietà meccaniche dei diversi materiali presenti all'interno della zona d'interesse,
- iii. la forma geometrica delle discontinuità (l'interfaccia letto roccioso/sedimenti soffici),

è ora possibile effettuare l'esperimento al calcolatore e simulare quello che accadrebbe se un sisma di magnitudo pari a 6.0 (di poco inferiore al recente terremoto dell'Aquila) si generasse lungo la faglia di Belledonne, nella zona nord orientale del bacino. Nella figura 3 vengono presentate alcune istantanee delle onde sismiche (in questo caso la velocità del terreno in direzione ortogonale alla faglia) propagate durante una delle simulazioni. L'intero filmato può essere visionato presso il sito del codice di calcolo GeoELSE. Come si vede nella figura 3, le onde si propagano in entrambe le direzioni (nord-est e sud-ovest), ma il loro comportamento è profondamente diverso: il

fronte a nord-est rimane pressoché inalterato e continua ad avere forma circolare, quello a sud ovest incontra i materiali soffici del bacino alluvionale e modifica la sua forma. Da un lato l'onda è costretta a rallentare all'interno di questi depositi producendo un'amplificazione del moto del suolo, dall'altro la forma tridimensionale del bacino innesca una serie di riflessioni e rifrazioni complicando notevolmente il campo di moto.

L'effetto appena descritto è ben noto e spesso viene ricordato l'esempio di città del Messico dove nel 1985 un sisma originatosi a circa 250 km di distanza produsse effetti devastanti nella capitale messicana. Il fenomeno è facilmente spiegabile ai nostri giorni ed è del tutto analogo a quanto appena mostrato per Grenoble: Città del Messico è infatti costruita su di un terreno che un tempo ospitava un lago, successivamente prosciugatosi. I materiali soffici poggiano a loro volta su strati rigidi e si comportano come un amplificatore naturale del moto. Un ulteriore esempio è fornito dai dati ottenuti durante il terremoto dell'Umbria-Marche del 1997 (Figura 4). Il primo sismogramma è stato registrato su materiali rocciosi (stazione GBB) mentre il secondo su materiali soffici (stazione GBP). Si noti che la distanza tra la sorgente del sisma e le stazioni di registrazione è pressoché identica (i dati sono disponibili presso il sito [15]), mentre lo scuotimento risulta superiore in ampiezza e molto più prolungato nel tempo. Quest'ultimo effetto è dovuto alle riflessioni delle onde sismiche che restano "intrappolate" all'interno del bacino.

Lo studio del bacino di Grenoble non si è esaurito con la simulazione di un solo ipotetico evento; grazie ai calcolatori disponibili presso il "Consorzio Interuniversitario Lombardo per l'Elaborazione Automatica" (CILEA, [16]) si è potuto si-



FIGURA 3

Istantanee delle onde sismiche (componente della velocità ortogonale alla faglia) in tre successivi istanti temporali

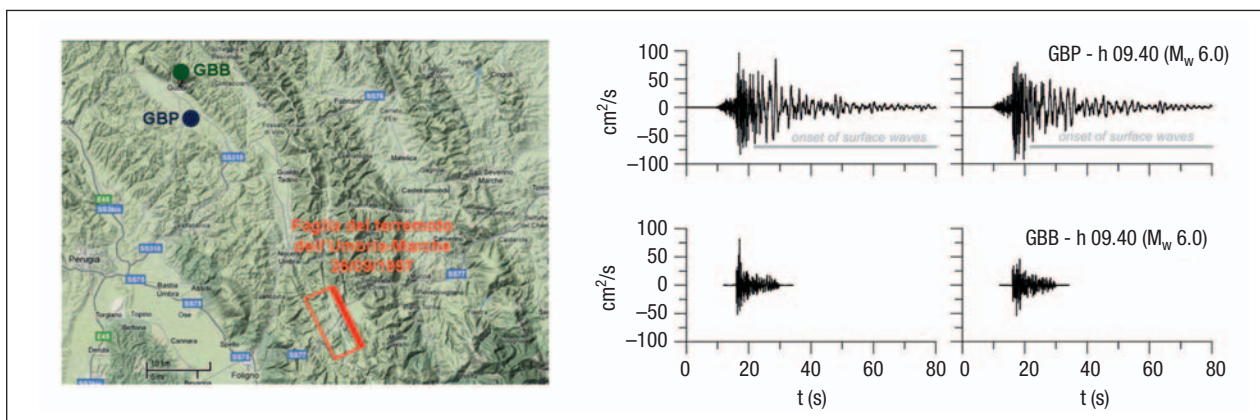


FIGURA 4

Posizione delle stazioni di Gubbio Piana (GBP) e Gubbio (GBB) durante l'evento sismico del 26 settembre 1997 (terremoto dell'Umbria-Marche). La prima stazione è collocata all'interno del bacino alluvionale su sedimenti soffici mentre la seconda su materiale rigido. A lato sono mostrate le registrazioni in accelerazione della componente EW e NS per l'evento principale delle ore 9.40 nelle due stazioni

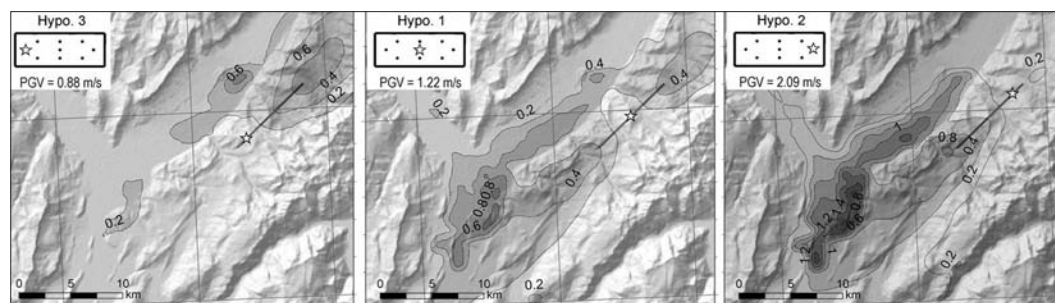


FIGURA 5

Mappe della velocità massima registrata durante le tre simulazioni con diversa posizione dell'ipocentro lungo la faglia di Belledonne. Si noti l'impressionante differenza tra la simulazione "Hypo3" e quella "Hypo2". Entrambe sono equiprobabili per il livello di conoscenza attuale

mulare lo scuotimento al quale sarebbe sottoposta la città di Grenoble se si verificassero una serie di eventi tutti fisicamente ammissibili ma profondamente diversi tra loro. I risultati di alcune di queste simulazioni vengono presentati nella figura 5. In particolare si è analizzato l'effetto della posizione dell'ipocentro (punto di nucleazione della frattura) rispetto al valore massimo di scuotimento registrato durante l'intera analisi. Le mappe mostrano molto chiaramente che questo parametro gioca un ruolo fondamentale: se l'ipocentro è collocato a sud-ovest lungo la faglia (Hypo3) la rottura si sviluppa prevalentemente a nord-est e di conseguenza il bacino alluvionale non è soggetto a scuotimenti eccessivi, viceversa se l'ipocentro è collocato a nord-est lungo la faglia (Hypo2) notiamo che lo scuotimento è nettamente superiore nel bacino di Grenoble, con conseguenze facilmente immaginabili per l'abitato. La situazione

intermedia (Hypo1) è quella in cui l'ipocentro è collocato al centro della faglia dando luogo ad una rottura denominata "bilatera". Riassumendo, queste analisi ci permettono di studiare la fisica dell'evento e conoscere in anticipo l'effetto che alcuni parametri avranno sullo scuotimento che si verificherà. L'allineamento tra l'ipocentro, la traccia della faglia e l'abitato di Grenoble gioca un ruolo fondamentale per uno studio di rischio sismico. In modo analogo possono essere esaminati altri parametri e valutata la loro influenza. Ciascuna delle simulazioni presentate è stata fatta girare in circa 8 h su 64 processori, presso il cluster parallelo del CILEA denominato LAGRANGE. Chi fosse interessato alle caratteristiche della macchina, può visitare il sito del consorzio. Ovviamente può sorgere il più che legittimo dubbio sull'attendibilità di queste simulazioni. I recenti dati raccolti durante il sisma dell'Aqui-

la hanno parzialmente confermato che la modellazione sismica della valle di Grenoble può essere considerata attendibile e può fornire una base molto utile per lo studio del rischio sismico della città.

5. CONCLUSIONI

Per concludere questo breve “viaggio” che ci ha portato a descrivere la risposta sismica di una valle alpina, due sono le riflessioni da mettere in evidenza.

La prima riflessione parte dalla domanda che è stata posta in una recente pubblicazione da due famosi ricercatori: “abbiamo visto il peggio?” [17], intendendo con questo se sia già stato osservato o meno lo scuotimento più intenso ammissibile in un certo sito.

Il dubbio sorge spontaneo visto che non è nemmeno passato un secolo da quando venne registrato (13 Marzo 1932, terremoto di “Long Beach”) il primo sismogramma attraverso uno strumento appositamente progettato per questo compito. Da allora, le reti sismometriche si sono decisamente sviluppate su tutto il pianeta e alcune hanno oggi una densità rimarchevole come le reti giapponesi K-NET e KiK-net (rispettivamente in funzione dal 1996 e 1997). Ciononostante, la conclusione dei ricercatori è che sia improbabile che tutto quello che si può vedere, sia già stato visto. In altre parole, è molto probabile che i livelli massimi di scuotimento non siano ancora stati registrati.

A tale proposito le simulazioni numeriche, se opportunamente calibrate sulla base di dati reali, possono fornire delle preziose indicazioni su quello che in un futuro ci potremmo aspettare. Lo studio di scenari deterministici prodotti attraverso metodi numerici si sta diffondendo con crescente importanza e diversi progetti prevedono l'impiego di questo tipo di analisi per completare lacune e carenze nelle basi di dati di sismogrammi osservati. Il progetto TeraShake [18], ShakeOut [19] negli Stati Uniti o il progetto Cashima [20] in Europa sono entrambi esempi concreti di come la comunità scientifica sia sempre più attenta nei confronti di questi metodi e, in questa ottica, il ruolo dei centri di calcolo e di macchine ad alte prestazioni sarà sempre più importante.

Ovviamente, esistono anche dei forti limiti: per

effettuare delle simulazioni credibili si devono possedere una serie di dati che purtroppo al momento sono disponibili solo in un numero molto limitato di siti nel mondo (per esempio il bacino di Los Angeles, Tokyo, Taipei o in Italia il bacino alluvionale di Gubbio).

In ogni caso, la strada è tracciata e i problemi sopra menzionati possono essere superati solo grazie ad una stretta collaborazione tra i diversi istituti ed enti di ricerca, nazionali ed internazionali, mettendo queste informazioni a disposizione gli uni degli altri.

La seconda riflessione ha un carattere più generale: come si è osservato all'inizio del testo, lo sviluppo del calcolo automatico ha avuto impulso durante la seconda guerra mondiale, e ha registrato stimoli importanti negli anni della guerra fredda e della gara per la “conquista” della luna. Perfino le guerre, fra tanti orrori e atrocità, danno talvolta occasione a progressi tecnici importanti. Si può dire qualcosa di simile anche per il terremoto. La protezione civile rappresenta oggi la civiltà in uno dei suoi aspetti migliori. In linea di principio sembra ragionevole immaginare una difesa globale contro il terremoto, estesa a tutta l'area della prevenzione.

La modellazione della valle di Grenoble è un esempio significativo in questa ottica. Essa mostra la possibilità di una mappatura territoriale, mirata a valutare in modo sistematico il rischio sismico associato a ciascuna località. Questa a sua volta è la base indispensabile per la progettazione dei fabbricati nuovi e anche per l'adeguamento di quelli esistenti, in modo da sfruttare pienamente le possibilità offerte dalla tecnologia edilizia attuale (per esempio, strutture disaccoppiate con l'uso di smorzatori, oppure materiali compositi fibro-rinforzati).

Un piano di miglioramento della sicurezza sismica impone di ottimizzare l'allocazione delle limitate risorse disponibili secondo criteri razionali e quindi presuppone un'analisi del rischio quanto più accurata possibile, sia in termini statistici che deterministici. Scienza e tecnica rendono oggi disponibili gli strumenti per un'azione sistematica di questo genere; i risultati potrebbero essere paragonabili, a lungo termine, a quelli ottenuti nei secoli scorsi per il miglioramento delle condizioni igieniche e per la prevenzione dei grandi contagi.

Da ultimo preme sottolineare che, per tutti coloro che si occupano di rischio sismico, la speranza è creare una vera e propria cultura del rischio, cioè che vi sia coscienza da parte della popolazione del rischio sismico a cui è soggetta e che questo porti in un futuro possibilmente non troppo lontano a prestare maggiore attenzione ad una problematica ancora oggi spesso posta in secondo piano rispetto ad altre priorità.

Bibliografia

- [1] Istituto Nazionale di Geofisica e Vulcanologia (INGV): www.ingv.it
- [2] McGuire R.K.: Probabilistic seismic hazard analysis: Early history. *Earthquake Engineering & Structural Dynamics*, Vol. 37, n. 3, 2008, p. 329-338.
- [3] Cornell C.A.: Engineering seismic risk analysis. *Bulletin of the Seismological Society of America*, Vol. 58, n. 5, 1968, p. 1583-1606.
- [4] Pacific Earthquake Engineering Research Center: NGA Database: <http://peer.berkeley.edu/nga/>
- [5] Global Earthquake Model (GEM): www.globalearthquakemodel.org
- [6] Faccioli E., Maggio F., Paolucci R., Quarteroni A.: 2D and 3D elastic wave propagation by a pseudo-spectral domain decomposition method. *Journal of Seismology*, Vol. 1, 1997, p. 237-251.
- [7] Komatitsch D., Vilotte J.P.: The spectral element method: an efficient tool to simulate the seismic response of 2D and 3D geological structures. *Bull. Seism. Soc. Am.*, Vol. 88, 1998, p. 368-392.
- [8] Chaljub E., Komatitsch D., Vilotte J.-P., Capdeville Y., Valette B., Festa G.: Spectral Element Analysis in Seismology, in *Advances in Wave Propagation in Heterogeneous Media*. Edited by Ru-Shan Wu and Valérie Maupin, *Advances in Geophysics, Elsevier*, Vol. 48, 2007, p. 365-419.
- [9] Codice di calcolo GeoELSE: <http://geoelse.stru.polimi.it>
- [10] Stupazzini M., Paolucci R., Igel H.: Near-Fault Earthquake Ground-Motion Simulation in the Grenoble Valley by a High-Performance Spectral Element Code. *Bulletin of the Seismological Society of America*, Vol. 99, n. 1, 2009, p. 286 - 301.
- [11] Zienkiewicz O., Taylor R.L.: *The finite element method*. McGraw-Hill, London, Vol. 1, 1989.
- [12] QUEST (QUAntitative Estimation of Earth's Seismic Sources and STRucture): <http://www.quest-itn.org/>
- [13] Laboratoire de Géophysique Interne et Tectonophysique (LGIT) <http://www-lgit.obs.ujf-grenoble.fr/>
- [14] The CUBIT Tool Suite: <http://www.cubit.sandia.gov/>
- [15] ITACA - Italian ACcelerometric Archive: <http://itaca.mi.ingv.it/ItacaNet/>
- [16] Supercomputing at CILEA: www.supercomputing.it
- [17] Strasser F.O., Bommer J.J.: Strong Ground Motions - Have We Seen the Worst? *Bulletin of the Seismological Society of America*, Vol. 99, n. 5, 2009, p. 2613-2637.
- [18] Progetto TeraShake: <http://scec.usc.edu/research/cme/groups/terashake>
- [19] Bielak J., Graves R.W., Olsen K.B., Taborda R., Ramírez-Guzmán L., Day S.M., Ely G.P., Roten D., Jordan T.H., Maechling P.J., Urbanic J., Juve G.: *The ShakeOut earthquake Scenario: Verification of three simulation sets, Submitted for publication*. 2009.
- [20] Pitilakis K.: EuroSeisTest report for the Cashima project, Technical report, Department of Civil Engineering, Aristotle University of Thessaloniki, Greece, Report distributed to the participants of the Cashima-Euroseistest Benchmark. 2008.

Alcuni riferimenti sulla sismologia in generale per eventuali approfondimenti

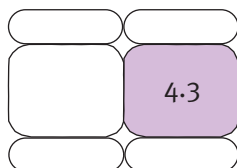
- [21] Faccioli E., Paolucci R.: *Elementi di Sismologia applicata all'Ingegneria*. Pitagora Editore, (Italiano), 2005, p. 255.
- [22] Boschi E., Dragoni M.: *Sismologia*. UTET. (Italiano), 1999.
- [23] Aki K., Richards P.: *Quantitative Seismology. Theory and Methods*, Freeman, San Francisco. (Inglese), 1980.

MARCO STUPAZZINI ha conseguito il dottorato in Ingegneria sismica e geotecnica presso il Politecnico di Milano nel 2004, è stato Post-Doc presso la Ludwig Maximilians Universität di Monaco di Baviera nel quadro del progetto di ricerca europeo SPICE. Dal 2008 al 2009 ha preso servizio come ricercatore di Scienza delle Costruzioni presso il Dipartimento di Ingegneria Strutturale del Politecnico di Milano. Dal 2009 lavora presso la Munich RE, società leader mondiale nel settore riassicurativo, come sviluppatore di modelli per il rischio sismico. I suoi interessi di ricerca riguardano lo studio della risposta sismica all'interno di bacini alluvionali o domini tridimensionali complessi, il calcolo ad alte prestazioni, lo studio degli effetti sismici di "campo vicino" e l'interazione dinamica terreno-struttura.
E-mail: MStupazzini@munichre.com



NANOTECNOLOGIE, CHIMICA TEORICA E CHIMICA COMPUTAZIONALE

Gian Franco Tantardini
Rocco Martinazzo
Simone Casolo



Le nanotecnologie ci danno la possibilità di creare materiali innovativi per svariate applicazioni, quali per esempio l'elettronica e la medicina. La capacità di creare e manipolare oggetti di dimensioni nanometriche cresce ad un ritmo impressionante, tuttavia la sperimentazione da sola non è sufficiente a comprendere i fenomeni fisici che avvengono su questa scala. La chimica teorica e la chimica computazionale offrono oggi la possibilità di studiare a livello molecolare questi sistemi, spiegare le osservazioni sperimentali e prevedere strutture con proprietà interessanti per nuove applicazioni.

1. INTRODUZIONE

Le leggi fisiche che sono state scoperte nel corso dei secoli hanno permesso e permettono tuttora di descrivere e comprendere il mondo macroscopico che ci circonda. Queste leggi sono basate sull'osservazione dei fenomeni fisici tramite gli esperimenti e sulla loro interpretazione. Galileo formulò le leggi del moto studiando la caduta dei corpi, così come Newton dedusse la sua legge di gravitazione universale dall'osservazione del moto dei pianeti nel cielo. Alla nostra scala di osservazione, il mondo sembra seguire leggi deterministiche, secondo le quali è possibile prevedere il comportamento di un sistema fisico in modo esatto, noto il suo stato ad un istante di tempo iniziale. Tuttavia, rivolgendo l'attenzione alla struttura microscopica della materia, si rende necessario un nuovo approccio, per certi versi "probabilistico", che è quello proprio della meccanica quantistica. Occorre abbandonare l'idea che l'osservazione non disturbi l'"osservato": se per studiare il moto di un pallone da calcio possiamo trascurare l'effetto della luce che è necessaria per renderlo visibile, lo stesso non

può dirsi per un elettrone. Secondo la meccanica quantistica, ogni corpo ha una diversa natura, dipendentemente da come lo si osserva; può essere particella ma anche onda. La sua evoluzione nel tempo è descritta da leggi deterministiche ma queste danno solo informazioni probabilistiche circa il risultato delle osservazioni. Ad una scala comparabile con le dimensioni atomiche è quindi necessario utilizzare leggi fisiche diverse da quelle del mondo macroscopico. Queste leggi furono formulate agli inizi del secolo scorso da una generazione di grandi fisici, che cominciò ad indagare la struttura microscopica della materia e ad intuire le enormi potenzialità che essa offre.

La "nuova" teoria è oggi alla base della chimica teorica: una disciplina che coniuga la millenaria conoscenza chimica della materia e delle sue trasformazioni, con la matematica e i moderni metodi di calcolo scientifico. Oggi i chimici teorici dispongono degli strumenti necessari per *determinare*, attraverso il *calcolo scientifico*, le proprietà fisiche (elettriche, magnetiche ecc.) e chimiche (per esempio la reattività) di molecole e materiali: informazio-

ni che permettono il *design* di nuovi prodotti nanotecnologici per le più svariate applicazioni. Lo strumento fondamentale allo scopo è l'*equazione di Schroedinger* (riquadro 1) - un'equazione alle derivate parziali per la co-

siddetta *funzione d'onda* - che sostituisce (generalizza) l'equazione di Newton nella descrizione del mondo fisico microscopico. Questa è un'equazione molto simile a quella di D'Alembert per le onde classiche e descrive

RIQUADRO 1 - L'equazione di Schroedinger e i programmi di chimica computazionale

L'equazione fondamentale da risolvere per studiare la materia a livello microscopico è l'equazione di Schroedinger, una equazione per la funzione d'onda Ψ alla derivata prima nel tempo che coinvolge anche le derivate parziali rispetto a tutte le coordinate del sistema fisico in esame. La sua soluzione è quindi estremamente complicata e, eccezion fatta per pochi sistemi modello, può essere ottenuta solo numericamente attraverso un certo numero di approssimazioni.

La più importante consiste nel separare quello che viene chiamato il "problema elettronico" dal moto dei nuclei, eliminando in un certo senso le variabili lente (le coordinate dei nuclei). Queste ultime, successivamente, possono essere studiate in approssimazione classica usando le equazioni di Newton (eventualmente correggendola per gli effetti quantistici) visto che i nuclei, soprattutto se pesanti, hanno un comportamento quasi classico, e questo semplifica enormemente il problema almeno per le applicazioni in oggetto.

Cosa diversa succede per gli elettroni che non possono mai essere considerati oggetti classici. Il problema elettronico consiste nel trovare le soluzioni *stazionarie* per il moto degli elettroni nel campo di potenziale generato dai nuclei fissi ed, eventualmente, da campi esterni. Essa quindi è una equazione ad autovalori-autofunzioni, $\hat{H}\Psi = E\Psi$ dove $\hat{H}$ è un operatore differenziale lineare che contiene le derivate parziali rispetto alle coordinate elettroniche. In principio - ma anche in pratica secondo alcuni approcci - l'equazione può essere ridotta in forma algebrica introducendo un insieme di funzioni "base" e trasformata in un problema ad autovalori-autovettori, dove i vettori sono i coefficienti di espansione della funzione d'onda Ψ nell'insieme di base introdotto. Poiché l'interesse è generalmente limitato allo stato fondamentale, di più bassa energia (detto *ground-state*), la ricerca della soluzione può essere effettuata con tecniche iterative per gli autovalori di modulo massimo. Le corrispondenti matrici, che possono arrivare ad avere (miliardi x miliardi) elementi, spesso vengono calcolate "on-the-fly" ed usate all'occorrenza in algoritmi estremamente parallelizzati. Accanto a questo approccio "brutale" vi è un gran numero di teorie, metodi e tecniche che cercano di ottimizzare la base per la traduzione algebrica del problema. Essi sono stati sviluppati in oltre 50 anni di vita della Chimica Teorica, ed hanno portato ad un numero incredibile di successi.

Tuttavia, nonostante l'enorme sviluppo *hardware* cui abbiamo assistito negli anni, ancora oggi non saremmo in grado di studiare con questo approccio problemi interessanti per le nanotecnologie. Il motivo è semplice: se per un elettrone abbiamo bisogno di M funzioni di base, per N elettroni (e N può essere qualche migliaio) avremmo bisogno di M^N funzioni di base, che possono essere solo parzialmente ridotte sfruttando certe proprietà di simmetria. Per di più M dipende anche dalle "dimensioni" molecolari, e cresce con la complessità del sistema. La soluzione a questo problema di carattere esponenziale viene da un lampo di genialità che è valso un premio Nobel per la chimica (W. Kohn, 1998). L'idea, nell'aria sin dagli anni Trenta, è che, almeno limitatamente al *ground-state*, si può sostituire la funzione d'onda Ψ con una funzione *molto* più semplice, la densità elettronica ρ (funzione solo di tre coordinate spaziali). L'equazione cui deve soddisfare ρ non è nota esattamente ma fortunatamente può essere scritta in forma approssimata facendo assunzioni ragionevoli. La conseguente teoria, nota come *Density Functional Theory* (DFT), sebbene presenti per certi versi carattere di "modello", fornisce una descrizione fisica formalmente rigorosa e risultati molto accurati. La semplicità dell'approccio è tale per cui, computazionalmente parlando, invece di risolvere l'equazione di Schroedinger per N elettroni interagenti si risolve un'equazione (nota come di "Kohn-Sham") in cui ciascun elettrone si muove nel campo medio generato dagli altri elettroni e soggetto ad un ulteriore "potenziale" non-classico che tiene conto della cosiddetta correlazione nel moto degli elettroni interagenti, oltre che del principio di Pauli. In pratica, anche l'equazione di Kohn-Sham è ridotta in forma algebrica, attraverso l'introduzione di un insieme di funzioni di base o l'uso di una griglia equivalente, e viene quindi risolta con metodi e tecniche simili a quelle accennate in precedenza. Nel caso di griglie di Fourier, molto utilizzate nel campo dello stato solido per sfruttare appieno la periodicità, l'uso della *Fast Fourier Transform* (FFT) è una pratica assai diffusa, specialmente sulle moderne architetture parallele. Con il DFT è possibile studiare oggi strutture che contengono centinaia e centinaia di atomi, all'accuratezza necessaria (chimica) per determinarne correttamente le loro proprietà.

Ovviamente, questo è ancora insufficiente per la maggior parte delle nanostrutture (già un cubo di 10 nm di lato conterrebbe quasi un milione di atomi), ma fortunatamente in molti casi si può fare ricorso a teorie e modelli semplificati che catturano l'essenza del problema. Nel caso delle nanostrutture di carbonio, ad esempio, il metodo *Tight-Binding* si rivela ragionevolmente accurato e fortemente predittivo. In questo metodo, per ciascun atomo di carbonio, vengono considerati solo alcuni elettroni e fra questi si trascura l'interazione reciproca. Il risultato è che si possono studiare ad esempio difetti casualmente o ordinatamente distribuiti su "campioni" confrontabili con quelli prodotti in laboratorio, e si possono quindi fare previsioni circa la loro influenza sulle proprietà elettroniche della nanostruttura.

Citiamo qui alcuni fra gli applicativi *software* oggi a disposizione del chimico teorico di maggior successo come GAUSSIAN (www.gaussian.com, il cui nome ricorda l'utilizzo di funzioni di base gaussiane), GAMESS (www.msg.chem.iastate.edu/games, un esempio interessante di software libero e *open source*), COLUMBUS (www.univie.ac.at/columbus) e MOLPRO (www.molpro.net). Questi programmi sono "Quantum Chemistry Packages" cioè, sebbene molti di loro contemplino anche algoritmi DFT, essi furono scritti inizialmente per risolvere direttamente l'equazione di Schroedinger. Fra i *software* dedicati al DFT sono sicuramente da menzionare VASP (cms.mpi.univie.ac.at/vasp), SIESTA (www.icmab.es/siesta), DACAPO (www.camd.dtu.dk/Software.aspx), ESPRESSO (www.quantum-espresso.org) e OCTOPUS (www.tddft.org/). Per ultimo, anche se non di diretto interesse per l'argomento qui trattato, è sicuramente da citare il *package* MCTDH (www.pci.uni-heidelberg.de/cms/mctdh.html), un altro codice gratuito e *open source* molto efficiente che consente di risolvere l'equazione di Schroedinger *dipendente dal tempo* per un numero arbitrario di particelle, per le quali l'operatore Hamiltoniano viene definito da *input* con una notevole libertà.



l'evoluzione di un certo stato iniziale, definendo così la funzione d'onda ad ogni istante. Applicando a quest'ultima le leggi della meccanica quantistica, è possibile ottenere tutte le informazioni sul comportamento fisico e chimico del sistema. Di particolare interesse sono le soluzioni stazionarie che, sotto determinate condizioni, esistono per certi valori (discreti) dell'energia. Queste soluzioni descrivono i possibili stati stabili nel tempo della materia; al tempo stesso vengono definiti i possibili valori (quantizzati) dell'energia.

Le funzioni d'onda sono solitamente oggetti molto complessi. In generale, sono funzioni delle coordinate di tutti i costituenti (elettroni, nuclei) del sistema in studio, il che rende l'equazione di Schroedinger impossibile da risolvere in modo esatto per sistemi realistici. Delle molte approssimazioni possibili, è fondamentale quella di Born-Oppenheimer, secondo cui è possibile separare l'equazione in un problema elettronico e uno nucleare. In effetti, il moto dei nuclei e quello degli elettroni avviene su scale di tempo molto diverse, visto che i primi possono essere diverse migliaia di volte più pesanti degli ultimi: è ragionevole quindi considerare i nuclei come fissi nello spazio e focalizzarsi sui soli elettroni e poi, se necessario, studiare il moto dei nuclei nel "campo di potenziale medio" generato dagli elettroni. Fortunatamente, per molte proprietà (esempio, elettriche e magnetiche) è sufficiente considerare le configurazioni più stabili dei nuclei e lo stato elettronico stazionario di più bassa energia corrispondente a tale configurazione.

La scala di lunghezze alla quale è necessario tenere conto del comportamento quantistico della materia è quella del *nanometro*. Il nanometro (nm) corrisponde ad un milionesimo di metro, è 75000 volte più piccolo dello spessore di un capello, più o meno corrisponde al diametro di una doppia elica di DNA¹: strutture di dimensioni fino un centinaio di nanometri sono soggette alle "strane" leggi della meccanica quantistica e la disciplina che si occupa di

studiarle e produrle prende il nome di *nanoscienza*. Le *nanotecnologie* rappresentano quella branca della nanoscienza che sfrutta appunto gli effetti dovuti alle piccolissime dimensioni per realizzare materiali con proprietà particolari, non ottenibili altrimenti².

Le nanotecnologie possono trovare oggi applicazioni nei più svariati campi come l'elettronica (nuovi sviluppi a livelli nanometrici per i computer), le biotecnologie (sinergia tra ingegneria su scala nanometrica e biologia), i materiali (creazione di nanostrutture e nanoparticelle per produrre nuovi materiali) ecc., che possono essere di interesse per la medicina, la produzione di energia, la catalisi di reazioni chimiche ecc., ma anche per prodotti di uso comune che sono già entrati nelle nostre case come articoli sportivi, tessuti tecnologici, cosmetici ecc.. Recentemente, è stata però l'elettronica a beneficiare maggiormente delle nanotecnologie: sono infatti basati su nano-materiali gli schermi OLED e le memorie *flash* di tipo NAND e DRAM, per citare qualche esempio. In aggiunta, la miniaturizzazione sempre più spinta dei circuiti integrati non potrà non tenere conto, a breve, delle particolarità del trasporto elettronico su scala nanometrica, e non vi è dubbio che in un prossimo futuro la cosiddetta *nanoelettronica* governerà il traffico sulle *mother-board* dei nostri *computers*. Ancora più avveniristica è l'opportunità (per ora solo teorica) che la meccanica quantistica offre di costruire nuovi paradigmi di computazione - il cosiddetto *quantum computing* - in cui il *bit* di informazione classico (binario) è sostituito dal cosiddetto *qubit*, generica unità elementare (per esempio lo *spin* di un elettrone) che obbedisce alle leggi della meccanica quantistica. I vantaggi sarebbero enormi: basti pensare che se *N bit* classici esistono in "soli" 2^N stati possibili, i *qubit* hanno *infiniti* stati a disposizione. Dei diversi nanomateriali che possono essere utilizzati nella nanoelettronica, la nostra ricerca si è indirizzata sui materiali a base di carbonio. Il carbonio è infatti un materiale molto ver-

¹ Per confronto, ricordiamo che la dimensione media di un atomo, quindi la lunghezza dei legami chimici, è dell'ordine di decimi di nanometri (Angstrom), mentre la più piccola forma vivente (il batterio del genere *Mycoplasma*) ha un diametro di 200 nm.

² È interessante osservare che le straordinarie proprietà di certi manufatti ceramici medievali sono legate alla dimensione nanometrica delle particelle di argento che in esse si trovano finemente disperse grazie alla particolare tecnica di lavorazione usata.

satite, economico ed *environmentally friendly*. Possiede caratteristiche simili a quelle del silicio, ma può essere facilmente preparato in forma di nanostrutture adatte alla fabbricazione di *transistor*, capacitori ecc. e, soprattutto, può offrire straordinarie proprietà elettroniche che stanno suscitando l'interesse di una comunità scientifica sempre più ampia, seguito a ruota da quello dei giganti dell'elettronica. Qui di seguito cercheremo di dare uno sguardo al futuro, pensando all'utilizzo di quell'elemento, il carbonio, che la Natura ha già eletto a fondamento della vita.

2. NANOSTRUTTURE A BASE DI CARBONIO

Il primo *transistor* fu realizzato negli anni '50 nei laboratori della società americana di telefonia *AT&T Bell Labs* (dove AT&T sta per *American Telephone and Telegraph Company*), ed era basato su cristalli di germanio. Solo in seguito si scelse di utilizzare un altro semiconduttore, il silicio, che è ancora oggi la base di tutti i sistemi microelettronici. Il silicio si trova nella tavola periodica nello stesso gruppo del germanio e quindi ne condivide le caratteristiche fondamentali; a differenza del suo predecessore, tuttavia, è un materiale molto più facile da rendere estremamente puro come richiesto dalle applicazioni in elettronica. Dal primo *transistor* ad oggi, la tecnologia ha consentito di miniaturizzare sempre più i componenti elettronici a base di silicio – e conseguentemente di migliorare enormemente le velocità di calcolo delle CPU – ma questo processo è ormai prossimo alla fine. Il silicio cristallino *non* è più stabile quando si trova in cristalli più piccoli di circa 10 nm, trasformandosi in un materiale amorfo, inutiliz-

zabile. La soluzione naturale a questo problema è quella di passare ad un altro elemento dello stesso gruppo chimico: il carbonio, l'elemento più piccolo del gruppo. Non sorprende quindi che tra le molteplici nanostrutture in studio e le loro possibili applicazioni, quelle a base di atomi di carbonio abbiano giocato il ruolo più importante nell'ultimo decennio di ricerca in nanotecnologia.

Il carbonio è un elemento chimico assai versatile, che esiste in diverse forme allotropiche (Figura 1). Alcune di esse sono tridimensionali, come per esempio il diamante (Figura 1 A) o la grafite (Figura 1 B), che pur essendo fatte entrambe di solo carbonio hanno proprietà completamente diverse (riquadro 2). Altre, più interessanti per le applicazioni nanotecnologi-

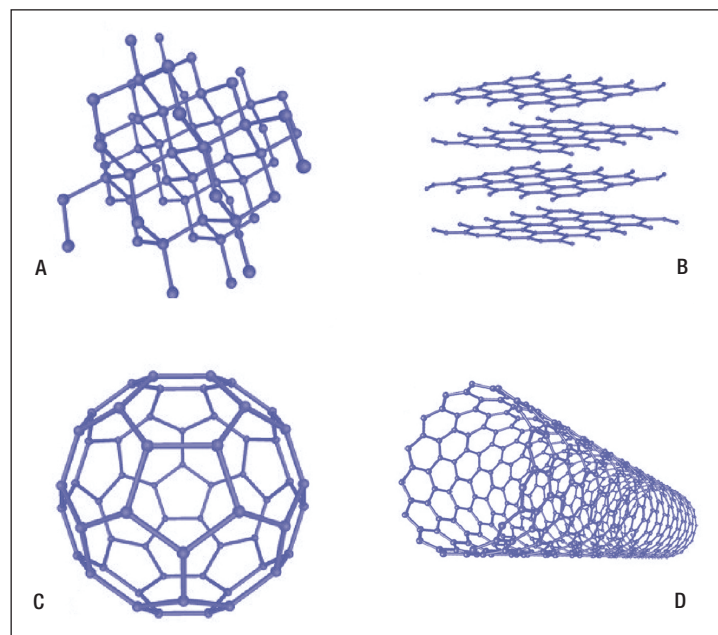


FIGURA 1

Alcune tra le strutture possibili del carbonio elementare: A - diamante; B - grafite; C - un fullerene e D - un nanotubo

RIQUADRO 2 - Le diverse facce del carbonio

Il carbonio ha 4 elettroni di valenza quindi forma sempre 4 legami con i suoi vicini. Le strutture elementari (allotropi) che esso forma differiscono solo per la disposizione geometrica dei vicini (si veda Figura 1 per alcune delle strutture più comuni). Nel diamante (Figura 1 A) gli atomi si trovano ai vertici di un tetraedro (il vecchio bricco del latte a quattro facce triangolari), con un atomo al centro. In questo modo il carbonio forma una struttura durissima, isolante sia termico che elettrico, e perfettamente trasparente che conosciamo sotto il nome di diamante. Nella grafite (Figura 1 B) gli atomi sono organizzati in piani che interagiscono solo debolmente fra loro. Su ogni piano ogni atomo di carbonio ha tre vicini disposti ai vertici di un triangolo equilatero, coi quali forma due legami semplici e un legame doppio. La struttura ad esagoni che si viene a creare, insieme alla libertà di scelta del *partner* per il doppio legame, è un tema dominante in chimica. Il più semplice esempio di tale struttura è il singolo anello, noto come benzene. Visto che il grafene (ciascun singolo foglio nella Figura 1 B) - opportunamente tagliato e reincollato - può essere trasformato in un fullerene (Figura 1 C) o in un nanotubo (Figura 1 D), il benzene può essere considerato il padre di tutte queste strutture. La "teoria della risonanza" del benzene (che lo descrive come sovrapposizione di più strutture chimiche che differiscono per l'arrangiamento dei doppi legami) spiega delle straordinarie proprietà del grafene.



che, hanno invece dimensionalità ridotta. Strutture sferiche di carbonio (in un certo senso “zero-dimensionali”) in cui gli atomi di carbonio occupano i vertici di un icosaedro troncato (Figura 1 C) (struttura simile ad un pallone da calcio) sono note come *fullereni* e valsero ai loro scopritori (Curl, Kroto e Smalley) il premio Nobel per la chimica nel 1996. Esse hanno diverse applicazioni, in particolare come superconduttori.

Se invece gli atomi di carbonio sono legati tra loro in modo da formare un “tubo” (in un certo senso “monodimensionale”) si parla di *nanotubo* (Figura 1 D): particolari strutture con proprietà (esempio, elasticità e resistenza) che le rendono adatte alla costruzione di materiali compositi. Per esempio, la bicicletta del vincitore del Tour de France 2006 aveva un telaio fatto da una resina composita contenente nanotubi, che coniuga un'estrema leggerezza con una elevata resistenza meccanica (peccato però che il vincitore - Floyd Landis - sia stato poi eliminato per *doping*! [1]). Nanotubi di carbonio di diversa forma, diametro e orientazione degli atomi possono essere facilmente sintetizzati in laboratorio, per esempio tramite scariche elettriche generate su elettrodi di grafite. Queste caratteristiche strutturali determinano le proprietà elettriche dei nanotubi, rendendoli metallici o semiconduttori, quindi potenzialmente adatti alla fabbricazione di *transistor*. Nel 2003 è stato realizzato un *transistor* costituito da un nanotubo collegato ad elettrodi di palladio che ha mostrato una transconduttanza 10 volte superiore a quella dei sistemi tradizionali a base di silicio cristallino (MOS) [2]. Poco tempo dopo sono stati presentati interi circuiti integrati a base di nanotubi; questi sono (sorprendentemente) totalmente flessibili e trasparenti al punto da suggerire una loro prossima applicazione nella tecnologia dei *touch screens* e delle celle fotovoltaiche.

Salendo di dimensionalità, arriviamo al singolo “foglio” di grafite, di un atomo di spessore, noto come *grafene*, la cui scoperta nel 2004 [3] è da molti considerata una vera rivoluzione. Sebbene immaginato già a partire dagli anni '50, la sua scoperta ha sfatato l'idea che materiali bidimensionali perfettamente ordinati non potessero esistere a causa delle naturali fluttuazioni termiche. Geim e collaboratori, dell'Università di Manchester, non solo isola-

rono il grafene (con del semplice nastro adesivo, un buon “microscopio” e... tanta pazienza!) ma dimostrarono anche che mostra l'effetto di campo elettrico, costruendo di fatto il primo *Field Effect Transistor* (FET) basato su grafene [3]. L'idea che anche i metalli - se ridotti a film monoatomici - mostrassero tale effetto trovava una verifica sperimentale in quello che per certi versi un metallo lo è, il grafene. Quello che non era ovvio all'epoca è che il grafene è molto più di un tradizionale materiale.

3. GRAFENE: IL MATERIALE DEL NUOVO MILLENNIO

Se la grafite ha segnato indiscutibilmente il millennio appena trascorso (con l'invenzione della matita, datata 1564) il grafene ha tutta l'aria di giocare da padrone per quello in corso. Il suo aspetto più straordinario è la struttura elettronica che ne fa non solo il primo vero “gas elettronico bidimensionale” (*two-dimensional electron gas*, 2DEG), ma anche l'unico “pseudo-relativistico”. In un foglio di grafene, come nei nanotubi, i singoli atomi di carbonio sono legati tra loro in modo da formare un reticolo a nido d'ape ed ognuno di loro partecipa con 4 elettroni: tre impegnati nei legami chimici con gli atomi vicini e uno libero di muoversi da atomo ad atomo. Questa particolare situazione, unitamente alla alta simmetria della struttura, porta ad avere elettroni che viaggiano come se non avessero massa: è come se il grafene fosse percorso da dei fotoni carichi, che possono trasportare corrente come gli elettroni, ma che mantengono tutte le proprietà tipiche delle particelle soggette alla teoria della relatività “ristretta” di Einstein (nella sua formulazione quantistica). Grazie a questa caratteristica unica sono stati osservati nel grafene effetti fisici mai visti prima e che possono essere sfruttati per realizzare nuovi componenti per la nanoelettronica [4, 5]. Per di più il materiale è strettamente bidimensionale, e quindi tali proprietà sono fortemente sensibili all'ambiente chimico-fisico circostante (presenza di eventuale supporto isolante, *stress* meccanici, presenza di molecole ecc.).

Già molto tempo prima che il grafene fosse stato prodotto in laboratorio, diversi gruppi di ricerca si erano dedicati allo studio di questo

materiale. Le prime trattazioni teoriche erano basate su di un modello molto semplice, chiamato “*tight binding*” (TB). Il modello TB consiste nel descrivere l'interazione dell'unico elettrone (per atomo di carbonio) libero di muoversi come se questo “saltasse” da un atomo all'altro lungo tutta la struttura del grafene. Ovviamente questo è un modello molto semplificato, ma sorprendentemente è già sufficiente a descrivere le caratteristiche principali della struttura elettronica. In particolare, il risultato più interessante è che l'energia elettronica dipende *linearmente* dal “momento” e quindi gli elettroni si muovono con velocità *costante* (indipendentemente dalla loro energia) come succede ad esempio per i fotoni. Per contro, nei semiconduttori tradizionali l'energia è funzione quadratica del momento, una situazione questa che assomiglia a quella di particelle non-relativistiche. Nel grafene puro la conduzione di corrente è minima visto che, strettamente parlando, non vi sono stati di conduzione a disposizione, ma basta sottoporre il materiale ad un potenziale esterno (cosiddetto di “*gate*”) o appoggiarlo su un substrato elettron-ricco o elettron-povero, perché si osservi vera conduzione. La corrente può essere trasportata tanto dagli elettroni quanto dalle “buche” (che rappresentano “carenza” di elettroni) e, in entrambi i casi, ha le stesse straordinarie proprietà di cui sopra, visto che c'è perfetta simmetria tra i due tipi di portatori di carica (in analogia alla simmetria tra particelle e antiparticelle della fisica delle alte energie). Le conseguenze di tutto ciò sono enormi e portano ad effetti fisici del tutto nuovi che sarebbe impossibile trattare in modo esaustivo in questo articolo. Citiamo solamente il trasporto balistico della corrente elettrica su scala micrometrica: cioè la relativa insensibilità del moto degli elettroni agli ostacoli che trovano lungo il loro percorso (impurezze, difetti della struttura atomica del grafene ecc.), che quindi viaggiano velocissimi (a circa 1/100 della velocità della luce), quasi indisturbati lungo il reticolo di atomi di carbonio. Questa caratteristica permetterebbe la realizzazione di *transistor* e microprocessori ad altissime prestazioni, da impiegare nei futuri computers. In effetti, è recente la notizia della fabbricazione di un *transistor* grafenico con una frequenza massima di lavoro pari a 100 GHz (IBM) [6]. Se

è vero che le frequenze di *clock* delle CPU sono ~10 volte più piccole di quelle di lavoro dei transistor che contengono, c'è da aspettarsi comunque un notevole passo in avanti già da questo risultato.

4. GRAFENE: LA MATERIA PRIMA

Quanto detto circa le straordinarie proprietà del grafene sarebbe poco utile se non fosse possibile modificare queste proprietà a seconda delle applicazioni, cioè disegnare strutture “grafeniche” che enfatizzino un aspetto o l'altro a seconda dell'interesse. In questo processo di *material design*, la chimica teorica (che si intreccia qui con la fisica e la chimica dello stato solido) non solo spiega fenomeni già osservati ma, con l'aiuto della teoria e del calcolo numerico, è in grado di suggerire le modifiche opportune da effettuare sul materiale su scala nanometrica affinché si abbia una certa proprietà. Fra queste, per esempio, la realizzazione di un vero *transistor* basato sul grafene richiederebbe l’“apertura” di un *gap* nella struttura a bande del materiale visto che, strettamente parlando, il grafene è un semiconduttore a *zero gap*. Il *gap* è necessario se si vuole “spegnere” totalmente la corrente ed avere un rapporto tra il segnale “on” e quello “off” sufficientemente elevato (diversi ordini di grandezza) per essere utile in applicazioni logiche [7, 8]. Sono molte le proposte teoriche, basate sull'intuito e il calcolo numerico, per la creazione di un *band gap* nel grafene. Alcune utilizzano la approssimazione TB mentre altre, più accurate, fanno uso della teoria del funzionale della densità elettronica (DFT).

Una, forse la più semplice, consiste nel tagliare il foglio di grafene in strisce sottili di qualche nm di larghezza (Figura 2), chiamate appunto *nanoribbon* (riquadro 3 a p. 44). Nei *nanoribbon* gli elettroni non possono muoversi liberamente in tutte le direzioni, ma sono naturalmente costretti dalla larghezza del *ribbon*. Questa limitazione induce il *bandgap*, che sarà tanto più grande quanto più stretto è il *ribbon*, come si può facilmente dedurre con qualche considerazione fisica. Ciò che è meno ovvio (e che la teoria ed il calcolo mostrano indiscutibilmente) è come la morfologia del bordo del *ribbon*, e quindi la direzione di taglio del foglio di grafene, sia fondamentale per la struttura elettronica [9]. Nel

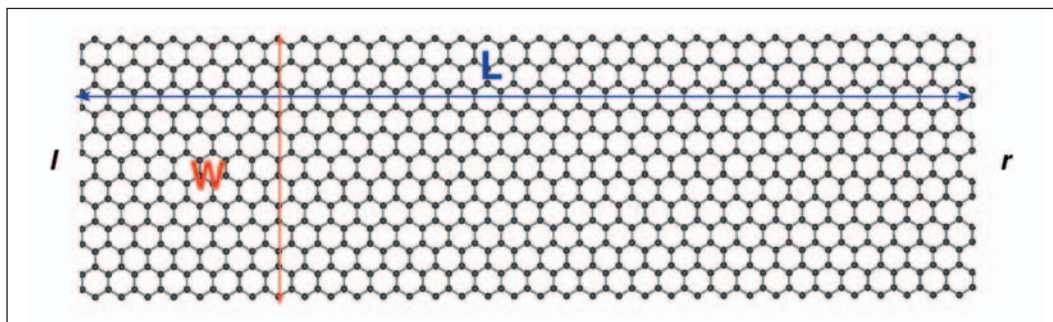


FIGURA 2

Un nanoribbon di larghezza W e lunghezza L . Anche indicati gli ipotetici contatti (l, r) che ne fanno un transistor a effetto di campo (anche noti come source e drain)

RIQUADRO 3 - Nanoribbon di carbonio e altre nanostrutture

Un nanoribbon (Figura 2) è una striscia di grafene, caratterizzata da due dimensioni, la larghezza W e la lunghezza L ($L \gg W$). Quest'ultima definisce la lunghezza del "canale" che si viene a creare quando la nanostruttura è collegata a due contatti (l e r nella figura) e un campo elettrico (di *gate*) è applicato ortogonalmente alla pagina.

I nanoribbons hanno proprietà diverse non solo a seconda di W e L ma anche dipendentemente dal tipo di linea lungo la quale è fatto il taglio. Nella figura qui sopra, la larghezza W è la distanza tra i due bordi con conformazione a "zig-zag"; nanoribbon "armchair" si ottengono invece quando W separa i bordi verticali in figura. Il primo caso non è ottimale per l'apertura di un gap di banda ma è molto interessante per le sue proprietà magnetiche: la teoria prevede infatti (e l'esperimento parzialmente conferma) ordine ferromagnetico su ciascun bordo zig-zag, e ordine antiferromagnetico tra un bordo e l'altro. Il caso armchair, diversamente, rappresenta la situazione migliore per l'apertura di un gap di banda, che risulta proporzionale a $\sim W^{-1}$.

La figura 3, invece, riporta un grafene "decorato", in cui atomi di carbonio (in nero) sono sostituiti da atomi estranei (blu) o sono semplicemente eliminati (a creare quella che si chiama una "vacanza"). Gli atomi estranei possono essere per esempio Boro e/o Azoto, analogamente a quanto avviene nei comuni semiconduttori drogati; le vacanze possono estendersi a formare isole più o meno circolari e sperimentalmente si possono realizzare con notevole precisione con le moderne tecniche nanolitografiche. La decorazione qui mostrata è perfettamente ordinata e costituisce un superreticolo, che si sovrappone alla struttura ordinata degli atomi di carbonio. Il superreticolo mostrato in figura 3 definisce quello che possiamo chiamare un "supergrafene".

Decorazioni come vacanze o singoli atomi che si legano covalentemente agli atomi di carbonio danno luogo ad interessanti proprietà magnetiche. Nella figura 4, per esempio, è riportata la "densità di magnetizzazione" originata da un singolo difetto (un atomo di Idrogeno - in verde - legato all'atomo di Carbonio al centro della figura): si vede che le regioni di spazio a magnetizzazione maggiore (rosso) si localizzano su un atomo di C ogni due, in accordo con la teoria della risonanza. Più difetti possono "compensare" o enfatizzare la magnetizzazione, dipendentemente dal loro arrangiamento.

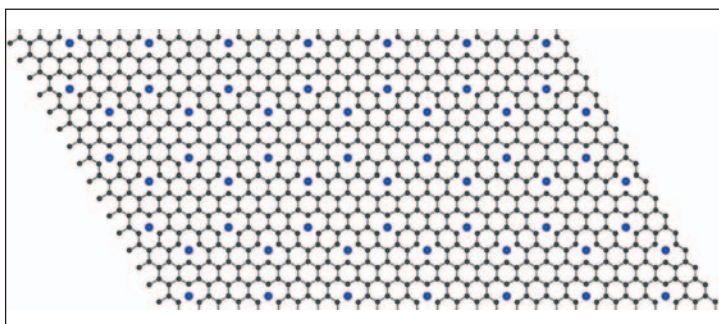


FIGURA 3

Un grafene decorato. La struttura va pensata estendersi indefinitamente nelle due dimensioni (cioè, in pratica, su lunghezze molto maggiori della distanza media tra le "decorazioni")

peggiore dei casi, quando la morfologia non è adatta all'apertura di un gap di banda, si ha una magnetizzazione spontanea dei bordi! E questo potrebbe essere utilizzato per separare la cor-

rente elettrica nelle sue due componenti di *spin* (ciò che sta alla base della cosiddetta *spintronica*) applicando semplicemente un campo elettrico esterno [10].

Recentemente, il progresso delle tecniche nanolitografiche ha permesso di verificare la correttezza delle previsioni teoriche, attraverso misure accurate della struttura a bande di nanoribbons di dimensioni controllate. E sono stati realizzati pure i primi FET basati sui nanoribbons, sebbene le loro prestazioni (in termini di rapporti *on/off*) non siano ancora del tutto soddisfacenti come nel caso dei nanotubi a causa della difficoltà nel controllare il taglio [7, 8, 11].

Tra le alternative, la "decorazione" del grafene che si ottiene adsorbendo atomi o molecole sulla sua superficie, oppure rimuovendo atomi di carbonio, è una proposta molto interessante (Figura 3). Tanto più che le tecniche nanoli-

tografiche odierne consentono di creare nella trama del grafene buchi di diametro controllato (fino a 1 nm) e periodicamente arrangiati a distanze piccole come 2-3 nm. La teoria e i calcoli mostrano come le proprietà elettroniche del grafene possano cambiare in funzione del tipo di “decorazione” e dei suoi parametri geometrici. Molto recentemente, ad esempio, abbiamo mostrato come certe “superstrutture” (che possiamo chiamare “supergrafeni”) che si ottengono creando dei fori circolari sul grafene in modo che formino un altro reticolo a nido d’api, abbiano delle proprietà molto interessanti, che coniugano quelle del grafene originario con la necessità di avere un *gap* nella struttura a banda [12].

I vantaggi della decorazione sono enormi. Sotto alcune condizioni, le proprietà elettroniche nei dintorni del “difetto” introdotto sono uniche, al punto che il grafene può mostrare *magnetizzazione* (Figura 4). L’esistenza di stati elettronici quasi-localizzati (e non schermati) è un risultato notevole previsto qualche anno fa dalla teoria e dal calcolo numerico [13] e, molto recentemente, verificato sperimentalmente [14] con il microscopio ad effetto *tunnel* (*Scanning Tunneling Microscope*, STM). Grazie a questo strumento non solo si ottengono immagini con risoluzione *atomica* dei fogli di grafene ma si possono anche fare “misure” di magnetizzazione e di conducibilità locale ad una risoluzione confrontabile con la precedente. Le possibilità applicative offerte dal grafene appaiono al momento innumerevoli, e la teoria e il calcolo numerico ogni giorno forniscono idee nuove e materiale per potenziali nuove applicazioni. Questo è possibile perché la facilità con la quale si riescono a studiare “*in silico*” tante nuove nanostrutture compensa abbondantemente le difficoltà intrinseche nella loro realizzazione (che quindi si può limitare a quelle strutture considerate più interessanti). Ironia della sorte, quindi, il silicio sta giocando un ruolo chiave anche in quella ricerca che lo vuole sostituire.

Bibliografia

- [1] CNET News, http://news.cnet.com/Carbon-nanotubes-enter-Tour-de-France/2100-11395_3-6091347.html?tag=fd_carsl
- [2] NEC, <http://www.nec.co.jp/press/en/0309/1901.html#chu1>

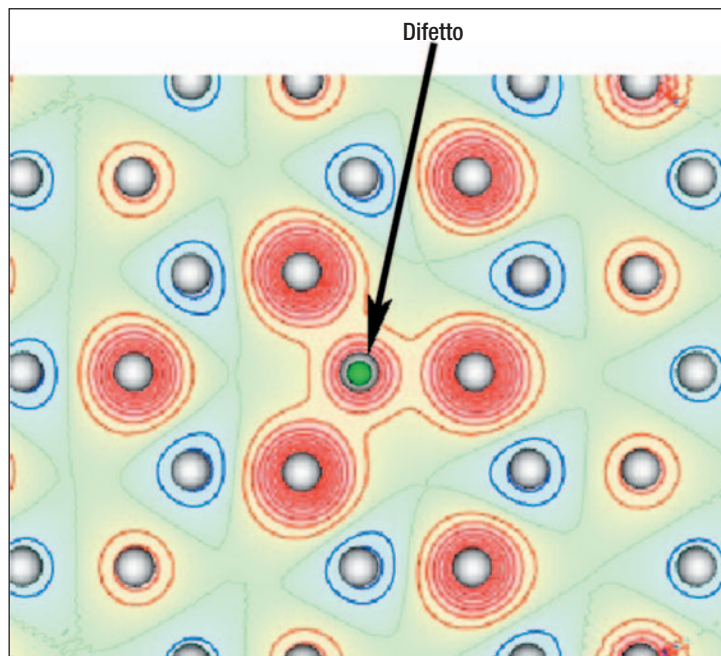


FIGURA 4

La densità di magnetizzazione indotta da un difetto (in verde) al centro della figura. In rosso le regioni a magnetizzazione maggiore

- [3] Novoselov K.S., Geim A.K., Morozov S.V., Jiang D., Zhang Y., Dubonos S.V., Gregorieva I. V., Firsov A.A.: Electric-field effect in atomically thin carbon films. *Science*, Vol. 306, 2004, p. 666.
- [4] Geim A.K., Novoselov K.S.: The rise of graphene. *Nature Materials*, Vol. 6, 2007, p. 183.
- [5] Castro Neto A.H.: The Carbon New Age. *Materials Today*, Vol. 13, 2010, p. 12.
- [6] Lin Y.-M., Dimitrakopoulos C., Jenkins K.A., Farmer D.B., Chiu H.-Y., Grill A., Avouris P.: 100-GHz Transistors from Wafer-scale Epitaxial Graphene. *Science*, Vol. 327, 2010, p. 662.
- [7] Avouris P., Chen Z., Perbeinos V.: Carbon based electronics. *Nature Nanotechnology*, Vol. 2, 2007, p. 605.
- [8] Schwierz F.: Graphene transistors. *Nature Nanotechnology*, in stampa (doi:10.1038).
- [9] Son Y.-W., Cohen M.L., Louie S.G.: Energy Gaps in Graphene Nanoribbons. *Physical Review Letters*, Vol. 97, 2006, 216803.
- [10] Son Y.-W., Cohen M.L., Louie S.G.: Half-metallic graphene nanoribbons. *Nature*, 2006, Vol. 444, p. 347.
- [11] Han M.Y., Özyilmaz B., Zhang Y., Kim P.: Energy Band-Gap Engineering of Graphene Nanoribbons. *Physical Review Letters*, Vol. 98, 2007, 206805.
- [12] Martinazzo R., Casolo S., Tantardini G.F.: Symmetry-induced bandgap opening in graphene superlattices. *Physical Review B*, Vol. 81, 2010, 245420.

[13] Pereira V.M., Guinea F., Lopes dos Santos J.M.B., Peres N.M.R., Castro Neto A.H.: Disorder Induced Localized States in Graphene. *Physical Review Letters*, Vol. 96, 2006, 036801.

[14] Ugeda M.M., Brihuega I., Guinea F., Gómez-Rodríguez J.M.: Missing Atom as a Source of Carbon Magnetism. *Physical Review Letters*, Vol. 104, 2010, 096804.

GIAN FRANCO TANTARDINI è Professore Ordinario di Chimica Fisica presso l'Università degli Studi di Milano, Docente del Corso di Laurea Triennale in Chimica e del Corso di Laurea Magistrale in Scienze Chimiche, membro del Consiglio Direttivo della Scuola di Dottorato in Scienze e Tecnologie Chimiche, Direttore del Dipartimento di Chimica Fisica ed Elettrochimica, membro della Commissione di Ateneo per la Ricerca Scientifica e il Trasferimento Tecnologico, ed è stato membro del Consiglio di Amministrazione. È stato Docente di Chimica Fisica presso l'Università degli Studi dell'Insubria e Presidente del Gruppo Interdivisionale di Chimica Computazionale della Società Chimica Italiana.

Svolge la sua attività di ricerca nel campo della Chimica Fisica Teorica e Computazionale.

E-mail: gianfranco.tantardini@unimi.it

ROCCO MARTINAZZO è dal 2004 Ricercatore universitario presso il Dipartimento di Chimica-Fisica ed Elettrochimica dell'Università degli Studi di Milano. Ottenuto il Dottorato di Ricerca in Scienze Chimiche presso la stessa Università nel 2002, ha passato lunghi periodi di ricerca presso l'Università degli Studi "La Sapienza" di Roma e l'Università di Potsdam. Si interessa di processi elementari all'interfaccia gas-solido e allo sviluppo di teorie e metodi computazionali in ambito quantistico.

E-mail: rocco.martinazzo@unimi.it

SIMONE CASOLO ha conseguito il titolo di Dottore di Ricerca in Scienze Chimiche nel 2009 presso l'Università degli Studi di Milano, svolgendo una tesi di chimica teorica/computazionale sull'interazione idrogeno-grafene, in collaborazione con l'Università di Oslo. È Assegnista di Ricerca presso il Dipartimento di Chimica-Fisica ed Elettrochimica e svolge la sua attività su derivati del grafene.

E-mail: simone.casolo@unimi.it

MODELLAZIONE AL COMPUTER DI BARCHE A VELA DA COMPETIZIONE

Lo straordinario sviluppo della capacità di calcolo unitamente a quello di codici numerici, ha proiettato molto rapidamente la ricerca nel campo delle applicazioni nautiche verso l'utilizzo di nuovi strumenti di indagine e acquisizione dati. La diffusione di nuovi programmi di calcolo adatti allo studio di imbarcazioni nasconde dietro a sé lo sviluppo di nuove tecnologie e conoscenze, frutto del lavoro di ricerca condotto negli ultimi anni in vari ambiti quali la matematica numerica, la visualizzazione scientifica e il calcolo parallelo. In questo articolo vengono introdotte e discusse le potenzialità di questi nuovi approcci.

1. INTRODUZIONE

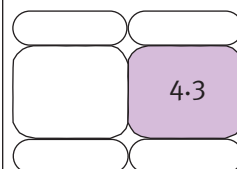
Il potenziamento dei moderni sistemi di calcolo ha permesso un inaspettato sviluppo della ricerca scientifica e quindi delle sue applicazioni anche nel settore della nautica. Programmi sempre più sofisticati e precisi possono girare su un semplice *laptop* e hanno costi contenuti e accessibili a tutti. Sia i piccoli cantieri di imbarcazioni da crociera sia i disegnatori di lussuosi mega-yacht, si stanno velocemente fornendo dei più recenti strumenti di progettazione e gestione dei processi. La diffusione di nuovi programmi di calcolo nasconde dietro a sé lo sviluppo di nuove tecnologie e conoscenze, frutto del lavoro di ricerca condotto negli ultimi anni in vari settori scientifici tra cui l'analisi numerica, il *software design* e la visualizzazione scientifica (Figura 1).

Come avviene in molti altri settori, per esempio quello dell'*automotive*, anche nella nautica, una delle maggiori opportunità per la ricerca scientifica risiede nelle competizioni sportive in grado di muovere ingenti capitali che giustificano importanti investimenti nella ricerca. Inoltre, regate come l'*America's Cup* (AC) sono

precedute da periodi di preparazione pluriennali in cui vere e proprie rivoluzioni tecnologiche segnano nuovi limiti e riferimenti tecnici. Nel 2007, per la prima volta nella storia della progettazione di imbarcazioni per l'AC, un team (che è risultato poi essere quello vincitore) ha dichiarato che la principale scelta nella strategia di progettazione è stata fortemente guidata da studi effettuati con tecniche numeriche di fluidodinamica computazionale (CFD). Questa affermazione segna una significativa evoluzione nell'uso di questo tipo di tecniche, storicamente relegate ad analisi di complemento o di verifica, che oggi mostrano finalmente un livello di affidabilità tale da permettere di essere considerate come primario strumento di indagine per lo studio e la progettazione di imbarcazioni a vela da competizione. La progettazione di un'imbarcazione a vela coinvolge sia problematiche strutturali che fluidodinamiche, ovvero sia lo studio della struttura sia lo studio del comportamento (dinamico) dei fluidi in cui la barca si muove. Questi due aspetti sono quanto mai intrecciati tra loro nel caso di una barca; tuttavia in questa pubblicazione



Ignazio Maria Viola
Raffaele Ponzini



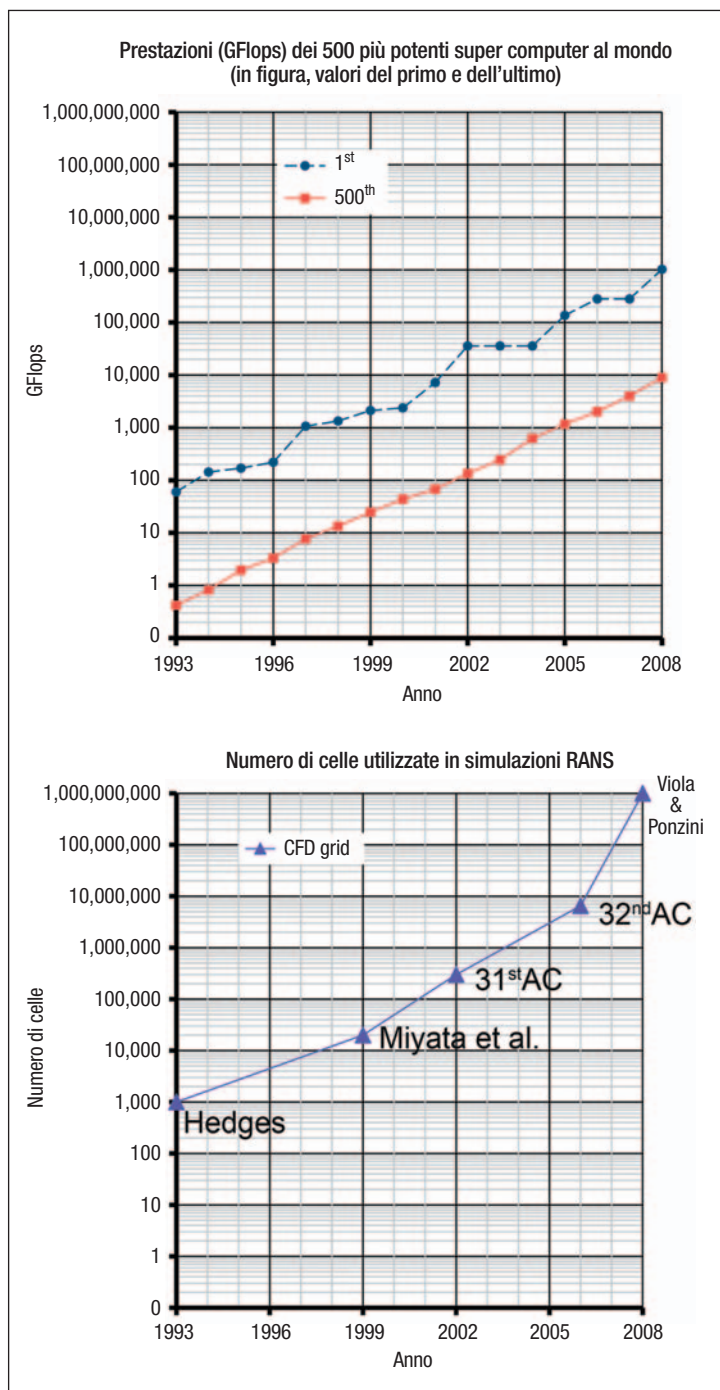


FIGURA 1

Comparazione tra andamento dello sviluppo tecnologico in termini di risorse di calcolo disponibili in GFlops (classifica top500.org) e numero di celle utilizzate nella modellazione CFD di imbarcazioni da regata

prenderemo in considerazione solo lo studio dell'aerodinamica di un'imbarcazione a vela. Più in dettaglio, in questo articolo faremo riferimento a studi di aerodinamica tramite modelli CFD semplificati cioè atti a valutare caratteristiche del

piano velico in condizioni statiche, trascurando i movimenti dell'imbarcazione e le variazioni di intensità e direzione del vento nel tempo. Per lo studio dei piani velici, modelli CFD semplificati permettono di calcolare l'equivalente delle grandezze che possono essere misurate tramite la galleria del vento, con livelli di accuratezza e di flessibilità superiori, e con costi inferiori a quelli sperimentali. Affinché i risultati dei modelli CFD possano essere considerati affidabili, e quindi utilizzati per lo studio di qualsiasi configurazione dell'imbarcazione, è necessario effettuare delle validazioni tra valori misurati in galleria e valori calcolati dai modelli in condizioni equivalenti. Come verrà discusso nel seguito, per avere dei calcoli accurati è necessario pagare un prezzo in termini di risorse computazionali. Questo articolo analizza il divario esistente tra risorse necessarie e risorse disponibili nello studio della fluidodinamica dei piani velici tramite modelli CFD. L'articolo, dopo aver riassunto l'evoluzione storica delle applicazioni numeriche in ambito velico, analizzerà brevemente le differenti tipologie di modellazione CFD possibili in questo ambito per poi discutere più in dettaglio quanto ottenuto all'interno di un progetto di collaborazione scientifica in atto tra la sezione calcolo del CILEA e il Yacht Research Unit di Auckland (NZ). Nella conclusione verranno tracciate alcune indicazioni circa possibili sviluppi futuri di questo tipo di applicazioni.

2. CICLO CAE E PROGETTAZIONE NAUTICA

Gli anni settanta hanno rappresentato un momento di grande sviluppo della dinamica delle imbarcazioni a vela. Nel mondo accademico, Marchay [1] all'Università di Southampton realizzava le prime prove in galleria del vento su piani velici e, contemporaneamente, Milgram al MIT applicava i primi codici a pannelli che aveva utilizzato in precedenza su profili alari [2, 3], pubblicando pochi anni dopo uno scritto sui coefficienti aerodinamici delle imbarcazioni a vela [4]. Negli stessi anni, nel mondo dell'industria, Gentry, della Boeing Company, applicava codici a pannelli sviluppati nel settore aeronautico alle imbarcazioni classe Coppa America, tra cui il 12-metri Courageous che difese la coppa nel 1974 e nel 1977, Freedom nel 1980, e Liberty nel 1983 [5]. Anche per le appli-

cazioni idrodinamiche, gli anni settanta sono stati un momento di fondamentale sviluppo. A questo decennio si debbono le prime prove sistematiche nella vasca navale di Delft [6] e i primi modelli VPP realizzati al MIT, che univano modelli idro/aero-dinamici in grado di stimare la velocità e l'assetto dell'imbarcazione [7, 8]. Da allora fino alla fine del secolo, le tecniche numeriche sono cresciute in ombra rispetto a quelle sperimentali, ma l'iniziale svantaggio competitivo legato all'eccessivo onere computazionale per i calcolatori di cui si disponeva, è andato diminuendo con lo sviluppo tecnologico. La campagna di Coppa America del 2003 ha segnato il primo passo per un utilizzo massiccio delle tecniche numeriche insieme a quelle sperimentali, soprattutto in ambito idrodinamico. Infatti, il calcolo numerico permette nei mesi prossimi alla competizione, di mettere a punto le ultime modifiche in tempi più stretti rispetto a quelli di una prova sperimentale in vasca. Si pensi, ad esempio, alla scelta della pinna e del bulbo che può essere fatta poco prima della regata. Per questo motivo, l'unione dell'economicità dello strumento rispetto alla prova in vasca e della discreta accuratezza che si ottiene anche con tecniche che trascurano la viscosità, ha permesso all'idrodinamica numerica di essere la prima a rimontare il gap che la separava dalla prova sperimentale.

Nel campo dell'aerodinamica, Michael Richelsen scrisse negli anni ottanta i codici per la velieria North Sails "Flow" e "Membrain", rispettivamente un codice a pannelli inviscido per il calcolo della distribuzione di pressione e un codice a elementi finiti (FEM) per il calcolo strutturale. Le versioni aggiornate di questi codici, sempre riviste dallo stesso autore, sono tuttora lo strumento principale per il disegno delle vele di bolina di barca da regata. L'accuratezza e la semplicità di questi codici ne hanno fatto la fortuna. Tuttavia, la natura inviscida di "Flow" non consente di modellare la separazione, creando il presupposto per la diffusione dei codici CFD. Questi ultimi hanno tardato molto a essere largamente impiegati, sia per la complessità di utilizzo, sia per il costo computazionale che fino alla fine del secolo aveva un bilancio costo/beneficio troppo sfavorevole. Nella Coppa America del 2003, quasi tutti i consorzi si sono timidamente affacciati all'uti-

lizzo di questi codici, principalmente sfruttando la vera innovazione che portavano con sé: la possibilità di risolvere il campo di moto in un ampio volume di calcolo anziché solo sulla superficie della vela e della scia come nel caso dei codici inviscidi a pannelli. Per questo motivo, i codici CFD sono stati utilizzati come strumento di visualizzazione del flusso da associare alle prove in galleria del vento, oppure come strumento per lo studio del posizionamento reciproco tra due barche.

Nella Coppa America del 2007, l'alzarsi del livello competitivo, il maggiore sforzo di ricerca richiesto da una classe già largamente strizzata perché in uso dal 1992, i tempi serrati imposti dal sistema degli ACT della Louis Vuitton Cup, ha lanciato appieno l'utilizzo di codici numerici. I codici CFD non sono stati utilizzati solo come strumento di visualizzazione, ma come vero e proprio strumento di progettazione per le vele in andature portanti, per la progettazione delle appendici di carena e, in taluni consorzi, anche per la progettazione delle forme di carena superando i più che collaudati codici inviscidi.

Infine, nell'ultima Coppa America, che ha visto protagonisti dei giganteschi multi-scafo, uno dei quali con un'ala rigida assai più grande dell'ala di un Boeing 747, l'analisi sperimentale è stata rilegata alla validazione dei modelli CFD.

Oggi, i codici CFD sembrano destinati a essere sempre più utilizzati e lo sviluppo tecnologico è talmente veloce che risulta molto difficile prevederne il futuro. È facile immaginare che presto supereranno i codici inviscidi in molte applicazioni e che diverranno sempre più utilizzati nella progettazione di ogni forma, sia essa scafo o vela.

In queste considerazioni vi sono le motivazioni più profonde del presente articolo. È infatti opinione degli autori che la pressante pulsione all'utilizzo di tecniche CFD debba necessariamente partire da un'indagine attenta dei limiti di questo nuovo strumento di indagine. Anche per questa ragione si è deciso di discuterne qui non solo i confini "esterni", ovvero le frontiere dell'integrazione dei diversi aspetti della dinamica dell'imbarcazione, ma anche quelli "interni", ovvero i limiti e i pregi di una simulazione numerica che riproduca quanto misurabile in galleria del vento. In

questo articolo, pertanto, da una parte si vuole mettere in luce le potenzialità conoscitiva intrinseca nella modellazione CFD in termini di quantificazione degli aspetti fluidodinamici e della visualizzazione degli stessi che consentono al progettista delle vele di discernere quantitativamente configurazioni di progetto favorevoli e sfavorevoli, dall'altra parte si vuole discutere circa l'opportunità di utilizzare il calcolo numerico come alternativa all'attività sperimentale stessa.

Le simulazioni numeriche discusse nel seguito si riferiscono a geometrie di vele rilevate in galleria del vento con tecniche fotogrammetriche, durante la misura delle forze aerodinamiche. In questo modo è quindi possibile ottenere i dati, in termini di geometria e condizioni al contorno, necessari per sviluppare i modelli CFD e, successivamente, confrontare i risultati numerici con i dati sperimentali rilevati nelle stesse condizioni. Il ciclo di modellazione numerica di un problema di aerodinamica del piano velico di un'imbarcazione da regata è mostrato nella figura 2. I passaggi fondamentali rappresentati sono:

i. (box-blu) la fase di *pre-processing* in cui a partire dai dati ottenuti sul modello fisico (in galleria del vento) si ottiene un modello CFD costituito da una geometria (che verrà discretizzata) e dalle condizioni fluidodinamiche che si sono misurate agli estremi di quello che verrà considerato come dominio di calcolo;

ii. (box-verde) la fase di calcolo vero e proprio in cui le equazioni costitutive della fluidodinamica vengono risolte per il modello CFD precedentemente ottenuto sfruttando architetture di calcolo parallelo e interconnessioni ad alta banda passante;

iii. (box-arancio) la fase di *post-processing* in cui i valori del campo di moto e di pressioni così calcolati vengono processati per ottenere informazioni qualitative e quantitative tramite la visualizzazione scientifica utili alla progettazione.

Nel seguito dell'articolo si farà sempre riferimento a modelli CFD ottenuti in questo modo e su cui sono state realizzate un grande numero di simulazioni variando di volta in volta le condizioni di navigazione, il modello numerico adottato per risolvere le equazioni della fisica che governano questo problema, la risoluzione della discretizzazione spaziale adottata (espressa in termini di numero di celle) al fine di approfondire i limiti e le potenzialità di questo approccio applicativo.

La CFD applicata allo studio dei piani velici, si inserisce nell'ambito della fluidodinamica esterna a bassi numeri di Mach e medio/alti numeri di Reynolds, pertanto richiede la risoluzione delle equazioni di Navier Stokes (NS) con condizioni al contorno sulla velocità e sulla pressione. Le equazioni fondamentali di conservazione della massa, bilancio della quantità di moto e dell'energia interna sono risolte per il

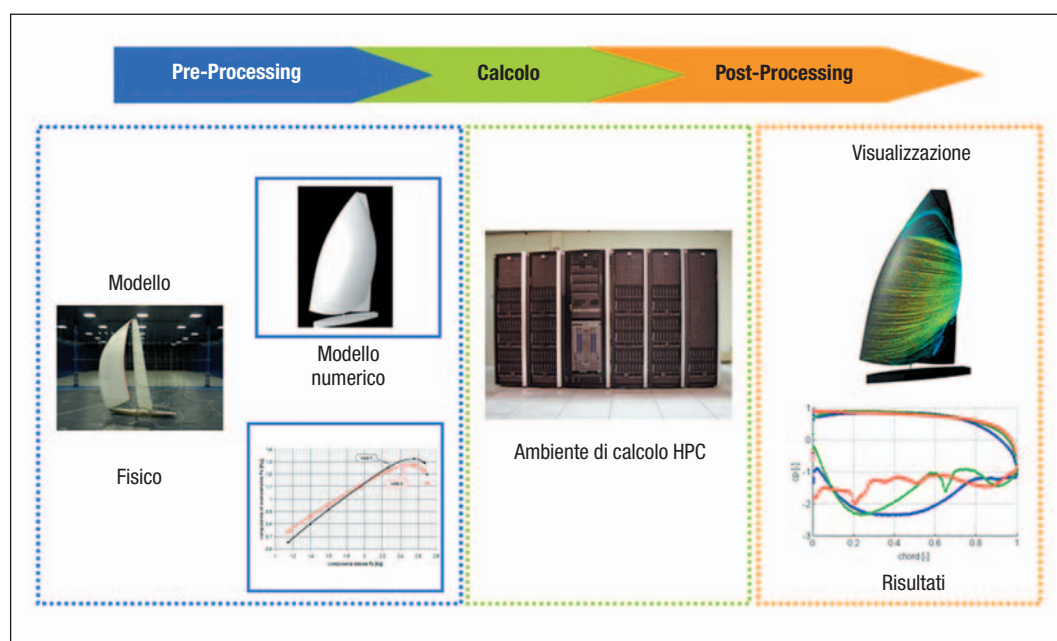


FIGURA 2
Ciclo di progettazione al calcolatore in ambito navale

fluido aria all'interno di una determinata geometria opportunamente discretizzata.

I bassi numeri di Mach consentono di ritenere ininfluyente le variazioni di densità, di conseguenza assumendo la viscosità costante, l'equazione dell'energia diventa disaccoppiata dalle altre equazioni e il campo di moto e di pressione è definito dalle sole prime due equazioni (conservazione della massa e della quantità di moto). Il sistema così ottenuto è risolvibile con le dovute condizioni iniziali e al contorno. Si osservi che le equazioni di Navier Stokes incomprimibili hanno la particolarità di perdere il significato fisico di pressione che compare solamente come gradiente di pressione. La prima delle due equazioni che deriva dalla conservazione della massa è detta *equazione di continuità*, la seconda che deriva dal bilancio di quantità di moto è detta *equazione del momentum* (riquadro a p. 62, *Equazioni di Navier Stokes* e, in particolare, le equazioni a.1 e a.3).

3. REYNOLDS

Le equazioni di NS incomprimibili possono essere riscritte esplicitando il numero adimensionale di Reynolds (Re). Da questa formulazione è possibile osservare che la soluzione del problema in esame è univocamente definita dalle condizioni al contorno e dal numero di Re , il quale esprime il rapporto tra le forze convettive e quelle viscosive diffusive. Nello studio dei piani velici in galleria del vento, i numeri di Re variano da 500.000 a 1.500.000 mentre al vero si hanno Re da 1.000.000 a 30.000.000. La differenza in termini di Re massimi raggiungibili in galleria del vento è un limite della metodologia sperimentale che non è possibile superare per motivazioni prettamente tecnologiche (resistenza dei materiali). La simulazione numerica non ha questo limite e, anzi, si presenta anche come un possibile strumento per scalare in vera grandezza i risultati ottenuti in galleria del vento.

3.1. DNS

Il sistema così descritto potrebbe essere discretizzato e risolto esattamente, in questo caso si parla di risoluzione diretta delle equazioni, in inglese *Direct Navier Stokes* (DNS). Tuttavia considerazioni energetiche ci imporrebbe-

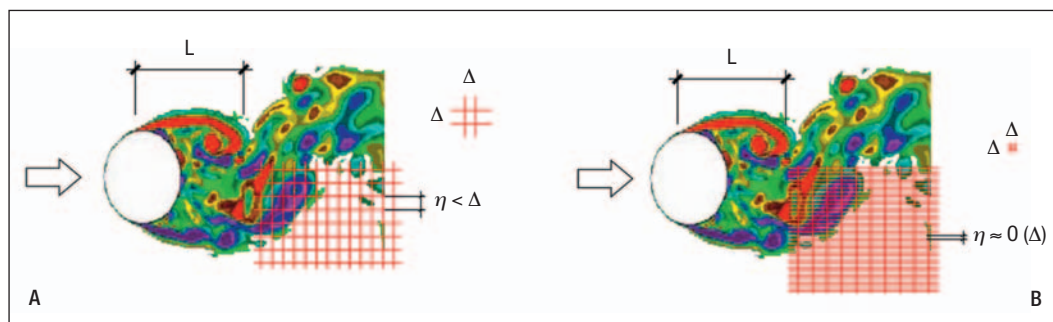
ro una discretizzazione eccessivamente onerosa per alti Re . Questo è uno dei punti fondamentali riguardo all'affidabilità delle simulazioni numeriche. Infatti, nella turbolenza 3D (al contrario della turbolenza 2D) il trasporto energetico avviene dalle strutture vorticosi più grandi, che sono della dimensione dell'oggetto immerso nel campo di moto, alle strutture vorticosi più piccole, che sono di dimensione tale che le strutture vorticosi sono dissipate per effetto viscoso. Pertanto, è necessario che la discretizzazione spaziale sia in grado di risolvere le scale più piccole ove l'energia è dissipata.

Analogamente, le scale temporali devono consentire di cogliere i tempi caratteristici delle strutture più grandi e delle strutture più piccole a cui avviene la dissipazione. Le strutture vorticosi possono essere considerate delle fluttuazioni di velocità e pressione associate a delle frequenze di fluttuazione, le quali sono sempre più alte al diminuire della dimensione spaziale della struttura vorticosi.

Utilizzando considerazioni dimensionali formulate la prima volta da Kolmogorov nel 1941 [9], si mostra che il numero di elementi N , e il numero di passi temporali T_s , necessari a cogliere le dimensioni a cui avviene la dissipazione viscosa sono esprimibili in funzione del numero di Reynolds. In particolare, N è dell'ordine di $Re^{3/4}$ e T_s risulta essere dell'ordine di $Re^{1/2}$. Considerando la necessità di discretizzare in tre dimensioni spaziali si ottiene per le quantità spaziali il valore di $Re^{9/4}$. In conclusione, il numero complessivo di gradi di libertà (gdl) descritti dal prodotto del numero di celle per ogni passo temporale che deve essere calcolato, risulta dell'ordine di $Re^{11/4}$.

Il flusso intorno a un cilindro è mostrato nella figura 3 (tratta da [10]). Si può notare come le strutture vorticosi tramite cui si immette energia nel sistema sono di dimensione dell'ordine di grandezza del diametro del cilindro (L). Queste vanno via via trasformandosi in strutture più piccole fino a scale di dimensioni pari a η , dove le strutture sono dissipate per effetto viscoso. La griglia di calcolo che consentirebbe la risoluzione diretta delle equazioni dovrebbe avere dimensione Δ dello stesso ordine di grandezza di η . In applicazioni nautiche, poiché η è molto piccolo rispetto ad L , questa condizione richiederebbe una griglia così fine

FIGURA 3
Strutture turbolente
nella scia di un
cilindro [10]



da avere un numero (ad oggi) irraggiungibile di elementi. Pertanto la dimensione di griglia Δ è tipicamente molto maggiore di η . Nelle applicazioni in esame, considerando l'ordine di grandezza dei numeri di Reynolds cui sono realizzate le prove in galleria del vento, $Re = O(10^5)$, il numero di gdl necessari alla risoluzione diretta delle equazioni risulta:

$$gdl = O(10^5)^{11/4} = O(10^{55/4}) = O(10^{14}) \quad (1)$$

All'attuale stato dell'arte non è possibile risolvere un così elevato numero di gdl . Ipotizzando di realizzare una simulazione stazionaria per diminuire il numero di gdl necessario, il numero di gdl si riduce significativamente. Infatti:

$$gdl = N = O(10^5)^{9/4} = O(10^{45/4}) = O(10^{11}) \quad (2)$$

ovvero 200 miliardi di celle. Al momento in cui questo articolo viene redatto, le simulazioni più all'avanguardia sono dell'ordine di poco superiori a un miliardo di celle ($gdl = 10^9$), utili quindi a simulare con questo approccio valori di Reynolds dell'ordine di 10^3 , infatti:

$$Re = O(10^9)^{4/11} = O(10^3) \quad (3)$$

Poiché ai Re di interesse non è possibile cogliere tutte le scale energetiche, diventa necessario modellare altrimenti gli effetti delle scale più piccole che non vengono colte da una discretizzazione più rada. Questo approccio risulta essere quello comunemente adottato in letteratura per ovviare a detti limiti computazionali. Tuttavia, come verrà spiegato nel seguito, questo è anche un approccio che introduce un livello di aleatorietà molto elevato e che impone una profonda fase di calibrazione del modello numerico prima di poterne fare uso con confidenza (vedi [11]).

3.2. Modelli mediat

Per determinare l'effetto delle scale più piccole si applica un operatore di media sul sistema di equazioni di NS. La differenza tra le equazioni primitive e quelle mediate rappresenta l'effetto delle scale scartate dall'operazione di media. Si osserva che il sistema di equazioni di NS che era risolvibile quando non mediato, dopo l'operazione di media presenta un'aggiuntiva incognita tensoriale (ovvero una matrice), nota come tensore (simmetrico) degli sforzi di Reynolds, composto da sei incognite scalari. Queste incognite rappresentano l'effetto sul campo medio delle componenti fluttuanti escluse dall'operatore di media. Il loro significato fisico dipende pertanto dalla definizione dell'operatore utilizzato. Queste grandezze non sono risolte, ma sono modellate con opportune equazioni euristiche che ne forniranno una stima.

Per avere nuovamente un sistema chiuso (ovvero risolvibile quando sono conosciute le condizioni al contorno e iniziali) come quello originale, è necessario introdurre tante equazioni quante le nuove incognite. Queste equazioni sono dette "modelli di turbolenza".

I modelli di turbolenza che sono comunemente utilizzati in questo tipo di applicazioni assumono l'ipotesi di Boussinesq, per il quale il tensore degli sforzi di Reynolds può essere modellato con un termine viscoso artificiale, indotto per via numerica nel modello fisico che si vuole modellare. I diversi modelli di turbolenza si distinguono per come modellano la cosiddetta "viscosità turbolenta".

3.3. Accuratezza e limiti computazionali

Come discusso nell'introduzione, al fine di studiare l'affidabilità dei vari possibili approcci di modellazione tramite CFD in questo ambito di progettazione, sono state svolte svariate analisi, riferendosi a dati di galleria del

vento per un'imbarcazione di Coppa America studiata in condizione stazionaria, al variare dei seguenti parametri di interesse:

i. modello di turbolenza;

ii. configurazioni geometriche (nominalmente angoli al vento dell'imbarcazione e regolazione delle vele);

iii. risoluzione della discretizzazione spaziale.

I risultati mettono in evidenza come, nonostante i modelli di turbolenza permettano di modellare con un limitato numero di *gdl* problemi che richiederebbero un numero di *gdl* molto oneroso, non sia sempre univoco il risultato che si ottiene. Per questo motivo, come detto, sono necessarie lunghe campagne computazionali di *tuning* del problema numerico al fine di valutare quale modello di turbolenza sia in grado di fornire risultati più simili alla galleria del vento in quelle specifiche condizioni. Viceversa, scegliendo strategie decisamente più onerose dal punto di vista computazionale, è possibile ottenere risultati più indipendenti dalle condizioni modellate. In accordo con le richieste teoriche sul livello di risoluzione spaziale della discretizzazione adottata, all'aumentare della dimensione della griglia di calcolo si ottengono risultati sempre più accurati.

Nell'agosto del 2008 si realizzava la simulazione più onerosa in ambito fluidodinamico con un codice di calcolo commerciale [12]. Un'imbarcazione di Coppa America con randa e *spinnaker* asimmetrico è stata modellata con una griglia composta da un miliardo di celle ($N = 10^9$) (Figura 4). Al crescere della risoluzione spaziale diminuiscono le dimensioni delle strutture vorticosi che sono escluse dall'operatore di media e che debbono essere modellate con il modello di turbolenza. Pertanto, in detta simulazione, per via dell'elevatissima risoluzione spaziale, nessun modello di turbolenza è stato adottato, considerando trascurabili le scale di dimensione inferiore a quella della griglia di calcolo. Per verificare l'effetto delle scale trascurate, diverse simulazioni sono state realizzate diminuendo la risoluzione spaziale fino alle dimensioni di griglia comunemente utilizzate. La differenza tra le forze aerodinamiche calcolate nelle diverse simulazioni, mostra l'effetto delle scale non risolte dalle griglie di calcolo. La differenza tra la forza aerodinamica calcolata con $N = 10^9$ è risultata quasi il 20% maggiore della forza calcolata con $N = 10^6$.

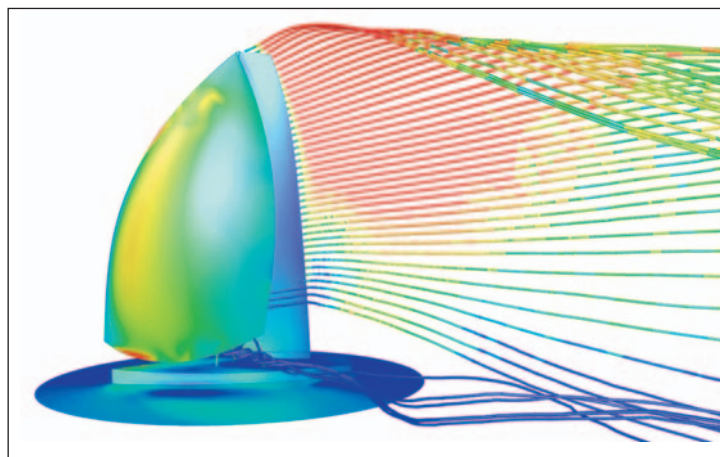


FIGURA 4

Visualizzazione delle linee di flusso colorate secondo le velocità, ottenute tramite un modello CFD da un miliardo di celle di un'imbarcazione a vela da regata

4. CONCLUSIONI

Pur ricordando che è sempre difficile prevedere lo sviluppo tecnologico e i suoi tempi, sembra lecito sperare che nei prossimi anni vi siano sufficienti motivazioni e volontà affinché si cerchi di integrare tutti i diversi aspetti dinamici dell'imbarcazione utilizzando tecniche proprie della fluidodinamica numerica e della realtà virtuale, realizzando un simulatore che consenta di progettare una vela ottimizzando l'unica variabile obiettivo del progetto: le probabilità di battere l'avversario su un determinato percorso.

È importante sottolineare che l'uso di tecniche numeriche non è al momento una completa alternativa alla prova sperimentale. Infatti, la prova sperimentale in galleria del vento oltre a fornire le forze prodotte da una particolare regolazione delle vele, permette di determinare la regolazione ottimale agendo sulle scotte, misurando in tempo reale le forze aerodinamiche che agiscono sull'imbarcazione. Al contrario, in ambito numerico, la ricerca della regolazione richiederebbe di accoppiare al calcolo aerodinamico in oggetto, un dedicato codice strutturale e un ottimizzatore, codici che si stanno diffondendo in questi ultimissimi anni e che sicuramente rappresentano il futuro dell'aerodinamica numerica delle vele in andature non di bolina. Tuttavia nel breve periodo non saranno in grado di sostituire l'attività sperimentale in galleria del vento.

Per il futuro si chiede a questi codici di muovere verso l'integrazione dei diversi aspetti della dinamica dell'imbarcazione. Già alcuni passi sono stati fatti in questa direzione: nel campo dell'idrodinamica, le simulazioni numeriche consentono di lasciare libero il movimento dell'imbarcazione e di risolverne il movimento insieme al campo di moto dei due fluidi aria e acqua, e di calcolare la forma della superficie libera (ovvero il moto ondoso generato dall'imbarcazione). Nel campo dell'aerodinamica, è possibile risolvere il campo di moto intorno a uno *spinnaker* tenendo conto dell'irregolarità nel tempo del vento incidente e di un moto imposto dell'imbarcazione. Queste potenzialità rendono il futuro della fluidodinamica numerica molto attraente ma mettono in luce una nuova criticità: la mancanza della verifica sperimentale. Sono state

realizzate negli ultimi quattro anni diverse simulazioni che accoppiano aero- e idro-dinamica in una sola simulazione, ma la mancanza di prove sperimentali non ha consentito di valutarne la bontà. Il paradosso di fronte al quale ci si trova in questo momento è che proprio quando la fluidodinamica numerica si trova a superare quella sperimentale, la prima si scopre dipendere fortemente dalla seconda per poter progredire.

Approfondimenti

Di seguito riportiamo per completezza una possibile trattazione introduttiva di una delle formulazioni tipicamente utilizzate per descrivere le equazioni di Navier-Stokes. Queste poche righe possono essere di ausilio per avere una visione più chiara di alcune parti del testo.

Equazioni di Navier-Stokes

Le equazioni di Navier-Stokes per fluidi incompressibili, ovvero per i quali la densità è costante e omogenea, sono dedotte dal principio della conservazione della massa e del bilancio della quantità di moto. La conservazione della massa si può esprimere nel seguente modo:

$$\nabla \cdot u = 0 \quad (\text{a.1})$$

dove: $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$

è l'operatore Nabla, u è il vettore velocità ed il punto ($\bullet$) esprime il prodotto scalare. Questa equazione è nota come *equazione di continuità*.

Il principio del bilancio della quantità di moto, che in formulazione Lagrangiana ha la nota forma $F = ma$, si può esprimere nella seguente forma dedotta dalla sua formulazione Euleriana:

$$\frac{\partial u}{\partial t} + (u \cdot \nabla) u = -\frac{1}{\rho} \nabla p + \nu \nabla^2 u \quad (\text{a.2})$$

dove compare la derivata temporale della velocità, il termine non lineare con l'operatore nabla, la variazione di pressione divisa per la densità (ρ), e il termine viscoso che contiene la viscosità cinematica (ν) e l'operatore Laplaciano

$$\nabla^2 = \text{div}(\text{grad}) \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right).$$

Questa equazione è nota come *equazione del momentum*.

Sostituendo i termini dimensionali dell'equazione del momentum con i seguenti termini non dimensionali:

$$\tilde{u} = \frac{u}{u};$$

$$\tilde{p} = \frac{p}{\rho u^2};$$

$$\tilde{x} = \frac{x}{L};$$

$$\tilde{t} = \frac{tu}{L};$$

l'equazione (a.2) diviene:

$$\frac{\partial \tilde{u}}{\partial \tilde{t}} + (\tilde{u} \cdot \nabla) \tilde{u} = -\nabla \tilde{p} + \frac{1}{\text{Re}} \nabla^2 \tilde{u} \quad (\text{a.3})$$

La quale mette in luce la dipendenza dal numero di Reynolds (Re) come discusso nel testo.

Bibliografia

- [1] Marchay C.A.: *Aero-Hydrodynamics of Sailing*. London: Adlard's Coles Nautical, 1979.
- [2] Milgram J.H.: *The analytical design of yacht sails*, *Trans. Soc. Naval Architect. Marine Eng.* SNAME, 1968, p. 118-160.
- [3] Milgram J.H.: *The aerodynamic of sails*. Proc. 7th Symp. Naval Hydrodyn, 1968, p. 1397-1434.
- [4] Milgram J.H.: *Sail Force Coefficients for Systematic Rig Variations*. SNAME Technical and Research Report: R-10, 1970.
- [5] Gentry A.E.: *The Aerodynamics of Sail Interaction*. Aient Interface III, The 3rd AIAA Symposium on Sailing, Nov. 20, 1971, Redondo Beach, California, US.
- [6] Gerritsma J., Onnink R., Versilius A.: *Geometry resistance and Stability of the Delft Systematic Hull Series*. The 7th HISWA Symposium, Amsterdam 1981.
- [7] Kerwin J.E.: *A Velocity Prediction Program for Ocean Racing Yachts*. Rep. 78-11 MIT, July 1978.
- [8] Kerwin J.E., Newman J.N.: *A Summary of the H. Irwing Pratt Ocean Race Handicapping Project*. SNAME 4th CSYS Annapolis 1979.
- [9] Kolmogorov A.N.: *Dissipation of energy in a locally isotropic turbulence*. *Doklady Akad. Nauk SSSR*, 32, 141, 1941. (English translation in: American Mathematical Society Translations 1958, Series 2, Vol. 8, p. 87, Providence R.I.).
- [10] Squires K.D.: *Detached Eddy Simulation of Turbulent Flows - Method and Applications*. Lecture notes, Von Karman Institute for Fluid Dynamics, Brussels, Belgium, March 13-16, 2003.
- [11] Viola I.M., Ponzini R., Rocchi D., Fossati F.: *Analysis of Aerodynamic Indices for Racing Sailing Yachts: a Computational Study and Benchmark on up to 128 CPU*. In the Proceedings of the 20th International Conference on Parallel Computational Fluid Dynamics (PARCFD 2008), 19th-22nd May 2008, Lyon, France.
- [12] Viola I.M., Ponzini R., Passoni G.: *Downwind Sail Aerodynamics: Large Scale Computing vs. Large Scale Wind Tunnel Test*. *Journal of Wind Engineering and Industrial Aerodynamics*. Submitted 2009.

RAFFAELE PONZINI ha una Laurea in Bioingegneria conseguita al Politecnico di Milano (2002) con una tesi di modellistica numerica applicata al campo dell'emodinamica (Analisi dei profili spaziali di velocità in un bypass coronarico da Y tramite un modello multiscala). Dal 2003 lavora all'interno del gruppo di calcolo ad alte prestazioni del CILEA dove è responsabile della gestione dei codici di fluidodinamica computazionale (CFD). Il suo campo di applicazione lavorativo include anche l'insegnamento di diversi linguaggi di programmazione (Python, C, C++) e di tecniche di programmazione ad oggetti. Nel 2007 ha conseguito un dottorato di ricerca in Bioingegneria presso il Politecnico di Milano (con lode) con una tesi dal titolo "*Computational modelling of local haemodynamics phenomena: methods, tools and clinical applications*". I suoi interessi di ricerca includono la modellistica multiscala, il processamento di immagini e la visualizzazione scientifica. È attualmente autore di 10 articoli (*peer-reviewed*) internazionali nel campo della biomeccanica computazionale e di numerose pubblicazioni a congresso nel campo biomedicale e della modellazione CFD in ambito navale.
E-mail: ponzini@cilea.it

IGNAZIO MARIA VIOLA, PhD, Ingegnere Navale, ricopre la posizione di *Lecturer in Naval Architecture* presso la *School of Marine Science and Technology* dell'Università di Newcastle (UK). Nato a Milano nel 1976, si laurea a pieni voti in Ingegneria Navale presso l'Università degli Studi di Genova nel 2001. Dopo una prima esperienza industriale di tre anni nella progettazione e costruzione di imbarcazioni (direttore della produzione di serie del cantiere MAS e coordinatore ufficio Ricerca e Sviluppo PRIMA), nel 2004 comincia una lunga collaborazione con la Galleria del Vento del Politecnico di Milano dove si interessa di fluidodinamica numerica e sperimentale. Si occupa dell'aerodinamica numerica per il consorzio Luna Rossa sfidante della Coppa America del 2007 e, nel 2008, completa il PhD in fluidodinamica applicata alle imbarcazioni. Lo stesso anno, in collaborazione con R. Ponzini (CILEA), realizza la prima simulazione numerica su una geometria complessa con una griglia di un miliardo di celle. Dal 2008 è membro del *Yacht Research Unit* dell'Università di Auckland (New Zealand), *scientific supplier* del consorzio di Coppa America *Emirates Team New Zealand*.

E-mail: im.viola@auckland.ac.nz
Web: www.ignaziovola.com