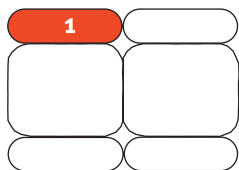




MEZZO SECOLO DI FUTURO L'INFORMATICA ITALIANA COMPIE CINQUANT'ANNI

Corrado Bonfanti



Si compiono cinquanta anni dall'esordio in Italia di due tecnologie che avrebbero cambiato il mondo: la televisione e l'informatica. Se la prima ebbe immediata e vasta risonanza a tutti i livelli, non altrettanto si può dire della seconda. Questa, infatti, prima di dilagare, si avviò in sordina, nell'ambito di una ristretta cerchia di specialisti. L'articolo ripercorre il periodo pionieristico dell'informatica italiana: il contesto culturale da cui originava, le diverse impostazioni e finalità delle iniziative che furono allora intraprese.

1. INTRODUZIONE

Cinquant'anni orsono, nel 1954, hanno fatto il loro esordio in Italia due tecnologie a forte impatto sociale: la televisione e il computer. Due ricorrenze importanti la cui coincidenza temporale, pur non implicando alcuna connessione, invita a qualche riflessione preliminare.

Le trasmissioni video della Rai ebbero inizio il 3 gennaio: in bianco e nero, su un solo canale, con appena qualche ora al giorno di trasmissione, con una copertura territoriale incompleta; era un avvio piuttosto rudimentale. L'enorme potenziale del nuovo *medium* non tardò comunque a esplodere, imponendosi come potente fattore di modificazione e di condizionamento della società, della cultura e del costume.

In quello stesso anno – e più esattamente nel

volgere di pochi mesi a cavallo tra il 1954 e il 1955 – a Milano, a Pisa e a Roma spuntarono anche i primi germogli dell'informatica in Italia¹. Tuttavia, mentre la TV, fin dall'inizio, poteva interagire facilmente e direttamente, anche se in maniera unidirezionale, con il pubblico di qualsiasi genere, l'avvento del computer ha riguardato inizialmente una ristrettissima cerchia di specialisti. L'opinione pubblica era coinvolta solo a livello emozionale: nasceva il mito del "cervello elettronico". Nonostante la successiva diffusione dell'elaborazione elettronica nei più disparati ambienti applicativi, la comunità degli informatici – pur crescendo numericamente e articolandosi in sottospecie specialistiche – è rimasta per almeno due decenni una categoria aliena, depositaria di saperi misteriosi e di poteri sotterranei. La percezione da parte del grande pubblico era affi-

¹ In questo lavoro, per motivi di comodità, il sostantivo "informatica" – con i suoi derivati – sarà usato liberamente nella sua accezione corrente anche se, essendo stato coniato appena nel 1962 dal francese Philippe Dreyfus della Compagnie des Machines Bull, possa apparire inappropriato quando riferito a vicende precedenti alla sua introduzione. Si avverte inoltre che, dove non espressamente segnalato, termini quali "computer", "elaboratore", "calcolatore" saranno considerati tra loro equivalenti seguendo l'uso comune.



data a fatti di cronaca commentati di volta in volta in chiave miracolistica o catastrofica e titolati all'insegna del "grande fratello" o del "cervellone"; per di più, l'attenzione si focalizzava sullo "strumento computer", lasciando in ombra il contesto culturale, professionale e industriale. A differenza di quanto avvenuto per la TV, l'informatizzazione di massa è, quindi, un fenomeno piuttosto recente che inizia con l'avvento del *personal computer* e prosegue sull'onda di Internet, debordando dai luoghi di lavoro specializzati e coinvolgendo capillarmente le singole persone e le famiglie. Accanto alla tele-dipendenza indotta da tv, anche la dipendenza da computer è oggi un fenomeno endemico, per non parlare del telefono cellulare. In termini più generali, si deve constatare che i microcomputer, i cellulari e gli stessi televisori, pur conservando ancora riconoscibili le rispettive funzioni originarie, altro non sono ormai – e sempre più saranno – che punti d'accesso alla "rete globale", quell'entità ubiqua e pervasiva che costituisce il supporto della società dell'informazione. Nello scenario primordiale del 1954/55, le cose erano ben diverse. In particolare, l'eventualità che il mondo dell'informatica e quello delle telecomunicazioni confluissero prima o poi nell'attuale ICT era contemplata solo nelle ipo-

tesi dei futurologi più audaci. Uno scenario in cui si assisteva anche ad altre importanti vicende come per esempio la *vexata quaestio* della identificazione/separazione tra cibernetica e informatica² oppure il temporaneo successo dei calcolatori elettronici analogici, i cosiddetti DDA (*Digital Differential Analyzer*)³. Sono temi di grande interesse che però non verranno approfonditi in questa sede avendo scelto di focalizzarci sulle origini e sui primi sviluppi in Italia dell'informatica in senso stretto; siccome, peraltro, molti passaggi della nostra storia "locale" rimarrebbero oscuri se non correlati agli sviluppi internazionali, anche a questi si dedicherà la dovuta attenzione.

2. L'ANNO ZERO DELL'INFORMATICA ITALIANA

Cosa successe, dunque, di così determinante da poter considerare il 1954/55 come "anno zero" dell'informatica in Italia? I fatti, in breve, furono i seguenti:

■ l'attivazione del Centro di calcoli numerici al Politecnico di Milano, dotato di un CRC 102A acquistato presso la NCR (*National Cash Register*) negli Stati Uniti⁴; è stato, in assoluto, il primo calcolatore elettronico a programma registrato che abbia mai operato in Italia;

² Su tale questione, si vedano i contributi di Antonio Lepschy (Prefazione), di Anna Cuzzer (*La diffusione dell'informatica in Italia*) e di Vittorio Somenzi (*Cibernetica, informatica e filosofia della scienza*) pubblicati in {2}. Qui si accenna solo al fatto che in quel periodo, e non solo in Italia, "cibernetica" e "scienza dei calcolatori" erano da molti percepiti come sinonimi; la "teoria della comunicazione" non si sapeva poi bene dove piazzarla. A questo stato di cose, come si evince già dal titolo e sottotitolo, aveva certo contribuito il saggio 'fondazionale' di Norbert Wiener *Cybernetics, or control and communication in the animal and the machine* (1948) come pure il più popolare *The Computer and the Brain* di John von Neumann (1958). Ci volle una laboriosa maturazione per arrivare a delineare un quadro disciplinare meglio articolato in branche specialistiche e per rendersi conto che le "zone grigie" di confine o di sovrapposizione potevano essere coperte in maniera ottimale con un approccio collaborativo multispecialistico.

³ Si precisa che le macchine DDA, prodotte principalmente dalla Bendix Corporation, erano un ibrido analogico-numerico; esse furono adottate, anche prima del 1954, dalle Università di Napoli e di Bologna e dal Politecnico di Torino; il loro punto di forza era la risoluzione di sistemi di equazioni differenziali. Tutti i testi generali sul calcolo elettronico si aprivano a quel tempo con una sezione dedicata ai calcolatori analogici. Luigi Dadda mise prontamente in guardia i colleghi nei confronti della inevitabile obsolescenza del calcolo analogico, contrapponendovi i vantaggi del calcolo numerico (digitale). La questione non era di poco conto, e non solo da noi: si pensi che, oltre dieci anni dopo, un autorevole esperto di controlli automatici metteva ancora in guardia sul fatto che i calcolatori numerici sono "incapaci di eseguire delle vere integrazioni. Da questo fatto, associato con [la durata dell'elaborazione], consegue la "naturale" incapacità [dell'elaboratore numerico] di risolvere grandi sistemi di equazioni differenziali. È chiaro che si può raggiungere un compromesso aumentando la dimensione del passo d'integrazione. Ma d'altra parte questo può portare a conseguenze terribili." (Cfr. Elgerd O.E.; *Control System Theory*; McGraw-Hill Book Company, 1967, p. 513).

⁴ Il computer CRC 102A era stato realizzato dalla Computer Research Corporation, un'azienda *start-up* che la NCR aveva appena acquisito per entrare nel mercato dell'informatica. La CRC era stata costituita da un gruppetto di ex ufficiali che durante la guerra avevano partecipato a progetti militari di tecnologia avanzata. La storia della CRC è un caso esemplare di dinamica imprenditoriale che prefigura lo stile *silicon valley*. (Cfr. Eckdahl D.E., Reed I.S., Sarkissian H.H.; *West Coast Contributions to the Development of the General-Purpose Computer – Building Maddida and the founding of Computer Research Corporation*; IEEE Annals of the History of Computing, Vol. 25, n. 1, 2003, p.4-33).

la costituzione del CSCE (*Centro Studi Calcolatrici Elettroniche*) e l'inizio del progetto CEP (*Calcolatrice Elettronica Pisana*) all'Università di Pisa;

la costituzione, nella stessa Pisa, del LRE (*Laboratorio Ricerche Elettroniche*) dell'Olivetti;

l'acquisto di un calcolatore Ferranti Mark I*, prodotto dalla inglese Ferranti Ltd.⁵, da parte dell'INAC (*Istituto Nazionale per le Applicazioni del Calcolo*) a Roma; questa macchina divenne nota come FINAC.

Tutte queste iniziative avevano avuto una gestazione indipendente l'una dall'altra; il fatto che prendessero corpo quasi contemporaneamente testimonia, quindi, che i tempi erano maturi affinché anche il nostro Paese cominciasse a operare attivamente nel campo del calcolo elettronico. Un campo in cui, rispetto al contesto internazionale, non si partiva certo da posizioni d'avanguardia, dal momento che la fase pionieristica si era ormai conclusa, avendo avuto le sue sedi privilegiate – e molto reclamizzate – negli Stati Uniti e in Inghilterra⁶. Quasi nessuno, all'epoca, era, invece, a conoscenza di quello che il solitario Konrad Zuse aveva realizzato in Germania⁷.

Per l'Italia quindi, più che di un 'ingresso' nel mondo dell'informatica, si deve parlare

dell'inizio di una "rincorsa". Rincorsa la cui progressione risultò efficace e piuttosto rapida proprio grazie all'eterogeneità delle iniziative, ora ricordate, con le quali s'imboccavano, in maniera non conflittuale, tutte e tre le vie d'approccio al nuovo mondo dell'informatica: la via della "cultura applicativa", che corrispondeva alla vocazione e alla tradizione del Politecnico di Milano e dell'INAC; la via della "cultura progettuale", che caratterizzava l'iniziativa dell'Università di Pisa; la via della "cultura industriale", che era connaturata alla missione della florida multinazionale Olivetti. E si trattava di iniziative sostenute da professionalità di assoluta eccellenza – a livello sia di gruppo che di singole personalità – tanto che esse si configurarono immediatamente come altrettante "scuole" nelle quali si è formata la prima generazione degli informatici italiani.

Se gli inizi furono concomitanti, non altrettanto può dirsi degli esiti che sortirono da quelle iniziative. Le macchine del Politecnico e dell'INAC, le cui inaugurazioni furono eventi memorabili, erano state acquistate pronte per l'uso e cominciarono immediatamente a sfornare calcoli applicati ai più svariati problemi scientifici e tecnici, con importanti ricadute anche sugli operatori eco-

⁵ La Ferranti Ltd. di Manchester, fondata e posseduta da una famiglia di lontane origini venete, era specializzata nella fabbricazione di impianti per energia elettrica e, nel corso del conflitto, aveva prosperato anche sulle commesse militari per sistemi di controllo e radar. Era entrata precocemente nel settore dei computer grazie a una tipica operazione di trasferimento tecnologico dalla ricerca all'industria, sostenuta anche da cospicui finanziamenti governativi: i suoi computer erano, infatti, una derivazione diretta dei prototipi realizzati da Frederic Williams e Tom Kilburn nel laboratorio di fisica dell'Università di Manchester. Sui computer pionieristici prodotti in Inghilterra, si veda: Lavington S.; *Early British Computers* (Digital Press, 1980).

⁶ Si ricorda che la caratteristica fondamentale che distingue il computer moderno da tutti gli altri apparati di calcolo, consiste nell'essere "a programma registrato" (*Stored Program Computer*), cioè nella capacità di operare in base al programma registrato in un organo (non necessariamente elettronico) di memoria dinamica, comune per programmi e dati, dove il programma stesso può quindi modificare le proprie istruzioni nel corso dell'esecuzione. Questo è il "sugo" della "architettura di von Neumann", teorizzata nel *First Draft of a Report on the EDVAC (Electronic Discrete Variable Automatic Computer)* che von Neumann compose nel 1945 come consulente della Moore School of Electrical Engineering (Philadelphia) e della direzione di artiglieria dell'esercito statunitense. Il testo completo del *First Draft* è apparso a stampa per la prima volta in appendice al saggio di Nancy Stern *From ENIAC to UNIVAC; an appraisal of the Eckert-Mauchly Computers* (Digital Press, 1981, p.177-246).

⁷ La notorietà di Konrad Zuse e della sua straordinaria avventura concettuale, tecnologica e poi anche imprenditoriale cominciò a diffondersi solo con il decennio 1960, e da allora fu un profluvio di riconoscimenti; anche l'Università di Siena gli ha conferito nel 1992 la laurea *honoris causa* e chi scrive ha avuto l'opportunità di suggerire tale iniziativa. L'azienda Zuse A.G., fondata nel 1949 con stabilimento nei pressi di Hünfeld nell'Assia, fu assorbita dalla Siemens nel 1969 e Zuse, libero ormai dai gravosi impegni manageriali, poté tornare ai "diletti studi" occupandosi specialmente di automi cellulari. L'opera di riferimento è la sua autobiografia *Der Computer – Mein Lebenswerk* pubblicata la prima volta nel 1970 (Verlag Moderne Industrie, München) e tradotta in italiano a cura di Mario G. Losano (*Zuse - l'elaboratore nasce in Europa*; Etas Libri, Milano, 1975). Maggiore diffusione, grazie alla notorietà ormai consolidata, ebbero successive edizioni anche in versione inglese (*The Computer – My Life*; Springer-Verlag, 1993).



nomici, in primo luogo nei settori dell'industria e delle opere di ingegneria civile⁸. I tempi di progettazione e di costruzione *ex novo* di macchine originali furono, invece, necessariamente più lunghi: la CEP fu pronta in una versione ridotta nel 1958 e la macchina completa nel 1960; i prototipi ELEA 9000 della Olivetti furono operativi dal 1957 e il primo esemplare di serie – il 9003 – venne consegnato nel 1959. E non si tratta di mera cronologia perché, nel lasso di appena qualche anno, non solo la tecnologia ma l'intero scenario avevano subito mutazioni profonde.

3. LE RADICI

Prima di tornare agli avvenimenti dell'"anno zero", è utile soffermarsi sulle radici più o meno remote da cui essi germogliarono.

Parlando di "calcolatori", è ovvio che la matematica applicata, intesa come soluzione numerica di problemi tecnici, fosse un terreno privilegiato per l'attecchimento della nuovissima tecnologia. Dalla fine dell'Ottocento fino al tragico sconvolgimento provocato dalle leggi razziali del 1938, la matematica italiana era assunta a prestigio internazionale in diversi settori di punta quali, ad esempio, la geometria algebrica. A causa di tanto successo nella ricerca "pura", gli ambienti accademici erano portati a snobbare

il calcolo numerico considerato appunto come "impuro", una sorta di contaminazione ingegneristica. Non sorprende, quindi, che proprio nell'ingegneria si sia sviluppata una diffusa attenzione a questi problemi. Un esempio per tutti è quello di Gino Cassinis al quale si deve un corposo trattato (di quasi 700 pagine) sul calcolo numerico⁹; citazione non casuale in quanto proprio il Cassinis, come rettore del Politecnico di Milano, sostenne con ferma determinazione la costituzione del Centro di Calcoli Numerici portando così a compimento un progetto che aveva coltivato fin dal 1941. Nello stesso ambiente del Politecnico, Ercole Bottani aveva scritto nel 1935 un articolo dal titolo, molto significativo, "*La matematica vista da un ingegnere*" e perseguì poi l'approccio analogico occupandosi di modelli fisici applicabili a problemi di calcolo¹⁰.

Date le premesse, può invece sorprendere il fatto che sia stato proprio un matematico – Mauro Picone, eccellente anche nella ricerca pura – a fondare, nel 1927, l'Istituto di Calcolo Numerico dell'Università di Napoli¹¹. Trasferito a Roma nel 1932, Picone riuscì a consolidare la sua iniziativa nell'ambito del CNR dando così vita all'INAC¹². Tralasciando i lavori di natura teorica, il successo e le ricadute di questa iniziativa sono testimoniati dal numero delle consulenze che l'istituto svolse per conto di aziende e

⁸ Più precisamente, la CRC 102A entrò in funzione nell'ottobre '54 e il Centro di Calcoli Numerici venne inaugurato il 31 ottobre dell'anno seguente; in proposito si veda quanto Luigi Dadda ha scritto nei suoi contributi *Il Centro di calcoli numerici e l'introduzione delle discipline informatiche al Politecnico di Milano e Ricordi di un informatico*, pubblicati rispettivamente in {1}, p.7-44 e in {2}, p.67-106; tali contributi saranno più volte richiamati nel seguito. La FINAC entrò in funzione nel giugno '55 e fu inaugurata il 13 dicembre dello stesso anno alla presenza del Presidente della Repubblica Giovanni Gronchi il quale, come si vedrà, è stato presente in altre analoghe circostanze.

⁹ *Calcoli numerici, grafici e meccanici*; Mariotti-Pacini, Pisa, 1928.

¹⁰ Cfr. Dadda L.; *Il Centro ecc.*; cit., p.10.

¹¹ A questa università apparteneva anche Ernesto Pascal, inventore di raffinati strumenti analogici, puramente meccanici, per la soluzione grafica di equazioni differenziali. Gli strumenti di Pascal, descritti nella memoria *I miei integrali per equazioni differenziali*, erano spesso citati nella letteratura internazionale specializzata. La memoria è compresa negli Atti della Reale Accademia delle Scienze Fisiche e Matematiche di Napoli, v. xv (1913), serie 2a, n.16.

¹² Per la figura di Picone in relazione alle vicende che qui interessano si veda il contributo di Enzo Aparo *Mauro Picone e l'Istituto Nazionale per le Applicazioni del Calcolo* pubblicato in {1}, p. 47-55. Aparo ricorda come uno dei problemi più impegnativi affrontato dall'INAC in quegli anni, siano stati i calcoli relativi al progetto della diga sul Vaiònt, eseguiti con un metodo simile a quello attuale degli elementi finiti; qui si vuole aggiungere che, sebbene nell'immaginario collettivo si sia fissato il 'crollo' di quella diga, essa è in realtà ancora in piedi, grigio monumento di una immane tragedia. Essa infatti era stata tanto ben 'calcolata' da resistere intatta al tremendo impatto della frana precipitata dal monte Toc, sovrastante l'invaso. Il tragico errore, indotto anche da pesanti interessi economici e rimasto sostanzialmente impunito, fu la sottovalutazione del rischio geologico che era peraltro ben evidente.

di enti pubblici: esse furono ben 1.492 nel periodo 1937-1964, con una media di quasi cinque al mese¹³. L'eccellenza dell'INAC è stata poi determinante per vincere, nel 1951, un'agguerrita competizione internazionale, guadagnando a Roma l'assegnazione della sede dell'ICC (*International Computation Centre*) costituito in seno all'Unesco¹⁴.

Con personaggi del calibro di Kurt Gödel, Alan Turing, Emil Post, Rudolph Carnap e Alonzo Church, la logica matematica ebbe uno straordinario rigoglio durante il decennio del 1930, venendo a formare un'importante premessa concettuale per lo sviluppo della *computer science*. A queste ricerche, l'Italia non fu in grado di dare alcun apporto sebbene Giuseppe Peano e i suoi allievi, all'inizio del secolo, vi avessero dato importanti contributi. Il che si spiega col fatto che la scuola di Peano fu praticamente ridotta al silenzio dal predominio assoluto

che la cultura neo-idealistica – propugnata dai pur grandi filosofi Benedetto Croce e Giovanni Gentile – esercitò sul nostro Paese nel periodo tra le due guerre mondiali, proprio mentre si dischiudevano i nuovi orizzonti del pensiero logico. La corrente neo-idealistica, per dirla in termini sbrigativi, derubricava la matematica e le scienze empiriche al ruolo di discipline strumentali e, quindi, sub-culturali, ed ebbe buon gioco nel rivendicare ai “filosofi professionisti” il monopolio sullo studio della logica, attenendosi a modelli speculativi ormai desueti. Una visione sostanzialmente oscurantista che procurò l'aborto delle ricerche in logica matematica¹⁵. Seri studi in questo campo riprenderanno, in Italia, solo nel secondo dopoguerra ad opera principalmente di Ludovico Geymonat e di alcuni dei suoi allievi che ne ripristinarono la dignità culturale e accademica¹⁶. Per quanto attiene più specificamente ai primi passi dell'informa-

¹³ In un articolo del 1953, ispirato dallo stesso Picone, si pone l'enfasi anche sull'aspetto economico. Vi si legge infatti: “l'INAC era giunto a versare annualmente nelle casse dello Stato [...] un utile pari a un quarto di miliardo d'oggi, ricavato dalle consulenze per ditte private [...]. Questa cifra è proprio quella che occorrerebbe ora, per l'acquisto di una sola delle macchine elettroniche, con cui nei Paesi più progrediti si è pervenuti a superare enormemente il rendimento delle calcolatrici meccaniche tuttora usate dal nostro Istituto” (Cfr. Sagredo; *Matematica e industria*; Civiltà delle macchine, Vol. a.1, n. 1, 1953, p.24-26).

¹⁴ Rimaste in lizza le candidature di Amsterdam e di Roma, si decise di affidarsi al giudizio di Herman H. Goldstine, autorevole personaggio statunitense la cui fama risaliva al progetto ENIAC e alla collaborazione con von Neumann a Princeton. Nella sua relazione si legge che “esaminando le svariate pubblicazioni di questo centro [cioè dell'INAC] si resta colpiti dall'immenso spazio che la direzione concede alle matematiche, e non si può non rimanere impressionati dall'ampiezza dei calcoli eseguiti al centro e dall'elevata qualità dell'analisi matematica che essi hanno comportato”. È anche interessante riprendere la frase finale della relazione: “In conclusione io raccomando, dal punto di vista tecnico, che il centro [qui si tratta dell'ICC] venga istituito a Roma [...] fermo restando naturalmente che questa raccomandazione deve essere anche esaminata dai punti di vista finanziario e diplomatico”. È immaginabile che i diplomatici, dal canto loro, siano stati ben felici di potersi appoggiare a una raccomandazione tecnica formulata in termini tanto netti da poter essere spesa come decisione finale in una contesa internazionale tanto delicata. Va anche ricordato che l'ICC - in seguito ridenominato IBI (*Intergovernmental Bureau for Informatics*) - ebbe vita non facile e che, constatate le radicali evoluzioni intervenute nel frattempo, l'Unesco considerò esaurita la sua missione e nel 1988 ne deliberò lo scioglimento. Per i passi citati in questa nota, cfr. le p. 56-58 del volume, curato da Picone, *Giudizi sull'opera trentennale dell'Istituto Nazionale per le Applicazioni del Calcolo* (Tipografia Pio X°, Roma, 1959), dove si è reperito un lungo stralcio della relazione di Goldstine, redatta in francese. Sulla laboriosa gestazione dell'ICC si veda la testimonianza dello stesso Goldstine, al capitolo 7 di *The Computer from Pascal to von Neumann* (Princeton University Press, 1972). Si ricorda qui anche la figura del matematico Aldo Ghizzetti al quale fu affidata la direzione tecnica dell'ICC; Ghizzetti subentrò a Picone nella direzione dell'INAC e fu inoltre presidente dell'AICA dalla sua fondazione (1961) fino al 1967.

¹⁵ Federico Enriques, grande matematico e razionalista convinto, si oppose coraggiosamente accettando il dibattito anche sul terreno propriamente filosofico. Riferendosi all'Enriques dei primi decenni del '900, Lucio Lombardo Radice scrive: “furono gli anni della partecipazione ai Congressi filosofici: il IV Congresso filosofico internazionale, tenutosi a Bologna, è presieduto da Federico Enriques. Furono gli anni del grande tentativo di Enriques di rompere il diaframma frapposto (e imposto) dall'idealismo di Croce e Gentile tra le due culture, gli anni [...] del rapporto vario e mosso colle grandi correnti epistemologiche italiane ed europee, dal formalismo logico di Peano al convenzionalismo di Poincaré all'empirio-criticismo di Mach.” (Passi citati dalla prefazione di Lombardo Radice alla ristampa anastatica – Zanichelli, 1982 – dell'opera: Enriques, F.; *Le matematiche nella storia e nella cultura, lezioni pubblicate a cura di Attilio Frajese*; Zanichelli editore, Bologna, 1938).

¹⁶ Si veda per esempio l'opera collettiva *Omaggio a Ludovico Geymonat; saggi e testimonianze* (Franco Muzzio editore, Padova, 1992).



tica teorica, basta poi fare il nome di Corrado Böhm¹⁷.

Quanto alle radici tecnologiche, occorre rammentare che l'elettronica non lineare (o impulsiva) è la tecnologia 'naturalmente' appropriata al calcolo logico e aritmetico; l'elettronica lineare, per converso, è stata alla base dei sistemi di telecomunicazione e tale è rimasta per lungo tempo. Si comprende, quindi, come il *know-how* maturato nelle tecnologie radar – anch'esse basate sull'elettronica impulsiva – sia risultato immediatamente riutilizzabile nello sviluppo dei primi computer americani e inglesi: quasi tutti gli ingegneri elettronici che s'impegnarono nella nascita del computer avevano infatti alle spalle valide esperienze nel campo del radar. Campo in cui gli anglosassoni avevano fatto investimenti enormi durante il conflitto, molto superiori a quelli fatti dai tedeschi, mentre in Italia si era rimasti quasi del tutto inerti. La miopia dei nostri Stati Maggiori lesinò le risorse al punto di impedire ogni esito significativo alle ricerche e alle realizzazioni speri-

mentali effettuate da Ugo Tiberio presso il RIEC (Regio Istituto Elettrotecnico e delle Comunicazioni) di Livorno e da Francesco Vecchiacchi del Politecnico di Milano come collaboratore delle industrie Marelli¹⁸.

Altre sparse isole di competenze affini alla nuova elettronica si raccordarono più o meno direttamente con gli esordi dell'informatica in Italia. Non è un caso che lo stesso Tiberio, appena approdato all'Università di Pisa nel 1954, lanciasse l'idea di costruire un elaboratore elettronico; idea che fu suffragata dall'autorevole suggerimento del premio Nobel Enrico Fermi, e si concretizzò nella costituzione del CSCE. Inoltre, la direzione del CSCE fu affidata al fisico Marcello Conversi il quale, pur non partecipando a tempo pieno alle attività tecniche, recava l'apporto di una lunga familiarità con l'elettronica impulsiva, maturata nelle ricerche sperimentali sui raggi cosmici: un filone di ricerca a cui Bruno Rossi, costretto poi a emigrare negli Stati Uniti, aveva contribuito anche con lo sviluppo di circuiti elettronici equivalenti alle "porte logiche"¹⁹.

¹⁷ Collaboratore dell'INAC fino dal 1953 e poi docente all'università di Roma, Böhm è stato il primo italiano ad acquisire una solida preparazione in materia: emigrato in Svizzera a causa delle leggi razziali, nel 1952 si era poi laureato al Politecnico Federale di Zurigo (ETH) con una originalissima tesi sui linguaggi algoritmici e i loro compilatori. Nell'Europa continentale dell'immediato dopoguerra l'ETH, con Rutishauser e altri, era il principale e forse unico centro universitario di ricerca in informatica; tra l'altro vi era installato il calcolatore Z4 che Konrad Zuse aveva dato in affitto (e si trattava di un precocissimo esempio di *business* informatico) dopo averlo rocambolescamente salvato dai disastri della guerra. Assieme a Giuseppe Jacopini, anch'egli dell'INAC, Böhm dimostrò più tardi un famoso teorema che costituì il fondamento teorico delle metodologie di programmazione strutturata (*Flow Diagrams, Turing Machines and Languages With Only Two Formation Rules*; Comm. of the ACM, Vol. 9, n. 3, 1966, p.366-371). Con l'occasione, è d'uopo rammentare che Jacopini aveva contemporaneamente raggiunto un risultato teorico ancora più radicale, descritto in *Macchina universale di von Neumann ad unico comando universale*; Calcolo, Vol.3, n. 1, 1966, p.23-29; il titolo di questo lavoro è di per sé espressivo.

¹⁸ Da notare che Tiberio e Vecchiacchi si erano entrambi formati proprio al RIEC, fondato e diretto da Giancarlo Vallauri al quale si deve, tra l'altro, la celebre 'equazione di Vallauri' che descrive il comportamento del triodo in regime lineare. È poi appena il caso di sottolineare le conseguenze dell'inerzia tecnologica degli alti comandi, particolarmente nefaste nella condotta della guerra sul mare. Sulla storia del radar italiano esiste una documentazione molto povera. Per di più, le poche fonti disponibili, tutte secondarie, non sempre concordano. Francesco Carassa presenta, infatti, il radar Lince come uno sviluppo (originale) dei laboratori Marelli con Vecchiacchi; un altro autore (Andrea Ottanelli) riferisce invece che il Lince, nella versione Lince "grande" altro non fu che un esatto clone del radar inglese catturato nei pressi di Tobruk nel corso della campagna d'Africa e che inoltre esso fu costruito a Pistoia nelle Officine Meccaniche San Giorgio ad opera di un consorzio a cui partecipavano la stessa San Giorgio (Genova), la Borletti (Milano) e la Galileo (Firenze). (Cfr. (1) Carassa F.; *La nascita dell'elettronica e delle telecomunicazioni: Vecchiacchi e i rapporti con l'industria* in: aa.vv.; *Il Politecnico di Milano nella storia d'Italia 1914/63*; Cariplo-Laterza, 1989, p. 501; (2) Ottanelli A.; *Il radar italiano*; Il coltello di Delfo, rivista di cultura materiale e archeologia industriale; n. 8, 1989, p.5-11). Per Tiberio e il radar navale italiano cfr. Calamia M., Palandri R.; *The history of the Italian radio detector telemetro* in: Burns R. (editor); *Radar development to 1945*; Peter Peregrinus Ltd. – IEE, Exeter, 1988, p.97-105. Va anche ricordato che Vecchiacchi, nel 1939, aveva per primo sviluppato la teoria quantitativa del multivibratore bistabile, il cosiddetto *flip-flop*.

¹⁹ Piuttosto che i lavori scientifici, di Rossi si segnala il libro divulgativo *I raggi cosmici* (Einaudi, Torino, 1971). Ai fisici delle particelle erano anche ben note le "scale binarie" che l'inglese Winn-Williams aveva introdotto fin dal 1932 per il conteggio automatico di "eventi" in rapida successione. Raffinando questa tecnica, Emilio Gatti presso il CISE (Centro Informazioni Studi ed Esperienze) realizzò un analizzatore di ampiezza d'impulsi per ricerche nucleari. Il CISE era stato costituito a Milano nel 1946, per iniziativa di Giuseppe Bolla, come polo di ricerca *high-tech* delle industrie che vi partecipavano.

Nel campo delle telecomunicazioni, lo stesso Vecchiacchi al Politecnico di Milano aveva sviluppato le tecniche di trasmissione a impulsi, antesignane dell'odierno digitale. Sistemi sperimentali di televisione, una tecnologia in parte simile a quella del radar, furono precocemente realizzati con il laboratorio Marelli e presentati alla fiera di Milano del 1939²⁰. Le ricerche sulla magnetostirazione condotte all'istituto di ultracustica erano potenzialmente utilizzabili per lo sviluppo di sistemi di memoria ma non sembra che siano state orientate in tal senso²¹. Come prologo immediato all'"anno zero", vanno anche ricordati i viaggi di studio e le partecipazioni a convegni internazionali che alcuni personaggi italiani compirono per approfondire le vaghe ma stimolanti notizie che giungevano d'oltreoceano e d'oltremarina. In questo, furono molto attivi i matematici dell'INAC, peraltro più attenti agli aspetti teorici e applicativi che non a quelli ingegneristici. Da questi contatti scaturirono una serie di pubblicazioni, apparse quasi tutte su *La Ricerca Scientifica*, che valsero a far circolare in Italia notizie più precise e di prima mano²². Particolarmente incisiva è stata l'opera di Bruno de Finetti – matematico e grande esperto di meccanografia applicata alle scienze statistiche e attuariali – il quale, nel 1952-53, collaborò strettamente con l'INAC in qualità di consulente del CNR. Fu molto influente il suo lungo e documentato articolo

"*Macchine 'che pensano' (e che fanno pensare)*", frutto di un viaggio negli Stati Uniti, con Piccone e Gaetano Fichera, durante il quale ebbe modo di esplorare le principali "culle" del computer²³. Egli, inoltre, introdusse l'INAC, che ne era completamente digiuno, all'uso di sistemi meccanografici "pre-informatici" per il calcolo scientifico e fu questo un fatto unico in Italia²⁴. L'Olivetti, dal canto suo, si era procurata un canale diretto grazie al piccolo laboratorio che aveva impiantato nel 1952 a New Canaan (Connecticut) con funzioni di osservatorio tecnologico.

Poco sopra si è accennato ai sistemi meccanografici e su di essi si avrà occasione di ritornare. Qui si deve però ricordare che proprio nell'elaborazione meccanografica dei dati affonda una delle radici della Linguistica Computazionale, una disciplina oggi matura e coltivata in tutto il mondo da schiere di studiosi. Tutto è cominciato in Italia per merito del padre gesuita Roberto Busa il quale concepì l'impiego delle schede perforate in un ambito di ricerca atipico e con metodologie del tutto originali rispetto alle consuete applicazioni della meccanografia. Fin dal 1946, egli mise mano alla più importante e pionieristica delle sue molte imprese: il monumentale *Index Thomisticus* che analizza esaustivamente, sotto il profilo linguistico, i dieci milioni di parole contenute nell'*Opera omnia* di Tommaso d'Aquino²⁵.

²⁰ La tesi di laurea di Dadda – 1945, relatore Vecchiacchi – verteva, appunto, sul progetto di un ponte radio a impulsi. (Cfr. Dadda L.; *Ricordi ecc.*; cit., p. 76)

²¹ Nel 1963, Giovanni De Sandre e Gastone Garziera che collaboravano con Pier Giorgio Perotto alla realizzazione dell'Olivetti P101, misero a punto la memoria a linea di ritardo magnetostirata e si trovarono a partire da zero, a cominciare dalla ricerca del materiale più adatto.

²² Scorrendo l'elenco delle pubblicazioni dell'INAC e isolando quelle del tipo ora considerato, si rilevano in particolare i nomi di Aparo, Dainelli, de Finetti, Kitz e Rodinò. (Cfr. *Pubblicazioni IAC 1927-1987*; IAC, Roma, 1989; si fa presente che nel 1960, all'atto del pensionamento del suo fondatore, l'INAC era stato ridenominato Istituto per le Applicazioni del Calcolo (IAC) "Mauro Piccone").

²³ L'articolo *Macchine che pensano ecc.* apparve su *Tecnica ed Organizzazione*; n. 3-4, 1952, p. 182-215. A un successivo viaggio di studio in Inghilterra, con Kitz e Rodinò dell'INAC-CNR, si riferisce l'articolo *Symposium on automatic digital computation*; *La ricerca scientifica*, a.23, n. 7, 1953, p.1 248-1259. Si veda anche Pitacco E.; *Bruno de Finetti: un matematico agli esordi dell'informatica* in {1}, p.151-168.

²⁴ Si veda l'articolo di de Finetti il cui titolo $y' = \sin kx - y^2$, alquanto ermetico, si chiarisce nel lungo sottotitolo: *con un impianto a schede perforate i tecnici dell'Istituto Nazionale per le Applicazioni del Calcolo hanno eseguito in meno di mezz'ora l'integrazione numerica di questa equazione differenziale* (*Civiltà delle macchine*; a.1, n. 3, 1953, p.31-33).

²⁵ L'altra radice della linguistica computazionale è quella delle ricerche sulla traduzione automatica, una zona di confine con l'intelligenza artificiale (Cfr. Calzolari N., Lenci A.; *Linguistica computazionale; strumenti e risorse per il trattamento automatico della lingua*; *Mondo Digitale*, a.3, n. 10, 2004, p.56-69). L'*Index*, che consta di settantamila pagine raccolte in 56 volumi, è oggi più agevolmente fruibile su CD. Tra gli innumerevoli scritti di (e su) padre Busa ci si limita a segnalare i suoi *The Annals of Humanities Computing: the Index Thomisticus* (*Computers and the Humanities*, Vol. 14, n. 2, 1980, p. 83-90) e Cinquant'anni a ... "*bitizzar*" parole, pubblicato in {1}, p.71-82.

4. LE MACCHINE 'COMPRATE'

Intorno al 1950, e non solo in Italia, il dilemma se fare o comprare un computer (*make or buy*) si poneva in termini sostanziali; dilemma che il Politecnico, grazie forse al sano pragmatismo dei milanesi, risolse rapidamente in favore del *buy* considerando assolutamente prioritaria la rapidità di messa in opera. Luigi Dadda poté, infatti, partire per la California con un mandato preciso e, trascorsi pochi mesi durante i quali partecipò alla costruzione della macchina e si addestrò alla sua programmazione, se ne tornò in patria con la stessa nave che trasportava la "sua" CRC 102A²⁶. All'INAC invece, la questione fu dibattuta per almeno un paio d'anni durante i quali si fece anche qualche modesto tentativo di realizzazione che, a quanto pare, non andò oltre un addizionatore con capacità di qualche cifra binaria. Allorché, sotto l'urgenza applicativa, si era ormai vicini alla decisione dell'acquisto, Mauro Picone, direttore dell'INAC, insistette affinché la Ferranti Ltd. valutasse come ultima spiaggia la possibilità che la macchina potesse essere almeno costruita in Italia²⁷. Va ricordato che una vicenda analoga, negli Stati Uniti, si era conclusa con esito esattamente opposto. Il *National Bureau of Standards* (NBS), che aveva anch'esso evidenti e urgenti interes-



Il calcolatore CRC-102A, installato al Politecnico di Milano

si applicativi, aveva aderito, fin dal 1946, all'iniziativa industriale di J.P. Eckert e J.W. Mauchly; questi celebrati pionieri, conclusa l'esperienza ENIAC alla Moore School di Philadelphia, erano stati i primi a sentire il profumo del business e avevano fondato la loro Electronic Control Company sulla base del progetto UNIVAC (*UNIVersal Automatic Computer*)²⁸. Resosi conto delle

²⁶ Il contributo di Dadda *Il Centro* ecc. reca in appendice la riproduzione del fascicolo di presentazione del centro di calcoli numerici con le caratteristiche tecniche del calcolatore. Per le caratteristiche tecniche della Ferranti Mark I^a, si veda Lavington S.; *Early* ecc.; cit.

²⁷ Alla pressante richiesta di Picone, la Ferranti oppose le oggettive difficoltà della cosa ma suggerì diplomaticamente qualche piccola concessione quale l'adozione di una telescrivente Olivetti per l'input/output. Stando a quanto Giorgio Sacerdoti ha raccontato a chi scrive, l'unica componente di produzione italiana fu alla fine l'impianto di raffreddamento a circolazione forzata d'aria. Per i risvolti dell'intera trattativa cfr. Bonfanti C.; *L'affare FINAC tra Manchester e Roma (1953-1955) ed alcuni documenti inediti ad esso relativi* negli atti del Congresso annuale AICA; Palermo, 1994, Vol. 1, p. 35-64. A conferma poi di quanto l'INAC abbia tentennato tra l'acquisto e la costruzione in proprio (o almeno in Italia), nell'articolo di Sagredo *Matematica* ecc., già citato, si legge che "è annunciata la costruzione di una calcolatrice elettronica per l'INAC presso la "Microlambda" di Napoli". La Microlambda era una vivace azienda *high-tech* che confluì poi nella Selenia. E de Finetti, reduce da un viaggio in Inghilterra, tra le "considerazioni applicabili al da farsi in Italia" scriveva: "il giudizio unanime di quanti si informavano della situazione in Italia e sentivano che avevamo intenzione di "comprare o acquistare" una macchina per l'INAC si esprimeva con l'esclamare (senza essere richiesti di un parere): "Build!", "You must build your computer yourselves!". Comunque, occorre che in un modo o nell'altro qualcosa si realizzi presto, perché si rischia di rimanere ultimi." (Cfr. de Finetti; *Symposium* ecc.; cit., p. 1249).

²⁸ L'ENIAC (*Electronic Numerical Integrator And Computer* (sic)) è passato alla storia come 'il primo' calcolatore elettronico del mondo. In realtà, pur essendo una macchina elettronica che faceva degli ottimi calcoli, non era ancora programmabile nel senso teorizzato dalla "architettura di von Neumann" e da allora accettato come requisito essenziale per ogni macchina che si volesse qualificare come "computer". L'ENIAC non era "il primo" neppure riguardo all'elettronica perché John Vincent Atanasoff, lavorando all'Iowa State College assieme al suo allievo Clifford Berry, aveva già realizzato nel 1942 un piccolo calcolatore elettronico (conosciuto come ABC - *Atanasoff Berry Computer*). Va da sé che neppure l'ABC possedeva una memoria programmabile e, per di più, risultava poco flessibile essendo stato concepito per un tipo specifico di calcolo matematico: la risoluzione di sistemi di equazioni lineari. È emerso successivamente che John Mauchly aveva mutuato proprio da Atanasoff le prime idee sul calcolo elettronico che furono alla base del progetto ENIAC. Su tutto questo, e sulla lunga contesa giudiziaria che ne seguì, si veda l'articolo di Allan R. Mackintosh *Il calcolatore di Atanasoff* (*Le scienze*; n. 242, 1988, p. 92-101) e il più ampio libro di C.R. Mollenhoff *Atanasoff, Forgotten Father of the Computer* (Iowa State University Press, Ames, 1988).

difficoltà che rallentavano il progetto industriale, l'NBS decise di costruirsi il computer in casa propria. Ne costruì anzi due, originali e diversi tra loro, che furono battezzati SWAC e SEAC²⁹; ed ebbe ragione e successo perché “quelli fatti in casa” furono pronti rispettivamente nel maggio e nel luglio 1950 mentre l'UNIVAC fu consegnato al primo cliente appena un anno dopo³⁰.



Il sistema FINAC, installato all'Istituto Nazionale del Calcolo di Roma

La fase dei computer “degli scienziati per gli scienziati” stava ormai cedendo il passo a quella dei computer “dell'industria per il mercato”.

Il fatto che i calcolatori acquistati dal Politecnico e dall'INAC³¹ fossero “pronti per l'uso” va poi chiarito alla luce dello stato dell'arte dell'epoca. La pura e semplice installazione, ad esempio, richiese un paio di settimane a Milano e, in quel di Roma, si protrasse per qualche mese³². Inoltre, nessuno dei produttori era in grado di fornire sul posto un servizio di manutenzione, laddove la componentistica richiedeva interventi quasi giornalieri per diagnosticare, localizzare e riparare i guasti circuitali; i maggiori problemi venivano dalle valvole termoioniche che erano 4.000 nella FINAC e “appena” 600 nel CRC 102A. Era dunque inevitabile - e accettato come normale - il fatto che il cliente dovesse disporre di un gruppo di esperti elettronici che svolgessero queste mansioni sulla base di una capillare conoscenza dell'*hardware*. E se questo fu un problema relativamente minore per il centro di calcoli del Politecnico che operava nel proprio ambiente naturale dell'ingegneria elettrica, costrinse, invece, l'INAC ad arruola-

²⁹ SWAC e SEAC sono rispettivamente acronimi di *Standard Western Automatic Computer* (realizzato a cura dell'Institute for Numerical Analysis dell'NBS, ubicato a Los Angeles sulla *West Coast*) e di *Standard Eastern Automatic Computer* (inizialmente e significativamente denominato *Standard Interim Computer*; fu realizzato dal Machine Development Laboratory, ubicato a Washington D.C. sulla *East Coast* dove è la sede centrale dell'NBS). Le sottolineature in questa nota sono state aggiunte dall'autore per richiamare l'attenzione sull'allusività dei nomi prescelti.

³⁰ Cade qui a proposito anche un altro episodio del folklore informatico di quegli anni; lo stralciamo da una testimonianza del celebre Edsger W. Dijkstra (*A Programmer's Early Memories* in: Metropolis N., Howlett J., Rota G.-C. (editori); *A History of Computing in the Twentieth Century*; Academic Press, 1980, p. 563-573): “la macchina era però così inaffidabile da risultare praticamente inutilizzabile. Fu inaugurata ufficialmente con tutta la pompa che la circostanza richiedeva; a scanso di sorprese, quale programma dimostrativo da eseguire in quel frangente così emozionante, era stata scelta la stampa di una tabella di numeri [pseudo]casuali, e, per quanto io ricordi, quello è stato il programma più complesso che, malgrado gli eroici sforzi, quella macchina abbia mai eseguito correttamente.” Accanto a questa versione ufficiale, che è già abbastanza gustosa, è circolata anche una leggenda metropolitana secondo la quale il programma dimostrativo doveva, in realtà, calcolare un problema ben preciso ma l'inaffidabile ARRA (tale il nome della macchina) sfornò, invece, una serie di numeri vistosamente fantasiosi. Uno dei responsabili, con ammirevole prontezza di riflessi, si rivolse allora agli ospiti esclamando: “perfetto, ecco la nostra tabella di numeri casuali!” e questa volta senza “pseudo”. A distanza di tanti anni, non è ingeneroso sottolineare che questa scena si svolse nel 1952 presso il centro matematico di Amsterdam, proprio quel centro che era stato in concorrenza con l'INAC per l'assegnazione della sede dell'ICC da parte dell'Unesco. Da notare che ARRA era una macchina “fatta in casa”; i rischi che Dadda e Picone avevano scansato optando per il buy non erano, quindi, affatto remoti.

³¹ Si aggiunga che, per entrambi gli acquisti, è stata determinante l'assegnazione di finanziamenti ERP (*European Recovery Program*, meglio conosciuto come “piano Marshall”). Con l'ERP gli Stati Uniti sostennero vigorosamente la ripresa postbellica dell'Europa occidentale; ne trassero peraltro duraturi vantaggi in termini di presenza, e talvolta di dominio, nell'economia europea.

³² La macchina americana, grazie a un'ottima ingegnerizzazione e alle dimensioni più contenute, veniva consegnata in “unità” preconfezionate. Per quella inglese, che era molto più ingombrante, gli armadi metallici e l'elettronica dovevano, invece, essere spediti in componenti separati o, al massimo, in piccoli sottoassemblaggi; in tali condizioni, il montaggio si rivelò un'operazione molto complessa e altrettanto formativa.



re in fretta e furia un gruppo di giovanissimi ingegneri - tra i quali Giorgio Sacerdoti, Paolo Ercoli e Roberto Vacca - per integrare il proprio *staff* che, fino a quel momento, era formato esclusivamente da matematici, di norma poco inclini a maneggiare saldatori e oscillografi.

E ben presto si dovettero superare i confini della “semplice” manutenzione in quanto entrambi i calcolatori di cui si parla dovevano, in qualche modo, accrescere le loro prestazioni per adeguarsi a nuove necessità e per tenere il passo con quanto succedeva altrove. Gli ingegneri della manutenzione si trasformarono, quindi, in provetti progettisti e realizzatori di hardware destinato a integrare o modificare quello originario. Circostanza questa che valse ad attenuare di molto, sul piano pratico, la distanza tra l’esperienza del *make* e quella del *buy* e che fu uno dei più fruttuosi sviluppi delle iniziative di Milano e di Roma. Sviluppi che coinvolsero in prima persona Luigi Dadda e molti dei suoi colleghi e allievi del Politecnico; quanto all’INAC, basta ricordare che Giorgio Sacerdoti fu determinante nel definire l’architettura dei progetti industriali dell’Olivetti e che Paolo Ercoli fu tra i più influenti ideatori della macchina CINAC, di originalissima concezione. Ma di questo argomento si parlerà nel seguito di questo articolo.

4.1. Il software, protagonista inatteso

Sempre in tema di computer “pronti per l’uso”, è altrettanto rilevante il fatto che le macchine venissero fornite funzionanti ma incapaci di operare: per renderle produttive era necessario programmarle in codice macchina partendo dalla più desolante *tabula rasa*. Si doveva imparare “a menadito” il corredo delle istruzioni native, districandosi tra codici operativi, formato delle istruzioni, uso dei registri e addirittura sincronizzare lo svolgimento del

programma in base ai tempi di esecuzione delle operazioni di *input/output*³³. Il problema del software cominciava a manifestare la sua importanza cruciale anche se i tempi e i costi del suo sviluppo rimasero per lungo tempo sottaciuti e sottovalutati, con effetti talvolta catastrofici³⁴. Facendo un momentaneo salto temporale, si rammenta il successivo insorgere della “sindrome del 90 per cento”, una malattia endemica nei progetti di sviluppo di software e ancora non del tutto debellata, i cui sintomi sono il puntuale “quasi completamento” del progetto e l’indefinito protrarsi del suo completamento effettivo.

Era, quindi, inevitabile che il centro di calcoli del Politecnico e l’INAC si impegnassero pesantemente sul fronte del software; l’impegno maggiore fu dedicato alla messa a punto di biblioteche di programmi applicativi ciascuno dei quali risolveva una delle classi di problemi più ricorrenti nella pratica. Esempi tipici erano la soluzione approssimata di equazioni algebriche oppure di sistemi di equazioni lineari e il calcolo matriciale in genere. Tali programmi risultavano tanto meglio concepiti e tanto più efficaci in quanto potevano adattarsi a un calcolo specifico non solo, com’è ovvio, accettando i relativi dati di input ma anche utilizzando un sistema flessibile di parametri di personalizzazione. A un livello sottostante a quello dei programmi applicativi stava un repertorio di componenti software elementari dedicati a elaborare singole operazioni quali il calcolo delle funzioni elementari e trascendenti oppure l’ordinamento e la ricerca tabellare; perfino le operazioni aritmetiche in virgola mobile (*floating-point*) vennero eseguite via software fino a quando gli stessi utenti-progettisti non realizzarono le estensioni hardware che le rendevano disponibili come istruzioni native. L’uso di questi componenti elementari av-

³³ È appena il caso di ricordare che il progetto dell’architettura logica di un computer si identifica in buona sostanza con la scelta del corredo di istruzioni native e del formato di queste (sistema di indirizzamento). Le difficoltà con le operazioni di i/o erano dovute sostanzialmente alla mancanza delle funzioni di *interrupt*.

³⁴ Un servizio dell’epoca, con sensibilità piuttosto inusuale nell’ambiente giornalistico, illustra efficacemente il ‘peso’ del software: “occorre molto più tempo all’uomo per l’elaborazione matematica del problema che alla calcolatrice elettronica per il calcolo. Spesso una preparazione di una settimana sfocia in un ronzio di un’oretta ...” Il servizio prosegue dicendo “Confidiamo di aver demolito così – con quanto sopra – il sogno dello studente di portarsi a scuola una calcolatrice elettronica tascabile per il compito in classe di matematica. Almeno per ora: perché non si può mai sapere ...” e Dadda commenta: “Quel giornalista fu ottimo profeta!” (Cfr. Dadda L.; *Il centro* ecc.; cit., p. 12).

venne in un primo tempo ricopiandoli e incorporandoli a mano nel programma principale; un metodo talmente scomodo e rudimentale che fu di stimolo a sviluppare tecniche via via più sofisticate (passaggio di parametri, rilocabilità, rientrabilità) per automatizzare l'ingaggio dei componenti elementari in forma di sottoprogrammi (*subroutine*) o di macro³⁵. Man mano che si andava avanti, ogni installazione doveva poi costruirsi i suoi programmi di servizio, gli embrioni dei sistemi operativi che raggiunsero un livello soddisfacente solo dopo il 1960.

Il fatto poi che questi patrimoni di software fossero sviluppati in linguaggio macchina o in uno equivalente, rendeva impossibile lo scambio di programmi tra utilizzatori di computer diversi per marca o per modello e rendeva addirittura traumatica la conversione quando si fosse deciso di passare a un nuovo computer³⁶. Se molti computer, in tutto il mondo, furono mantenuti in esercizio ben oltre il limite di obsolescenza tecnologica e di economicità d'impiego, lo si dovette proprio all'enormità dello sforzo di conversione del loro corredo di software, al quale veniva così implicitamente riconosciuto un valore prevalente rispetto all'hardware³⁷. Un passo deci-

sivo verso la compatibilità del software rispetto all'hardware venne compiuto in seguito con l'invenzione dei linguaggi standard - detti anche linguaggi "evoluti", "di alto livello" e "orientati al problema" (per contrapposto a quelli "orientati alla macchina" come gli assembler) - i più famosi dei quali nacque e si imposero come standard di fatto; solo in seguito, constatata l'importanza cosmica del problema, se ne occuparono commissioni e organismi di standardizzazione riconosciuti a livello internazionale³⁸. Si trattò di una autentica "rivoluzione copernicana" che aprì una nuova epoca al modo di essere dell'informatica: il suo valore più fondante non risiede tanto nell'aver fornito ai programmatori di tutto il mondo dei linguaggi più vicini al modo di esprimersi degli umani - in particolare modo degli umani di specie anglofona - quanto nell'aver affrancato per sempre il software dal giogo di asservimento all'hardware³⁹. Per chi non la dovesse progettare, costruire o comunque "lavorarci dentro", la macchina reale veniva relegata al ruolo di scatola nera; il vero interlocutore diventava la macchina virtuale, una entità immateriale definita dalla sintassi e dalla semantica del linguaggio standard.

³⁵ Il primo e molto influente testo sulle metodologie per costruire una biblioteca di software (composta da programmi matematici e di servizio) è stato scritto da Maurice V. Wilkes, David J. Wheeler e Stanley Gill (*The preparation of programs for an electronic digital computer - With special reference to the EDSAC and the use of a library of subroutines*; Addison-Wesley, Cambridge, 1951). Malgrado lo sforzo di generalizzare i metodi, era inevitabile che il tutto (e in primo luogo i numerosi esempi di programmi di cui si forniva il listato) rimanesse vincolato a una specifica macchina e al suo linguaggio nativo, come espressamente avvertito nel sottotitolo. Wilkes aveva diretto il progetto e la costruzione dell'EDSAC (*Electronic Delay Storage Automatic Calculator* all'Università di Cambridge e divenne subito uno dei più autorevoli *guru* dell'informatica in Inghilterra. Per la sua importanza, il libro ora citato è stato incluso tra le 16 opere di rilevanza storica fondamentale che sono state selezionate dal CBI (*Charles Babbage Institute*) per essere ristampate, tra il 1984 e il 1992, e rese così meglio fruibili agli studiosi "moderni"; una notevole impresa editoriale compiuta con la collaborazione di MIT Press, di Tomash Publishers e di prestigiosi *editor*. Il CBI opera dal 1977 presso l'Università di Minneapolis come centro di raccolta e di studio per la documentazione della storia dell'informatica; gli è ormai riconosciuto il ruolo di punto di riferimento internazionale in questo settore specialistico e per noi prezioso.

³⁶ I linguaggi mnemonici di tipo *assembler* furono introdotti abbastanza precocemente e alleviarono il lavoro di programmazione. Essi però ricalcavano il linguaggio macchina e non portavano alcun beneficio nel problema che qui interessa.

³⁷ La FINAC, rimasta in uso per quasi dieci anni, fu un caso notevole di longevità. (Cfr. Ercoli P.; *Otto anni di attività del calcolatore dell'Istituto nazionale per le applicazioni del calcolo* negli atti del 1° congresso AICA; Bologna, 1963, p. 149-158).

³⁸ Tale fu la storia dei linguaggi più famosi: il FORTRAN (*FORmula TRANslator*) e il COBOL (*COmmon Business Oriented Language*), introdotti rispettivamente nel 1954-55 - per iniziativa dell'IBM (John Backus) - e intorno al 1960 - per iniziativa del CODASYL (*COmmittee on DATA SYstems Languages*), sponsorizzato dal DoD (Department of Defense) statunitense. Il riferimento classico è: Wexelblat R.L. (editor); *History of Programming Languages*; Academic Press, 1981. Delle vicende italiane in questo settore si sono occupati Stefano Crespi Reghizzi e Luigi Petrone: *Linguaggi e teoria dei linguaggi*; in {1}, p. 85-101.

³⁹ Rimaneva comunque necessaria un'interfaccia, un oggetto software che avesse come ingresso il "programma sorgente", scritto in linguaggio standard quasi-umano, e producesse in uscita il "programma oggetto", equivalente al primo ma espresso nel codice nativo di una specifica macchina. Anche se ci volevano tanti programmi traduttori (compilatori o interpreti che fossero) quanti erano i tipi di macchina destinati a ospitare un determinato linguaggio standard, la semplificazione era comunque enorme.



4.2. Ricadute e conseguenze

Tutte le problematiche a cui si è accennato nei due precedenti paragrafi hanno trovato puntuale riscontro nei pionieristici ambienti applicativi del Politecnico di Milano e dell'INAC, portando in breve tempo a risultati sulla cui importanza non si vuole insistere oltre. Si vuole, invece, sottolineare come dal coacervo delle attività pratiche indotte dalla semplice presenza di un computer "comprato", si sia passati rapidamente a una visione organica e disciplinare delle varie branche dell'ingegneria dei computer, dell'ingegneria del software, della matematica computazionale, della complessità di calcolo, dei linguaggi formali e via elencando. Tutte cose inaudite per l'Italia di quell'epoca e rispetto alle quali quegli stessi ambienti si posero come centri di irradiazione sia mediante le attività didattiche e l'interazione con il mondo delle imprese, sia grazie alla mobilità degli specialisti i quali, formati in quelle prime isole informatiche, andavano a inseminare nuovi centri di competenza.

Già a partire dal 1954, il Politecnico attivò corsi di programmazione e poi insegnamenti di calcoli numerici e di calcolatrici elettroniche. Si tennero corsi speciali e di perfezionamento ingaggiando come docenti anche esperti "prestati" dalle industrie e, non appena la legge lo consentì, nel 1960 fu istituito ufficialmente il corso di laurea in ingegneria elettronica. Non si può sottacere come Luigi Dadda, confortato dalla qualità dei suoi allievi-colleghi e da un clima generalmente collaborativo, sia stato l'infaticabile motore di questa evoluzione, guadagnandosi - assieme a pochi altri tra i quali certamente Alessandro Faedo - una posizione di prestigio e di influenza a livello nazionale e di autorevolezza negli organismi internazionali di cui ha fatto parte⁴⁰. A lui si devono, tra l'altro, le prime idee sulla costituzione dei consorzi universitari per l'informatica; e da quelle idee germinò verosimilmente la proposta della rete europea dei calcolatori, una sorta di ARPANET, adottata

come iniziativa comunitaria COST II. Ne sortì la EIN (*European Informatic Network*) che venne però chiusa nel 1978 a causa dell'opposizione dei monopolisti dei servizi telefonici. È comunque ovvio che non siano mancati ostacoli, controversie e incomprensioni. Tra i tanti e vivaci episodi che Dadda ha registrato nei suoi *Ricordi di un informatico*, più volte richiamati in questo lavoro, si ricorda la polemica che vide lui stesso e gli "informatici" contrapposti a Ruberti con i "sistemisti", riguardo all'impostazione dei corsi di laurea nelle facoltà d'ingegneria; Dadda l'ebbe vinta e si giunse alla sopra menzionata legge del 1960. E poi, in una tavola rotonda del 1971, la sua reazione verso il fisico Gilberto Bernardini il quale "non riteneva che l'informatica fosse una vera e propria disciplina scientifica" in quanto tra i suoi cultori non era "invalsa [...] l'abitudine a esprimersi per teoremi"⁴¹.

5. LE MACCHINE "INVENTATE"

Passando ora al versante del *make*, è opportuno chiarire subito le profonde differenze che intercorrono tra due maniere, due stili, di affrontare questa scelta: lo stile scientifico e lo stile industriale.

In linea di principio, e non solo nel settore dei computer, lo stile scientifico è caratterizzato da:

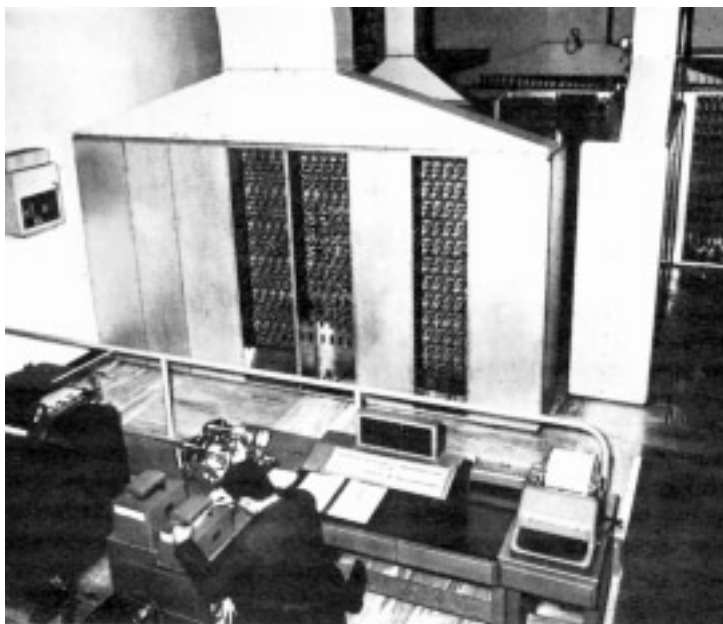
1. i risultati della ricerca sono valutati col metro del loro contributo alla conoscenza in generale, senza restrizioni che vadano a beneficio materiale di istituzioni o di singole persone;
2. il trasferimento tecnologico si attua mediante l'aperta diffusione delle informazioni (pubblicazioni, convegni, contatti interpersonali);
3. i manufatti tecnologici sono destinati a uso interno e tipicamente realizzati in esemplare unico.

I criteri generali dello stile industriale sono invece:

1. i risultati della ricerca vanno a beneficio dell'impresa e delle persone che vi partecipano e vi operano e il progresso scientifico generale è un'istanza residuale;

⁴⁰ Si ricorda che egli è stato a lungo rettore del politecnico di Milano di cui è professore emerito, ma non ci avventuriamo in una elencazione dei numerosi incarichi che ha ricoperto. Ci fa comunque piacere rammentare in questa sede che, nel triennio 1967-1970, egli è stato presidente dell'AICA e ne è socio onorario.

⁴¹ Cfr. Dadda L.; *Ricordi ecc.*; cit., p. 98 e p. 78.



La Calcolatrice Elettronica Pisana (CEP)

2. le informazioni sono “proprietarie” e devono essere protette verso la concorrenza;
3. il trasferimento tecnologico si realizza mediante la vendita (o altre forme di cessione remunerativa) e il suo valore consiste nei benefici derivanti dall'uso, presso gli acquirenti, delle tecnologie prodotte;
4. i requisiti qualitativi dei manufatti (*design* estetico ed ergonomico, bassi costi d'esercizio ecc.) e l'economia del processo produttivo condizionano le attività di ricerca & sviluppo. Tanto basta a rendere conto delle diversità tra l'avventura scientifica del CSCE e l'avventura industriale dell'Olivetti. C'era in comune

il fatto di partire da zero; due autentiche “sfide italiane” a cui arrise il successo⁴². Si comprende anche il motivo per cui, pur mantenendo viva la collaborazione, il CSCE e il laboratorio elettronico Olivetti perseguirono fin dall'inizio obiettivi divergenti: un calcolatore per uso scientifico interno (e fu la CEP) e un elaboratore da prodursi in serie per il mercato, orientato prevalentemente alle applicazioni gestionali (e fu l'ELEA 9003)⁴³. Le problematiche ora esaminate si manifestavano anche nei Paesi dell'Europa continentale molti dei quali avevano intrapreso, con il nostro stesso *gap* di 5-10 anni, la rincorsa agli inglesi e agli americani; sul terreno industriale solo la francese Bull era riuscita, non senza affanno, a tenersi al passo mentre i prodotti della Zuse A.G. rimanevano attestati sulla tecnologia elettromeccanica dei relè⁴⁴. Non mancano le analisi retrospettive sullo scenario europeo di quel periodo; argomento di grande interesse che però esula dal tema che ci si è qui proposti⁴⁵.

5.1. CEP

Con la costruzione dell'elettrosincrotrone di Frascati, anche la fisica italiana cominciava a confrontarsi con la *big science* e Pisa aveva sperato che quella grande impresa si realizzasse nel proprio territorio; delusa in questo disegno, cercò altre direzioni per mettere a frutto il consistente finanziamento (150 milioni di lire di allora) reso disponibile dagli enti locali (Province e Comuni) della stessa

⁴² Anche l'università di Padova (istituto di fisica) si propose di costruire un calcolatore elettronico. A questa idea, presto abbandonata, accenna Dadda nei *ricordi di un informatico*: “a Padova il fisico Francesco Piva alla fine degli anni cinquanta intraprendeva la costruzione di una macchina, per... istigazione degli statistici.” (p. 78). È da supporre che tra gli “statistici” padovani ci fosse Mario Volpato, personaggio ben noto nell'ambiente informatico. Rispondendo recentemente a una domanda postagli da chi scrive, lo stesso Dadda ha aggiunto che i fisici di Padova, in particolare Silvio Bezzi (cristallografo) e Carlo Panattoni, “volevano usare la macchina del Centro di calcolo del Politecnico ma non potevano pagare nulla (neanche la modesta tariffa stabilita per le università, Politecnico compreso) e decisero di costruirsi una macchina”. È opportuno precisare che l'uso del centro di calcolo del politecnico, un “centro servizi” a tutti gli effetti, rientrava nelle lesinate spese correnti mentre per la costruzione di una macchina si sperava evidentemente di attivare altri finanziamenti. Su queste storie “minori”, che sono state verosimilmente più di una, sarebbe auspicabile una ricerca più approfondita; tra di esse rientra per esempio la storia dei brevetti di Alessandro Boni, collaboratore dell'INAC. (Cfr. Bonfanti C.; *L'affare ecc.*; cit., p. 57 segg.)

⁴³ Le varianti tecnologiche che distinguono un calcolatore scientifico da uno gestionale non ledono il fatto che entrambe le specie rimangono intrinsecamente *general-purpose*.

⁴⁴ Il primo elaboratore di Zuse a tecnologia elettronica è stato lo Z 22, del 1955. Il suo collaboratore e amico Helmut Schreyer aveva elaborato un progetto elettronico già nel 1938 ma i 2.000 tubi termoionici (valvole) che si stimavano necessari rimasero per lungo tempo al di fuori della loro portata.

⁴⁵ Si rinvia senz'altro a una delle analisi di questo genere che si trova in: Aspray W.; *International Diffusion of Computer Technology, 1945-1955*; *Annals of the History of Computing*, Vol. 8, n. 4, 1986, p. 351-360.

Pisa, di Lucca e di Livorno. Si fece allora una scelta coraggiosa e innovativa: costruire un computer – o meglio, una “macchina calcolatrice elettronica”. Già se ne parlava, ma l'autorevole suggerimento del premio Nobel Enrico Fermi risultò decisivo⁴⁶. Le cose andarono avanti rapidamente e, il 18 aprile 1955, fu formalmente istituito il CSCE, affidandone la direzione a Conversi, direttore dell'istituto di fisica. Alla base dell'iniziativa stava una relazione esplorativa redatta da Alfonso Caracciolo di Forino, che fu discussa in un incontro con “un nutrito gruppo di studiosi di varie Università italiane” e a conclusione del quale si decise, tra l'altro, che “prendendo per modello una delle macchine attualmente funzionanti all'estero, occorre dar subito corso ad un progetto di larga massima che però non si limiti agli aspetti tecnici dell'impresa [...]”⁴⁷. Il CSCE fu articolato inizialmente in una sezione logico-matematica, diretta da Caracciolo, e in una di ingegneria-elettronica, diretta da Giovanni Battista Gerace. Nel frattempo (nel maggio del 1955), era stata stipulata la convenzione tra Università di Pisa e Olivetti e in seguito arrivarono altri fondi dal CNR, dall'INFN e dal CNRN⁴⁸. La “macchina ridotta” fu

pronta nel corso del 1958 e, malgrado il nome, era perfettamente in grado di supportare le attività “produttive” di calcolo e di sviluppo di software. A queste ultime, sotto l'impulso di Caracciolo, il CSCE si dedicò con grande attenzione ed eccellenti risultati⁴⁹; si sottolinea questo aspetto per contrapporlo a quanto avvenne, come si vedrà nel seguito, nel progetto ELEA.

Si è visto come, fin dall'inizio, si prendessero in considerazione, a modello, “le macchine attualmente funzionanti all'estero”; in effetti, i progettisti della CEP ebbero molti e benefici contatti con chi aveva cominciato prima di loro, specialmente con gli inglesi e, tra questi, Maurice Wilkes. Il fatto che la CEP, come più tardi l'Olivetti ELEA 6001, sia stata una macchina ‘microprogrammata’ risente evidentemente dell'esperienza di Cambridge⁵⁰. Ma la tecnologia sviluppata dal CSCE, la matrice logica di sequenza, fu una soluzione del tutto originale⁵¹. La memoria centrale della CEP, a nuclei di ferrite, era organizzata a parole di 36 bit, con capacità di 8 kparole, associata a una memoria di massa (tamburo magnetico) di capacità doppia; la CEP definitiva fu dotata anche di unità a nastro magne-

⁴⁶ Si veda il contributo di Franco Denoth *CEP (Calcolatrice Elettronica Pisana) – Dal CSCE all'IEI*, pubblicato in {1}, p.1 03-115. Denoth, che è stato uno dei pionieri del CSCE e poi direttore dell'IEI, è ritornato recentemente sull'argomento, con una relazione non pubblicata (per ora?) in occasione del convegno “1954/2004, 50 anni dalla prima macchina da calcolo” in cui sono intervenuti anche Luigi Dadda e Ivo De Lotto, attuale presidente AICA (Villa Monastero, Varenna (Lecco), 16 giugno 2004). Ricordiamo che Fermi, nel suo ultimo soggiorno in Italia, era intervenuto alla scuola estiva internazionale di Varenna e lì era stato consultato su tale questione dai fisici Marcello Conversi, Giorgio Salvini e Gilberto Bernardini. La famosa lettera di Fermi, che peraltro menziona solo Conversi e Salvini, è datata 11 agosto 1954; di essa si è vista la fedele trascrizione dattilografica (all'epoca non esistevano le fotocopiatrici) fattane dalla segreteria di Enrico Avanzi, rettore dell'università di Pisa, ma sarebbe interessante rintracciarne l'originale firmato. È curioso osservare come la lettera di Fermi esordisca con un cordiale “Caro Professore” mentre la risposta di Avanzi si apre con un deferente “Cara Eccellenza”; effettivamente, durante “il ventennio”, come membro dell'Accademia d'Italia, a Fermi spettava il titolo di “eccellenza” e lui ci scherzava volentieri sopra.

⁴⁷ Cfr. le “schermate” con cui Denoth ha illustrato la sua relazione al citato convegno di Varenna.

⁴⁸ È interessante leggere gli impegni che quest'ultima si assumeva: “destinare due ingegneri o fisici e due tecnici, di gradimento del Centro [...]; comunicare, nei limiti che la Società riterrà opportuni, informazioni di carattere tecnico scientifico [...] e] l'utilizzazione gratuita di eventuali brevetti; corrispondere all'università di Pisa un contributo annuo [...] di lire dieci milioni”. I due ingegneri dell'Olivetti furono V. Sabbadini e G. Cecchini; quest'ultimo rimase assegnato al CSCE fino al completamento della CEP.

⁴⁹ Per fare un solo esempio, si ricorda il compilatore FORTRAN, primo per l'Italia, sviluppato dal gruppo di Otello Giacomo Mancino; tale compilatore implementava alcune istruzioni non previste dal linguaggio standard e introdotte per sfruttare al meglio alcune caratteristiche dell'hardware CEP.

⁵⁰ Anche la strategia pisana di procedere a tappe, dove la prima tappa è stata la “macchina ridotta”, richiama l'esperienza inglese dell'Università di Manchester che aveva costruito in fretta e furia la sua molto spartana *baby machine*.

⁵¹ La tecnica della microprogrammazione anticipava quella del *firmware*, ma la matrice logica di sequenza andava oltre ed era una vera e propria *eprom*, con l'effetto che, riconfigurando la matrice, la CEP poteva all'occorrenza trasformarsi in una macchina diversa. I bit della matrice erano costituiti da piccoli cilindri di ferrite e, per spostarli comodamente tenendoli però infissi ai sostegni, si usò la ‘gomma pane’, quella per cancellare: una soluzione da autentici creativi!

5.2. Olivetti ELEA

Adriano Olivetti, in sintonia con i figli Roberto e Dino, aveva percepito l'opportunità, anzi la necessità strategica, che l'azienda entrasse nel campo dell'elettronica⁵⁶.

Sondato il terreno con l'osservatorio tecnologico di New Canaan, il primo passo operativo fu l'assunzione di Mario Tchou, un brillante scienziato trentenne di famiglia italo-cinese che lo stesso Adriano aveva conosciuto alla Columbia University. A lui fu affidato il compito di selezionare e di dirigere il gruppo di giovani ingegneri e fisici che dettero vita al laboratorio di ricerche elettroniche, ospitato in un primo tempo dall'istituto di fisica dell'università e insediato poi (inizio del 1956) a Barbaricina, un sobborgo di Pisa, in una villa il cui gradevole parco confinava con un allevamento di cavalli. La scelta di Pisa, dove non c'era nessun precedente insediamento dell'Olivetti, lascia supporre che, almeno in un primo momento, si avesse in mente un modello di collaborazione stretta tra università e industria; un modello che aveva illustri precedenti come ad esempio quello di Manchester.

Data l'assoluta novità per l'Italia di quel tipo di ricerche, il primo compito che Tchou assegnò ai suoi "ragazzi di Barbaricina" fu quello di mettersi a studiare, indirizzandoli anche con i suoi seminari, e poi anche a sperimentare; il tutto in un clima informale e creativo di sapore goliardico⁵⁷. Tchou, intanto, faceva la spola con Roma per consultarsi con Giorgio Sacerdoti, fintanto che questi lavorò all'INAC, sull'architettura generale del computer⁵⁸. Chiarite le idee, vennero assegnate mansioni più precise riguardanti i sottosistemi (memorie, governo periferiche, unità di calcolo, unità di controllo) che nella primavera del 1957 fu-



Il sistema ELEA 9003 della Olivetti



Il gruppo dei progettisti. Da sinistra a destra e dall'alto in basso: G. Calogero, F. Filippazzi, M. Tchou, R. Galletti, M. Grossi, S. Sibani, G. Sacerdoti, L. Borriello, S. Fubini, O. Guarracino, G.F. Raffo

rono completi e pronti per l'assemblaggio. Nel frattempo era, però, intervenuta una innovazione dirompente: il *transistor*. Tchou insistette con i vertici aziendali per adottare la nuova tecnologia anche a rischio di rallentare il progetto pur di uscire sul mercato con un

⁵⁶ Nella figura di Adriano Olivetti si riscontra un amalgama di intelligenza imprenditoriale e di utopia sociale talmente singolare che ne ha fatto uno dei "casi" più studiati nella storia dell'industria. I capisaldi del suo "credo" sono stati la solidarietà aziendale (Comunità) e l'urbanistica (Città dell'uomo); alcuni hanno ancora ben vivo il ricordo di come egli fu snobbato in politica e osteggiato dalla confindustria (dove era considerato poco meno che un sovversivo) e dalle organizzazioni sindacali nazionali (per le quali era solo un astuto paternalista e che mal sopportavano l'autonomia sindacale dei lavoratori dell'Olivetti). Si veda ad esempio il libro di Valerio Occhetto *Adriano Olivetti, industriale e utopista* (Cossavella editore, Ivrea, 2000; ristampa anastatica della prima edizione *Adriano Olivetti*; Arnoldo Mondadori editore, 1985) e quello più recente di Luciano Gallino e Paolo Ceri *L'impresa responsabile: un'intervista su Adriano Olivetti* (Edizioni di Comunità, Torino, 2001).

⁵⁷ Caratteristica di questo ambiente fu l'usanza di affibbiare nomignoli onomatopeici: Franco Filippazzi divenne così Flip (il cagnolino spaziale del popolare Etabetta) e Lucio Borriello fu invece Bor (con allusione al celebre fisico Niels Bohr).

⁵⁸ Giorgio Sacerdoti si era laureato a Roma in ingegneria elettrica con una tesi sui computer, la prima in Italia su questo avveniristico argomento; a tale riguardo egli ha raccontato come, per documentarsi sull'elettronica impulsiva, dovette rivolgersi ai fisici perché gli ingegneri ne parlavano solo "per sentito dire". Con l'occasione, ricordiamo che, venuto a mancare Tchou, Sacerdoti assunse la direzione tecnica della divisione elettronica; egli è stato poi presidente dell'AICA nel corso di tre mandati e ne è socio onorario.

prodotto di assoluta avanguardia⁵⁹. Mentre il prototipo a valvole termoioniche veniva completato (fu la “macchina zero”, poi installata nella sede centrale di Ivrea per automatizzare la gestione del magazzino) il prototipo transistorizzato fu portato a termine rapidamente, grazie anche alla concezione modulare del progetto, e installato presso la direzione commerciale di via Clerici a Milano per essere utilizzato come sistema dimostrativo per i clienti e per lo sviluppo di software; questo elegantissimo prototipo, come anche il 9003 di serie, recava la prestigiosa firma del *designer* Ettore Sottsass Jr. e fu inaugurato l’8 novembre 1959 alla presenza del presidente Gronchi, non nuovo a simili circostanze. Il primo esemplare della versione prodotta industrialmente, l’E-LEA 9003, fu acquistato dalla Marzotto alla fine del 1959 e installato l’anno seguente nei suoi stabilimenti tessili di Valdagno. Senza addentrarsi in particolari tecnici si vuole ricordare che, oltre a essere il primo elaboratore completamente transistorizzato, l’E-LEA 9003 introduceva una tecnica simile a quella degli *interrupt* che disimpegnava l’unità centrale dalle operazioni di ingresso/uscita consentendo la multiprogrammazione fino a tre livelli; possedeva una memoria centrale espandibile modularmente da 20.000 a 160.000 caratteri di sei bit più uno di parità⁶⁰.

Nel lasso di tempo tra l’inizio delle ricerche e l’esordio sul mercato, l’azienda aveva naturalmente la necessità di migliorare e di aggiungere valore alla propria linea di prodotti tradizionali, che comprendevano tra l’altro i sistemi meccanografici commercializzati dalla consociata Olivetti-Bull e le macchine elet-

trocontabili scriventi della serie Audit, particolarmente apprezzate nel settore bancario⁶¹. Queste furono dotate anche di un dispositivo per la registrazione automatica su banda di carta perforata, una soluzione di gran lunga più economica rispetto a un perforatore di schede. Le opportunità di innovazione relative ai prodotti Audit furono individuate in due direzioni:

1. il miglioramento delle prestazioni al livello delle operazioni “locali”, che si svolgevano tipicamente in agenzie disseminate sul territorio;

2. la capacità di alimentare direttamente un sistema meccanografico centrale, che veniva con ciò anch’esso valorizzato.

Entrambi gli obiettivi furono raggiunti grazie al laboratorio elettronico in cui, mentre era in corso il progetto E-LEA, furono realizzati due prodotti elettronici che furono rilevanti sul piano tecnologico e commerciale anche se possono apparire di minor conto a paragone della grande meta finale. In primo luogo, le prestazioni “locali” delle Audit vennero potenziate dotandole della UME (*Unità Moltiplicatrice Elettronica*). Il secondo obiettivo fu conseguito grazie al dispositivo elettronico CBS (*Convertitore Banda-Schede*), sfornato anch’esso dal laboratorio e il cui sviluppo fu affidato a Pier Giorgio Perotto. Le bobine di banda registrate in periferia dalle Audit venivano inviate al centro meccanografico e lì, mediante il CBS, convertite rapidamente in schede perforate; si eliminavano così tutte le operazioni intermedie basate sul tradizionale flusso di documenti cartacei che comportava costi aggiuntivi, spreco di tempo ed errori di trascrizione. Sebbene il semplice trasporto

⁵⁹ L’avvento del transistor fu anche alla base della costituzione della SGS (Società Generale Semiconduttori), azienda che, dopo molti passaggi di proprietà, è diventata l’odierna ST Microelectronics, tra i più forti produttori di microcircuiti. La società fu ideata da Adriano anche per verticalizzare la catena produttiva dei computer, rendendo così l’Olivetti autonoma nell’approvvigionamento della componentistica e, in quel momento, specificamente dei transistor. La Fairchild fu uno dei primi soci della SGS ed ebbe un ottimo naso nel portarsi in America Federico Faggin, giovane fisico che si era formato nella SGS. Di recente, Angelo Gallippi ha scritto la storia, che tuttora continua, dei successi di Faggin nella *silicon valley* (*Faggin, il padre del chip intelligente*; Adnkronos libri, Roma, 2002).

⁶⁰ Per una esposizione sommaria, ma esauriente, delle caratteristiche tecniche si vedano per esempio i contributi di Sacerdoti (*Il calcolatore Olivetti E-LEA 9003*) e di Galletti e Filippazzi (*Sicurezza nel tempo: fattore determinante nel progetto dei calcolatori elettronici*) apparsi in: *Contributi all’automazione aziendale* (Olivetti, 1959). Si vuole solo aggiungere che la memoria a nuclei magnetici (anellini di ferrite, ogni anellino rappresentava un bit), benché di recente introduzione, era già uno standard; l’E-LEA 9003 ne adottava però una versione innovativa, brevetto di Franco Filippazzi, che riduceva alla metà i circuiti di selezione normalmente necessari.

⁶¹ La Olivetti-Bull, una *joint venture* costituita nel 1949, assicurava in Italia la rete commerciale e di assistenza tecnica per i sistemi meccanografici prodotti dalla Compagnie des Machines Bull.



e maneggio di un numero elevato, spesso enorme, di bobine di banda perforata si rivelasse in pratica un'attività critica e non priva di seri inconvenienti, questo sistema – conosciuto come “meccanizzazione integrale” – fu applicato dall'Olivetti sia in ambiente meccanografico e sia avendo al centro elaboratori ELEA 9003 (come nel caso dell'RGS - Ragioneria Generale dello Stato) o Bull Gamma 60 (come al Credito Italiano) dotati nel frattempo di convertitori bandanastro magnetico. La “meccanizzazione integrale” rimase la soluzione più avanzata fino allo spirare del decennio 1960, con l'avvento dei sistemi a rete operanti in tempo reale.

Con l'avvio della produzione industriale cessarono i tempi eroici di Barbaricina e il laboratorio, nel 1958, si trasferì a Borgolombardo (Milano) per essere incorporato in una struttura articolata e complessa come la DEO (*Divisione Elettronica Olivetti*) che, a pieno regime, arriverà a contare circa 2.000 addetti, tra progettazione, stabilimenti e organizzazione commerciale. Una grave deficienza dell'ELEA 9003 è stata la estrema carenza di software, che pesò anche sul gradimento da parte dei clienti e nel confronto con i prodotti della concorrenza⁶². Al 9003 fece subito seguito l'ELEA 6001, una macchina orientata al calcolo scientifico ma prodotta anche, con minime varianti, nella versione “C” (Commerciale). Seguì, più tardi, lo sviluppo di un sistema di fascia media che vide, però, la luce con il marchio OGE (*Olivetti – General Electric*) e più tardi solo GE: era il GE 115, una macchina di grande successo che fu prodotta negli stabilimenti ex Olivetti ed esportata all'estero in migliaia di esemplari. E si arriva così alle dolenti note dell'acquisizio-

ne della DEO da parte della General Electric (1964); una vicenda in cui Roberto Olivetti, venuti a mancare Adriano e Mario Tchou, aveva da tenere a bada due fronti di ostilità già maturati ai tempi di Adriano: quello esterno, “politico” in senso lato, e quello interno dei “meccanici”. Dell'ostilità politica si è già detto in nota. Si aggiunga che l'Olivetti, come azienda privata e, quindi, non facente parte del feudo delle partecipazioni statali, non ebbe alcun sostegno dai vari governi dell'epoca, diversamente da quanto avveniva negli altri Paesi impegnati nella corsa al computer nei quali le varie forme di agevolazione all'industria nazionale erano pratica corrente (protezionismo, contributi alla ricerca, preferenza negli acquisti); fu anzi l'Olivetti a regalare un ELEA 9003 all'RGS, ente statale⁶³. Quanto ai “meccanici”, si ricorda che molta gente dell'Olivetti, affezionata alla tradizione dei prodotti – le cosiddette “galline dalle uova d'oro” – che avevano fatto le fortune dell'azienda, era culturalmente estranea all'avventura elettronica e ne constatava con malcelato rancore il peso finanziario. Questi due nodi vennero al pettine, virulenti e purtroppo vincenti, in concomitanza con la crisi finanziaria che a sua volta era da ricollegarsi al troppo grande esborso per l'acquisto dell'azienda Underwood negli USA e alla caduta dei mercati internazionali; viene spontaneo il confronto con gli esordi del '55 che erano avvenuti in pieno miracolo economico⁶⁴.

Da notare che quelli della General Electric si dimostrarono ferrei nel selezionare i prodotti da mettere in produzione, e se ne dovette accorgere specialmente la Bull con la quale essi conclusero una operazione analoga e concomitante⁶⁵.

⁶² A causa delle difficoltà col software, la multiprogrammazione, che era uno dei punti di forza del 9003, per molti utenti rimase nel libro dei sogni. La lezione fu imparata e, messa sotto pressione la squadra dei ‘softwaristi’ guidata da Mauro Pacelli, le cose andarono meglio col 6001.

⁶³ L'Olivetti elargì poi un ELEA 6001 all'International Computation Centre, la cui sede era ospitata dall'INAC; a causa della negligente parsimonia dei governi che vi partecipavano, l'ICC non poteva sostenere l'onere di un acquisto.

⁶⁴ Per queste vicende si rinvia al saggio, ormai classico, di Lorenzo Soria *Informatica: un'occasione perduta – La Divisione elettronica dell'Olivetti nei primi anni del centrosinistra*; Giulio Einaudi editore, Torino, 1979.

⁶⁵ L'Olivetti aveva in corso una collaborazione con l'INAC-CNR per la realizzazione di un computer molto originale basato su un'architettura *stack* che lo rendeva particolarmente efficiente nell'ospitare programmi compilatori, lasciando con ciò trasparire ‘lo zampino’ di Böhm. (Cfr. (1) Ercoli P.; *Il calcolatore INAC dal punto di vista sistemistico e programmatico*; Calcolo, Vol. 3, n. 4, 1966, p. 441-470; (2) Vittorelli V.; *The Olivetti-INAC computer*; ibid., p. 481-491). Paolo Ercoli, come già ricordato, era dell'INAC mentre Vittorelli lavorava nell'Olivetti; quel computer fu anche conosciuto sotto il nome di CINAC. La Olivetti – General Electric onorò la collaborazione fino al completamento della macchina ma, date le sue caratteristiche atipiche rispetto alla domanda del mercato, non le diede alcun seguito industriale. Si veda anche: Ercoli P.; *From FINAC to CINAC*; in {1}, p. 57-68.

Si può, quindi, ben dire che la messa in produzione del GE 115, oltre ad essere stata un ottimo affare per la General Electric, sia suonata come un esplicito riconoscimento all'eccellenza dei progettisti italiani. E si deve anche riconoscere che il successo del GE 115 sui mercati esteri ben difficilmente si sarebbe realizzato in assenza dei nuovi padroni; il macroscopico *handicap* costituito dall'incapacità (o sarebbe meglio dire impossibilità?) dell'Olivetti di collocare su tali mercati un prodotto, come il computer, radicalmente diverso da quelli tradizionali era stato chiaramente avvertito da Roberto che aveva cercato, infatti, di porvi rimedio. La formula giusta e vincente per il rilancio dell'elettronica Olivetti, quella dell'informatica "leggera", l'avrebbe trovata di lì a poco Pier Giorgio Perotto con la sua straordinaria Programma 101 che, prodotta in 40.000 esemplari, fu un successo mondiale; ma per la DEO i giochi erano ormai conclusi⁶⁶.

6. LA DIFFUSIONE DELL'INFORMATICA

Si è già accennato al fatto che negli anni intercorsi tra gli inizi (anno zero) e gli esiti concreti delle avventure CEP ed ELEA (1960 circa) lo scenario era profondamente mutato.

L'aspetto più vistoso di tale mutazione è stata senza dubbio la diffusione quantitativa degli elaboratori elettronici che vengono illustrati con dati estratti – e qui ridotti all'osso – da una approfondita analisi elaborata nel 1967 da Dadda e collaboratori⁶⁷.

La tabella 1 si commenta da sola: in Italia, dai due elaboratori del '55 si arriva in dieci anni alle soglie del migliaio! Si vuole solo richiamare l'attenzione sull'andamento dell'incremento percentuale annuo che appare molto simile tra Italia ed Europa Occidentale, tanto da far saltare all'occhio le eccezioni del '59 e del '61; mentre è verosimile che la prima sia una fluttuazione statistica su valori assoluti ancora troppo bassi – specialmente quello dell'Italia che è sotto il 100 – quella del '61 è certamente più significativa e si è tentati di metterla in relazione con il momento di maggior diffusione degli elaboratori Olivetti ELEA 9003 e 6001 il cui numero di esemplari venduti in Italia (e solo in Italia!) si stima essere stato rispettivamente di 40 e di 200. Questa possibile correlazione è una traccia indubbiamente tenue ma invita comunque a riflettere su cosa sarebbe accaduto se l'Olivetti, volente o nolente, non si fosse limitata al mercato nazionale.

Per l'analisi che è stata citata non erano disponibili dati relativi alla ripartizione per costruttore del numero di calcolatori installati. Si può, comunque, confermare che, a partire dal 1960 in poi, l'IBM assunse la posizione dominante, lasciandosi progressivamente sfuggire solo la fascia dei super-computer e quella dei mini, allorché questi ultimi comparvero sulla scena; la salace metafora di "Biancaneve e i sette nani" dipingeva fedelmente la si-

Anno	Italia		Europa Occ.	
	Numero	Δ%	Numero	Δ%
1955	2	--	29	--
1956	5	+150	63	+117
1957	10	+100	143	+127
1958	25	+150	336	+135
1959	55	+120	548	+63
1960	90	+63	866	+58
1961	200	+122	1.555	+80
1962	340	+70	2.625	+69
1963	510	+50	4.000	+52
1964	650	+27	5.605	+49
1965	850	+31	6.960	+24

TABELLA 1

Elaboratori elettrici installati

⁶⁶ Perotto ha lasciato un bellissimo ricordo della sua vicenda umana e professionale nel libro *Programma 101 - L'invenzione del personal computer: una storia appassionante mai raccontata* (Sperling & Kupfer, 1995). Il libro è di grande interesse anche perché ripercorre l'intera vicenda dell'elettronica Olivetti. Più in succinto, si veda il suo contributo *Olivetti, dalla P 101 in avanti* pubblicato in {1}, p. 205-218.

⁶⁷ Cfr. Dadda L., Lecchi P.E., Rossi C.; *Il settore dei calcolatori* negli atti del convegno FAST "La ricerca industriale e l'Italia di domani"; Alberto Mondadori Editore, Milano, giugno 1967, Vol. 1, p. 549-571. *Nei ricordi di un informatico*, già citati più volte, Dadda rammenta la difficoltà di reperire i dati quantitativi che occorre per quella relazione – e che nessuno fino allora si era preoccupato di raccogliere in Italia – finché non li rintracciò fortuitamente in un rapporto dell'olandese De Bruijn.

tuazione di fatto. La data del 1960 si riferisce precisamente all'introduzione sul mercato dell'IBM 1401, il primo elaboratore ad essere venduto in molte migliaia di esemplari surclassando così la concorrenza; da sottolineare che, per le sue caratteristiche di costo e di struttura, il 1401 sembrava fatto a posta per rendere appetibile la transizione dalla meccanografia all'elettronica, lasciando pressoché invariati i procedimenti dell'elaborazione sequenziale; se ne parlerà qui nel seguito.

Dalla tabella 2 si ha la conferma di come, dal punto di vista industriale e commerciale, il peso quantitativo degli utenti scientifici fosse divenuto rapidamente minoritario e questo induce a esaminare il fenomeno della transizione dalla meccanografia all'informatica: contrariamente all'opinione corrente, non si è trattato di un momento di frattura bensì di una fase evolutiva nello sviluppo delle tecnologie dell'informazione e specialmente del loro impiego sul campo. Dal punto di vista tecnologico, le macchine meccanografiche avevano progressivamente incorporato dosi crescenti di elettronica (unità di calcolo, lettori fotoelettrici di schede) e si era passati da una collezione di elementi separati (lettore di schede, selezionatrice, tabulatrice ecc.) a un sistema organicamente interconnesso come accadde attorno al 1950 per i vari modelli Gamma 3 della Bull e per il CPC (*Card Programmed electronic Computer*) dell'IBM. E l'elaboratore elettronico non soppiantò repentinamente la meccanografia; tutt'altro! Basti pensare che, in casa IBM, il fatturato globale derivante da computer arrivò a superare quello da sistemi meccanografici solo nel 1962 e la scheda perforata continuò a costituire il principale supporto di *data entry* fino a tutto il decennio 1970.

Ancora più importante, in termini di transizione evolutiva, il fatto che l'elaborazione meccanografica dei dati abbia preparato il terreno, e non solo in Italia, alla rapida diffusione dell'elaborazione elettronica nelle applicazioni gestionali d'interesse delle aziende: non

appena gli elaboratori elettronici furono disponibili a un rapporto costo/prestazioni paragonabile a quello dei sistemi meccanografici, le industrie fornitrici cominciarono a proporre il computer come sostituto fisiologico della meccanografia. Nell'applicare questa tattica di sostituzione evolutiva, le industrie fornitrici che potevano contare su una consistente base di clienti meccanografici sfruttarono appieno l'"effetto fedeltà" ed ebbero così un enorme vantaggio sulle concorrenti *new entry*⁶⁸. Tutto ciò era ben presente tanto alla IBM quanto alla Olivetti che si spartivano il mercato meccanografico italiano, pur con quote nettamente a favore della prima⁶⁹.

Né vanno sottaciute o rimosse le modalità con cui questa transizione coinvolse le migliaia di addetti ai centri meccanografici la cui preparazione di base non raggiungeva di norma il livello universitario e il cui bagaglio professionale si era formato sul campo sotto la guida, non disinteressata, del personale tecnico e commerciale dei fornitori. Situazione questa che si aggravò quando agli stessi addetti vennero affidate mansioni e responsabilità "informatiche" dando luogo alla classica struttura basata sulle figure di: capo centro, analista, programmatore, operatore, perforatrice⁷⁰. L'analisi di Dad-

Categoria	Numero	Percentuale*
1 - Università, istituti scientifici	53	8,8
2 - Enti pubblici governativi e locali	107	17,8
3 - Assicurazioni, enti previdenziali, banche	135	22,6
4 - Imprese	270	45,0
5 - Centri servizio privati	35	5,8
	600	100,0

* Italia, 1965, su un campione del 70%

TABELLA 2

Elaboratori elettronici per categorie di utenti

⁶⁸ In una pubblicità dell'epoca, l'elaboratore ELEA 4-115 della Olivetti - General Electric veniva espressamente raccomandato "per sostituire il centro meccanografico tradizionale, a costi corrispondenti". (Cfr. le inserzioni pubblicitarie apparse su alcuni numeri della rivista *Calcolo delle annate* 1965-66).

⁶⁹ Tra enti pubblici e aziende, nel 1950 esistevano in Italia già 140 centri meccanografici.

⁷⁰ È sintomatico il fatto che solo quest'ultima, la meno qualificata delle categorie, venisse declinata al femminile.

da toccava anche il fabbisogno di personale addetto ai centri elettronici il cui totale, aggregando le categorie ora elencate, si prevedeva che sarebbe arrivato nel 1975 a 150.000 addetti⁷¹. Il sistema formativo pubblico era tutto da inventare e le “isole informatiche” di cui ci si è occupati in precedenza erano di gran lunga insufficienti a dare una risposta quantitativamente adeguata, ancorché ne fossero richieste. I corsi erogati dai fornitori di hardware (e dalle scuole private che fiorivano su livelli qualitativi talvolta discutibili) rimanevano l'unico canale di aggiornamento professionale. Per di più, il retaggio dell'approccio meccanografico faceva inerzia contro lo sfruttamento ottimale delle potenzialità insite nell'elaborazione elet-

tronica: è stato solo con un lento ricambio generazionale che si è passati dalle “procedure” settoriali al concetto di sistema informativo integrato e dal software applicativo “fatto in casa” all'adozione di “pacchetti” e, dove appropriata, all'esternalizzazione.

Molte cose ci sarebbero ancora da dire. Visto però che si è ormai travalicato il limite cronologico del nostro tema principale, si deve considerare conclusa la rivisitazione di una storia così recente eppure così lontana dalla realtà attuale. Importante è rendersi conto che tutto è cominciato da lì, con gente che ha saputo guardare al futuro.

Bibliografia

⁷¹ Su questo argomento, che era di stringente attualità, l'AICA, sotto le presidenze di Luigi Dadda e di Giorgio Sacerdoti, si fece carico di una approfondita ricerca su “la preparazione del personale per l'elaborazione elettronica dei dati in Italia”; ricerca che, iniziata nel 1968, è proseguita con aggiornamenti annuali fino almeno al 1972.

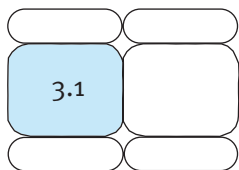
- [1] Bonfanti C. (a cura di): *Convegno internazionale sulla storia e preistoria del calcolo automatico e dell'informatica*. Siena, 10-11-12 settembre 1991; AICA, Milano, 1991.
- [2] Cuzzer A. (a cura di): *La cultura informatica in Italia; riflessioni e testimonianze sulle origini 1950-1970*. Bollati-Boringhieri, Torino, 1993.

CORRADO BONFANTI è responsabile del progetto AICA “Storia dell'informatica”. Nato a Tripoli (Libia) nel 1940, laureato in fisica, ha lavorato con IBM Italia e poi nel gruppo Finsiel: con Italsiel a Roma, con Insiel a Trieste e infine a Bucarest come direttore generale di Finsiel-Romania. Ha ricoperto incarichi universitari a Roma, Trieste e Bari. Da vent'anni si occupa di storia del calcolo automatico e dell'informatica.
corrado-bonfanti@hotmail.com



AUTONICA L'ELETTRONICA SULL'AUTO

Gianfranco Burzio



Autonica è un neologismo nato dall'unione dei termini Automobile ed Elettronica e indica l'applicazione dell'elettronica sull'auto: un impiego abbastanza recente rispetto allo sviluppo della tecnologia elettronica stessa. La maturazione della tecnologia, soprattutto con la microelettronica, ha consentito di raggiungere i requisiti richiesti con costi accettabili e ora l'elettronica sta rivoluzionando molti componenti dell'auto. Un processo ancora in forte sviluppo che porterà ad avere veicoli nel prossimo futuro con miglior comfort, efficienza e sicurezza.

1. INTRODUZIONE

L'automobile è indubbiamente una delle innovazioni tecnologiche che hanno cambiato il mondo. Nonostante abbia oltre un secolo, molte case automobilistiche hanno festeggiato questo traguardo negli ultimi anni (la FIAT ad esempio, nel 1999), è tutt'altro che una tecnologia matura. Anche il mercato, a dispetto delle apparenze, ha ancora enormi opportunità di crescita.

L'auto, infatti, è un bene del quale può godere oggi soltanto il 12% della popolazione mondiale. Il mercato dell'auto continua a crescere al ritmo del 2% medio annuo, per raggiungere, si prevede, una dimensione tra i 60 e 70 Milioni di veicoli prodotti annualmente nel 2010. Questo se non avvengono improvvise accelerazioni del mercato in Paesi come la Cina che potrebbero stravolgere tutte le previsioni.

Ma insieme al mercato sta crescendo, in modo molto più rapido, la competitività tra le aziende automobilistiche. Competitività che si gioca sui prezzi, sulla varietà di offerta di modelli e sulla disponibilità di nuovi contenuti di prodotto. Tutti questi fattori aumenta-

no i costi e riducono fortemente i margini di profitto.

Per rispondere alle diverse esigenze, non solo del mercato ma anche delle normative, tutte le case automobilistiche investono nello sviluppo di nuove tecnologie, per la massima parte basate sull'elettronica. La maggior parte delle innovazioni introdotte sull'auto si basano sull'elettronica.

Attualmente la percentuale di componentiistica elettronica nel valore dell'auto è intorno al 25%. Dieci anni fa era pari al 18% e si ritiene arriverà al 35% nel 2010 (fonte Bosch).

2. L'ELETTRONICA NELL'EVOLUZIONE DELL'AUTO

Dalla sua nascita, negli ultimi anni del XIX secolo, l'automobile si è evoluta per molti decenni soprattutto nel propulsore, in particolare con l'introduzione dei motori a elevato rapporto di compressione, resi possibili anche grazie alla evoluzione dei combustibili. La successiva, grande rivoluzione si è avuta con la introduzione dei microprocessori. Gli

ultimi modelli ormai dispongono di decine di centraline elettroniche di controllo (ECU, *Electronic Control Unit*) che controllano praticamente tutti i dispositivi (motore, freni, climatizzazione ecc.).

Centraline elettroniche controllano le serrature che si aprono con un telecomando, il tergicristallo che adatta la sua velocità alla intensità della pioggia, le luci che si accendono automaticamente entrando in un tunnel. Ogni dispositivo assume una propria intelligenza, rendendo automatiche molte operazioni.

Altre centraline si occupano della sicurezza. Ci permettono di avere un migliore controllo del veicolo, in condizioni di bassa aderenza (ABS, VDC) e, se proprio l'incidente accade, cercano di ridurne gli effetti azionando una serie di *airbag* intorno agli occupanti il veicolo, dopo aver teso le cinture di sicurezza per trattenerci perfettamente.

Inoltre, si preoccupano del nostro comfort, controllando il clima interno all'abitacolo, facendoci ascoltare la radio, indicandoci il percorso da seguire verso la destinazione.

La presenza di intelligenza su tutti i vari dispositivi consente di realizzare funzioni complesse, grazie alla interazione tra i vari dispositivi. Si ottengono così funzioni semplici ma utili, ad esempio l'attivazione automatica del tergicristallo, quando viene inserita la retromarcia e i tergicristalli sono attivi, oppure molto complesse, come il controllo della dinamica del veicolo (**VDC, Vehicle Dynamic Control**, oppure ESC, *Electronic Stability Control*) ottenuta mediante il controllo indipendente della frenata sulle quattro ruote e del motore.

Ottenere queste funzioni richiede che tutti i vari dispositivi siano tra loro interconnessi. Se su un veicolo moderno ogni connessione necessaria fosse realizzata, il numero di connessioni, e quindi di fili, sul veicolo sarebbe enorme. Fortunatamente l'elettronica ha fornito una soluzione anche a questo problema.

3. LA RIVOLUZIONE NELLE CONNESSIONI IL MULTIPLEX

Nel primo dopoguerra le vetture di gamma elevata potevano avere circa 70 collegamenti elettrici. Oggi ne troviamo anche 1600. Questo significa diversi chilometri di cavi.

Sarebbero addirittura moltissimi di più se negli ultimi anni non si fosse introdotto l'impiego di connessioni monofilari, che sostituiscono con un'unica coppia di fili moltissimi collegamenti.

Se al guidatore il veicolo appare sempre più dotato di nuove e interessanti funzioni, questo avviene anche grazie a questa rivoluzione, invisibile ai suoi occhi ma di enorme importanza: le connessioni tra i vari componenti non sono più multifilari (Figura 1), ovvero un filo per ogni segnale o comando, ma monofilari o multiplexate: su una coppia di fili viaggiano più informazioni o comandi (Figura 2 e 3). Come si vede dalla figura esistono due versioni del sistema monofilare:

■ le connessioni tra due centraline o componenti vengono sostituite da una singola

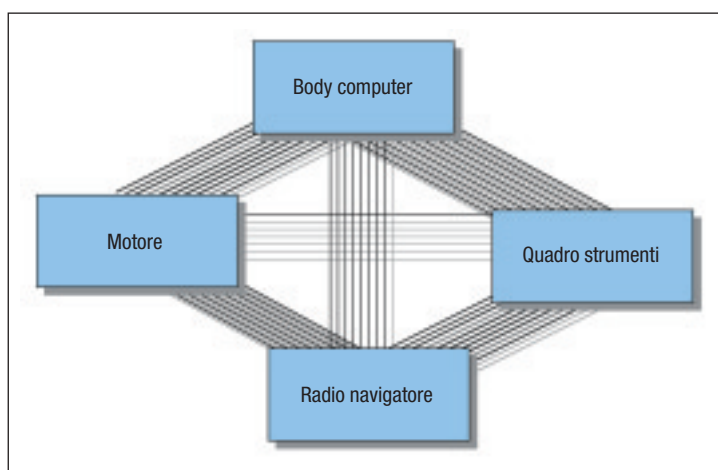


FIGURA 1
Schema Multifilare

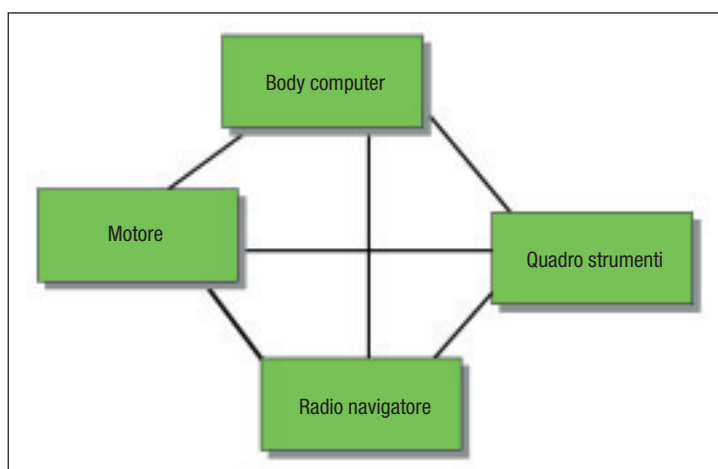


FIGURA 2
Schema Multiplex

Il sistema **VDC – Vehicle Dynamic Control** è uno dei più recenti, come ingresso sul mercato, sistemi di sicurezza attiva. È noto anche con il nome **ESC, Electronic Stability Control**. Questo sistema è diventato famoso alcuni anni fa, come soluzione per risolvere la tendenza a ribaltarsi da parte di una vettura, prossima a essere lanciata sul mercato, di una delle più prestigiose case automobilistiche del mondo.

La curva più in basso nella figura A mostra la traiettoria che il veicolo seguirà se la sua accelerazione laterale è compatibile con le condizioni di attrito tra i pneumatici e la strada; in questo caso, il mezzo segue fedelmente il moto desiderato dall'automobilista.

Se la strada è scivolosa, con un coefficiente di attrito basso, tale cioè da non consentire lo sviluppo della necessaria forza laterale, il raggio di curvatura della traiettoria aumenta, allontanandosi da quello che si avrebbe in condizioni di massima aderenza, fino ad arrivare a condizioni di instabilità, come mostra la curva più a sinistra.

Per controllare la traiettoria occorre in questo caso armonizzare le azioni laterali e longitudinali sulle quattro ruote in modo da non superare per nessuna di esse la massima forza di interazione possibile tra pneumatico e strada; a costo naturalmente di un aumento del raggio di curvatura della traiettoria del veicolo (curva intermedia).

L'angolo tra l'asse longitudinale del veicolo e un sistema di riferimento solidale alla strada è detto angolo di imbardata, la sua variazione (velocità di imbardata) misura la variazione di direzione che il veicolo sta compiendo. L'angolo che, invece, l'asse longitudinale crea con la traiettoria del veicolo viene invece detto "angolo di deriva". Le condizioni di aderenza stradale portano a valori limite di tali grandezze fisiche.

L'azione del VDC consiste nel controllare la velocità di imbardata e l'angolo di deriva del veicolo β , in modo da non superare tali limiti, ottenendo quindi la traiettoria centrale descritta in figura; come si può notare questa traiettoria si discosta da quella desiderata ma consente di mantenere il controllo direzionale del veicolo.

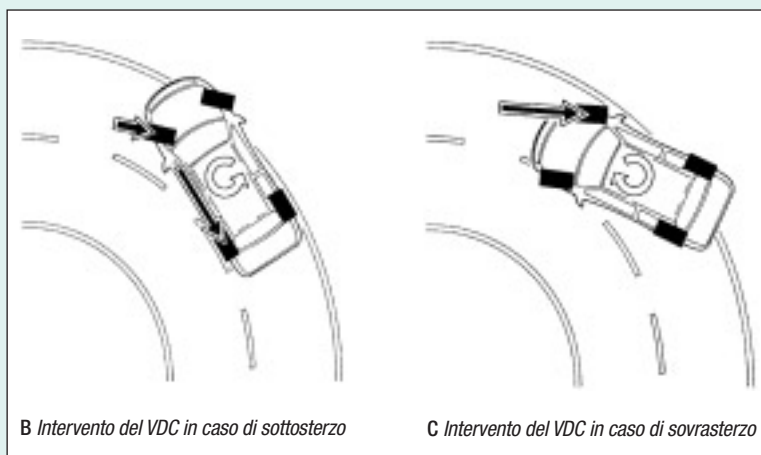
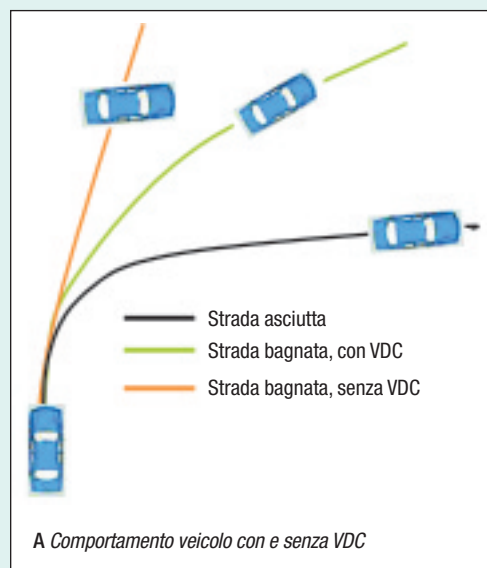
È da sottolineare però che il VDC interviene per correggere eventuali comportamenti sottosterzanti o sovrasterzanti della vettura anche in condizioni non critiche, prevenendo manovre da parte del guidatore che possono essere di difficile gestione. Il VDC interviene in modo diverso nei seguenti due casi (l'esempio si riferisce a una vettura con trazione anteriore).

Nel caso di comportamento sottosterzante dell'autoveicolo: la centralina verifica la presenza del sottosterzo (prevalenza della deriva sull'assale anteriore) e corregge il comportamento della vettura, frenando le ruote interne alla curva ed eventualmente riducendo la coppia motrice.

Il comportamento sottosterzante si ha quando il veicolo tende a curvare di meno rispetto all'angolo sterzo, sembra che la vettura "tiri dritto" (Figura B).

Nel caso del comportamento sovrasterzante: la centralina, in presenza di condizioni di sovrasterzo (prevalenza della deriva sull'assale posteriore), corregge il comportamento della vettura, frenando la ruota anteriore esterna alla curva al fine di creare un momento d'imbardata opposto eventualmente incrementato da un aumento della coppia motrice.

Il comportamento sovrasterzante si ha quando il veicolo tende a "sbandare", il posteriore sfugge all'esterno (Figura C).



coppia di fili, esiste una singola connessione per ogni coppia di centraline connesse tra di loro (multiplex con linee seriali punto-punto, Figura 2);

■ un unico cavo bifilare connette tra di loro tutte le centraline (Bus, Figura 3).

Questo cambiamento apparentemente semplice ha, invece, richiesto molti anni per essere diffuso. I motivi sono diversi.

1. Un sistema multifilare è apparentemente più affidabile: la rottura di un filo danneggia soltanto quella particolare informazione o comando. In un sistema monofilo la rottura potrebbe compromettere molte funzioni.

Questa maggiore affidabilità è soltanto apparente perché la probabilità di rottura di un filo è molto bassa, in confronto alla probabilità di guasti dei vari componenti e sensori. In que-

sto senso il sistema monofilare, più flessibile, consente, attraverso una diversa distribuzione dei segnali/comandi di recuperare eventuali guasti non gravi. Migliora, poi, la diagnostica e la tolleranza al funzionamento in ambienti elettromagneticamente perturbati.

2. Un sistema multifilare è meno costoso.

Inizialmente questo era vero, per l'elevato costo dei componenti rispetto al risparmio in termini di connessioni e cavi. Ora i componenti, ormai prodotti in volumi di milioni di pezzi, hanno costi molto ridotti. Più recentemente, per venire incontro a tali esigenze di costo limitato, sono state concepite e realizzate versioni molto semplici di bus (LIN) dove il costo dei componenti è estremamente modesto. Inoltre, con soluzioni unifilari è più semplice condividere le informazioni tra diverse centraline, evitando di dover disporre uno stesso sensore in più esemplari in diverse parti del veicolo.

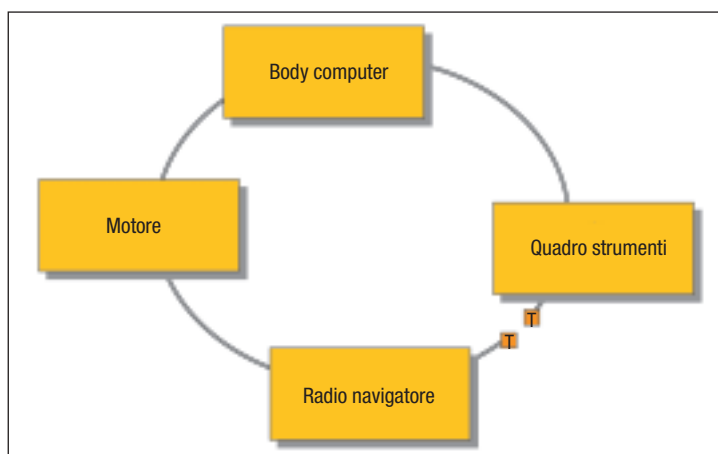
3. Un sistema monofilare è più complesso, maggiori costi di progettazione e sviluppo.

Con la standardizzazione delle soluzioni questo aspetto si è ribaltato, è molto più semplice mettere a punto un sistema a bus. Infatti, diventa possibile "simulare" un sistema di dispositivi molto complesso, semplicemente riproducendo il suo comportamento sul bus, e in questo modo mettere a punto, ad esempio, una centralina di controllo.

Molte delle difficoltà a introdurre un sistema a bus o multiplexato, risiedevano nel fatto che occorre definire uno standard unico, in modo da aumentare i volumi dei componenti e degli strumenti di sviluppo, riducendone, quindi, i costi e aumentandone le prestazioni. Due standard sono stati proposti: il CAN (*Controlled Area Network*) e il VAN (*Vehicle Area Network*). La maggiore forza del gruppo industriale che sosteneva il CAN (in prima linea Bosch, che lo ha concepito nel 1980) ha portato questo a prevalere rispetto al secondo, ormai usato solamente in alcuni veicoli francesi. Ora il **Bus CAN** è ampiamente diffuso su tutti i nuovi modelli di veicolo.

Il bus CAN è estremamente semplice e per questo molto affidabile e sicuro. Ma ha un limite importante, che ne pregiudica l'impiego in sistemi dove l'affidabilità deve essere assoluta.

Solamente il messaggio che ha la massima priorità ha la garanzia di poter essere tra-



smesso entro un determinato tempo (quello necessario affinché il messaggio in corso termini). Tutti i messaggi a priorità inferiore possono subire ritardi, non definibili a priori nel loro massimo valore.

Per evitare questo occorre una attenta progettazione dell'intero sistema di dispositivi connessi a un singolo bus, eventualmente predisponendo una centralina a compiti di diagnosi e di controllo che non avvengano situazioni che, di fatto, impediscono a una centralina che gestisce messaggi di priorità inferiore di riuscire a lavorare correttamente.

Ma diventa sempre più difficile e costoso eseguire questa progettazione, per la necessità di avere sempre più varianti di un modello di veicolo e di poter aggiungere nuovi componenti a un modello già sviluppato. Sta emergendo forte l'esigenza di connessioni che non abbiano, per loro natura, problemi nel caso si renda necessario rimuovere o aggiungere una centralina o aggiungere nuove funzioni a quelle esistenti.

Per questo si stanno proponendo nuove versioni del bus CAN, che ne migliorano le prestazioni ma, soprattutto, garantiscono a tutti i messaggi dei tempi di latenza (tempo tra la decisione di spedire un messaggio e la ricezione dello stesso dalla centralina destinataria) definiti e limitati nei loro massimi valori. Applicazioni molto critiche dal punto della sicurezza, come per esempio lo *steer-by-wire*, sistemi di sterzata dove non esiste un collegamento meccanico tra volante e ruote, ma un collegamento via bus tra la centralina elettronica che legge la posizione del volante e la centralina che comanda la sterzata del-

FIGURA 3

Schema a bus

Il **Bus CAN** è di tipo Multi-master, tutti i componenti connessi al bus (Nodi) sono equivalenti tra di loro e possono decidere autonomamente quando inviare un messaggio. Tutti i messaggi sono ricevuti da tutti i nodi presenti sul bus. I messaggi contengono un identificativo, ma non hanno nessuna informazione né del nodo trasmittente e neanche del nodo/i destinatario del messaggio. Sulla base del codice identificativo, i vari nodi stabiliscono se un dato messaggio è di loro interesse o meno (multi-cast). Ogni singolo messaggio (o meglio frame) è composto secondo la tabella.

STRUTTURA DEL MESSAGGIO CAN

Nome	N. bit	Valore	Note
SOF Start of Frame	1	1	Indica l'inizio del messaggio, permette la sincronizzazione dei vari dispositivi in lettura
Message ID	11	ID	Identificativo del messaggio
Request of Transmission	1	----	Vale 1 per un messaggio inviato, 0 per la risposta
Controllo-IDE	1	----	Messaggio standard (=0) o esteso (=1)
Controllo	1	----	Libero per futuri utilizzi
Ctr-lunghezza	4	----	Lunghezza messaggio successivo in byte (0-8)
Dati	0-64	Messaggio	I byte non utilizzati possono anche essere inseriti, con valore 0
CRC	15	----	Campo controllo correttezza dati messaggio
Delimitatore	1	0	
ACK-DEL	2	----	Il nodo trasmittente li imposta a zero, il ricevente pone il primo bit a 1 per indicare al nodo trasmittente di aver ricevuto correttamente il messaggio
EOF	7	0	Chiude il frame, i nodi riceventi possono utilizzarlo per segnalare problemi sulla ricezione
IFS	3	0	Garantisce un tempo minimo tra frame successivi
Idle bus	----	0	Indica che il bus è libero

Non essendo previsto un master che coordina e sincronizza l'invio di messaggi da parte dei vari nodi occorre un meccanismo di "arbitraggio", onde risolvere le situazioni di più centraline che contemporaneamente si attivano per inviare un messaggio.

Il meccanismo del CAN è di tipo CSMA/CA, *Carrier Sense Multiple Access /Collision Avoidance*.

Il Bus si trova in uno stato "ideale" quando nessuno sta trasmettendo. Ogni nodo può quindi capire (Carrier Sense) se il bus è disponibile o meno. Se lo è può iniziare immediatamente a trasmettere un messaggio. Se due o più nodi si attivano contemporaneamente succede una collisione, della quale non è possibile accorgersi immediatamente. La collisione viene risolta dal codice identificativo.

Infatti il bus ha due stati logici, alto-basso, dei quali uno (basso) è dominante. Se, quindi, due nodi pilotano il bus verso stati diversi, il bus si porta sullo stato basso. Se un nodo si accorge di non riuscire a pilotare il bus nello stato desiderato, interrompe la trasmissione e si pone in ricezione. In questo modo la collisione viene evitata.

Quindi riuscirà a trasmettere il nodo che sta trasmettendo un messaggio a maggiore priorità e la priorità è stabilita dal codice identificativo. In fase di progettazione occorre quindi distribuire attentamente i codici identificativi dei messaggi, in modo che le priorità siano correttamente attribuite.

Normalmente in una vettura esistono diverse reti CAN, in funzione dei diversi sottosistemi e delle velocità di trasmissione richieste. Esistono, infatti, tre diversi livelli:

- A – 10 kbit/s o inferiori
- B – da 10 a 125 kbit/s
- C – da 125 kbit/s a 1 Mbit/s

I livelli inferiori (A, B) sono utilizzati per sottosistemi non critici, ad esempio per la diagnostica, climatizzazione, comando dispositivi secondari ecc.. Il Can C viene tipicamente utilizzato dalla centralina controllo motore, freni e ABS, cambio ecc..

le ruote, saranno solo possibili, per l'affidabilità che richiedono, con queste nuove generazioni di bus.

Il Flex-Ray, standard derivato dal Can, proposto da una serie di case automobilistiche e di industrie di elettronica e microelettronica, è il principale candidato a diventare la soluzione per questa prossima generazione di bus. Questo tipo di bus si caratterizza per la sua estrema flessibilità, ideale quindi per poter gestire un futuro di modelli sempre più variegati nelle loro configurazioni. Può, quindi, passare da configurazioni con caratteristiche simili all'attuale CAN, ma comunque superiori in termini di velocità di trasmissione e robustezza, sino a configurazioni con elevata affidabilità e *fault tolerant*, cioè tolleranti ai guasti.

Nel Flex-Ray l'arbitraggio tra i diversi nodi usa un meccanismo di TDMA, *Time Division Multiple Access*. I vari nodi si dividono il bus in tempi diversi, secondo regole stabilite a priori. Ogni nodo ha, quindi, la garanzia che, entro un certo tempo, avrà la possibilità di inviare il proprio messaggio.

4. INNOVAZIONE TECNOLOGICA DEL VEICOLO

Il trend attuale vede l'innovazione tecnologica dell'auto impegnata soprattutto su tre fronti:

1. sviluppo di motorizzazioni che, senza rinunciare alle prestazioni, offrano consumi più contenuti e un minore impatto ambientale;
2. miglioramento della sicurezza di guida;
3. miglioramento del comfort e della qualità della vita a bordo del veicolo.

Per ragioni di spazio non si approfondisce il primo dei tre punti, nel quale, anche se in modo meno "visibile" per l'utente, l'Autonica ha consentito enormi sviluppi. Un esempio è il motore diesel: è un motore tradizionalmente "meccanico", non avrebbe neppure bisogno di elettricità alcuna per funzionare a regime, neanche le candele del motore a benzina. Grazie alla elettronica abbinata a sistemi di iniezione *common rail*, si è trasformato. Era un motore affidabile ma di scarse prestazioni, rumoroso e inquinante. Ora:

■ le prestazioni sono equivalenti se non superiori a quelle dei motori a benzina, la soglia dei 100 CV/litro è ormai prossima (ci sono motori sul mercato con potenze intorno agli

80 CV/litro); l'erogazione della potenza è però molto più fluida nel motore diesel, che ha coppie molto superiori ai bassi regimi;

■ le emissioni sono inferiori, fatta eccezione per le PM₁₀, per le quali sono comunque possibili soluzioni, con filtri particolari;

■ il rumore è contenuto, è sempre più difficile riconoscere un diesel da un benzina.

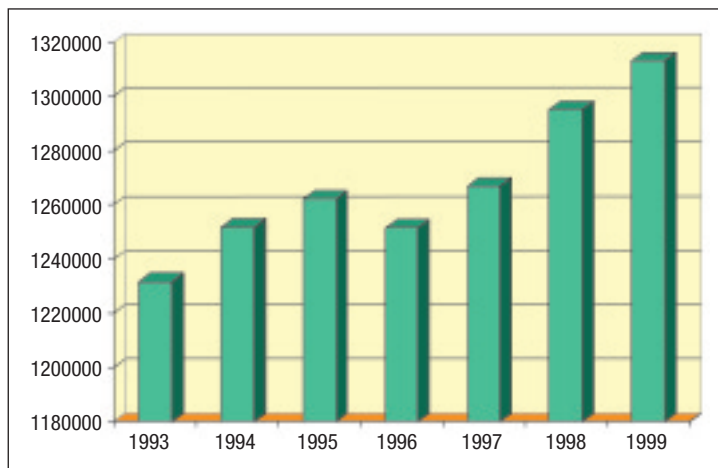
Questo sta portando il diesel a raggiungere, in Italia, il 50% del mercato, anche se la differenza di costo tra i combustibili diesel e benzina è ormai molto limitata e un veicolo diesel ha, salvo eccezioni dovute a politiche commerciali particolari, un costo superiore a un equivalente motore a benzina.

Per realizzare tali prestazioni i motori diesel più moderni del gruppo FIAT utilizzano la tecnologia Multijet, dove il combustibile viene iniettato nella camera di combustione mediante una serie di getti successivi, modulati opportunamente in funzione del regime del motore e la richiesta di potenza. Questi sofisticati controlli sono realizzati da centraline elettroniche dotate di notevole potenza di calcolo, diverse centinaia di MIPS, e sistemi operativi *real-time* appositamente realizzati o adattati.

5. LA SICUREZZA

Negli ultimi dieci anni molta ricerca è stata dedicata alla soluzione di uno dei maggiori problemi della circolazione su strada: la sicurezza. La sicurezza stradale è notevolmente aumentata negli ultimi anni, gli incidenti stradali continuano ad aumentare con l'aumentare del traffico (Figura 4), mentre le conseguenze degli incidenti, in particolare i morti,

FIGURA 4
Numero incidenti
stradali in Europa



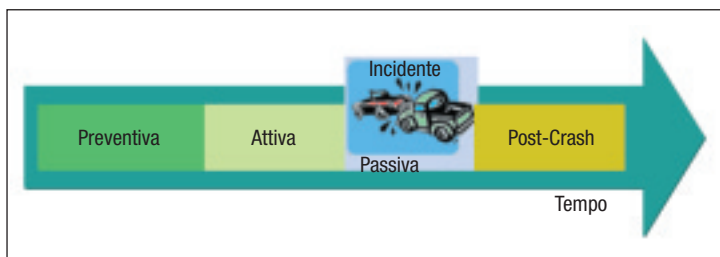


FIGURA 5
Tipi di sicurezza

diminuiscono. Comunque la cifra di circa 6000 morti all'anno in Italia rimane un tributo alla mobilità individuale ancora molto, troppo, pesante.

Negli incidenti stradali il fattore umano è la causa principale, oltre il 90% è causato da un non corretto comportamento di guida. Le case automobilistiche, a partire dalla fine degli anni '80 hanno iniziato a collaborare, nel programma di ricerca Europeo PROMETHEUS, al fine di migliorare la sicurezza stradale attraverso lo sviluppo di sistemi di sicurezza "preventiva" che, grazie a sensori che riescono a percepire il traffico circostante, aiutano il guidatore a evitare situazioni pericolose.

Nel migliorare la sicurezza è possibile intervenire su diversi aspetti (Figura 5):

- la sicurezza passiva: ridurre le conseguenze di un incidente;
- la sicurezza attiva: permettere al guidatore il massimo controllo del veicolo, al fine di evitare l'incidente o ridurne le conseguenze;
- la sicurezza preventiva: evitare che si creino le premesse per un incidente;
- l'intervento *post-crash*: ridurre le conseguenze di un incidente attraverso un tempestivo intervento dei mezzi di soccorso.

I miglioramenti di questi ultimi decenni, in termini di sicurezza stradale sono stati ottenuti soprattutto sulla sicurezza passiva e attiva. Esistono norme precise che fissano i requisiti minimi da rispettare per un veicolo, su urti standard sia frontali che laterali. Tranne alcuni veicoli di gamma bassa, ormai tutti i veicoli dispongono, di serie, di airbag guidatore e ABS. I trend futuri prevedono di intervenire su tutti i vari aspetti della sicurezza.

5.1. Sicurezza passiva

La tendenza futura sarà nel modulare l'intervento dei sistemi di sicurezza passiva, cinture e airbag, in funzione dell'incidente e degli occupanti il veicolo. L'elettronica dovrà fornir

re soluzioni affidabili ai due problemi critici:

■ rilevare con assoluta sicurezza un incipiente incidente, con un anticipo tale da permettere un intervento di tipo meccanico, non pirotecnico, dei sistemi di ritenuta (sensori pre-crash);

■ rilevare la presenza, posizione e caratteristiche degli occupanti, per modulare correttamente l'intervento dei sistemi airbag.

Aumenterà, inoltre, l'attenzione agli incidenti che coinvolgono gli utenti della strada più vulnerabili: pedoni e ciclisti. La progettazione dei veicoli, soprattutto nella parte frontale, dovrà in futuro tenere conto anche di questi aspetti, per superare apposite normative in fase di definizione. Su questo aspetto l'auto-nica sarà utile nello sviluppo di soluzioni attive: per esempio, cofani motore che si sollevano leggermente in caso di urto con un pedone, per attutire un probabile successivo urto della testa del pedone sul cofano stesso.

5.2. Sicurezza attiva

Dopo l'ABS, sistema che impedisce alle ruote di bloccarsi in frenata e, quindi, permette di mantenere il controllo della direzione del veicolo anche in situazioni di frenata di emergenza, si stanno diffondendo i sistemi di controllo della dinamica del veicolo.

Un guidatore non esperto, di fronte a un ostacolo improvviso, sterza bruscamente con una manovra che spesso può portare a perdere il controllo del veicolo (sbandata, testa coda ecc.) con conseguenze imprevedibili. Grazie a una azione di frenatura sulle singole ruote, una centralina elettronica permette di aumentare la capacità del veicolo di modificare la propria traiettoria, senza sbandare. Il guidatore si trova, quindi, a disporre di una capacità di controllo del veicolo che anche un pilota esperto, su un veicolo normale, non avrebbe.

La ricerca si sta ora focalizzando sulla tecnologia *Drive-by-wire*: eliminando ogni collegamento meccanico tra gli organi di controllo della dinamica del veicolo (ruote, sterzo, sospensioni, motore, cambio, freno e acceleratore) e i comandi (pedale freno e acceleratore, volante) sarà possibile realizzare veicoli con 4 ruote indipendenti, sterzanti e frenanti, sotto il completo controllo di una centralina elettronica.

Il guidatore esprimerà la sua volontà attraverso i comandi, che saranno eseguiti dal vei-



colo sfruttando al meglio la dinamica possibile: se occorre sterzare a destra per andare a sinistra (e in condizioni di bassa aderenza questo può accadere) il veicolo lo farà senza che il guidatore nemmeno se ne accorga.

Il livello di affidabilità delle centraline preposte al controllo di tali funzioni dovrà essere ovviamente assoluto. Nei sistemi attuali, ABS e VDC, è possibile, qualora si diagnostichi un malfunzionamento, anche minimo, disabilitare il controllo: si ritorna ad avere un veicolo normale, ma pur sempre in grado di frenare e sterzare. Nei futuri sistemi Drive-by-Wire questo non sarà più possibile: i sistemi dovranno avere una affidabilità assoluta, o quantomeno consentire un utilizzo degradato ma pur

sempre funzionale, che non metta a rischio di incidente gli occupanti del veicolo.

5.3. Sicurezza preventiva

Le prime applicazioni di questi sistemi sono nel controllo automatico della velocità. Grazie a radar a microonde o laser, il veicolo riesce a percepire la presenza di un veicolo più lento nella corsia di marcia e rallentare in modo da adeguarsi alla velocità del veicolo che precede. Questa funzione (detta **Cruise Control Adattativo**) è già presente sul mercato da alcuni anni: per esempio è disponibile come optional su FIAT Stilo e Lancia Thesis.

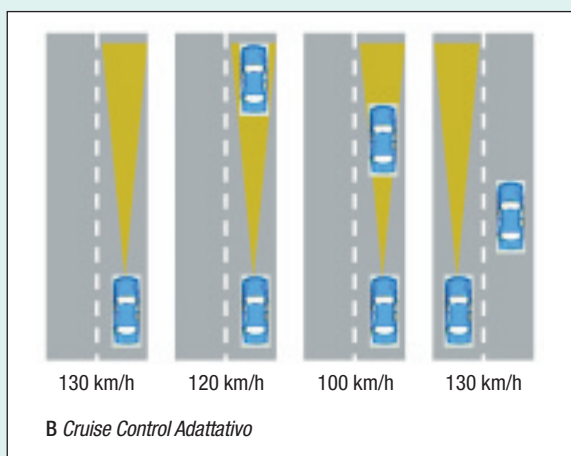
Nel controllo laterale del veicolo sistemi sensoriali basati su telecamere possono aiutare

Durante i lunghi viaggi in autostrada, con poco traffico, controllare la velocità del veicolo, mantenendola al valore desiderato, diventa un compito noioso per il guidatore, oltre che stancante. Per questo molti veicoli dispongono del sistema "Cruise Control", che provvede a mantenere la velocità del veicolo al valore impostato dal guidatore. Quando il veicolo ha raggiunto la velocità desiderata il guidatore attiva il sistema e tale velocità viene mantenuta automaticamente. Il pedale acceleratore può essere rilasciato. Nel caso il guidatore desideri temporaneamente aumentare la velocità lo può fare liberamente azionando il pedale dell'acceleratore. Quando il guidatore lo rilascia, il veicolo torna automaticamente alla velocità precedentemente impostata. Nel caso di azionamento sul freno il sistema si disinserisce. Ma il guidatore dispone di un comando "Resume" per riattivarlo e il veicolo accelera automaticamente per raggiungere nuovamente la velocità precedentemente impostata. Questo sistema è molto comodo ma soltanto in condizioni di traffico modesto, oppure molto fluido e con veicoli che hanno tutti quanti velocità molto simili. Queste condizioni sono comunissime, ad esempio, negli Stati Uniti, dove infatti il "Cruise Control" è ormai di serie su quasi tutti i veicoli. In Europa, dove tali condizioni sono molto più infrequenti, il "Cruise Control" non ha avuto molto successo di mercato. È disponibile di serie o come *optional* solo sui veicoli di alta gamma.

In condizioni di traffico intenso o con veicoli che viaggiano a velocità differenti il "Cruise Control" perde i suoi vantaggi: il guidatore si trova spesso a dover frenare, per poi riattivare il sistema. Con il **Cruise Control Adattativo** questo non accade più: grazie a un sensore radar (Figura A), che percepisce la presenza di un eventuale veicolo più lento nella stessa corsia di marcia, il dispositivo adatta la velocità al veicolo che precede, mantenendo una adeguata distanza di sicurezza.



A Sensore radar per Cruise Control Adattativo (Bosch)



Quando il veicolo che precede cambia corsia, oppure il guidatore si sposta per sorpassarlo, il sistema rileva che la strada è nuovamente libera e accelera automaticamente per tornare alla velocità impostata. In assenza di veicoli nella corsia di marcia il dispositivo si comporta esattamente come un "Cruise Control" normale (Figura B). Con il Cruise Control Adattativo è quindi possibile viaggiare con una velocità di crociera impostata automaticamente anche nelle nostre affollate autostrade; solamente se la velocità scende al di sotto di un predeterminato valore (tipicamente 50 km/h) il sistema si disinserisce. Recentemente sono stati introdotti sul mercato anche sistemi in grado di funzionare a bassa velocità e gestire, quindi, anche la marcia in coda (Stop&Go), a singhiozzo.

In questo caso, il veicolo si arresta e riparte seguendo il comportamento del veicolo che precede. Questo dispositivo ha sollevato un curioso problema: occorre gestire la situazione dove, tra un arresto e la successiva ripartenza, il guidatore ha deciso di scendere dal veicolo! Il veicolo potrebbe ripartire automaticamente, senza controllo. Per evitare questo viene disposto un sensore di presenza del guidatore, oppure ogni riavvio viene "autorizzato" dal guidatore attraverso un pulsante al volante.

il guidatore a mantenere la corsia di marcia ed evitare uscite di corsia. Una segnalazione, acustica o tattile (per esempio vibrazione al volante) viene generata quando il sistema rileva una imminente, non segnalata, uscita dalla corsia di marcia. Irisbus, azienda italo-francese produttrice di autobus, fornisce un sistema che aiuta in questo modo l'autista a mantenere il mezzo nella corsia di marcia e, inoltre, ad accostare con la massima precisione il mezzo durante le fermate.

Sono allo studio sistemi anche per il controllo automatico dello sterzo (mantenimento automatico della corsia, Lane Keeping). I costruttori automobilistici sono molto prudenti su questa funzione, che potrebbe essere interpretata non come un ausilio ma come una vera e propria guida automatica del veicolo. Quindi questa funzione non potrà probabilmente essere introdotta sul mercato finché non sarà possibile avere una assoluta affidabilità sia nel riconoscimento della corsia di marcia che nel rilievo ostacoli eventualmente presenti sulla stessa.

Altri esempi di funzioni di ausilio alla guida disponibili sul mercato o in fase avanzata di sviluppo sono i seguenti.

❑ **Visione notturna.** Telecamere operanti nell'infrarosso consentono al guidatore di avere una migliore percezione in condizioni di bassa visibilità, di notte e con la nebbia.

❑ **Copertura angolo cieco.** Gli specchietti retrovisori laterali soffrono dell'angolo cieco, area laterale non visibile da parte del guidatore, se non girando la testa. Mediante telecamere e centraline elettroniche di elaborazione immagine è possibile avvisare il guidatore della presenza di un veicolo in fase di sorpasso.

❑ **Ausilio nelle manovre di parcheggio.** Sensori di parcheggio sono già ampiamente diffusi su molti veicoli. In futuro, una centralina elettronica rileverà lo spazio tra due veicoli e, se sufficiente, ci aiuterà nella manovra controllando lo sterzo.

5.4. Sicurezza post-crash

A incidente avvenuto diventa di importanza vitale che l'informazione sia tempestivamente trasmessa alle centrali di soccorso, per predisporre un tempestivo intervento di ambulanze, vigili del fuoco e polizia. Centraline

a bordo veicolo, protette e dotate di alimentazione autonoma, possono eseguire in automatico tale chiamata, tramite un telefono cellulare GSM.

Esempi di questo servizio già avvengono, attraverso centrali di soccorso private e per gruppi specifici di utenti con esigenze particolari, per esempio trasporto valori. Affinché questo servizio si diffonda occorre però che siano definiti degli standard, a livello Europeo, sulla modalità con cui la chiamata viene eseguita.

Esistono in questo senso iniziative a livello Europeo di standardizzazione. In particolare si dovrebbe confluire su un unico numero per le chiamate di emergenza: il 112.

6. IL MIGLIORAMENTO DELLA QUALITÀ DELLA VITA E DEL COMFORT DI VIAGGIO

Su questo aspetto sono molte le innovazioni introdotte sul veicolo negli ultimi decenni. Per esempio il climatizzatore è passato da accessorio per le auto di lusso a dispositivo di serie anche su veicoli di gamma medio-bassa.

Ma ci soffermeremo soprattutto su un aspetto che promette di modificare sostanzialmente il modo nel quale ci muoveremo in futuro, la telematica.

La disponibilità di sistemi di comunicazione in grado di funzionare su un veicolo in movimento, assieme alla possibilità di conoscere con precisione la posizione del veicolo, sono due tecnologie che stanno consentendo l'introduzione di moltissime funzioni a supporto del guidatore. Prima di illustrare brevemente le principali funzioni che la telematica permette, è utile presentare le tecnologie che le rendono possibili. Si tralascia per ragioni di spazio di illustrare la telefonia cellulare, tecnologia molto nota e diffusa.

Sulla telefonia cellulare è importante solo accennare al fatto che la diffusione delle nuove tecnologie, come GPRS e UMTS, hanno rimosso un importante ostacolo per le applicazioni telematiche su veicolo. Infatti, molti servizi telematici sono erogati lungo tutto un viaggio e, quindi, *per molti minuti o alcune ore, anche se la quantità di informazioni scambiate è piuttosto limitata*. È ovvio, quindi, una connessione il cui costo è basato sul



tempo di connessione è poco adatta per questo tipo di applicazioni. Il GPRS, in cui il *costo di comunicazione si basa sul volume dei dati trasmessi*, è invece molto più adeguato.

6.1. La localizzazione satellitare

La localizzazione satellitare è una tecnologia sviluppata negli anni settanta, con scopi puramente militari: occorre una tecnologia che consentisse di guidare missili intercontinentali sul bersaglio stabilito con una precisione di pochi metri. Questo doveva servire per distruggere i bunker atomici blindati del nemico: con una precisione minore i missili, anche se a testata nucleare, potevano essere inefficaci. Solo dopo un grave incidente aereo, causato dalla imprecisione degli strumenti di bordo del velivolo, gli Stati Uniti accettarono di aprire il servizio anche ad impieghi civili, ma a determinate condizioni.

In particolare, il GPS (*Global Positioning System*) degli Stati Uniti prevedeva la funzione della *Selective Availability*: il segnale disponibile per uso civile era degradato in modo tale da non consentire precisioni superiori ai 100/150 m. La degradazione avveniva inserendo un errore casuale, impossibile da recuperare se non con tecniche differenziali.

Lo sviluppo di questo sistema è avvenuto in piena guerra fredda, inevitabile quindi che anche l'allora Unione Sovietica sviluppasse un sistema analogo, chiamato GLONASS, con prestazioni analoghe al sistema americano. Ma dopo la caduta del muro di Berlino il sistema è andato in disuso, e quindi oggi è solamente utilizzabile il sistema GPS. Gli Stati Uniti hanno da diversi anni rimosso la *Selective Availability*, per cui il GPS è ora utilizzabile al pieno delle possibilità previste per l'impiego civile.

Questa situazione di monopolio degli Stati Uniti, che crea una potenziale dipendenza in tutto il mondo, ha portato l'Europa a decidere di dotarsi di un proprio sistema di localizzazione satellitare chiamato, in onore del grande scienziato italiano, Galileo. Il sistema è in fase di realizzazione e si prevede l'avvio del servizio per l'anno 2008. Anche la Cina e Israele hanno recentemente aderito alla iniziativa Europea, garantendo il loro supporto tecnico ed economico.

Il sistema Europeo funziona in modo analogo

al sistema GPS, su frequenze molto vicine: sarà, quindi, possibile realizzare ricevitori GPS-Galileo combinati, senza eccessivi aggravii di costo, con molti vantaggi dal punto di vista delle prestazioni e della copertura in ambito urbano. Infatti, per poter eseguire il calcolo della propria posizione occorre "vedere" almeno quattro satelliti. Per vedere si intende che esiste una linea diretta tra il ricevitore e il satellite in orbita, non ostruita da alcun ostacolo (edifici, montagne ecc.). La probabilità di "vedere" quattro satelliti ovviamente aumenta se la flotta di satelliti considerata è composta non solo da 24 (GPS) o 27 (Galileo) ma bensì da 51 satelliti, la somma delle due flotte.

7. SERVIZI TELEMATICI

La combinazione delle due tecnologie, la comunicazione cellulare e la localizzazione satellitare, permette la realizzazione di un grande numero di cosiddetti "servizi telematici", a supporto di una guida più confortevole ed efficiente.

Alcuni esempi sono presentati nei paragrafi seguenti.

7.1. Navigazione

Spesso, quando si viaggia con qualcuno verso una destinazione non familiare, viene spontanea la domanda "conosci la strada?". In caso di risposta negativa diventava necessario consultare mappe e stradari, oppure affidarsi alle informazioni ricavabili dai passanti. In entrambi i casi è elevato il rischio di allungare la strada e il tempo di percorrenza.

I navigatori satellitari, oggi, ci vengono in aiuto: basta digitare la destinazione desiderata e loro ci guidano verso questa, senza rischio di perderci o allungare il percorso. Il navigatore satellitare si basa sulla localizzazione satellitare, che permette di conoscere la posizione del veicolo e la sua direzione, e sulla conoscenza dettagliata della mappa stradale, opportunamente "digitalizzata". Tramite queste informazioni il navigatore è in grado di calcolare il percorso da seguire per raggiungere la destinazione desiderata e quindi suggerirci, nel momento opportuno, la manovra da svolgere, ad esempio svoltare a sinistra al prossimo incrocio. Questi suggerimenti vengono

forniti attraverso un display sulla plancia del veicolo e, per ridurre la distrazione dalla guida, anche mediante un messaggio vocale. Si dispone di una sorta di "guida elettronica".

I sistemi di navigazione sono stati un grosso successo nei primi anni novanta, sul mercato giapponese. In Giappone le città non hanno, per la maggior parte, un sistema di vie denominate con numeri civici ordinati che indicano i vari edifici presenti ai lati delle vie stesse. Le case sono divise in quartieri e hanno numeri successivi, secondo l'ordine con cui vengono costruite. Spiegare, quindi, a una persona che non conosce la zona come raggiungere un determinato indirizzo è un compito molto complesso e gli stessi tassisti spesso non escono mai dalla zona della città a loro nota. I sistemi di navigazione satellitare hanno, quindi, risposto a un forte bisogno, bisogno che i sistemi alternativi (mappe cartacee per esempio) non risolvevano completamente. Il mercato dei sistemi di navigazione in Giappone ha avuto uno sviluppo rapidissimo, soprattutto con soluzioni "after market", vale a dire sistemi installabili su qualsiasi veicolo. Già alla fine degli anni novanta i sistemi di navigazione iniziavano a utilizzare la tecnologia DVD per memorizzare le mappe, essendo un CD non più sufficiente a contenere le mappe dell'intero Giappone, o anche solo della sola capitale Tokio. Per fare un confronto, le mappe del territorio italiano stanno ancora oggi comodamente su un singolo CD. Un'altra curiosità: sempre per ovviare alla mancanza di un sistema di indirizzi stradali, anche i telefoni fissi sono memorizzati nelle mappe, per cui esiste una corrispondenza diretta tra numero di telefono e la sua posizione GPS, anche in altezza (numero del piano).

I sistemi di navigazione satellitare si stanno affermando anche in Europa e sono ormai un "must" per i veicoli di gamma alta. Questa diffusione ha permesso anche uno sviluppo delle "mappe digitali": in Italia è coperto ormai tutto il territorio nazionale, come rete stradale, e circa il 70-80% della popolazione come copertura urbana.

7.2. Informazioni su punti di interesse

Le mappe digitali possono contenere lo stradario, ma anche molte informazioni relative a cosa si trova "lungo" le strade. Grazie a un navigatore satellitare è possibile chiedere di

raggiungere la farmacia oppure il distributore più vicino, magari lungo la strada che ci porta verso la destinazione desiderata.

Però la farmacia potrebbe essere chiusa, come anche il distributore. Si potrebbe pensare di mettere gli orari di apertura sul CD, ma le farmacie cambiano continuamente il loro turno di chiusura notturno o festivo, mentre il CD si può cambiare una volta all'anno o poco più. La comunicazione per via cellulare viene in aiuto: possiamo inviare la nostra posizione a un centro servizi e questo ci ritorna le informazioni aggiornate sull'apertura degli esercizi. Possiamo, quindi, recarci, senza il dubbio di trovare le saracinesche abbassate, verso la farmacia selezionata.

Grazie al collegamento con il Centro Servizi è possibile anche fare a meno della mappa digitale su CD: scelta la destinazione, il Centro Servizi può trasmetterci una porzione della mappa digitale, esattamente quella che serve per seguire il percorso calcolato verso la direzione. È questa mappa sarà sempre aggiornata, anche con gli ultimi cantieri di lavoro per la costruzione della metropolitana.

7.3. Informazioni sul traffico

Anche se conosciamo la strada per arrivare alla destinazione, difficilmente sappiamo prevedere quanto traffico incontreremo e questo è un'altra frequente causa di ritardi nell'arrivo alla destinazione. Quando arriviamo, con ore di ritardo per un incidente in autostrada, spesso ci diciamo: "se lo avessi saputo sarei passato da....".

Disponendo di aggiornate informazioni sul traffico, un centro servizi, collegato al nostro veicolo, può informarci tempestivamente di problemi di traffico presenti sul percorso che stiamo seguendo, permettendoci quindi, se abbiamo alternative, di considerarle.

Questo è uno dei servizi, comprensibilmente, più desiderati dagli utenti, sempre più spesso alle prese con traffici congestionati. La sua realizzazione dipende però dalla disponibilità di dati sul traffico precisi e aggiornati. Molto spesso, soprattutto nel caso di incidenti, una congestione si crea in pochi minuti, e, una volta rimossi i veicoli incidentati, in pochi minuti si può risolvere. Il rilievo tempestivo di un problema nel flusso del traffico è ancora irrisolto.

Traffico	Condizioni traffico, aggiornamento automatico lungo il percorso richiesto
Mete e Indirizzi	Ogni indirizzo diventa automaticamente impostato sul navigatore di bordo Ricerca e trasmissione di un indirizzo privato o di aziende Alberghi. Compresa prenotazione Ristoranti. Sulla base della cucina. Prenotazione Spettacoli. Cinema e teatri con orari e programmazione Locali. Utilità. Bancomat, ospedali, farmacie, aeroporti, officine, stazioni di servizio ecc. Musei e mostre. Monumenti e luoghi di interesse turistico
Info e News	Previsioni meteo Orari di voli, treni e traghetti Attualità e Borsa Valori
Assistenze	Assistenza stradale 24 h al giorno, 365 giorni l'anno, in tutta Europa Invio di una officina mobile o carro attrezzi. Auto sostitutiva Assistenza sanitaria 24 h al giorno, 365 giorni l'anno in tutta Europa Consulenza medica, invio di una ambulanza, prenotazione di un Centro Ospedaliero

TABELLA 1

Servizi telematici offerti nel pacchetto Connect

Anche su questo aspetto la tecnologia sul veicolo potrebbe fornire la soluzione: è il veicolo stesso che, automaticamente, segnala i dati sul traffico del tratto che sta percorrendo. Questo avviene semplicemente trasmettendo a una centrale di raccolta dati il percorso che si è appena seguito e la velocità con cui lo si è percorso.

Ovviamente, questa soluzione richiede che i veicoli che fungono da sensori del traffico (*floating probe car*) siano numerosi affinché: sia elevata la probabilità che uno di tali veicoli si trovi coinvolto in una congestione, nei primi minuti nella quale si forma;

grazie a segnalazioni multiple si possano filtrare rallentamenti dovuti a problemi del singolo veicolo o alla volontà del guidatore.

Recenti studi e simulazioni dimostrano che la percentuale dei veicoli "sensori" necessaria dipende dal livello di traffico medio: su tratti di strada molto trafficati un 2% può essere sufficiente, per strade poco trafficate sono necessarie percentuali del 15% o più.

In Italia è già attivo, caso unico a livello Europeo, un servizio di informazioni sul traffico che si basa sulle informazioni ricavate da circa 25.000 veicoli "sensori" (infotraffic bConnect) (Tabella 1). Le informazioni ottenute sono diffuse in tempo reale a tutti i

clienti del servizio e attraverso notiziari via radio (RDS *Radio Dimensione Suono e Radio Capital*).

7.4. Pagamenti sulla base dell'utilizzo

Poter conoscere la posizione del veicolo in ogni suo spostamento porta a immaginare nuove modalità con cui vengono erogati molti servizi connessi con l'utilizzo del veicolo stesso.

Un primo esempio è l'assicurazione: stanno nascendo le prime formule "pay-per-use": il premio è costituito da una parte fissa, minima, e una parte variabile, calcolata in funzione del numero dei chilometri percorsi e dove sono stati percorsi. Per le assicurazioni permette di rendere più preciso il calcolo del rischio e premiare quegli utenti che utilizzano molto poco il veicolo. Le tecnologie GPS-GSM consentono poi una serie di accorgimenti tali da rendere quasi impossibile una frode assicurativa.

Su questo schema stanno nascendo altre iniziative: il governo olandese aveva lanciato un progetto, ora sospeso, per rendere tutte le tasse sul veicolo proporzionali all'utilizzo del veicolo stesso. In Germania, dove le autostrade sono gratuite, sistemi telematici sono alla base di un progetto per intro-

durre il pagamento di pedaggi sui soli veicoli industriali.

A livello Europeo, si vuole stimolare una rapida introduzione di sistemi di localizzazione su tutti i veicoli di futura produzione. I benefici, sul traffico e sulla sicurezza, sono indubbi. Una direttiva è stata recentemente approvata dal Parlamento Europeo: propone una graduale sostituzione dei sistemi di pedaggio basati sui caselli, noti colli di bottiglia nei flussi di traffico, con sistemi elettronici tipo Telepass oppure di tipo telematico. I secondi dovrebbero, secondo questa direttiva, diventare quelli di riferimento a partire dal 2012 in poi.

Ma per il cittadino Europeo le notizie non sono del tutto positive: la direttiva stimola questo rapido passaggio anche per rendere disponibile un nuovo strumento per il controllo traffico: il *road pricing*. Le strade più congestionate, anche in ambito urbano, potranno essere facilmente rese "a pagamento" senza richiedere infrastrutture o provocare interruzioni al flusso di traffico. Avremo quindi una sorta di "tassometro" anche sul nostro veicolo. Esempi di *road pricing* esistono già: per entrare a Londra occorre pagare cinque sterline al giorno, indipendentemente dalla percorrenza (Figura 6). La te-

lematica consentirà soluzioni nelle quali il pagamento avviene sulla base dell'effettivo utilizzo della rete stradale urbana.

8. CONCLUSIONI

L'elettronica ha rivoluzionato le vetture moderne sia nei sistemi di motopropulsione che in una serie di sistemi funzionali a una guida più efficiente, sicura e confortevole. La ricerca sta dimostrando che esistono ancora molti sviluppi possibili:

- nel settore della motopropulsione, accanto allo sviluppo di motori che utilizzano combustibili meno inquinanti (gas naturale, idrogeno), notevoli sviluppi saranno possibili con le tecnologie ibride (motori elettrici combinati con motori a combustione interna), che consentono notevoli riduzioni di consumi ed emissioni, senza pregiudizio nelle prestazioni complessive del veicolo;

- nel settore dei sistemi di controllo del veicolo i sistemi Drive-by-wire consentiranno di realizzare veicoli nei quali le quattro ruote potranno essere controllate indipendentemente in tutti i parametri: coppia motrice, azione frenante, sterzata, sospensioni. I miglioramenti possibili in termini di capacità di controllo della dinamica del veicolo sono notevoli, ma anche massima dovrà essere l'affidabilità di tali sistemi;

- nella sicurezza si diffonderà l'utilizzo dei sistemi di assistenza alla guida, al fine di ridurre in modo "preventivo" situazioni che possono portare a incidenti stradali;

- nella telematica si assisterà a un continuo incremento delle informazioni disponibili dall'infrastruttura: traffico e condizioni stradali ad esempio. Questo permetterà di rendere il traffico veramente "intelligente", centri di controllo del traffico che dialogano con i veicoli, conoscono la loro posizione e destinazioni desiderate e instradano ogni veicolo secondo criteri che ottimizzano complessivamente l'utilizzo delle infrastrutture stradali disponibili. I veicoli coopereranno, quindi, tra di loro e con l'infrastruttura per raggiungere livelli assoluti di sicurezza e minimi intasamenti di traffico. Sembra fantascienza, ma i programmi di ricerca Europei, Giapponesi e Statunitensi dei prossimi anni hanno queste soluzioni tra i loro principali obiettivi.



FIGURA 6

Cartello stradale che a Londra indica l'inizio della zona a pagamento "Congestion charging"

E la privacy? Vivremo in un mondo dove il "Grande Fratello" saprà sempre dove andiamo? L'attenzione a questa tematica dovrà essere in futuro sempre maggiore, non tanto o solo per aspetti di privacy personale, ma soprattutto per gli aspetti di sicurezza, sia dei dati personali ma anche della incolumità fisica delle persone. Recentemente, negli USA, sono riusciti a copiare il dispositivo a infrarossi che consente alla polizia di controllare i cicli semaforici: chiunque poteva generarsi una propria "onda verde semaforica". Oltre agli ovvi problemi di traffico che questo può provocare, sono possibili anche usi impropri tali da causare incidenti stradali. Massima attenzione dovrà, quindi, essere data allo sviluppo di soluzioni che siano assolutamente sicure, rendendo impossibile un utilizzo improprio o dannoso. Se eviteremo questi rischi le nuove tecnologie consentiranno di rendere gli spostamenti su strada sempre più sicuri, efficienti,

confortevoli, poco inquinanti e meno stressanti. Tutto sommato, quindi, più divertenti e potremo riscoprire un piacere che in passato avevamo ma ora sembra si sia perso: il piacere di guidare. E l'Autonica si occupi di tutto il resto!

Bibliografia

- [1] Yilin Zhao. Vehicle Location and Navigation Systems. Artech House 1997.
- [2] R.K. Jurgen Automotive Electronics Handbook 1999.
- [3] Automotive Handbook Robert Bosch 1995.
- [5] FISTERA First Report on Key European Technology Trajectories 2003.
- [6] Elliot D. Kaplan: Understanding GPS Principles and Applications Artech House Publishers.
- [7] AA.VV.: Proceedings ITS World Congress Madrid 2003.
- [8] Automotive Electric/Electronics Handbook Robert Bosch SAE.

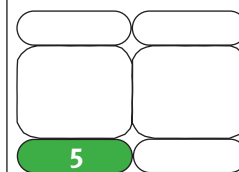
GIANFRANCO BURZIO è laureato in Ingegneria Elettrotecnica. Specializzazione in Automazione Industriale. Dopo una breve attività nel campo dello sviluppo di software gestionale entra nel Centro Ricerche FIAT nel 1982, dove lavora allo sviluppo di applicazioni della visione artificiale per ispezione, controllo qualità e guida robot. Nel 1989 diventa responsabile delle applicazioni sul veicolo di sistemi di visione artificiale, nell'ambito del programma di ricerca Europeo PROMETHEUS. Diverse applicazioni vengono sviluppate dal suo gruppo per migliorare la sicurezza e il comfort di guida: sensore di sorpasso optoelettronico, sistema di mantenimento corsia e sistemi sensoriali anticollisione; alcune di queste sono ora disponibili sul mercato. Nel 1995 diventa anche responsabile delle attività nel campo della informativa di bordo e della telematica applicata su veicolo. Ora è responsabile del coordinamento e dello sviluppo delle attività di ricerca della Divisione Tecnologie del Centro Ricerche FIAT (linee di business verso i Settori FIAT e finanziamenti pubblici, progetto Torino Wireless). gianfranco.burzio@crf.it



IL COSTO DELL'IGNORANZA INFORMATICA NELLA SANITÀ

Una ricerca condotta da AICA e SDA Bocconi nel 2003 sul costo dell'ignoranza informatica per il nostro Paese ha concluso che la perdita di produttività derivante dall'inadeguata preparazione delle risorse umane sull'uso appropriato dell'ICT, risulta dell'ordine dei 15 miliardi di euro ogni anno. Successivamente il gruppo di ricerca ha rivolto la sua attenzione su un'area di grande rilevanza sociale qual è la Sanità, cercando di capire se, anche in questo comparto, l'ignoranza informatica determina conseguenze simili.

Pier Franco Camussone



1. INTRODUZIONE

Per avere una prima indicazione del grado di efficienza del sistema sanitario italiano, un indicatore molto importante è rappresentato dal livello della spesa informatica globale annua in questo settore. Infatti, un valore elevato degli investimenti in questa tecnologia lascia intendere un elevato grado di automazione dei processi di gestione e di erogazione dei servizi, ed è, quindi, un indicatore iniziale del grado di efficienza del sistema sanitario nel suo complesso. Inoltre, le tecnologie informatiche sono ormai essenziali anche in campo clinico, per migliorare la diagnosi e rendere più efficaci le cure, grazie ad apparecchiature con una forte componente informatica. È, quindi, evidente che è di grande interesse conoscere il livello della spesa informatica rispetto alla spesa globale del sistema sanitario, perché si tratta di un indicatore significativo del livello di efficienza e di efficacia del sistema sanitario.

Il calcolo della spesa informatica della sanità in Italia è utile anche come elemento di confronto (*benchmarking*), per comparare il gra-

do di informatizzazione (e, quindi, di efficienza) raggiunto dal sistema sanitario italiano con quello conseguito da altri Paesi, con cui è abituale e significativo un confronto. In tal modo, ci si può rendere conto della posizione relativa del nostro sistema sanitario rispetto a quella di Paesi più avanzati, o supposti tali. Il confronto avviene misurando il rapporto:

$$\frac{\text{spesa informatica}}{\text{spesa sanitaria}}$$

che si riscontra nel settore della sanità italiana e comparando tale indice con quanto calcolato similmente negli altri Paesi.

Per ottenere questo risultato si deve prima misurare, o stimare, l'entità globale della spesa sanitaria e poi valutare la componente di tale spesa destinata all'acquisto e all'impiego delle risorse informatiche.

2. LA SPESA SANITARIA IN ITALIA

La spesa sanitaria in Italia è oggettivamente difficile da determinare perché in parte, è ori-

ginata da strutture pubbliche e in parte, è derivante da attività che si svolgono nel settore privato. In particolare, per questa ultima componente non è facile avere dati certi, in quanto le differenti realtà che vi operano (ospedali, laboratori, centri di assistenza, medici e operatori sanitari) non sono riconducibili a strutture di riferimento comuni, che possono fornire una visione unitaria delle attività e dei costi corrispondenti.

Nel caso della componente pubblica il problema è più semplice: il Ministero della Salute sovrintende alla parte pubblica della spesa sanitaria nazionale, mentre le regioni amministrano tale spesa assegnando le risorse agli ospedali regionali e alle altre strutture territoriali, come i centri di dialisi, gli stabilimenti idrotermali, i centri di salute mentale, i consultori materno - infantili, i centri diurni psichiatrici le residenze sanitarie assistenziali (RSA) e le "case protette" [4].

Una seconda difficoltà è rappresentata dalla elevata articolazione del settore sanitario. Per esempio, se si considera la parte pubblica di tale sistema, numerosi sono gli operatori che ne fanno parte. Oltre al già menzionato Ministero della Salute, si devono prendere in considerazione gli Assessorati Regionali che determinano le politiche di gestione della spesa sanitaria in ciascuna regione e che, pertanto, influenzano il tipo e il livello dei servizi ai cittadini. Poi si devono analizzare i bilanci di ASL¹ e AO², che rappresentano la struttura fondamentale per l'erogazione dei servizi sanitari ai cittadini. Ma non basta: sono da prendere in considerazione anche altri importanti operatori, quali gli *Istituti di Ricovero e Cura a Carattere Scientifico*³ (IRCCS), che dipendono direttamente dal Ministero della Salute, i Policlinici Universitari, le strut-

ture di medicina di base (medici di famiglia), le già menzionate strutture territoriali e la cosiddetta "spesa farmaceutica" a carico dello Stato. Il comportamento individuale di tutti gli operatori, ma anche le interazioni che si sviluppano tra di loro, determinano il livello e la qualità dei servizi sanitari ai cittadini.

Anche il settore privato è altrettanto complesso e articolato: esso comprende IRCCS di diritto privato, ospedali privati accreditati o meno, laboratori di diagnosi, centri di terapia specializzata, operatori per il recupero e l'assistenza ai malati, medici privati, e così via. L'estrema articolazione ed eterogeneità dei soggetti, che fanno parte del sistema sanitario, rendono enormemente difficile la rilevazione dei costi del sistema sanitario italiano. Un'ulteriore difficoltà nella valutazione della spesa sanitaria è rappresentata dal fatto che, vista la pluralità dei soggetti che la determinano, i sistemi di misurazione della spesa adottati sono molteplici e spesso non coerenti tra loro.

Pervenire a una stima globale della spesa sanitaria nel nostro Paese comporta, quindi, sia l'acquisizione di dati da un numero rilevante di fonti, sia un successivo consistente lavoro di armonizzazione di quanto rilevato. Utilizzando differenti approcci ed esaminando le fonti più autorevoli che forniscono dati sulla spesa informatica in Italia [11, 12, 16] si trovano valori non molto dissimili circa la spesa globale (Tabella 1). La differenza tra i valori calcolati dai diversi Enti non supera il 5% e, quindi, può essere spiegata con l'imprecisione, o l'approssimazione dei sistemi di rilevazione, o con le differenze nei criteri di contabilizzazione adottati. In conclusione, si può ritenere che, in Italia, la spesa per prestazioni sanitarie, nel 2001, sia stata nell'ordine di 100 miliardi di euro. Tre quarti di tale importo è stato sostenuto dallo Stato e il rimanente 25% è stato pagato dai privati. A una conclusione analoga perviene anche il Centro di Ricerca sulla Gestione dell'Assistenza Sanitaria dell'Università Bocconi (CERGAS) [17].

Si tratta di una somma adeguata o insufficiente? Per rispondere a questo quesito è possibile confrontare il livello di spesa italiano con quello di altri Paesi sviluppati (Tabella 2). L'impressione che se ne ricava è che il nostro Paese sia in linea con Svezia, Danimarca e Olanda, Paesi che hanno un li-

¹ Azienda Sanitaria Locale: struttura responsabile della gestione di enti come ospedali, laboratori, centri di terapia ecc. che erogano servizi nel territorio di sua competenza.

² Azienda Ospedaliera: ospedale che non fa parte di una ASL ma che ha una entità giuridica autonoma in ragione della sua dimensione, o della sua importanza.

³ Si tratta di istituti di ricerca a carattere scientifico di grande importanza, ma che erogano anche servizi sanitari ai cittadini.

Spesa sanitaria in Italia
 (Fonte: OECD)

2001

	€	% del PIL
Spesa sanitaria pubblica (75% del totale)	76.869.261.000	6,3%
Spesa sanitaria privata	25.623.087.000	2,1%
Spesa sanitaria totale	102.492.348.000	8,40%

NB: Valore del PIL = 1.220.147.000.000 €

Spesa sanitaria in Italia
 (Fonte: Ministero Economia e Finanze ed ISTAT)

2001

	€	% del PIL
Spesa sanitaria pubblica (75% del totale)	73.112.073.213	6,0%
Spesa sanitaria totale	24.370.691.071	2,0%
Spesa sanitaria totale	97.482.764.284	8,0%

NB: Valore del Pil = 1.218.534.553.548 €

TABELLA 1
*La spesa sanitaria
in Italia (dati
aggregati)*

vello di *welfare* spesso indicato come un modello da raggiungere.

È interessante notare come il nostro Paese, che ha un sistema pubblico abbastanza sviluppato spenda meno di Stati Uniti, Canada e Svizzera, dove la componente pubblica del sistema è meno rilevante, mentre spende di più rispetto al Regno Unito che ha un sistema pubblico simile al nostro, ma gestito centralmente e unitariamente dal *National Health Service* (NHS), che risulta essere con circa un milione e trecentomila dipendenti la seconda azienda al mondo in assoluto in termini di dipendenti. Si ricorda in proposito che, con la regionalizzazione e la aziendalizzazione del sistema sanitario italiano, i dipendenti del Servizio Sanitario Nazionale (SSN), nel 2001, erano poco meno di 650.000 [13].

3. I COSTI INFORMATICI DEL SISTEMA SANITARIO ITALIANO

Come già sottolineato, i servizi sanitari sono erogati in Italia sia dal settore pubblico che dai privati. Il *sistema pubblico della sanità* è costituito, in termini logici, da:

■ Ministero della Salute che sovrintende (programmazione e controllo) all'intero sistema sanitario nazionale.

	1999	2000	2001
Australia	8,7	8,9	9,2
Austria	7,8	7,7	7,7
Belgio	8,5	8,6	9,0
Canada	9,1	9,2	9,7
Danimarca	8,5	8,3	8,6
Finlandia	6,9	6,7	7,0
Francia	9,3	9,3	9,5
Germania	10,6	10,6	10,7
Grecia	9,6	9,4	9,4
Islanda	9,5	9,3	9,2
Irlanda	6,2	6,4	6,5
Italia	7,8	8,2	8,4
Giappone	7,5	7,7	8,0
Olanda	8,7	8,6	8,9
Portogallo	8,7	9,0	9,2
Spagna	7,5	7,5	7,5
Svezia	8,4	8,4	8,7
Svizzera	10,7	10,7	11,1
Gran Bretagna	7,2	7,3	7,6
Stati Uniti	13,0	13,1	13,9

TABELLA 2
Spesa sanitaria nazionale in percentuale sul PIL. (Fonte: OECD)

0

Istituti di Ricerca e di Cura a Carattere Scientifico (IRCCS) che dipendono direttamente dal Ministero.

Assessorati delle Regioni e Province autonome (per quanto concerne l'amministrazione delle spesa sanitaria locale e le politiche sanitarie locali).

Aziende Sanitarie Locali (ASL), che erogano servizi sul territorio tramite strutture (Ospedali, laboratori, centri di assistenza ecc.) gestiti direttamente.

Aziende Ospedaliere (AO), ovvero grandi complessi ospedalieri dotati di autonomia gestionale.

Policlinici Universitari, che operano come case di cura e come strutture di specializzazione universitaria.

Strutture territoriali di assistenza (centri di dialisi, stabilimenti idrotermali, centri di salute mentale, consultori materno infantili ecc.).

Medici di base, o medici di famiglia.

Fanno, inoltre, parte del settore pubblico le farmacie che appartengono a enti pubblici locali.

Non tutto il personale di queste entità è, come si è visto, a carico del Servizio Sanitario Nazionale, che comprende in senso proprio le sole unità i cui dipendenti sono in carico al bilancio del SSN. Si tratta, in particolare, di ASL e AO, mentre per gli IRCCS e i Policlinici il personale è in parte a carico del *Ministero della Università e della Ricerca* (MIUR). I medici di base giuridicamente non sono dipendenti del sistema sanitario pubblico, ma professionisti con un contratto specifico.

Similmente, il sistema di erogazione dei servizi appartenente al settore privato è formato da:

Istituti di ricerca e di cura a carattere scientifico di diritto privato.

Ospedali e case di cura privati accreditati.

Ospedali e case di cura privati non accreditati.

Strutture territoriali private (laboratori, centri di terapia specializzata, consultori, ecc.).

Medici (liberi professionisti).

Farmacie private.

Il settore sanitario è, quindi, costituito da una molteplicità di attori, come illustrato in tabella 3, i quali svolgono compiti differenti quali l'indirizzo e il coordinamento della politica sanitaria, l'erogazione di servizi sanitari, la ricerca,

il presidio del territorio con funzioni di assistenza e di primo intervento, la distribuzione di farmaci al dettaglio, e così via. È, quindi, evidente che ogni attore utilizza l'informatica con finalità specifiche, che derivano dal ruolo che l'operatore svolge nel sistema sanitario.

Per esempio, nel caso del Ministero della Salute, che ha funzioni di indirizzo e di governo del sistema sanitario italiano, l'informatica è usata per migliorare la conoscenza del sistema, tenere sotto controllo le diverse sue componenti e formulare piani e programmi di sviluppo. Si tratta delle tipiche funzioni d'uso dell'*Information and Communication Technology* (ICT) che si riscontrano ai vertici di grandi imprese e che si sintetizzano nell'espressione "sistema informativo direzionale". Nelle AO, invece, l'informatica è impiegata per rendere più efficiente il processo di gestione del malato e più efficace la cura del medesimo. Pertanto, accanto a un uso tradizionale dell'ICT, per migliorare la gestione amministrativa dell'ospedale e del processo di erogazione del servizio, si prospetta anche un impiego di questa tecnologia per migliorare i processi clinici (anamnesi, diagnosi e terapia), che possono trarre vantaggio dal ricorso ad apparecchiature informatizzate. Infine, Istituti di Ricerca e Policlinici universitari, oltre alle applicazioni dell'ICT richieste dagli ospedali, si attendono anche un supporto per le attività di ricerca e di studio che essi svolgono.

Il caso dei medici di base, poi, è molto specifico. Essi possono usare l'ICT per archiviare i dati dei propri pazienti, tenere conto della loro storia clinica (anamnesi), registrare le terapie prescritte e gli esiti ottenuti. Inoltre, possono usare l'informatica per accedere a patrimoni di conoscenze mediche che possono essere di grande aiuto nella loro pratica professionale. Queste basi di conoscenza, mantenute costantemente aggiornate dai centri di ricerca, riportano i protocolli di diagnosi e di cura per moltissime malattie e disfunzioni e mettono in grado il medico di base di operare con un livello di professionalità superiore. Infine, quando il medico deve prescrivere cure o terapie, un sistema informatizzato, collegato in rete con ASL e AO permette di indirizzare il paziente verso la struttura libera più idonea. Anche le prescrizioni

Livello della struttura	N° operatori	Spesa sanitaria	Costi informatici annuali €	Costi informatici/spesa sanitaria
Sistema Pubblico				
1 Ministero	1	970.000.000	29.479.000	
IRCCS	15	1.570.000.000	9.420.000	
2 Regioni	20		102.331.416	
3 ASL totali	197	61.062.061.000	274.779.275	
a H Gestiti direttamente	566			
b AO	100			
c Policlinici Universitari	10			
d Strutture territoriali				
4 Medici di base	54.226	48.803.400		
5 Farmacie	1.291	11.837.866.194	2.323.800	
		76.869.927.194	475.716.891	0,62%
Sistema privato				
1 IRCCS	16	2.000.000.000	16.000.000	
2 Ospedali privati accred.	530	2.000.000.000	16.000.000	
3 Ospedali privati non accred.	102			
4 Strutture territoriali				
5 Medici	75.000	9.600.000.000	67.500.000	
6 Farmacie	15.435	12.023.000.000	27.783.000	
		25.623.000.000	127.283.000	0,50%
	Totale	102.492.927.194	602.999.891	0,59%

La spesa informatica nella sanità rappresenta il 2,9% della spesa informatica italiana

TABELLA 3

La spesa sanitaria (disaggregata) e i costi informatici del sistema sanitario italiano (2001). (Elaborazione su dati del Ministero della salute, del CNIPA, di Federfarma, OASI/CERGAS, e rilevazioni ad hoc)

farmaceutiche possono essere comunicate direttamente a una farmacia, dove il paziente può passare a ritirarle in seguito.

Nel caso delle farmacie, l'informatica serve a gestire in modo puntuale le scorte dei medicinali, segnalare consumi crescenti e formulare previsioni di *stock-out* prima che si verifichino, consentendo di formulare tempestivamente richieste di rifornimento. Anche i farmacisti come i medici di base possono usare l'ICT per accedere a patrimoni remoti di cono-

scenze farmacologiche in modo da migliorare il servizio alla clientela, che si rivolge loro per suggerimenti e consigli. L'informatica serve, inoltre, per semplificare le procedure di gestione della farmacia: registrando ogni operazione si possono poi produrre relazioni e statistiche sull'uso dei farmaci, come richiesto nel caso di farmaci il cui uso deve essere strettamente monitorato. Se poi i pazienti sono dotati di carta sanitaria, come previsto nel progetto in corso di attuazione presso la Regione

Lombardia, si può alimentare il Sistema Sanitario Nazionale direttamente dalla farmacia, con i dati del paziente e della relativa medicina consegnata, in modo da semplificare notevolmente il controllo della spesa farmaceutica, rendendo possibile conoscere per ciascun soggetto la spesa farmaceutica a carico dello Stato e quella da lui sostenuta direttamente. In tabella 3 sono evidenziati i costi informatici sostenuti dai diversi operatori appartenenti a ciascun livello della struttura del Sistema Sanitario Nazionale⁴.

I risultati del lavoro di ricerca indicano che la spesa informatica nel settore della sanità italiana ammonta a oltre 600 milioni di euro l'anno. Tale importo rappresenta lo 0,59% della spesa sanitaria e più precisamente lo 0,62% nel caso della sanità pubblica e lo 0,50% nel caso della componente privata del sistema sanitario. Si tratta di una percentuale non ele-

vata, se comparata con quanto si riscontra negli altri settori economici. Si ricorda in proposito che, in Italia, la spesa informatica annuale media di un'azienda del settore manifatturiero ammonta al 1,14% del fatturato [1], mentre nel settore dei servizi commerciali il corrispondente valore è dello 0,80% [2].

In termini macroeconomici, il settore dei servizi sanitari, in Italia, ha le seguenti dimensioni: esso dà lavoro a circa 1.280.000 di persone, ovvero impiega il 4,6% della forza lavoro disponibile nel nostro Paese, mentre le sue attività economiche rappresentano l'8% del PIL italiano. Si tratta, quindi, di una realtà consistente e importante nel contesto produttivo del nostro Paese. La spesa informatica generata da questo settore è però inferiore rispetto a quella di altri settori, anche se raffrontata a essi in termini relativi piuttosto che assoluti (Tabella 4). Si tratta di "un gigante" economi-

TABELLA 4
Dimensione di alcuni settori economici e spesa informatica relativa.
(Fonte: Elaborazione SDA su dati ISTAT)

Settore	% del PIL	% addetti su totale forza lavoro italiana	% della spesa informatica globale italiana
Industria	26%	31,8%	29,0%
Trasporti e Comunicazioni	7%	5,5%	8,0%
Commercio	16%	15,9%	9,0%
Finanza	15%	3,4%	24,0%
Sanità	8%	4,6%	2,9%
PA (escluso sanità)	5%	8,8%	7,0%
Servizi	13%	6,9%	5,0%
Altri settori	10%	23,1%	15,4%
	100%	100,0%	100,3%

⁴ Per il calcolo sono stati adottati i seguenti accorgimenti. La spesa informatica del Ministero della Sanità è stata ricavata dalla "Relazione annuale 2002: Stato dell'informatizzazione nella Pubblica Amministrazione" edita dal Centro Nazionale per l'Informatica nella Pubblica Amministrazione (già AIPA). La spesa informatica sostenuta dagli Assessorati Sanità delle regioni è stata desunta dal "Rapporto sullo stato dell'informatizzazione delle Regioni e delle Province Autonome" dell'AIPA. Il valore globale per le 20 Regioni è stato ottenuto estrapolando i valori misurati da AIPA su un campione di 10 Regioni. I costi informatici delle ASL sono stati calcolati come segue: la ricerca SDA Bocconi - AICA ha rilevato nel campione di indagine il rapporto medio esistente tra costi informatici e ricavi. Tale rapporto è stato applicato a tutte le ASL estrapolando la situazione del campione all'universo della popolazione. Per le Aziende Ospedaliere, i Policlinici Universitari e gli IRCCS è stato rilevato il costo informatico annuale sostenuto in una realtà emblematica dal punto di vista rappresentativo della intera popolazione. Il valore riscontrato è stato poi estrapolato all'intero insieme dei soggetti. Nel caso dei medici la ricerca ha rilevato un costo medio pro-capite di circa 900 € l'anno (300 per l'Hw, 300 per il Sw e 300 per le Telecomunicazioni), che è stato poi estrapolato all'intero insieme. Per le farmacie è stato stimato un costo annuo di circa 1.800 € (800 per l'Hw, 500 per il Sw e 500 per le Telecomunicazioni) che è stato estrapolato all'intero universo dei soggetti.

Per gli ospedali privati si è ipotizzata una spesa informatica, in relazione ai ricavi, un po' superiore a quella riscontrata per gli ospedali pubblici, essendo queste imprese private tradizionalmente meglio disposte verso questa forma di investimento.



co che però, comparato con gli altri settori, appare “un nano” dal punto di vista informatico. Il livello di informatizzazione di questo settore è enormemente lontano rispetto a quello della finanza o dell’industria, anche se è evidente il vantaggio che una maggiore informatizzazione potrebbe determinare sia per i processi amministrativi che per quelli clinici.

Se poi ci si compara con il livello di informatizzazione che il settore sanitario ha raggiunto in altri Paesi evoluti la sensazione di arretratezza ne esce confermata. Secondo più fonti la situazione dell’informatizzazione del settore sanitario italiano è ben al di sotto dei livelli riscontrabili in altri Paesi sviluppati. Buccoliero e Marsilio, nel loro studio, [1] affermano che negli altri Paesi europei più evoluti il livello della spesa informatica rispetto alla spesa sanitaria ammonta al 2%. Mentre secondo il Gartner Group [8], la spesa informatica annuale nelle aziende sanitarie americane oscillerebbe in un intervallo compreso tra 1,68% e 2,92% del loro fatturato.

4. GLI ADDETTI DEL SETTORE SANITARIO ITALIANO

Il settore sanitario è un comparto ad alta intensità di lavoro, che in Italia impiega circa 1.280.000 addetti, molti dei quali con un livello di professionalità assai elevato. Rispetto ad altri comparti economici si tratta di una forza lavoro assai significativa, che equivale, come già ricordato, al 4,6% del totale dei lavoratori disponibili nel nostro Paese. Una porzione di questi addetti fa parte del sistema pubblico della sanità mentre un’altra frazione opera nel settore privato, come si può vedere dalla tabella 5, in cui la struttura della forza lavoro è stata scomposta sulla base di indicazioni fornite dal Ministero della Salute [14] e da altri Istituzioni [5, 11].

Il settore pubblico è costituito fondamentalmente dal *Servizio Sanitario Nazionale* (SSN), che è stato ormai in gran parte regionalizzato, demandando alle Regioni la gestione del Servizio a livello locale. A tal proposito, il Ministero della Salute così descrive la struttura e la consistenza del SSN italiano (Ministero della Salute 2003): “*Il personale dipendente del SSN è costituito dal personale delle Aziende Sanitarie Locali (ASL), com-*

preso quello degli Istituti di cura a gestione diretta (421.752 unità) e dal personale delle Aziende Ospedaliere (216.881 unità). Nel 2001 tale personale ammonta a 648.633 unità e risulta così ripartito:

- 67% nel ruolo sanitario (medici ed infermieri),
- 21% nel ruolo tecnico,
- 10% nel ruolo amministrativo,
- 0,2% nel ruolo professionale.

Nelle strutture di ricovero delle ASL, nell’ambito del ruolo sanitario, il personale medico è costituito da 46.225 unità (4,5 medici ogni 10 posti letto) e quello infermieristico da 120.310 unità (11,6 infermieri ogni 10 posti letto e 233 giornate di degenza per infermiere); il rapporto fra infermieri e medici, a livello nazionale, si attesta sul valore di 2,6 infermieri per ogni medico. Nelle Aziende Ospedaliere (AO) operano 38.148 medici e 94.144 unità di personale infermieristico; tali strutture dispongono mediamente di circa 5,2 medici ogni 10 posti letto e di 12,8 infermieri ogni 10 posti letto”.

Una delle caratteristiche di questo settore è l’elevata scolarizzazione e professionalità della maggior parte degli addetti. Il personale in possesso di laurea in campo medico, o farmaceutico, sfiora le 420.000 unità, le dimensioni del personale infermieristico in senso stretto sono dello stesso ordine di grandezza (460.000 addetti), mentre il personale tecnico supera le 230.000 unità, e gli addetti di provenienza amministrativa sono oltre 150.000.

5. GLI UTENTI DELL’INFORMATICA NEL SETTORE SANITARIO

In relazione al ruolo ricoperto nel sistema sanitario, il personale che lavora in questo settore può essere suddiviso nelle seguenti categorie:

- il personale medico dedicato alla cura dei pazienti, nelle strutture sanitarie;
- il personale infermieristico dedicato alla cura dei pazienti nelle strutture sanitarie;
- il personale con funzioni tecniche (addetti di laboratorio, specialisti di apparecchiature ecc.);
- il personale amministrativo che svolge compiti impiegatizi di coordinamento, di controllo e di amministrazione;

Sistema Pubblico									
Livello della struttura	Operatori	N° operatori	Personale					Personale stimato da ricerca	Personale secondo Eurostat
			Medici	Inferm.	Tecnico	Ammin.	Totale		
1	Ministero	1				2.220	2.220		
	IRCCS	15	3.000 20%	7.200 48%	3.100 21%	1.700 11%	15.000		
2	Regioni	20							
3	ASL totali	197	38.200 22%	58.200 33%	37.000 21%	41.200 24%	174.600		
a	H Gestiti direttamente	566	46.250 18%	145.500 57%	54.000 21%	11.500 4%	257.250		
b	AO	100	38.150 18%	112.000 52%	49.100 23%	17.550 8%	216.800		
c	Policlinici Universitari	10	4.950 22%	10.200 45%	4.000 18%	3.400 15%	22.550		
d	Strutture territoriali	nd	7.470 18%	15.355 37%	8.715 21%	9.960 24%	41.500		
4	Medici di base	54.220	54.220				54.220		
5	Farmacie	1.291	5.810				5.810		
			198.050	348.455	155.915	87.530	789.950	789.950	790.000
Sistema Privato									
1	IRCCS	16	3.000 20%	7.250 48%	3.296 22%	1.687 11%	15.233		
2	Ospedali privati accred.	530	14.484 25%	19.434 33%	15.277 26%	8.898 15%	58.093		
3	Ospedali privati non accred.	102	2.897 25%	3.887 34%	3.055 26%	1.780 15%	11.619		
4	Strutture territoriali	nd	56.884 22%	86.667 33%	55.097 21%	61.352 24%	260.000		
5	Medici liberi professionisti (compresi dentisti)	75.000	75.000				75.000		
6	Farmacie	15.435	69.458				69.458		
			221.723	117.237	76.726	73.716	489.402	489.402	490.000
Totale addetti settore sanitario								1.279.352	1.280.000

Addetti SSN
648.650

TABELLA 5

Gli addetti del settore sanitario in Italia (2001). (Elaborazione su dati del Ministero della salute, del CNIPA, di Federfarma, Eurostat, Regioni)

■ medici di medicina generale (MMG), ossia i medici di famiglia, e i medici liberi professionisti⁵,
■ il personale delle farmacie.

La consistenza di ciascuna categoria è riportata in tabella 5. Non tutte le categorie professionali sono interessate in ugual misura dall'utilizzo delle tecnologie informatiche. Nelle strutture aziendali (ASL e AO) di minore dimensione quasi la metà del personale è, attualmente, interessata dall'uso degli

⁵ Tra questi ultimi sono da ricordare i medici odontoiatri che sono circa 35.000.

strumenti informatici, mentre nelle aziende di maggiore dimensione questa percentuale tende a decrescere (35-38%). Va però osservato che l'uso o meno della strumentazione dipende per talune categorie (personale medico e infermieristico) dalla propensione, o capacità individuale, a utilizzare direttamente la tecnologia piuttosto che lasciare ad altri (*staff*) il compito di interagire con l'informatica. La categoria dei medici, per esempio, ha trovato enormi vantaggi nell'uso di Internet per accedere a informazioni medico-farmaceutiche aggiornate e, pertanto, ha cambiato le proprie abitudini circa l'aggiornamento e la ricerca di informazioni scientifiche.

Un secondo importante risultato emerso dalla ricerca riguarda il tipo di strumentazione informatica usata dagli utenti. L'indagine ha messo in evidenza che ben tre quarti degli utenti attuali impiegano strumenti di informatica individuale (*word processor, browser* di accesso a Internet, fogli di lavoro elettronici, posta elettronica e così via), però la grande maggioranza di questi utilizzatori possiede solo conoscenze di base e non si può definire un'"utenza esperta".

Per quanto concerne le strutture sanitarie e ospedaliere (Ospedali pubblici e privati, Policlinici, IRCCS), si può, quindi, calcolare che attualmente quasi 340.000 persone usino l'informatica per necessità di lavoro, però circa 200.000 di queste non hanno una solida preparazione al riguardo come evidenziato in tabella 6.

Si può, quindi, concludere che nelle strutture ospedaliere italiane, e nelle corrispondenti ASL, vi siano quasi 200.000 addetti che impiegano l'ICT senza una solida preparazione e che necessitano, quindi, di interventi formativi; mentre dei restanti 430.000, che attualmente non usano questa tecnologia, si potrebbe discutere se debbano, o meno, in un prossimo futuro disporre di strumenti informatici (e in che misura) per il proprio lavoro.

La ricerca ha, inoltre, stabilito che l'utenza informatica attuale risulta costituita in gran parte dal personale amministrativo che usa questa tecnologia per finalità gestionali e contabili, come illustrato in tabella 7.

Per quanto riguarda gli strumenti di informa-

Addetti ospedalieri	771.000		100,00%
Utenti informatici	337.800		43,81%
Utenti di informatica individuale	250.300		32,46%
di cui non esperti		197.260	25,58%
di cui esperti		53.040	6,88%

TABELLA 6

L'utenza informatica nelle strutture ospedaliere pubbliche e private

Categorie di addetti	
Personale amministrativo	49%
Personale tecnico e infermieristico	24%
Personale medico	25%
Altre figure professionali	2%
	100%

TABELLA 7

L'utenza informatica nelle strutture ospedaliere pubbliche e private

tica individuale maggiormente impiegati dagli utenti informatizzati la ricerca ha segnalato che il livello di utilizzo raggiunto da Internet e dalla posta elettronica ha ormai superato quello di molti pacchetti di *office automation* disponibili da più lungo tempo.

6. LA SITUAZIONE TRA I MEDICI DI MEDICINA DI BASE

Per quanto riguarda, invece, i medici di medicina generale (medici di base), è stata svolta un'indagine su un campione rappresentativo, costituito da circa 800 medici, ripartito tra nord, centro e sud dell'Italia, e formato da specialisti appartenenti sia ad aree cittadine, sia alla provincia. L'età media degli intervistati è risultata essere di 50 anni, ed è altresì risultato che il 76% degli intervistati usa strumenti informatici per la propria professione mentre il 7,5% ha dichiarato la propria intenzione di farlo quanto prima. La quasi totalità dei medici che usano strumenti informatici, utilizza pacchetti *software* specifici per la gestione delle visite ambulatoriali e per la cura dei pazienti. Questi pacchetti si possono reperire

sul mercato, anche se, in qualche caso, sono le aziende farmaceutiche che ne fanno dono ai medici. Mentre, per quanto riguarda gli strumenti di informatica individuale maggiormente impiegati dagli utenti informatizzati, la ricerca ha segnalato, anche in questo caso, la grande popolarità di Internet e la buona diffusione dei pacchetti di office automation (Figura 1).

Molto interessante è stato scoprire dai medici di medicina generale quali sono le "funzioni d'uso" per cui essi si servono della tecnologia informatica (Figura 2). Gli impieghi più diffusi dell'informatica si riferiscono alla creazione e al mantenimento delle cartelle cliniche e alla stampa delle prescrizioni. In seconda battuta viene l'accesso alle banche

dati via Internet e solo un quarto degli utenti si serve della tecnologia per gestire lo studio, o per produrre dei report (statistiche).

Che Internet abbia assunto un ruolo rilevante nell'attività dei medici di medicina generale (collegamento a banche dati) è confermato anche da una indagine empirica svolta dal *Boston Consulting Group* (BCG) nel 2002 su 606 medici europei, costituito per il 50% da medici di base e per l'altro 50% da medici specialisti (cardiologi, dermatologi, ginecologi ecc.) che ha fornito risultati del tutto in linea con quelli ottenuti sul campione italiano. Un'ulteriore preziosa informazione riguarda le circostanze in cui avviene la navigazione in rete da parte dei medici europei (indicate in Tabella 8), che dimo-

FIGURA 1
La diffusione degli strumenti di informatica individuale tra i medici di medicina generale

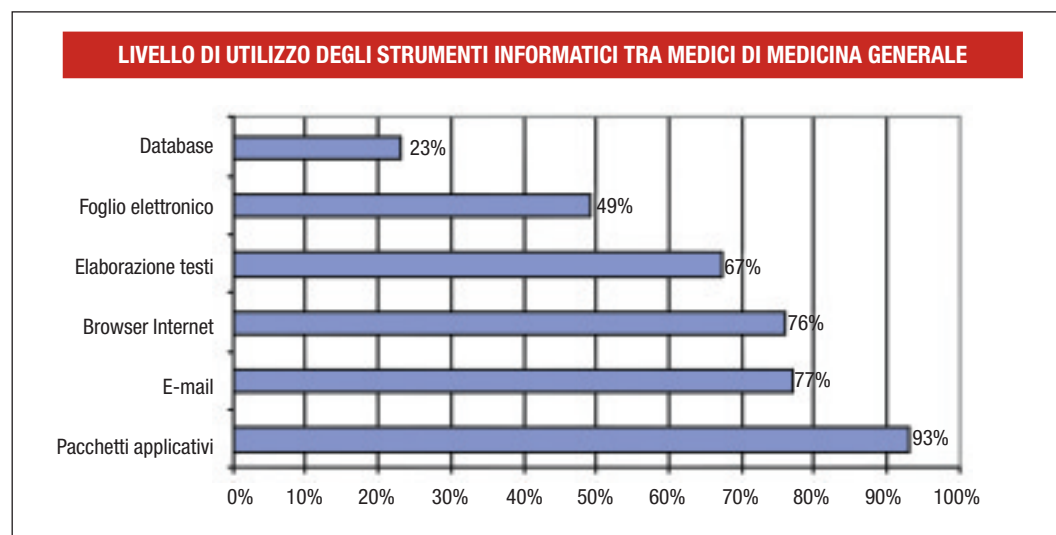
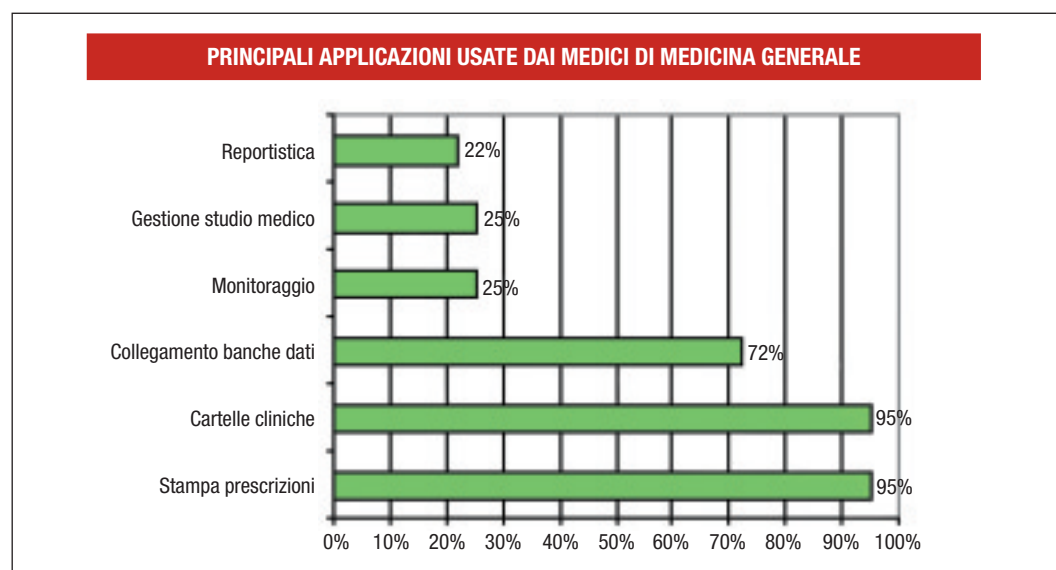


FIGURA 2
Finalità d'uso dell'informatica da parte dei medici di medicina generale



strano come l'accesso ad Internet per ragioni professionali assuma l'aspetto di un processo di aggiornamento e come tale il medico lo affronti nei momenti in cui il lavoro gli consente delle pause.

È, quindi, difficile stabilire un rapporto tra uso dell'informatica e aumento della produttività. Il medico "lavora" anche nei momenti in cui non è costretto a farlo in base ai vincoli imposti dal contratto di lavoro. È altrettanto evidente che la ricerca e il ritrovamento di informazioni utili per la propria professione migliorano il rendimento e l'efficacia della pratica medica quotidiana, con grandi vantaggi sia per il professionista che per il paziente.

Molto interessante è anche conoscere in base a quali elementi i medici selezionano le fonti informative (*Web Site*) cui collegarsi (Tabella 9).

È evidente, dai dati forniti dai medici, che una gran parte di essi ha fatto di Internet il canale preferenziale, piuttosto che complementare,

per l'acquisizione di informazioni medico-farmaceutiche.

7. IL COSTO DEL TEMPO IMPRODUTTIVO NELL'USO DELL'INFORMATICA

L'informatica, come molte altre tecnologie, produce vantaggi in termini di efficienza e di produttività del lavoro, ma talvolta la sua complessità determina problemi di utilizzo. Durante la ricerca è stato chiesto agli intervistati se perdessero, o meno, del tempo per superare difficoltà nell'uso dell'informatica, oppure se incappassero in problemi concernenti l'impiego di questa tecnologia che facessero loro "perdere del tempo". La risposta è stata positiva: sia i medici di medicina generale che il personale medico ospedaliero e infermieristico hanno segnalato questo fenomeno.

Come ricordato inizialmente, la Scuola di Direzione Aziendale e l'AICA hanno svolto nel 2003 una ricerca sul costo dell'ignoranza informatica nel nostro Paese [3]. In tale lavoro è stato affrontato il problema della stima del tempo perso in modo improduttivo durante l'uso degli strumenti informatici e del costo nascosto che ne deriva per le aziende⁶. I medesimi criteri di misura del tempo perso per ignoranza informatica sono stati applicati alla realtà costituita dal sistema sanitario del nostro Paese, procedendo nel modo seguente (Tabella 10). Il personale del settore sanitario italiano è stato ripartito nelle sei principali fasce professionali che lo contraddistinguono. Per ciascuna fascia è stato indicato il numero degli addetti e la percentuale di essi che utilizza l'informatica. La ricerca ha, inoltre, rilevato, sulla base delle dichiarazioni degli utenti, quanto tempo settimanalmente viene speso davanti ai computer e quanto viene perso in modo improduttivo, per problemi derivanti dall'uso della tecnologia, da parte di ciascuna categoria di utenti. Per esempio, dai dati rilevati è risultato che l'80% del personale medico delle strutture sanitarie utilizza l'informatica, e mediamente

Medici che si collegano per scopi professionali¹

Dopo il lavoro	86%
Nel week-end	71%
Tra le visite dei pazienti	46%
Durante la visita di un paziente	16%

¹ (ammesse risposte multiple)

TABELLA 8

Quando i medici si collegano a Internet

Fonte della segnalazione

Materiale scritto	27%
Segnalazione di colleghi	15%
Motori di ricerca	12%
Promotori di medicinali	9%
Portali sanitari	5%
Esperienza personale	5%

TABELLA 9

Come i medici vengono a conoscenza dei Siti di loro interesse

⁶ Si veda l'articolo pubblicato su Mondo Digitale Anno II, n. 2 giugno 2003.

Personale del sistema sanitario	N° addetti	Utilizzatori di informatica	Tempo perso annualmente (giorni)	Costo annuale pro capite (€)	Costo giornaliero (€)	Costo tempo perso annualmente pro capite (€)	Costo totale tempo perso annualmente (€)
Personale medico strutture sanitarie	280.405	224.324 80%	2,04	90.740	412,45	839,35	188.285.066
Personale infermieristico	476.887	95.377 20%	6,11	36.296	164,98	1.008,22	96.161.592
Personale tecnico	232.059	46.412 20%	7,94	29.316	133,25	1.058,63	49.133.063
Personale impiegatizio	160.504	144.454 90%	17,42	35.900	163,18	2.842,08	410.549.237
Medici di base	54.226	41.212 76%	7,53	83.400	379,09	2.853,16	117.583.658
Personale delle farmacie	75.267	75.267 100%					
Totale	1.279.348						861.712.616

Spesa Sanitaria 102.492.927.194

Spesa informatica 602.999.891

Costo dell'improduttività 861.712.616

Ignoranza/spesa sanitaria 0,84%

TABELLA 10

Costo del tempo improduttivo (perso) nell'uso delle tecnologie informatiche

spende 21 h settimanali con il computer. Se si assume che l'orario lavorativo contrattuale prevede 36 h settimanali di lavoro, ne consegue che il personale medico lavora con il computer per il 58% del proprio tempo. L'indagine ha anche rilevato che settimanalmente i medici dichiarano una perdita di tempo di 20 min (0,33 h) per problemi derivanti dall'uso della tecnologia. Ne risulta che per ogni giorno lavorativo lo 0,9% dell'orario di lavoro è assorbito, o meglio perduto, dalla gestione di queste disfunzioni. Un calcolo analogo è stato fatto per le altre figure professionali. In particolare, i medici di medicina generale hanno dichiarato una perdita di tempo di circa 3 h la settimana per problemi derivanti dagli strumenti di informatica, di cui il 43% imputabili non alla tecnologia, ma alla propria impreparazione informatica (1,31 h). Nella tabella 10 è stato riportato il calcolo del valore economico annuale di tali "perdite di

tempo", che rappresenta sostanzialmente il costo della improduttività derivante da problemi nell'uso dell'informatica. Per tale calcolo ci si è basati sul costo aziendale delle singole categorie di addetti come risulta da rilevazioni riportate nel rapporto OASI 2003 e dai dati ISTAT e CNEL [9, 11].

8. CONCLUSIONI

Dai dati raccolti ed elaborati dalla ricerca si è giunti alla conclusione che, per quanto si è potuto calcolare, il costo della improduttività (tempo perso) ammonta per l'intero sistema sanitario nazionale (pubblico e privato) a oltre 850 milioni di euro l'anno. Tale valore è addirittura superiore alla spesa informatica dell'intero settore sanitario e rappresenterebbe, se fosse messo in evidenza, lo 0,84% della spesa sanitaria dell'intero Paese. Si tratta, evidentemente, di una componente

della spesa sanitaria assolutamente persa, ovvero improduttiva. Il quesito che nasce da questo risultato potrebbe essere espresso nei seguenti termini: *conviene incrementare gli investimenti informatici (spesa informatica) nel settore sanitario italiano, o non converrebbe sforzarsi di ridurre l'ignoranza informatica il cui costo è addirittura più elevato della spesa informatica?*

Non ci si può illudere di recuperare il costo nascosto dell'ignoranza informatica nella sua interezza con interventi di formazione, o con la creazione di strutture di assistenza all'utenza che riducono il tempo perso improduttivamente. È però evidente che un intervento di formazione diffuso e capillare può ridurre grandemente, come si vedrà in un articolo pubblicato prossimamente su questa rivista, questa "perdita" occulta.

Bibliografia

- [1] Buccoliero L., Marsilio M.: Lo stato delle dotazioni informatiche e delle strategie ICT, 2003, Cap. 19, in Camussone P.F., Biffi A.: I sistemi informativi nelle aziende manifatturiere italiane. *Il Sole 24 Ore*, 1992.
- [2] Camussone P.F., Biffi A.: I sistemi informativi nelle aziende della distribuzione commerciale. *Il Sole 24 Ore*, 1994.
- [3] Camussone P.F., Occhini G.: *Il costo dell'ignoranza nella società dell'informazione*. ETAS 2003.
- [4] Cantù E., Carbone C. 2003: in Pessina E.A., Cantù E.: *Rapporto Oasis 2003: L'aziendalizzazione della sanità in Italia*. EGEA 2003
- [5] Federfarma: *La presenza delle farmacie sul territorio*. 2003.
- [6] EUROSTAT: *Health statistics: Key data on health*. 2002.
- [7] Gomolski B., Grigg J., Potter K.S.: *2001 IT spendine and staff survey results*. Gartner, 2001.
- [8] Gomolski: *2002 IT spendine and staff survey results*. Gartner, 2002.
- [9] CNEL: *Retribuzioni orario e costo del lavoro*. 2003.
- [10] CNIPA: *Lo stato dell'informatizzazione nella Pubblica Amministrazione*. Relazione Annuale 2002.
- [11] ISTAT: *Salute e sanità*. 2000.
- [12] Ministero dell'Economia e delle Finanze: *Relazione Generale sulla Situazione Economica del Paese*. 2002.
- [13] Ministero della Salute: *Annuario statistico del Servizio Sanitario Nazionale*. Direzione Generale del Sistema Informativo, Ufficio di Statistica, Dicembre 2002.
- [14] Ministero della Salute 2003a: *Compendio del Servizio Sanitario Nazionale 1998-2001*. Direzione Generale del Sistema Informativo, Ufficio di Statistica, 2003.
- [15] Ministero della Salute 2003b: *Attività gestionali ed economiche delle ASL e AO: Annuario statistico del servizio sanitario nazionale, Anno 2002*. Direzione Generale del Sistema Informativo; Ufficio di Statistica, Ottobre 2003.
- [16] OECD: *Health Data*. 2003.
- [17] Pessina E.A., Cantù E.: *Rapporto Oasis 2003: L'aziendalizzazione della sanità in Italia*. EGEA 2003.

PIER FRANCO CAMUSSONE Professore di "Organizzazione e sistemi informativi" presso l'Università di Trento. Direttore dell'Area Sistemi Informativi della Scuola di Direzione Aziendale (SDA) della Bocconi. Membro di comitati scientifici di diverse riviste (tra cui *Economia e Management*, *Mondo Digitale*). Autore di numerose pubblicazioni.
pfcamussone@unibocconi.it

NANOTECNOLOGIE E ICT: POTENZIALITÀ E PROSPETTIVE

Il rapido sviluppo delle nanotecnologie apre vie alternative alla miniaturizzazione dei circuiti integrati schiudendo nuovi orizzonti al settore ICT. Queste prospettive sono inquadrare in questo articolo evidenziando concetti, metodiche e strumenti peculiari delle nanotecnologie, mettendo in luce i materiali più promettenti e innovativi. Si riportano, inoltre, alcuni esempi di nanodispositivi e nanostrutture particolarmente idonee alla realizzazione di nuovi paradigmi computazionali.

1. COSA SONO LE NANOTECNOLOGIE?

Il termine “nanotecnologie” sta acquistando sempre maggiore popolarità: esso è, infatti, impiegato per descrivere una varietà di campi di ricerca e sviluppo, spesso di carattere interdisciplinare, entro i quali ci si confronta con strutture aventi dimensioni caratteristiche inferiori a 100 nm. Si potrebbe dire che riguarda lo studio e la manipolazione di “oggetti piccoli”, da 0.1 a 100 nm. La scala spaziale caratteristica delle nanotecnologie è allora il *nanometro* (un milionesimo di millimetro, tre ordini di grandezza inferiore rispetto al *micron* che è l'unità di riferimento tradizionale per la microelettronica). Per dare un'idea si consideri che è confrontabile con la larghezza del DNA (circa 2.5 nm) ed è la lunghezza di una catena lineare costituita da 6 atomi di carbonio.

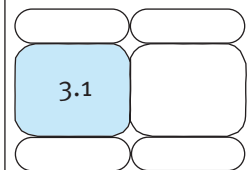
Il “nanomondo” è allora popolato da oggetti come atomi, molecole, “macchine molecolari” come il *ribosoma* (un organo cellulare che produce proteine nel corpo umano) che possono opportunamente essere assemblati in

“nanostrutture”: gli obiettivi principali delle nanotecnologie sono appunto realizzare, studiare e sfruttare le nanostrutture.

L'intuizione che si potesse giungere a manipolare e posizionare addirittura singoli atomi e molecole a questa “nanoscala” risale storicamente al fisico teorico Richard Feynman, quando, nel 1959, espose una famosa relazione dal titolo “There's Plenty of Room at the Bottom” [7] (ovvero “C'è abbondanza di spazio là sotto”) al congresso annuale della American Physical Society. In quella occasione affermò che “i principi della fisica non sono contro la possibilità di manipolare le cose un atomo alla volta...è un qualcosa che può essere fatto”, e per illustrarne l'impatto proponeva di scrivere l'intero contenuto dei 24 volumi dell'Enciclopedia Britannica sulla punta di uno spillo! Tuttavia è solo negli anni '80 che si registra il vero impulso alle nanotecnologie con l'invenzione del primo microscopio a effetto *tunnel* (STM, *Scanning Tunneling Microscope*) da parte di Binnig e Rohrer mediante il quale si riesce a “osservare” la materia fino ai singoli atomi [4]. Nel 1990,



Ermanno Di Zitti
Davide Ricci
Daniele D. Caviglia





una superficie di oro formando forti legami covalenti con essa. Controllando il movimento della punta sulla superficie si può sfruttare una goccia d'acqua come canale per far migrare le molecole dalla punta al campione, ottenendo un processo analogo alla scrittura con una penna a inchiostro. All'interno delle nanotecnologie esiste, però, una metodologia alternativa che sta sempre più emergendo negli ultimi anni e che di fatto

è quella più caratterizzante: costruire dal basso verso l'alto (*bottom-up*). Rappresenta il tentativo di costruire entità complesse sfruttando le capacità di **autoassemblamento** o di **autoorganizzazione** dei sistemi molecolari. È pertanto un approccio di tipo chimico o biologico, potenzialmente in grado, a differenza delle tecniche fotolitografiche, di creare strutture tridimensionali complesse a basso costo e in grande quantità.

Autoassemblamento: verso un "kit" di costruzione dal piccolo

Per poter seriamente prendere in considerazione da un punto di vista commerciale la possibilità di realizzare calcolatori operanti su scala atomica o molecolare, si rende necessario mettere a punto metodi di fabbricazione a basso costo di componenti su scala nanometrica con tolleranze a livello atomico.

Le tecniche più promettenti per raggiungere questo obiettivo si basano su proprietà a livello atomico e molecolare spesso definite "auto-organizzazione" o "autoassemblamento".

In questo contesto, si può intendere autoassemblamento come la costruzione automatica di sistemi ordinati o di strutture complesse a partire da blocchi elementari attraverso un processo guidato dalle caratteristiche fisico-chimiche dei blocchi e dalla minimizzazione dell'energia complessiva del sistema, processo che termina con il raggiungimento dell'equilibrio termodinamico.

È possibile distinguere essenzialmente tre diversi tipi di autoassemblamento, ciascuno legato al tipo di costituenti elementari, vale a dire atomi, molecole e aggregati (*cluster*): autoassemblamento chimico, fisico e colloidale.

L'auto-assemblamento chimico interviene nell'ordinamento su scala molecolare di composti aventi architetture atomiche esattamente predefinite per la realizzazione di strutture funzionali con un livello dimensionale superiore.

Una caratteristica attraente di tale tecnica consiste nel fatto che, controllando la sintesi chimica delle unità elementari, si possono sintonizzare la struttura e le proprietà del sistema ordinato risultante. Le strutture autoassemblate chimicamente annoverano i seguenti vantaggi:

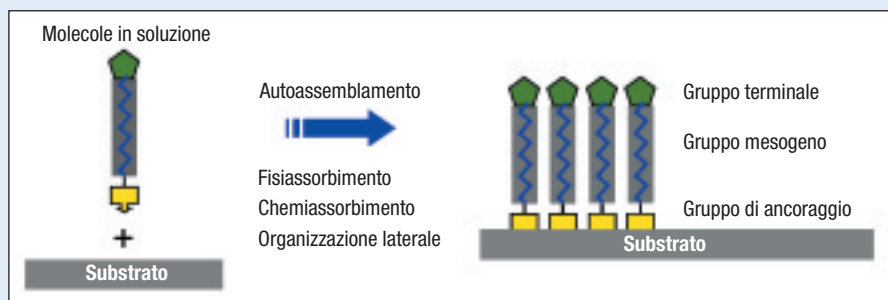
- ✓ sono facili da preparare e veloci da fabbricare a partire da soluzioni di molecole assemblanti (processo altamente parallelo);
- ✓ sono ordinate a livello molecolare e "robuste" in diverse condizioni di utilizzo;
- ✓ sono termodinamicamente stabili, si formano spontaneamente e tendono, quindi, a rigettare difetti;

Un esempio di autoassemblamento chimico è illustrato nella figura.

L'autoassemblamento fisico viene spesso riferito al comportamento auto-organizzante a livello nanometrico di materiali inorganici durante processi di deposizione (come l'epitassia da fascio molecolare, MBE, *Molecular Beam Epitaxy* o la deposizione in fase vapore, CVD, *Chemical Vapour Deposition*) o trattamenti successivi all'accrescimento. In netto contrasto con gli obiettivi dei processi tradizionali, che mirano a ottenere strati il più possibile continui e uniformi, queste tecniche sfruttano il comportamento dei materiali quando vengono depositati in condizioni disomogenee, quando cioè "imperfezioni" indotte ad arte su scala atomica nel substrato (vacanze atomiche, gradini, dislocazioni, disallineamenti reticolari ecc.) guidano la localizzazione o l'aggregazione delle specie chimiche depositate successivamente, dando luogo a proprietà nuove e utili ai fini delle applicazioni nanoelettroniche. Da questo tipo di processi sono state create strutture come i *quantum dot* (a tal proposito si rimanda alla lettura del riquadro su questo argomento) e *atomic wire*.

Inoltre, è da puntualizzare come, dall'associazione (contestuale o sequenziale) dell'autoassemblamento fisico con quello chimico, è stato possibile creare strutture *mono*-, *bi*- e *tri*-dimensionali su scala nanometrica aventi proprietà innovative. Gli esempi sono numerosi, quelli più noti e significativi riguardano i nanotubi di carbonio (si veda il riquadro sull'argomento) e i "nanofili" di silicio (*silicon nanowire*), ma vale la pena citare anche nanotubi di boro, di disolfuro di molibdeno, cluster metallici da aerosol ecc..

In ultimo, si cita quello che si potrebbe definire **l'autoassemblamento colloidale**, capace di produrre nanoparticelle o aggregati (*cluster*) di diversi materiali aventi dimensioni caratteristiche comprese tra 1 e 20 nm, e che manifestano proprietà di tipo ibrido, che si collocano a cavallo tra le proprietà dei singoli atomi e le proprietà di tipo "bulk" di una struttura a stato solido. In questo ambito, si ottengono significative proprietà quantistiche e di confinamento della carica per la realizzazione di nanodispositivi. Esempi significativi sono la fabbricazione di nanoparticelle semiconduttrici, conduttrici, organo-metalliche, ferromagnetiche. Partendo poi da queste si possono ottenere, su di una scala dimensionale superiore, sistemi organizzati di nanoparticelle in matrice organica isolante che possono essere parte integrante di vere e proprie nanoarchitetture autoassemblate (si veda la sezione 5.2).



Le nanotecnologie *bottom-up* nascono dal progetto di molecole o aggregati molecolari (eventualmente sintetizzati *ad hoc*) che hanno la capacità di autoassemblarsi o autoorganizzarsi in strutture di ordine più elevato (quali strutture mesoscopiche o macromolecolari). Un tale approccio può essere molto efficace impiegando le molecole che siano maggiormente adatte ad autoassemblarsi spontaneamente sotto l'azione di uno specifico agente chimico o fisico, come la variazione di pH, la concentrazione di uno specifico soluto, o l'applicazione di un campo elettrico. I meccanismi fisici che producono l'autoassemblamento, ovvero le forze pilotanti che spingono le molecole ad autoassemblarsi in strutture organizzate, sono dovuti alla termodinamica e alle interazioni competitive di tipo molecolare che comprendono forze idrofobiche/idrofiliche, legami a ponte d'idrogeno e interazioni di van der Waals, che agiscono in modo da minimizzare gli stati energetici per diverse configurazioni molecolari. È importante tenere presente che l'obiettivo diventa, allora, quello di progettare sistemi capaci di autoassemblarsi in strutture macroscopiche di ordine più elevato che presentano le proprietà di tipo chimico e/o fisico desiderate, caratteristiche del comportamento collettivo e non di quello delle singole molecole o particelle costituenti. Attualmente, questa metodologia è chiamata, "autoassemblamento guidato" (*directed self-assembly*). Il riquadro sull'autoassemblamento

illustra le principali metodiche impiegate a tal fine.

Le due metodologie descritte, che trovano applicazione nelle nanotecnologie, non sono in contrapposizione, anzi attualmente lo sforzo maggiore è ricercare tra esse la sinergia più adeguata per l'applicazione desiderata. Un esempio è la cosiddetta litografia "sofice" (*soft-lithography*) [25] in cui si usa uno "stampo" realizzato con tecnologia a fascio di elettroni per poi trasferire pattern di molecole autoassemblate sulle zone desiderate del substrato.

La figura 1 riassume quanto detto e illustra come nel corso del tempo le due metodologie convergano nelle nanotecnologie da ambiti molto diversi. Si può allora comprendere come la "visione" delle nanotecnologie sia interdisciplinare e tanto articolata. Si è così coniato il termine "nanoscienze", che costituisce il punto di incontro di discipline diverse che vanno dalla fisica quantistica alla chimica supramolecolare, dalla scienza dei materiali alla biologia molecolare, e rappresentano una realtà ormai affermata nel mondo della ricerca.

2.2. Lavorare a "nanoscala"

Nell'approccio peculiare *bottom-up* si parte dall'introduzione di tipologie di nanoblocchi elementari (i "mattoni" che costituiscono un sistema) che presentano proprietà fisiche innovative a seconda della dimensione, della morfologia e della composizione chimica.

A questo proposito, nasce l'esigenza di comprendere le proprietà fondamentali delle nuove nanostrutture in modo da poter ricavare le descrizioni funzionali di ogni "nanocella". È così possibile mettere in relazione ogni nuovo concetto di nanodispositivo/nanocella con quelli noti e resi disponibili dalla tecnologia dei circuiti integrati in modo da poterli combinare efficacemente per progettare sistemi integrati di tipo innovativo. La figura 2 illustra la strutturazione di un tale approccio, dove i concetti di progetto architeturale e assemblamento gerarchico sono situati all'estremo opposto rispetto alla metodologia di progettazione strutturata VLSI, *Very Large Scale Integration*. Per quest'ultima, infatti, l'approccio top-down si traduce in una specifica di tipo algoritmico-comporta-

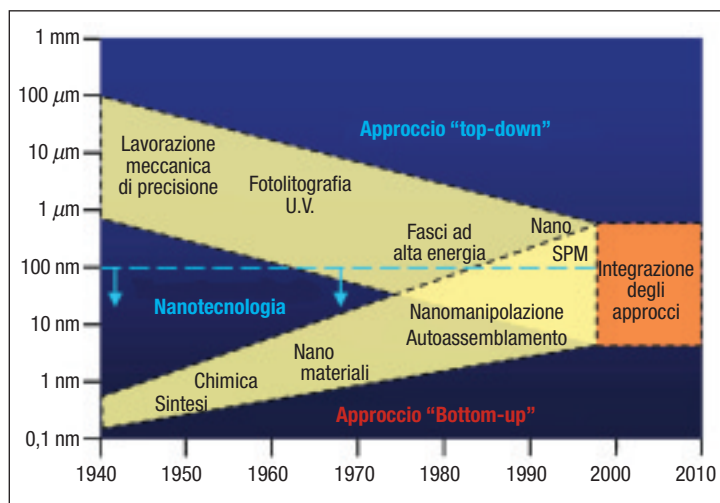


FIGURA 1

La sintesi tra metodologie top-down e bottom-up presenti nelle nanotecnologie

mentale che gerarchicamente si scompone nei livelli funzionali e strutturali (librerie di celle) che a loro volta a livello più basso si traducono nei livelli logici e circuitali fino ad arrivare alla realizzazione fisica vera e propria (*layout fisico*). È necessario allora comprendere preliminarmente le conseguenze di lavorare alla nanoscala: che cosa accade alle proprietà di un oggetto man mano che ne diminuiscono le dimensioni fino ad arrivare a quelle del nanometro?

Un parametro importante è il rapporto tra l'Area della superficie e il Volume dell'oggetto. Negli oggetti macroscopici, che hanno un piccolo rapporto Area/Volume, le proprietà chimiche e fisiche sono essenzialmente determinate dalla struttura del solido (*bulk*). Negli oggetti più piccoli, con un elevato rapporto Area/Volume, le caratteristiche della superficie diventano determinanti, influenzando fortemente le proprietà ottiche, l'adesività, la bagnabilità, l'attrito, la reattività chimica ecc.. Per comprendere numericamente su quale scala dimensionale ciò avviene, si introduce il concetto di "dispersione", ovvero il rapporto tra il numero di atomi della superficie rispetto al numero di atomi totali in una particella. Si osservi l'andamento della dispersione al variare del raggio della particella riportato in figura 3: al di sotto di un raggio pari a 1 nm, si può notare come il numero di atomi della superficie diventi confrontabile al numero di atomi totali, indicando chiaramente come la fisica e chimica delle nanotecnologie siano dominate dalle proprietà delle superfici.

È da tenere, allora, presente che uno dei maggiori ostacoli allo sviluppo delle applicazioni delle nanotecnologie risiede nelle inerenti scale di tipo spaziale e temporale secondo le quali si svolgono gli eventi chimico/fisici di nostro interesse. Esse, infatti, variano rispettivamente dal nanometro a centinaia di nanometri e dal nanosecondo al femtosecondo¹ (se si considerano le oscillazioni dei legami atomici). Essendo i blocchi molecolari manipolati su una scala così ridotta, i substrati

hanno livelli di organizzazione spaziale e/o temporale su svariati ordini di grandezza, con livelli di ordine inferiore annidati su livelli di ordine più elevato (per esempio, 6 ordini di grandezza per un dispositivo che sia strutturato su scala nanometrica ma con una struttura macroscopica sul millimetro). Pertanto, al fine di studiare ed esplorare la ricchezza e la complessità di questi sistemi, si richiedono strumenti altamente sofisticati, sia di tipo teorico che sperimentale.

2.3. Quali strumenti?

La visualizzazione, caratterizzazione e manipolazione di materiali e dispositivi a scala nanometrica richiedono sofisticate tecniche quantitative e di *imaging* con risoluzioni spaziali che vanno dal micron fino sotto il livello

FIGURA 2

La visione delle nanotecnologie: un paradigma per costruire dal basso verso l'alto

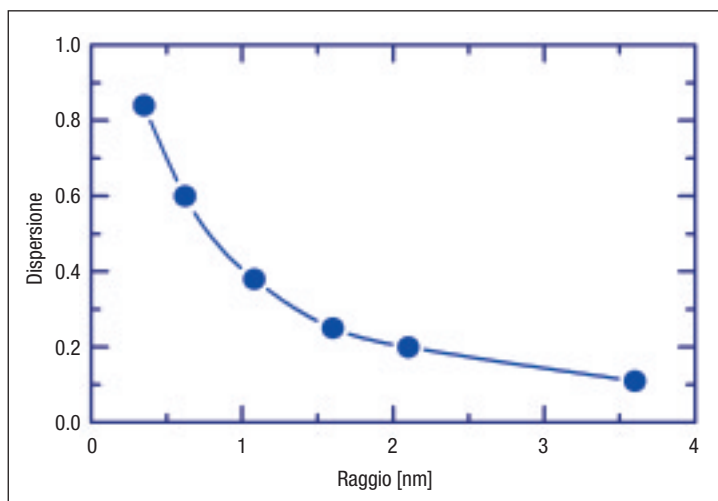
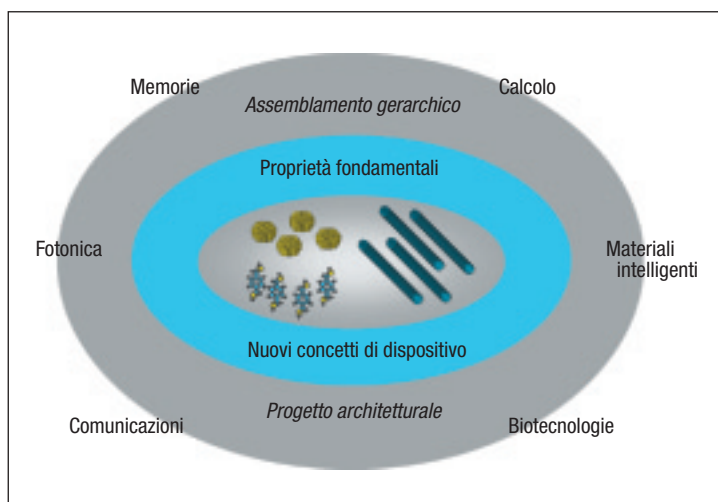


FIGURA 3

Andamento della dispersione al variare del raggio di una nanoparticella

¹ Un femtosecondo, 10^{-15} s, è 6 ordini di grandezza inferiore al nanosecondo, risultando in una unità estremamente piccola.

molecolare. Le informazioni ottenute alle diverse scale dimensionali possono così essere messe in relazione tra loro per poter comprendere le proprietà delle strutture a nanoscala e come queste possano essere convenientemente sfruttate a livello macroscopico. Esempi importanti di strumenti/tecniche al momento disponibili sono:

■ sorgenti di luce di sincrotrone altamente focalizzate ($1-2\ \mu\text{m}$) per la diffrazione dei raggi X e le tecniche relative in grado di fornire dettagliate informazioni di tipo strutturale sondando direttamente il posizionamento degli atomi [10];

■ le microscopie a scansione elettronica (SEM, *Scanning Electron Microscopy*), a trasmissione (TEM, *Transmission Electron Microscopy*) e a sonda di scansione (SPM, *Scanning Probe Microscopy*), che consentono visualizzazioni topografiche bi- e tridimensionali di strutture a nanoscala;

■ tecniche di monitoraggio *in situ* che consentono di monitorare e valutare l'assemblamento e la crescita dei blocchi costitutivi elementari, quali la riflessione da diffrazione di elettroni ad alta energia (RHEED, *Reflection High-Energy Electron Diffraction*) [19]. Il principio sottostante è indirizzare un fascio di elettroni ad alta energia verso un campione a una certa angolatura, e rilevare come esso rimbalzi indietro dopo l'interazione con la struttura atomica del materiale sotto test. Dalla conformazione dei diversi pattern di elettroni riflessi che si producono su uno schermo a fosfori si possono evincere informazioni sulla struttura atomica del campione.

3. I NANOMATERIALI

Lo sviluppo delle tecnologie dell'informazione negli ultimi quarant'anni è avvenuto impiegando per i componenti attivi materiali inorganici, tra i quali il silicio ha assunto il ruolo di protagonista. Da un lato, si prevede che tale ruolo guida permanga ancora nel breve e medio termine, dall'altro esiste la concreta possibilità che il silicio possa essere progressivamente rimpiazzato da materiali di tipo organico per realizzare dispositivi particolari. La situazione che si va delineando per il lungo termine riguarda piuttosto l'impiego di silicio e materiali organici per lo sviluppo di tecnologie di tipo ibrido organico-silicio. Il fatto che rende così interessante l'utilizzo di materiali organici è la loro grande varietà e disponibilità, la possibilità di sintetizzare o funzionalizzare molecole con proprietà *ad hoc* per usi specifici e i bassi costi di produzione. Il fattore economico è, infatti, quello che appare essere la vera limitazione allo sviluppo della tecnologia dei circuiti integrati al silicio (Seconda legge di Moore [3]), per cui la sfida è sviluppare una tecnologia a basso costo basata sull'impiego di nanomateriali. Questo campo è attualmente di competenza della "nanochimica" che comprende la sintesi di nanoparticelle e di materiali nanocristallini nonché la fabbricazione di strutture policristalline a singola fase o multifase aventi dimensione caratteristica dei "grani" da 1 a 100 nm almeno in una dimensione. A seconda della dimensionalità si definiscono alcune tipologie tipiche (Tabella 1) quali:

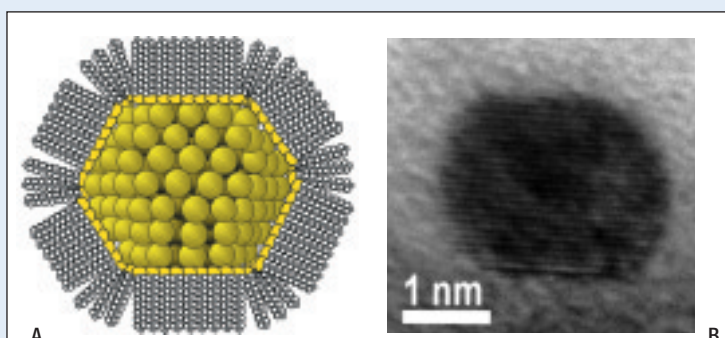
TABELLA 1
Classificazione
dei nanomateriali
al variare
della dimensionalità

Nanomateriali tipici	Dimensione	Materiali
(0-D) Nanocristalli e cluster (quantum dot)	diametro 1-10 nm	Metalli, semiconduttori, materiali magnetici
Altre nanoparticelle	diametro 1-100 nm	Ossidi ceramici
(1-D) Nanowire	diametro 1-100 nm	Metalli, semiconduttori, ossidi
Nanotubi	diametro 1-100 nm	Carbonio
(2-D) Matrici di nanoparticelle	svariati nm^2 - μm^2	Metalli, semiconduttori, materiali magnetici
Superfici e film sottili	Spessori 1-1000 nm	Materiali vari (organici e non)
(3-D) Strutture tridimensionali	diversi nm nelle 3 dimensioni	Metalli, semiconduttori, materiali magnetici

I nanocristalli

I nanocristalli sono aggregati cristallini il cui diametro arriva al massimo fino a qualche nanometro e sono costituiti da qualche decina a qualche migliaio di atomi. Possono essere sia nanocristalli semiconduttivi come quelli di silicio e germanio usati tipicamente in una matrice di SiO_2 , che nanocristalli di tipo colloidale. Questi ultimi possono essere sintetizzati a partire da materiali metallici come oro, argento o cobalto, da semiconduttori come solfuro o seleniuro di cadmio, arseniuro di gallio e da isolanti come ossido di ferro o di titanio. I nanocristalli colloidali sono dispersi in un solvente e devono essere stabilizzati in modo che non si formino agglomerati dovuti a fenomeni di *coalescenza*¹. Di particolare interesse sono i nanocristalli colloidali costituiti da nanoparticelle metalliche incapsulate da un monostrato di molecole organiche, per la loro capacità di dare origine a strutture autoorganizzate a una, due o tre dimensioni. Infatti, volendo realizzare strutture reticolari di nanocristalli, il processo può essere ottimizzato per modulare le proprietà di aggregazione e di idrofobicità in modo da ottenere la spaziatura desiderata tra i singoli nanocristalli. A titolo

di esempio si riporta nella figura a sinistra (Figura A) la rappresentazione di una nanoparticella di oro di nucleo (*core*) di diametro 2.4 nm ricoperta da un monostrato di dodecanetiolo, la cui molecola presenta una struttura con la forma di un ottaedro troncato contenente 459 atomi di oro. Nella figura a destra (Figura B) viene mostrata, invece, l'immagine corrispondente ad alta risoluzione ottenuta al microscopio a trasmissione elettronica (TEM). La qualità del processo di produzione dei nanocristalli è determinata principalmente dalla sua capacità di dare luogo a una distribuzione della loro dimensione la più stretta possibile. A tal fine, o si impiegano tecniche altamente complesse per la sintesi diretta di nanocristalli monodispersi oppure, una volta sintetizzati con metodiche più semplici, si devono approntare dei processi di selezione dei nanocristalli con svariate tecniche (*vapor transfer*, cromatografia) fino a ottenere la distribuzione desiderata.



A Schema rappresentativo di una nanoparticella di oro (diametro del nucleo 2.4 nm) ricoperta da un monostrato di dodecanetiolo (per un diametro complessivo di circa 5 nm)
B Immagine TEM ad alta risoluzione di un nanocristallo. La struttura cristallina della nanoparticella risulta evidente dalle frange di interferenza elettronica

¹ Fenomeno di aggregazione originato da fenomeni diffusivi superficiali per i quali piccoli agglomerati in fase dispersa si fondono tra loro per costituire aggregati più grandi o, al limite, una fase continua.

I nanocristalli (si veda il riquadro sull'argomento);

I strutture a filamento;

I strutture stratificate (monostrato o multistrati) o lamellari;

I materiali nanostrutturati (materiali i cui elementi costitutivi quali particelle, aggregati o cavità hanno almeno una dimensione inferiore a 100 nm).

Le prime tre sono quelle più interessanti per applicazioni nell'ambito delle tecnologie dell'informazione mentre i materiali nanostrutturati possono essere utilizzati sia sotto forma di manufatti o polveri finemente divise sia come pellicole sottili (*thin film*) per rivestire superfici di materiali convenzionali.

Questi materiali, a prima vista, non differiscono dai materiali strutturati su scala convenzionale (per esempio un pezzo metallico o ceramico dove le dimensioni lineari dei "grani" costituenti possono variare tipicamente tra 0,01 e 0,1 mm), ma le loro proprietà vengono modificate manipolando le loro strutture a livello delle particelle che la compongono. Per fare alcuni esempi, le ceramiche nanostrutturate sono molto più dure e resistenti di quelle a grana grossa, metalli a nanofase hanno proprietà meccaniche migliori di quelli convenzionali, e la granulometria nanometrica presente in alcuni cosmetici rende molto più efficace l'azione di schermo nei confronti delle radiazioni solari.

Ricordando quanto si è discusso in precedenza sull'elevato rapporto tra superficie/interfaccia e volume, i nanomateriali possono essere progettati in modo da sfruttare specifici effetti legati alla dimensione (controllata) delle particelle presenti.

Si possono così evidenziare i fatti più salienti

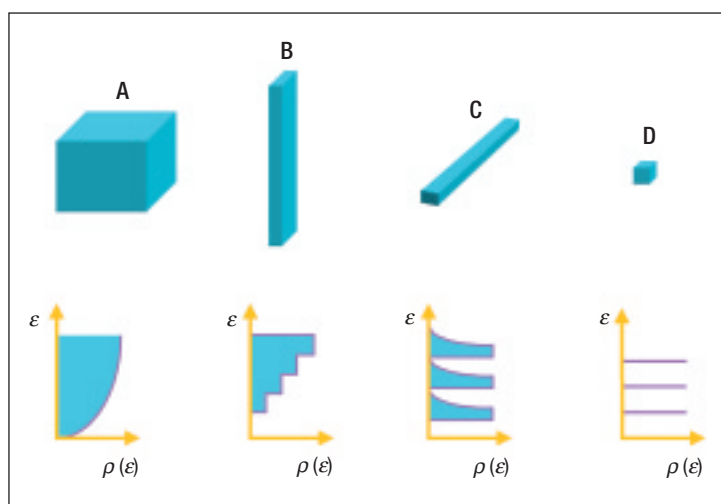
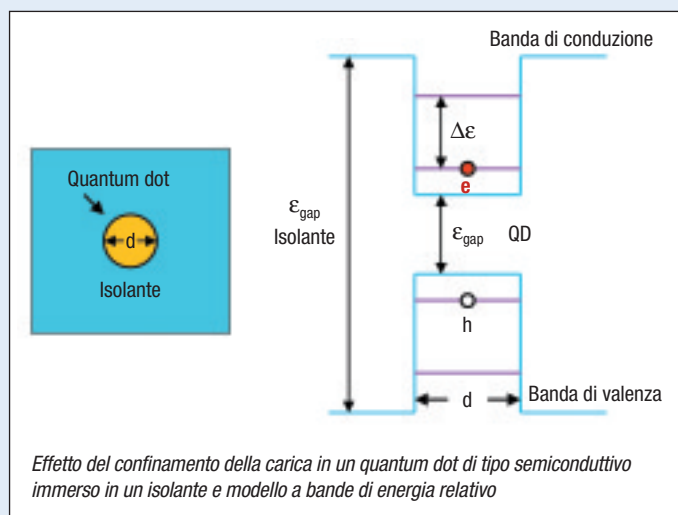


FIGURA 4

Densità di stati e confinamento quantico per diverse tipologie di nanostrutture. A Strutture 3-D tipo bulk; B pozzi quantici (2-D); C fili quantici (1-D); D punti quantici (0-D)

Quantum dot

A causa della loro piccola dimensione, che è dello stesso ordine di grandezza della lunghezza d'onda di De Broglie di elettroni e lacune a temperatura ambiente, gli stati energetici dei portatori di carica nei nanocristalli sono quantizzati. Per nanocristalli sferici in cui elettroni e lacune sono confinati in tutte e tre le dimensioni il movimento dei portatori è completamente governato dalla meccanica quantistica, ed è per questo motivo che i nanocristalli sono chiamati quantum dot. Analogamente al modello della buca di potenziale, solitamente associato alle *quantum well*, si ottengono livelli elettronici discreti $\Delta\epsilon$ la cui spaziatura è funzione (inversamente proporzionale al quadrato) della dimensione d del nanocristallo. Come evidenziato nella figura, si può in questo modo modulare la variazione di energia associata alla ricombinazione di un elettrone della banda di conduzione con una lacuna della banda di valenza, potendo ottenere emissione di fotoni alla lunghezza d'onda desiderata. Un quantum dot ha anche un'altra caratteristica, chiamata *energia di caricamento*, analoga all'energia di ionizzazione di un atomo: rappresenta l'energia di repulsione elettrostatica dovuta alla carica all'interno del dot. Per tali similitudini con gli atomi reali i quantum dot sono anche detti *atomi artificiali*.



in relazione alle applicazioni possibili per le tecnologie dell'informazione.

□ L'effetto di *confinamento quantico* dovuto alla diversa forma della funzione densità di stati elettronici (Figura 4) genera una discretizzazione dei livelli energetici degli elettroni nelle relative nanostrutture 2-D (pozzi quantici o *quantum well*), 1-D (fili quantici o *quantum wire*) e 0-D (punti quantici o *quantum dot*) e trova i principali sbocchi applicativi nella progettazione di eterostrutture² a semiconduttore, nell'optoelettronica e nell'ottica non lineare.

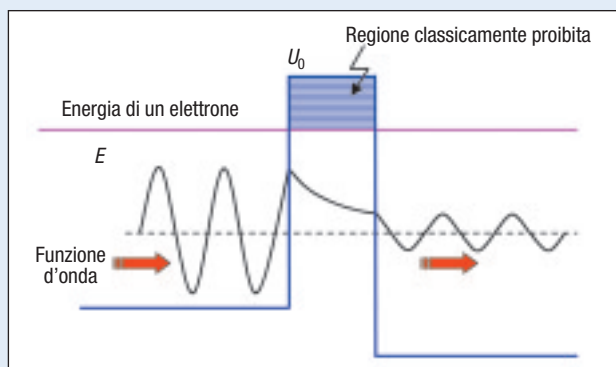
□ Le proprietà di *tunneling* (si veda il riquadro sull'**effetto tunnel**) possono essere modulate variando la lunghezza delle catene molecolari di tipo isolante disposte tra nanoparticelle/molecole conduttive; in questo modo, analogamente al caso delle eterostrutture realizzate con semiconduttori composti, si possono realizzare dispositivi e nanostrutture quantistiche.

□ A differenza dei materiali macroscopici che presentano domini magnetici multipli, tante piccole particelle ferromagnetiche formano un singolo dominio magnetico, potendo così realizzare nuovi tipi di memorie.

² Strutture cristalline formate dalla deposizione stratificata di regioni di semiconduttori a diverso gap come per esempio sequenze di GaAs e GaAlAs.

L'effetto tunnel

Secondo la fisica classica una particella di energia E inferiore all'altezza U_0 di una barriera di potenziale non può penetrarla in quanto la regione interna alla barriera è classicamente proibita. La natura quantistica di una particella incidente la barriera, descritta da una *funzione d'onda* (per esempio, un elettrone), fa sì che se ne osservi un decadimento di tipo esponenziale all'interno della barriera, che dà luogo a una probabilità finita di penetrazione, apprezzabile per spessori dell'ordine del nanometro. Tale fenomeno è chiamato effetto tunnel.



4. NANOELETTRONICA E NANODISPOSITIVI

Le particolari proprietà fisiche e quantistiche rese disponibili dai nanomateriali hanno consentito il rapido sviluppo della nanoelettronica che attualmente riguarda lo sfruttamento delle proprietà di trasporto elettronico sia nelle nanostrutture a stato solido che in quelle di tipo molecolare o ibrido. Data la vastità e la complessità dell'argomento, che ha portato alla formulazione di una *roadmap* per la nanoelettronica [24] analoga a quella per le tecnologie dei semiconduttori, si vuole qui porre l'accento solo sui filoni principali che contengono gli archetipi dei materiali e i relativi dispositivi emergenti lasciando da parte l'evoluzione delle tecnologie MOS (*Metallo-Ossido-Semiconduttore*), con le quali è prodotta una grandissima parte dei circuiti integrati oggi sul mercato (e delle relative nanotecnologie che fanno riferimento al mondo del silicio), che rimangono comunque il riferimento con il quale confrontarsi. In tale contesto i nanocristalli [1] e i nanotubi di carbonio (nanostrutture unidimensionali a filamento) sono di particolare interesse per le tecnologie dell'informazione a causa della loro versatilità e flessibilità di impiego. I nanocristalli sono direttamente associabili alle strutture quasi zero-dimensionali dette quantum dot (si veda il riquadro), che, presentando livelli energetici quantizzati e spaziali in funzione della dimensione caratteristica, possono essere impiegati per esempio per emettere o assorbire una data lunghezza d'onda della luce variando il diametro del *dot*. Più pic-

colo è il nanocristallo, più grande è la spaziatura dei livelli, e il conseguente maggior *energy gap* consente di lavorare a lunghezze d'onda inferiori. Per esempio, un nanocristallo di CdSe di 2.5 nm ha fluorescenza sul verde, mentre uno di 7 nm la presenta sul rosso. Oltre che per applicazioni di tipo optoelettronico, tali quantum dot si prestano a essere impiegati innanzitutto per la realizzazione di memorie non convenzionali [24], in quanto un dot o un cluster di dot può essere impiegato per realizzare memorie non volatili³. Tali dispositivi sono così un esempio significativo di convergenza di tecnologie microelettroniche esistenti con nanotecnologie autoassemblanti.

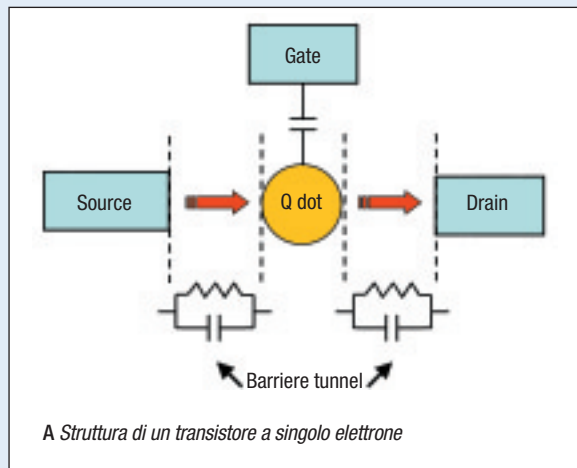
Un'altra applicazione importante riguarda la realizzazione di dispositivi a singolo elettrone [17] (si veda a tal proposito il riquadro sulla **Coulomb blockade e i dispositivi a singolo elettrone**), particolarmente attraenti in quanto conservano la scalabilità fino al livello atomico, potendo controllare il movimento anche di un solo elettrone. L'idea sottostante a questi dispositivi è quella di sfruttare un quantum dot come una piccola "isola" conduttiva separata da due elettrodi mediante sottili zone isolanti attraverso le quali un elettrone possa passare per effetto tunnel. In

³ La presenza di una carica (elettroni) aggiunta dall'esterno può impedire il passaggio di corrente in un dispositivo, creando così due possibili condizioni di risposta a seguito dell'applicazione di una tensione di lettura, in modo analogo a quanto accade in una comune memoria non volatile a tecnologia convenzionale (memoria *flash*).

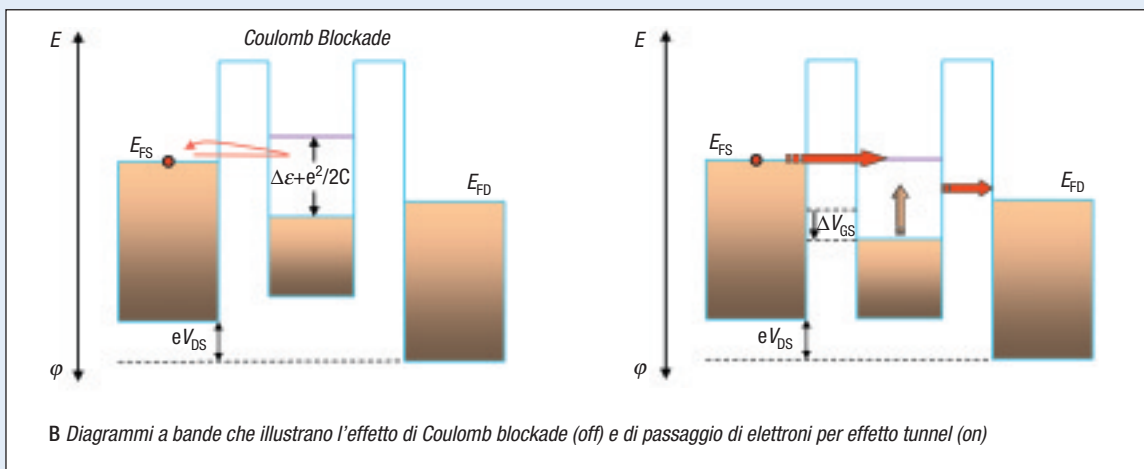
La Coulomb blockade e i dispositivi a singolo elettrone

In figura A è rappresentata la struttura tipica di un transistor a singolo elettrone. Si tratta di un'isola conduttiva (Q dot) separata da due elettrodi (source e drain) mediante due barriere tunnel e accoppiata a un elettrodo di controllo (gate) mediante una capacità. L'energia di caricamento dell'isola vale $q^2/2C$, dove q è la carica dell'elettrone e C è la capacità totale tra l'isola e il resto del sistema, e rappresenta un gap di energia tra lo stato energetico dell'ultimo (N -esimo) elettrone e il primo successivo stato disponibile ($N+1$). Se la dimensione dell'isola è sufficientemente piccola da far sì che questo valore diventi molto superiore all'energia termica nel sistema (kT , ove k è la costante di Boltzmann e T la temperatura assoluta) ne consegue che nessun elettrone può passare per effetto tunnel nel dot o dal dot. Il valore di capacità necessario per osservare questo effetto a temperatura ambiente ($T = 300$ K) deve essere molto inferiore a 1 aF ($1 \text{ aF} = 10^{-18} \text{ F}$), corrispondente a una dimensione caratteristica di 10 nm. Per questo motivo un quantum dot realizzato con un nanocristallo di diametro dell'ordine del nanometro risulterebbe adeguato allo scopo, mentre con strutture a stato solido di dimensioni caratteristiche intorno a 100 nm si dovrebbe lavorare in condizioni di temperatura prossime allo zero assoluto, sotto 1 K.

Il diagramma a bande mostrato in figura B mostra come, nonostante la presenza di una polarizzazione V_{DS} , non ci siano nell'isola stati elettronici disponibili all'effetto tunnel, in quanto i soli stati liberi si trovano al di sopra dell'energia degli elettrodi. Questa condizione è detta *Coulomb blockade*. Tuttavia, se viene applicata una tensione a un elettrodo esterno di gate in modo da modificare la profondità del potenziale elettrostatico per accoppiamento capacitivo (in modo classico, in quanto lo spessore del dielettrico deve essere sufficiente da non consentire il passaggio di elettroni per effetto tunnel) fino a far allineare il livello libero per l'elettrone $N+1$ con quello dell'elettrodo di source, si verifica il passaggio di un elettrone alla volta attraverso l'isola. Per questo motivo il dispositivo conseguente è chiamato transistor a singolo elettrone.



A Struttura di un transistor a singolo elettrone



B Diagrammi a bande che illustrano l'effetto di Coulomb blockade (off) e di passaggio di elettroni per effetto tunnel (on)

figura 5 è mostrato lo schema di controllo di quello che può essere considerato un condensatore "quantistico": una tensione V_g applicata al gate fa sì che un elettrone del source guadagni sufficiente energia per passare nell'isola per effetto tunnel e così caricarla negativamente. A questo punto, si genera una forza repulsiva di tipo coulombiano che impedisce il passaggio di un secondo elettrone. Questo effetto è detto Coulomb blockade. L'utilizzo di questo effetto nelle memorie flash (a tal proposito si veda anche la nota precedente) potrebbe rivelarsi particolar-

mente attraente non solo per il fatto che esso consente di conservare più a lungo la carica immagazzinata, ma anche nella prospettiva di memorizzare più bit su una stessa cella (aumentando nello stesso tempo la densità di integrazione).

Dallo schema visto, aggiungendo una seconda giunzione tunnel che consenta il passaggio di elettroni da una parte all'altra del dispositivo, deriva il transistor a singolo elettrone (SET, *Single Electron Transistor*, figura A nel riquadro) e tutti i dispositivi che ne conseguono. Anche in questo caso si prefigurano tecno-

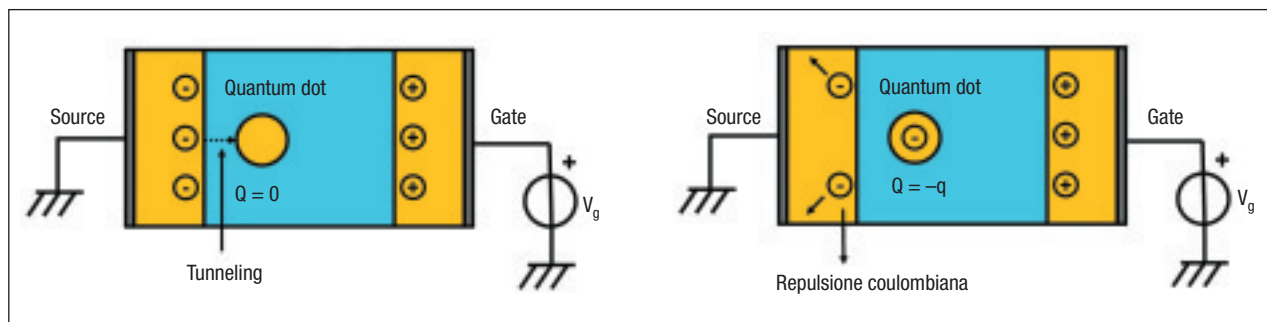


FIGURA 5

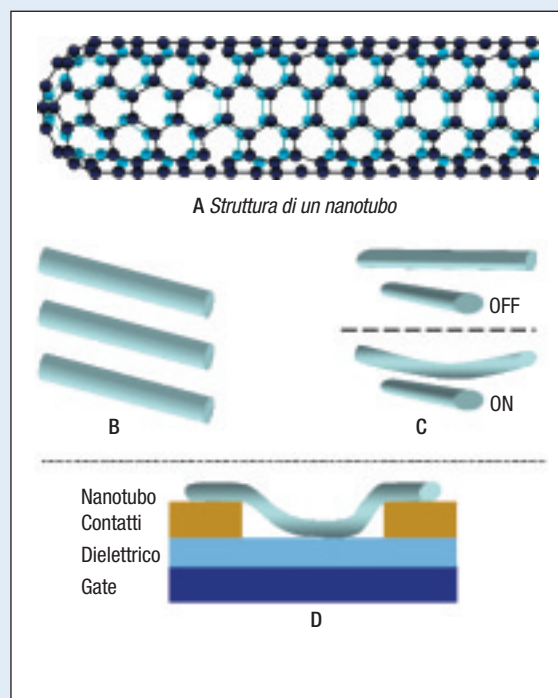
Schema concettuale di un condensatore "quantistico" che illustra la fase di carica di un quantum dot per effetto tunnel e l'effetto di blocco coulombiano dovuto alla repulsione elettrostatica

Nanotubi di carbonio per l'elaborazione dell'informazione

I nanotubi di carbonio sono nanostrutture filamentari scoperte nel 1991 da S. Iijima [13] studiando il materiale depositato su un catodo durante la sintesi dei fullereni. La scoperta di questa prima famiglia di molecole, terzo stato cristallino del carbonio, era già valsa ai ricercatori Kroto, Heath e Smalley il premio Nobel per la Chimica. I fullereni sono strutture di forma sferoidale, la più nota delle quali è il C_{60} (dove i 60 atomi di carbonio sono disposti esattamente come la geometria dei vertici degli esagoni e pentagoni di un pallone da calcio tradizionale). Iijima osservò che oltre ai fullereni durante il processo di sintesi in realtà si erano formate una ulteriore varietà di strutture mai osservate prima, strutture aventi un elevato rapporto lunghezza-diametro, i nanotubi.

I nanotubi sono strutture molto interessanti, che affascinano per le loro dimensioni (diametri dell'ordine di pochi atomi e lunghezze fino a qualche micron), per la loro forma, di fatto monodimensionale, e per le loro notevoli proprietà fisiche. Possono essere pensati come un singolo foglio di grafite (grafene) arrotolato per formare un cilindro, molto lungo e sottile, terminante con una semisfera di fullerene (Figura A). Una tale struttura corrisponde a un nanotubo a singola parete (SWNT, *Single-Wall carbon NanoTube*), mentre strutture multiple coassiali di SWNT originano i nanotubi a pareti multiple (MWNT, *Multi-Wall carbon NanoTube*).

In figura sono mostrati alcuni possibili impieghi dei nanotubi per le tecnologie dell'informazione: B Nanowire; C Unità di memoria attraverso "incroci"; D Transistore a effetto di campo a nanotubo, a configurazione di tipo backgate (avente gate in basso)



logie di tipo ibrido SET-MOS dove i transistori MOS provvedano a fornire più adeguati valori di guadagno e di pilotaggio di corrente, punti deboli dei SET. Potendo però usare nanocristalli al posto di quantum dot realizzati con tecnologie a stato solido si avrebbe il vantaggio della dimensione per usufruire di più facili condizioni operative. Infatti, la dimensione caratteristica del dot dell'ordine di pochi nanometri (1-3 nm) consente di far funzionare i dispositivi a temperatura ambiente, mentre avendo a disposizione quantum dot dell'ordine dei 100 nm si è costretti a lavorare a temperature estremamente basse a causa dei bassi

valori dell'energia di caricamento che deve restare significativamente superiore all'energia termica dell'ambiente, kT .

4.1. Dal silicio al carbonio: i nanotubi

È sorprendente constatare come dalla scoperta dei nanotubi di carbonio (CNT, *Carbon NanoTubes*) appena un decennio fa [13] si siano innescate così tante sinergiche attività di ricerca su queste nanostrutture (si rimanda al riquadro: **Nanotubi di carbonio per l'elaborazione dell'informazione**). I nanotubi hanno proprietà elettroniche particolari, potendo esibire un comportamento

metallico o semiconduttivo a seconda del loro diametro e dell'angolo secondo cui risulta avvolta la struttura esagonale del singolo foglio di grafite (grafene) che li compone (proprietà di *chiralità*). L'impiego naturale dei nanotubi metallici è la realizzazione di fili ultrasottili (*nanowire*), fatto particolarmente importante in quanto si potrebbero risolvere due fondamentali problemi nell'attuale miniaturizzazione dei circuiti, ovvero costruire piccole interconnessioni e con esse trasportare una grande quantità di corrente senza vaporizzare i fili stessi.

Come semiconduttori possono, invece, essere parte integrante di un transistor e dare origine a dispositivi di tipo ibrido carbonio-silicio compatibili con le attuali tecnologie microelettroniche. Inoltre, rispetto ad altri dispositivi di tipo quantistico (per esempio, diodi/transistori a tunneling risonante) o molecolare, una nanoelettronica basata sui nanotubi appare più promettente nell'ottica di costruire un computer. In tale ottica vale la pena considerare i seguenti aspetti:

▣ **Conduttori/interconnessioni:** un nanotubo è un dispositivo monodimensionale che conduce essenzialmente in superficie e presenta migliori proprietà di trasporto di carica rispetto ad una struttura a semiconduttore dove esse sono determinate dalla struttura cristallina di bulk. Infatti, in un nanotubo risulta considerevolmente ridotta la probabilità di *scattering* (rallentamento del movimento dei portatori di carica dovuto alle collisioni con la struttura cristallina), che aumenta la *mobilità* dei portatori. Il comportamento è, allora, quello di un eccellente conduttore adatto a realizzare le interconnessioni tra dispositivi. Si stima, inoltre, che un nanotubo a singola parete (SWNT) possa sopportare densità di corrente di tre ordini di grandezza superiori a quella del rame o dell'oro! Inoltre, un nanotubo è estremamente robusto dal punto di vista meccanico e, potendo essere reso corto fino a circa 0.4 nm [20], lo spazio per le interconnessioni può essere molto aumentato.

▣ **Interruttori/diodi:** dato che l'ampiezza della banda proibita (*bandgap*) di un nanotubo dipende dal suo diametro (maggiore è il diametro, più piccolo è il *bandgap*), mettere

insieme nanotubi di diverso diametro può dare origine a un diodo. Un modo alternativo è poi quello di drogare opportunamente un nanotubo realizzando così una giunzione pn.

▣ **FET/transistori:** è stata dimostrata la possibilità di realizzare transistori a effetto di campo con nanotubi, detti CNFET, *Carbon Nanotube Field Effect Transistors* [18]. Tali dispositivi, dal punto di vista della velocità intrinseca, appaiono superiori ai transistori MOS convenzionali per il valore effettivo di mobilità, almeno 20 volte più grande di quella del silicio. La progettazione della struttura è simile a quella di un MOSFET convenzionale, rimpiazzando il canale con un nanotubo. È anche possibile realizzare strutture di tipo *backgate* come illustrato nella figura D del riquadro a p. 13.

▣ **Unità di memoria:** i nanotubi possono formare "incroci" (*crossbar*) mediante l'applicazione di un campo elettrico. Dopo la rimozione di tale campo, sono le forze molecolari a mantenere la connessione, e si ottiene così un meccanismo per la memoria.

Si può allora comprendere come i nanotubi possano svolgere nell'ambito della nanoelettronica un ruolo simile - almeno per alcuni aspetti - a quello svolto nella microelettronica dal silicio, una volta che la loro tecnologia giunga a maturazione. Restano, infatti, alcuni problemi importanti dei nanotubi da risolvere, quali:

- ▮ controllo della chiralità;
- ▮ controllo della direzione di crescita e il loro posizionamento;
- ▮ resistenze di contatto piuttosto alte ($k\Omega$);
- ▮ tendenza a crescere piuttosto lunghi (mm).

5. LE NANOARCHITETTURE: VERSO NUOVI PARADIGMI COMPUTAZIONALI

I circuiti integrati convenzionali sono costituiti essenzialmente da dispositivi a commutazione e da interconnessioni: sono queste ultime, in realtà, a determinare l'occupazione di area sulla superficie del silicio, e quindi le dimensioni complessive del *chip*. Al diminuire delle dimensioni caratteristiche della tecnologia, esse costituiscono un fattore che sempre più limita le prestazioni dei circuiti. Questa visione, tipicamente le-



gata ai processi di miniaturizzazione, può essere superata nell'ambito delle potenzialità offerte dalle nanotecnologie se si cambia il modo di pensare un circuito a nanoscala, per il quale le tecniche di fabbricazione fanno riferimento a varie forme di autoassemblamento, coniando così il termine "autoassemblaggio" di circuiti. Esistono varie opportunità e sono state formulate diverse proposte in merito, non sempre però legate a un concetto di computer di tipo "general-purpose".

Prima che esistesse l'attuale disponibilità di singoli processori estremamente potenti, il tentativo di migliorare le prestazioni computazionali dei sistemi di elaborazione ha diretto la ricerca sulle architetture di tipo concorrente e/o ad elevato parallelismo nell'ottica della realizzazione VLSI. A partire dai concetti generali delle macchine SIMD e MIMD (*Single/Multiple Instruction Multiple Data*) sono state sviluppate architetture [11] quali *array* di processori, *array* sistolici e a fronte d'onda per la mappatura efficiente di specifici algoritmi. Di pari passo è proseguita la ricerca su modalità alternative per eseguire algoritmi prendendo spunto dalla biologia e dalle neuroscienze, introducendo paradigmi quali il DNA computing, le reti neurali artificiali e gli automi a celle [23]. Quanto però possono essere adatti i dispositivi nanoelettronici alla realizzazione di siffatte architetture? Nell'attuale scenario della nanoelettronica i noti concetti architettureali devono essere verificati o modificati. La verifica è necessaria in quanto i nuovi dispositivi nanoelettronici possiedono proprietà qualitativamente nuove in termini di dimensioni fisiche, dissipazione di potenza, guadagno, numero di stati stabili, frequenza di *clock* e soprattutto livello di integrazione. Lo sforzo sarà quello di ricercare la mappatura ottimale degli algoritmi sui nanosistemi disponibili o ideare nanosistemi innovativi per la soluzione degli algoritmi desiderati, dovendo contenere in entrambi i casi i costi rispetto alle soluzioni tradizionali per poter essere competitivi. Nel prosieguo del capitolo si vogliono fornire alcuni esempi significativi di nanoarchitetture prototipali in cui i nuovi concetti delle nanotecnologie possono prendere forma.

5.1. Nanoarchitetture guidate dalle interconnessioni

Al posto di considerare un computer come un insieme di interruttori connessi da fili, si potrebbe pensare, nell'ottica della progettazione, di vederlo come due fitti insieme di fili paralleli e perpendicolari tra loro la cui interconnessione può essere abilitata o meno da dispositivi dotati di una opportuna memoria⁴ posti a cavallo tra essi. L'idea non è nuova in quanto prende spunto dalle architetture regolari di tipo crossbar (Figura 6), può però trovare una realizzazione molto efficiente sfruttando in modo naturale le potenzialità (e anche i limiti) delle tecniche di autoassemblaggio; basti pensare a piccoli dispositivi molecolari che possano abilitare le interconnessioni a loro volta realizzate mediante nanowire (a stato solido) o nanotubi. Un approccio simile [5] è stato impiegato da un gruppo di ricercatori della Hewlett-Packard e dell'Università di California, Los Angeles, nell'ottica di sviluppare a scopo di ricerca un computer di tipo riconfigurabile che può tollerare la presenza di difetti facendo ricorso alla ridondanza. L'architettura, chiamata Teramac [9], sfrutta un monostrato di molecole bistabili (rotaxani) attivabili con impulsi elettrici che costituiscono le giunzioni e rappresenta un prototipo concettuale di realizzazione nanoelettronica a basso costo. Esso comprende 10^6 unità logiche pilotate a una frequenza rela-

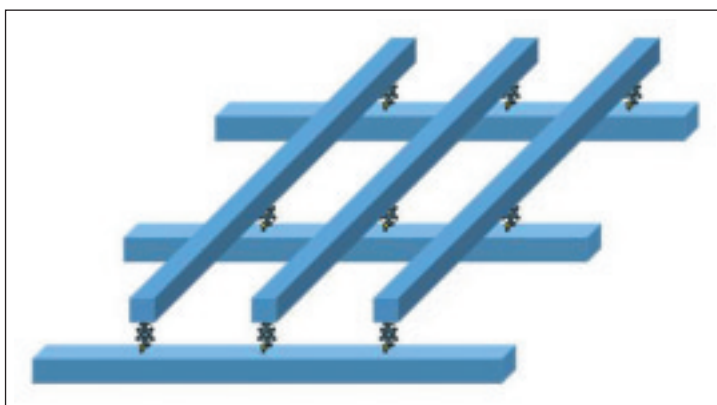


FIGURA 6

Architettura crossbar realizzata con nanowire e dispositivi bistabili molecolari

⁴ Tipicamente a due stati stabili, da cui il termine "bistabile".

tivamente bassa di 1 MHz, quindi con una potenza di calcolo equivalente a 10^{12} operazioni al secondo (da cui il nome).

Una più nota e analoga realizzazione dal punto di vista concettuale è la memoria RAM magnetica (MRAM) che si basa sullo spin dell'elettrone per ottenere l'effetto di bistabilità [8], ma si trovano altri esempi, tra cui uno significativo mediante nanotubi [21].

5.2. Nanoarchitetture a connettività locale

Un altro modo per costruire nanoarchitetture efficienti è sfruttare la *località* delle interconnessioni. Un tale approccio prende spunto dai sistemi molecolari e biologici e, come si è accennato in precedenza, il notevole *background* culturale ereditato può favorire l'impiego di opportune nanostrutture per la mappatura di classi specifiche di algoritmi. Svariate nanostrutture che dan-

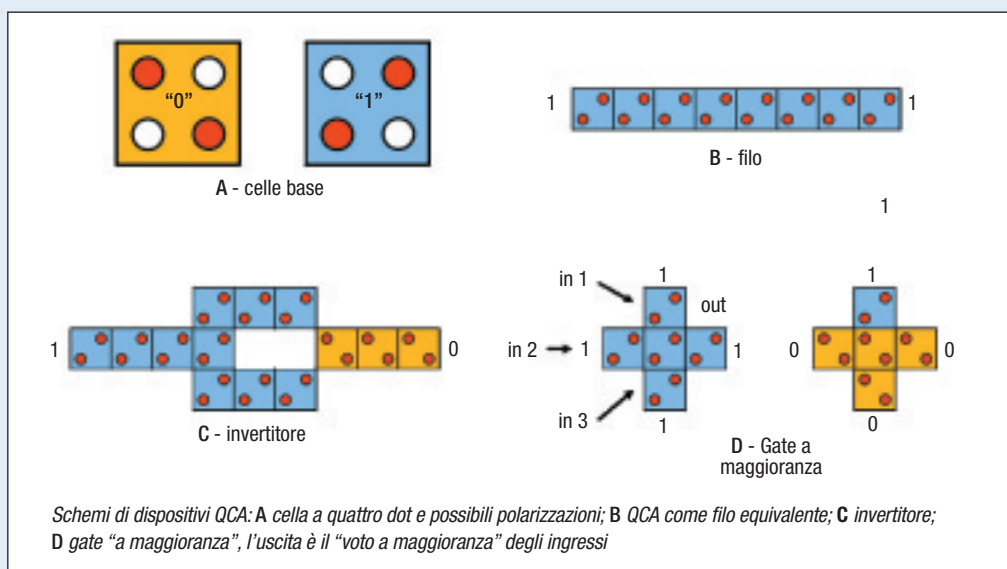
no origine a dispositivi quantistici possono essere usate per questo tipo di realizzazione, ma quelle che si prestano naturalmente a questo tipo di funzionalità sono i quantum dot.

5.2.1. AUTOMI A CELLE QUANTICI

Le configurazioni regolari di quantum dot conducono, in via naturale, al paradigma degli automi a celle, e sono pertanto chiamati Automi a Celle Quantici (QCA, *Quantum Cellular Automata*) (si veda il riquadro su **automi a celle**). L'idea è piuttosto recente e descrive un modo intelligente di realizzare automi a celle utilizzando le nanotecnologie [16], che risulta anche attraente per il fatto che l'elaborazione non dipende dalle interconnessioni né dalle correnti elettriche. Nella teoria dell'informazione e nella cibernetica gli automi a celle costituiscono

Automi a celle quantici

Un automa a celle quantico (QCA) può essere pensato come una rete infinita di piccoli e identici automi finiti, o celle, connessi uniformemente e sincronizzati. La cella elementare di un QCA consiste di quattro (o cinque) quantum dot disposti su un quadrato e separati da barriere tunnel, e comprende due elettroni (extra) che possono spostarsi per tunneling all'interno della cella ma senza lasciarne i confini. A causa dell'interazione colombiana, questi elettroni si dispongono in configurazione a minima energia su una delle due diagonali, definendo in questo modo due "polarizzazioni" associate ai due stati logici binari (Figura A). Se viene applicato un campo elettrico per variare la polarizzazione della prima cella, si provoca la propagazione dell'informazione alle celle vicine, quindi, a tutto il sistema. L'idea fondamentale che sta pertanto alla base del funzionamento dei QCA è che lo stato energetico di un insieme di elettroni, inizialmente in uno specifico stato base, viene alterato in conseguenza della variazione delle condizioni al contorno. Se la variazione negli stati di energia è condotta in modo adiabatico il risultato finale sarà un nuovo stato base che dipende solo dalle nuove condizioni al contorno. In definitiva, in questo paradigma la computazione è associata all'evoluzione verso lo *stato base* a minima energia del sistema. I dispositivi QCA si costruiscono allora disponendo le celle in un *layout* appropriato; alcuni esempi sono riportati in figura B-D, ma sono stati progettati circuiti più complessi come semplici microprocessori.



un modello computazionale molto importante, soprattutto per le applicazioni rivolte alla fisica (comportamento intrinseco distribuito e auto-organizzazione), alle scienze della vita e alle scienze sociali. La macchina di Turing è, inoltre, un caso particolare di automa a celle.

I QCA si presentano come un'alternativa radicale al transistor per l'elaborazione su nanoscala, tuttavia molta ricerca dovrà essere svolta per ripensare le nanoarchitetture appropriate per questo tipo di impostazione. Da un lato, esiste un'intensa attività di tipo teorico-simulativo, dall'altro la proposta di utilizzare strutture a stato solido di tipo singolo-elettone in condizioni di temperatura molto bassa (*criogeniche*), in quanto per lavorare a temperatura ambiente i singoli quantum dot dovrebbero comunque essere di dimensioni inferiori a 5 nm. Tali dimensioni potrebbero consentire di immagazzinare dati in celle più piccole di 25 nm × 25 nm.

5.2.2. MATRICI DI QUANTUM DOT

Una matrice di quantum dot ognuno dei quali interagisce con i vicini e con un substrato attivo (Figura 7) può implementare un concetto più avanzato di elaborazione. L'accoppiamento con barriere tunnel e la non linearità presente nel substrato (per esempio, RTD, *Resonant Tunneling Diode* ovvero diodi a tunneling risonante) originano un interessante comportamento dinamico [22] in quanto la rete nel suo complesso possiede stati stabili multipli analogamente a una rete neurale artificiale (per esempio, una rete di Hopfield), distaccandosi così dai QCA. Tali reti potrebbero implementare efficacemente elaborazioni di immagini a basso livello e memorie associative. Queste strutture potrebbero essere realizzate sfruttando l'autoassemblamento di isole metalliche quali nanoparticelle d'oro opportunamente spaziate le une dalle altre da molecole organiche, mentre il substrato attivo potrebbe contenere eterostrutture di semiconduttori composti. A titolo di esempio la figura 8 mostra un'immagine ottenu-

ta al microscopio ad effetto tunnel di un superreticolo⁵ di nanocluster d'oro aventi il nucleo di diametro pari a $1.6 \text{ nm} \pm 0.23 \text{ nm}$ e incapsulati in una matrice organica, depositato su terrazze di oro eteroepitassiale. Non appare particolarmente problematico produrre tali superreticoli su substrati tipo GaAs, mentre la vera sfida è realizzare l'insieme di (nano)elettrodi con l'adeguato pilotaggio dei segnali.

5.3. Quantum computing

Il quantum computing [2] è un paradigma computazionale radicalmente diverso, che sfrutta le proprietà di interferenza degli stati "correlati" (ovvero di tipo *entangled*) delle particelle quantistiche per consentire a ogni

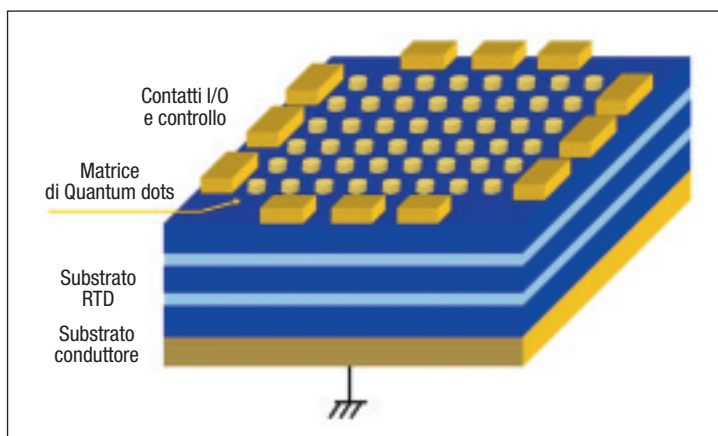


FIGURA 7

Archetipo di circuito nanoelettronico realizzato con matrici di quantum dot

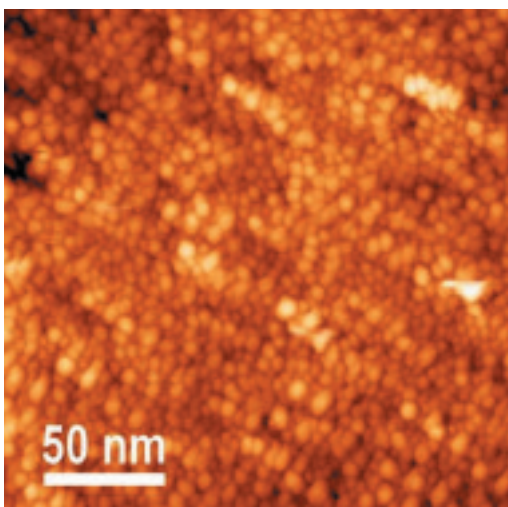


FIGURA 8

Immagine STM di un superreticolo di nanocluster d'oro in matrice organica depositati su substrato d'oro eteroepitassiale

⁵ Struttura periodica che si inserisce sopra una precedente struttura cristallina.

bit di informazione quantistica (detto *qubit*⁶) presente nel sistema di essere intimamente legato a ogni altro qubit.

Esistono anche qui alcune proposte di utilizzo di nanostrutture per il quantum computing. Per esempio, matrici di quantum dot analoghe a quelle viste in precedenza (ma realizzate confinando elettricamente gli elettroni mediante *split gate* su eterostrutture di AlGaAs e GaAs) possono essere adatte all'implementazione di qubit attraverso i singoli dot (ogni dot contenente un solo elettrone rappresenta un qubit) o coppie di quantum dot. In questo caso, risulta più appropriata l'informazione dello spin degli elettroni rispetto agli stati energetici del dot in quanto il tempo di decoerenza è superiore di vari ordini di grandezza. Tuttavia, la realizzazione di qubit [12] è estremamente difficoltosa dal momento che dovrebbero formare un nanosistema chiuso senza effetti di interferenza con il mondo macroscopico e con l'ambiente circostante (bagno termico) e la tecnologia è davvero ai primi passi.

6. OLTRE IL CMOS: COSA C'E' DIETRO L'ANGOLO?

Nel capitolo precedente sono state presentati svariati esempi di nanoarchitetture innovative per l'elaborazione dell'informazione, si vuole adesso proporre una sintesi di quelle che possono essere in prospettiva futura direzioni alternative alle tecnologie-guida CMOS (*Complementary-MOS*). Vale la pena puntualizzare che comunque per queste ultime, sotto la spinta delle nanotecnologie, miglioreranno materiali (per esempio, materiali ad alta costante dielettrica), processi (di tipo SOI, *Silicon On Insulator*, o basati su composti silicio-germanio detti *strained silicon*) e dispositivi ultrascalati MOS: le più recenti stime Intel valutano che cambiamenti drastici non dovrebbero essere necessari fino ad arrivare a transistori di 10 nm di lunghezza di canale (recenti proiezioni

stimano che questa dimensione potrebbe essere raggiunta nel 2011), mentre la comparsa di nanotecnologie radicalmente nuove per applicazioni commerciali è prevista intorno al 2020.

Volendo allora offrire criticamente un panorama di quali approcci concettuali si possono prefigurare in prospettiva di una tale svolta, si potrebbero indicare i seguenti filoni:

■ **nanowire a semiconduttore**: possono implementare interconnessioni e funzioni logiche e sono i dispositivi a nanoscala più probabilmente disponibili a breve termine;

■ **elettronica basata su nanotubi**: l'ibridizzazione con la tecnologia MOS appare già adesso la futura evoluzione delle tecnologie microelettroniche;

■ **circuiti nanoelettronici di tipo quantistico**: è di fatto un approccio "trasversale", quantum dot e dispositivi quantistici riscuotono un interesse sempre crescente anche se non appare banale passare dalla scala mesoscopica a quella macroscopica con l'adeguata consistenza del segnale;

■ **elettronica molecolare/organica**: rimpiazzare i transistori a effetto di campo con dispositivi molecolari e/o autoassemblati appare una sfida considerevole ma possibile [14], per vincere la quale occorre risolvere il problema critico dei contatti;

■ **quantum computing**: obiettivo affascinante e di grande impatto, ma decisamente una ricerca a lungo termine il cui successo è anche subordinato alla risoluzione di problemi tecnici quali decoerenza, correzione di errori, segnali di ingresso e uscita; richiede, inoltre, il funzionamento criogenico.

Un confronto tra vantaggi e svantaggi di questi approcci è, infine, riassunto in tabella 2. Sebbene non si possa ancora individuare con chiarezza quale tecnologia possa succedere alla microelettronica a stato solido, il rapido sviluppo delle nanotecnologie registrato degli ultimi anni (anche indotto dai considerevoli finanziamenti) ha fatto crescere le relative aspettative.

⁶ Un qubit è l'informazione contenuta in un sistema quantistico a due stati, come i diversi stati di polarizzazione di un fotone, la direzione di spin degli elettroni, gli stati energetici interni di un atomo o di un quantum dot.

Nanotecnologia	Vantaggi	Inconvenienti/problemi aperti
Nanowire a semiconduttore	Facilità di manipolazione (drogaggio, contatti ohmici ecc.) Compatibilità con le tecnologie microelettroniche	Layout multistrato
Elettronica basata su nanotubi	Ottimi conduttori/interconnessioni Semplicità nello sviluppo di dispositivi a commutazione Compatibilità con le tecnologie microelettroniche	Difficoltà di controllo della chiralità Controllo del posizionamento Alte resistenze di contatto
Circuiti nanoelettronici di tipo quantistico	Autoassemblaggio guidato Elaborazione non convenzionale	Funzionamento criogenico
Elettronica molecolare/organica	Capacità di autoassemblamento Funzionamento a bassa potenza	Creazione di strutture di interconnessione Interfacciamento col mondo macroscopico
Quantum computing	Elaborazione parallela di tipo massivo e non convenzionale Azione a distanza	Funzionamento a temperature prossime allo zero assoluto Decoerenza Gestione dei segnali di ingresso/uscita

7. CONCLUSIONI

Le ricerche sulle nanotecnologie hanno aperto prospettive di grande interesse per lo sviluppo delle tecnologie dell'informazione al di là dei limiti dei processi di fabbricazione attualmente impiegati. Costituiscono, inoltre, uno stimolo a realizzare nuove classi di dispositivi nanoelettronici in grado di sfruttare effetti quantistici per un'elaborazione dell'informazione di tipo non convenzionale. Se nel futuro più immediato il silicio appare ancora il protagonista, l'innovazione introdotta dalle nanotecnologie costituisce un elemento concreto di complemento nonché di estensione allo *scaling* e alla legge di Moore. In particolare, le possibilità offerte dai sistemi autossemblanti di tipo organico delineano uno scenario in cui si prefigurano architetture computazionali di tipo ibrido dove il silicio può costituire la piattaforma ideale per un'integrazione di tipo eterogeneo. Le nanotecnologie rappresentano, pertanto, una svolta epocale nelle tecnologie di fabbricazione in quanto al concetto classico di miniaturizzazione si affianca quello di autoassemblaggio guidato dove i due metodi top-

down e bottom-up non sono in contrapposizione, ma possono trovare una efficace sinergia in quanto proprio le infrastrutture legate alle tecnologie microelettroniche possono alimentare l'innovazione delle nanotecnologie. Affinché si possano progettare e sfruttare in modo efficace le nanostrutture è comunque necessario lo sviluppo delle nanoscienze di base e, conseguentemente, poter disporre di una adeguata modellistica delle nanocelle che possono costituire i processori elementari di nuove architetture computazionali. Un approccio di tipo interdisciplinare sarà condizione necessaria e chiave di volta per realizzare i computer del futuro.

Bibliografia

- [1] Andres R.P., et al.: *The design, fabrication and electronic properties of self-assembled molecular nanostructures*. The Handbook of Nanostructured Materials and Nanotechnology, Nalwa HS (Ed.), Academic Press, San Diego, 1998.
- [2] Angeleri E.: Quantum computing: sogno teorico o realtà imminente? *Mondo Digitale*, n. 1, 2003, p. 36-50.
- [3] Baldi L., Cerofolini G.: La legge di Moore e lo sviluppo dei circuiti integrati. *Mondo Digitale*, n. 3, 2002, p. 3-15.

TABELLA 2

Confronto tra strategie alternative di realizzazione di nanosistemi di elaborazione

- [4] Binnig G., Rohrer Gerber Ch., Weibel E.: Surface studies by scanning tunneling microscopy. *Phys. Rev. Lett.*, Vol. 49, 1982, p. 57.
- [5] Chen Y., Jung G., Ohlberg D., Li X., Stewart D.R., Jeppesen J.O., Nielsen K.A., Stoddart J.F., Williams R.S.: Nanoscale molecular-switch crossbar circuits. *Nanotechnology*, Vol. 14, 2003, p. 462-468.
- [6] Eigler D.M., Schweizer E.K.: Positioning single atoms with a scanning tunneling microscope. *Nature*, Vol. 344, 1990, p. 524.
- [7] Feynman R.P.: There's plenty of room at the bottom. *Caltech's Engineering and Science*, Vol. 23, 1960, p. 22-36. <http://www.zyvex.com/nanotech/feynman.html>
- [8] Gallagher W.J., Kaufman J.H., Papworth S.S., Scheuerlein R.E.: *Magnetic memory array using magnetic tunnel junction devices in the memory cells*. U.S. Patent 5 640 343 (IBM), June 1997.
- [9] Heath J.R. et al.: A Defect-Tolerant computer architecture: opportunities for nanotechnology. *Science*, Vol. 12, 1998, p. 1716-1721.
- [10] Hellemans A.: X-rays find new ways to shine. *Science*, Vol. 277, 1997, p. 1214-1215.
- [11] High-Performance VLSI Signal Processing Innovative Architectures and Algorithms, Vol. 1, Algorithms and Architectures. Ray Liu HJ, Kung Yao K (Eds), Wiley-IEEE Press, 1997.
- [12] Hofmann E.: Dal computer classico a quello quantistico: realizzabilità e potenziali applicazioni. *Mondo Digitale*, n. 4, 2003, p. 63-75.
- [13] Iijima S.: Helical microtubules of graphitic carbon. *Nature*, Vol. 354, 1991, p. 56.
- [14] Joachim C.: Bonding more atoms together for a single molecule computer. *Nanotechnology*, Vol. 13, 2002, p. R1-R7.
- [15] Lee K.B., Park S.J., Mirkin C.A., Smith J.C., Mrksich M.: Protein. Nanoarrays generated by dip-pen nanolithography. *Science*, Vol. 295, 2002, p. 1702-1705.
- [16] Lent C.S., Isaksen B.: Clocked molecular quantum-dot cellular automata. *IEEE Trans. Electron Devices*, Vol. 50, 2003, p. 1890-1896.
- [17] Likharev K.: Single electron devices and their applications. *IEEE Proceedings*, Vol. 87, 1999, p. 606-632.
- [18] Martel R., Derycke V., Appenzeller J., Wind S., Avouris Ph.: *Carbon nanotube Field-effect transistors and logic circuits*. Proceedings of the 39th Conference on Design Automation, 2002, p. 94-98.
- [19] Neave J.H., Joyce B.A., Dobson P.J., Norton: *Appl Phys A*, Vol. 31, 1983, p.1.
- [20] Qin L.C., Zhao X., et al.: The smallest carbon nanotube. *Nature*, Vol. 408, 2000, p. 50.
- [21] Rueckes T., Kim K., Joselevich E., Tseng G.Y., Cheung C., Lieber C.M.: Carbon nanotube-based

nonvolatile random access memory for molecular computing. *Science*, Vol. 289, 2000, p. 94-97.

- [22] Roychowdhury V.P., Janes D.B., Bandyopadhyay S.: Nanoelectronic architecture for boolean logic. *Proceedings of the IEEE*, Vol. 85, 1997, p. 574-587.
- [23] *Solutions to Parallel and Distributed Computing Problems: Lessons from Biological Sciences*. Zomaya A.Y., Ercal F., Olariu S. (Eds), Wiley, 2000.
- [24] *Technology Roadmap for Nanoelectronics*. Editor: R. Compagno (European Commission), Luxemburg, Office for Official Publications of the European Communities, 2001.
- [25] Xia Y., Rogers J., Paul K., Whitesides G.: Unconventional methods for fabricating and patterning nanostructures. *Chem. Rev. Ed.*, Vol. 99, 1999, p. 1823-1848.

ERMANN DI ZITTI ingegnere elettronico nel 1980 e dottore di ricerca nel 1986, è Professore Associato dal 1992 presso la Facoltà di Ingegneria dell'Università di Genova dove insegna "Nanotecnologie" oltre a vari corsi di elettronica. La sua attività di ricerca, svolta presso il Dipartimento di Ingegneria Biofisica ed Elettronica, ha riguardato inizialmente i dispositivi a semiconduttore di potenza per proseguire nei sistemi VLSI di tipo parallelo, sfociando infine nel settore della nanoelettronica e delle nanotecnologie. È coautore di circa 50 pubblicazioni internazionali. dizitti@unige.it

DAVIDE RICCI si è laureato in Fisica nel 1989 e ha conseguito il Dottorato di Ricerca in Ingegneria Elettronica e Informatica presso l'Università di Genova. Svolge la sua attività presso il Dipartimento di Ingegneria Biofisica ed Elettronica dal 1989. Specialista in Microscopia a Sonda di Scansione, si è interessato della messa a punto di nuove tecniche di indagine submicroscopiche di superfici di sensori e sistemi biologici. Da alcuni anni, l'attività si incentra sulla realizzazione di nuovi dispositivi per l'elaborazione dell'informazione basati su nanostrutture autoassemblate accoppiate a dispositivi realizzati tramite la microelettronica tradizionale. È coautore di circa 50 pubblicazioni internazionali. davide.ricci@dibe.unige.it

DANIELE D. CAVIGLIA si è laureato in Ingegneria Elettronica nel 1980 e si è specializzato in Ingegneria Informatica nel 1982 presso l'Università di Genova. Fa parte del Dipartimento di Ingegneria Biofisica ed Elettronica dal 1984. È professore ordinario di Elettronica presso la Facoltà di Ingegneria della stessa Università dal 2000. La sua attività di ricerca ha riguardato la modellistica di transistori MOS, la progettazione VLSI, la realizzazione di architetture neurali con autoapprendimento, in particolare in tecnologia CMOS a bassa potenza. I suoi interessi sono ora nel settore della progettazione circuitale a radiofrequenza e nella realizzazione di dispositivi a nanotubi di carbonio. caviglia@dibe.unige.it

ICT E DIRITTO

Rubrica a cura di

Antonio Piva e 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.



Il diritto d'autore nella rete

1. INTRODUZIONE

Quando nel 1455 Johann Guttenberg inventò la stampa a carattere mobile, le opere letterarie ebbero una diffusione fino a quel tempo inimmaginabile grazie ai costi di produzione dei libri fortemente ridotti. Allora, per la prima volta, si parlò dei diritti dell'autore, distinti da quelli dell'editore, e si riconobbe al contenuto (ovvero, il testo, chiamato *corpus mysticum*) una tutela indipendente dal suo supporto (ovvero, la pagina, detta invece *corpus mechanicum*). Fu una rivoluzione giuridica e culturale i cui effetti sono ancora presenti ai nostri giorni.

Oggi, a 5 secoli di distanza, l'informatica e la telematica hanno rimesso in discussione il diritto d'autore, costringendo a ridiscuterne i fondamenti, a valutarne i limiti, nonché a ricercare efficaci tutele dinanzi ai nuovi attacchi.

La tecnologia digitale, infatti, ha permesso la "smaterializzazione" del supporto dell'informazione, ovvero di scrivere un testo come il presente articolo, ad esempio, mediante un computer senza bisogno della carta, ma anche (e soprattutto) di trasformare un brano musicale, un'immagine fotografica, una sequenza cinematografica, e così via, in un *file* di semplice gestione, memorizzazione e riproduzione da parte dell'elaboratore. In questo modo, risulta assai agevole duplicare un'opera in numerose copie, tutte perfettamente identiche, ovvero copiare a basso costo una canzone o un film (i prezzi dei masterizzatori sono, ormai, accessibili a chiunque), con una qualità pressoché identica all'originale; tanto più che perde perfino significato la stessa distinzione tra originale e copia, trattandosi di *bit* "uguali" a quelli di partenza!

D'altro lato, lo sviluppo di Internet ha consentito la fruizione globale dell'opera digitalizzata, che potrà essere facilmente comunicata e diffusa, ovvero commercializzata, in tutta la rete.

Il risultato di questo progresso tecnologico ancora una volta si ripercuote di riflesso e in modo significativo sulle scienze sociali, economiche e giuridiche, ossia sui comportamenti delle persone (scaricare un file musicale in formato compresso MP3), sulle strategie commerciali (l'autore può vendere *on line* la propria opera senza necessità di intermediazione) e sulle scelte legislative.

Sotto quest'ultimo profilo, negli ultimi 15 anni (a partire dalla direttiva 91/250/CEE relativa alla tutela del *software*) ha avuto un ruolo determinante la spinta dell'Unione Europea al fine di evitare differenze e incertezze normative negli Stati membri.

2. LA LEGGE SUL DIRITTO D'AUTORE

Le disposizioni comunitarie in materia di diritto d'autore sono state recepite puntualmente anche nell'ordinamento italiano, mediante una serie di decreti¹ che hanno aggiornato la legge 22 aprile 1941 n.633 (cosiddetta *Legge sul diritto d'autore*, l.d.a.).

¹ Cfr. d.lgs. 6 maggio 1999, n. 169 "Attuazione della Direttiva 96/9/CE relativa alla tutela giuridica delle banche dati"; d.lgs. 2 febbraio 2001, n. 95 "Attuazione della direttiva 98/71/CE relativa alla protezione giuridica dei disegni e dei modelli"; d.lgs. 9 aprile 2003, n. 68 "Attuazione della Direttiva 2001/29/CE sull'armonizzazione di taluni aspetti del diritto d'autore e dei diritti connessi nella società dell'informazione".

Riquadro 1

La Legge sul diritto d'autore tutela:

1. le opere letterarie;
2. le opere musicali;
3. le opere coreografiche e pantomimiche;
4. le opere delle arti figurative;
5. i disegni e le opere dell'architettura;
6. le opere dell'arte cinematografica;
7. le opere fotografiche;
8. i programmi per elaboratore;
9. le banche di dati;
10. il disegno industriale (*design*).

Il diritto d'autore (Riquadro 1) sorge con la creazione stessa dell'opera (purché sia nuova e originale) e si sostanzia in due categorie fondamentali di facoltà:

1. diritti morali (sono imprescrittibili e non possono essere ceduti a terzi neppure dietro compenso):

- venire riconosciuto autore dell'opera;
- mantenerla inedita;
- opporsi a modifiche.

2. diritti patrimoniali (si estinguono in 70 anni dalla morte dell'autore e possono essere oggetto di cessione), ossia il diritto di sfruttare economicamente l'opera mediante:

- riproduzione in più esemplari;
- comunicazione e distribuzione;
- noleggio e prestito;
- diffusione via etere, cavo o satellite.

In caso di violazione sono previsti idonei strumenti volti a ottenere, anche in via d'urgenza, la rimozione dello stato di fatto (ad esempio, chiusura di un sito Internet), l'inibizione delle attività illecite (per esempio, divieto di mettere *on line* un'opera protetta), nonché il risarcimento del danno patito.

A questa tutela di tipo civile (disciplinata negli artt. 156 e segg. l.d.a.) si affiancano le sanzioni penali di cui agli artt. 171 e segg. l.d.a., cioè pene detentive e/o pecuniarie a seconda della gravità del reato commesso.

Ciò premesso, si esaminano, nel seguito, alcune fattispecie alle quali può applicarsi la normativa sul diritto d'autore nella rete Internet.

3. INTERNET E LINK

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 legisla-

zione 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"*.

Come noto, le pagine web sono realizzate tramite un linguaggio denominato HTML (*Hyper-text Markup Language*) che il *web browser* (per esempio, Internet Explorer, Netscape Navigator, Opera ecc.) 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.

Fermo restando che il *link* sotto il profilo tecnico non richiede necessariamente la cooperazione del sito di destinazione o di quello dal quale si estraggono le informazioni, si è soliti operare alcune distinzioni tra le tipologie di collegamenti (Riquadro 2).

Riquadro 2

Link esterno (*surface linking*): si collega alla *home page* di un altro sito.

Link interno (*deep link*): si collega a una pagina interna di un altro sito.

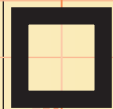
Framing: all'interno della cornice del sito di partenza viene caricata la pagina di un altro sito.

Inlinking o *embedded link*: cliccando nel simbolo di collegamento, si inseriscono nel sito di partenza informazioni di un altro sito, senza però abbandonare la pagina.

Il link esterno è giuridicamente lecito, anzi di solito viene gradito per l'evidente traffico in entrata che può generare; pertanto, non è necessario il consenso del titolare del sito collegato che, comunque, può sempre impedirlo tecnicamente.

La legittimità del link interno è, invece, assai dibattuta in quanto (in assenza di una norma che lo permetta o lo vieti espressamente) alcuni hanno sostenuto che l'"aggiramento" della *home page* possa costituire una violazione della proprietà intellettuale provocando un danno al titolare del sito (per esempio, la diminuzione dei proventi pubblicitari).

Oltre alla legge sul diritto d'autore, contro tale pratica sono state invocate anche le norme sulla concorrenza sleale e sullo sviamento della clientela; pertanto, il link interno non sarebbe permesso senza il consenso dell'interessato (comunque la giurisprudenza su questo punto è ancora oscillante).



Pare riscontrarsi maggiore uniformità tra gli interpreti nel ritenere certamente illecito il *framing* (inclusione del contenuto del sito all'interno della cornice di un altro sito) senza che l'avente diritto abbia previamente espresso il proprio consenso.

A margine di queste considerazioni, si segnala la necessità che il consenso al framing e al link interno rispettino la forma scritta. Infatti, partendo dal presupposto che le pagine web siano opere dell'ingegno di carattere creativo e, quindi, siano tutelate dal diritto d'autore, ci si deve riferire all'art.110 l.d.a. il quale prevede appunto che *“la trasmissione dei diritti di utilizzazione deve essere provata per iscritto”*.

4. LA MUSICA ON LINE

Il mercato che ha risentito in maniera assai rilevante della rivoluzione tecnologica, è senza dubbio quello discografico e ciò non solo grazie al tradizionale sistema di duplicazione e distribuzione di file musicali protetti, ma soprattutto grazie ad alcune invenzioni ormai celeberrime.

La prima è, senza dubbio, il file in formato MP3, abbreviazione di MPEG1-Layer3 (*Moving Picture Experts Group*), ossia un algoritmo di compressione audio che codifica file audio e video in formato digitale compresso.

La tecnologia MP3 si richiama ai principi di psicoacustica, eliminando tutti i suoni che non sono udibili dall'orecchio umano e quelli che risultano coperti da altri più forti; in sostanza, viene effettuata una “compressione selettiva” eliminando tutte le frequenze non percepibili: in tale maniera, a parità di qualità del suono, si ottiene una notevole riduzione delle dimensioni del file e ciò consente una trasmissione via Internet decisamente più agevole.

In secondo luogo, il software “Napster”, frutto di una tecnologia ideata dal californiano Shawn Fanning e diventata ben presto un vero fenomeno di costume, ha permesso lo scambio gratuito della musica in rete per anni, mediante l'utilizzo combinato di un programma per lo scambio di contenuti su Internet, di una messaggeria istantanea e di alcuni motori di ricerca.

Questo sistema, perfettamente funzionante fi-

no alla chiusura del suo *server* centrale a opera dei giudici di San Francisco, ha fondato la propria fortuna sulla condivisione delle informazioni (e delle canzoni) attuato mediante lo scambio *peer to peer* ossia da punto a punto, attraverso il diretto collegamento tra i singoli PC di ciascun utente connesso in rete, metodo perfezionato e tuttora impiegato in molte applicazioni (WinMX, Kazaa, Gnutella ecc.).

Proprio la condivisione tra utenti di opere protette dal diritto d'autore *“... anche mediante programmi di condivisione di file tra utenti [...] mediante reti e connessioni di qualsiasi genere”* è stata quest'anno oggetto di un contestato intervento legislativo (il cosiddetto Decreto Urbani) del quale si parlerà più avanti.

5. MISURE TECNOLOGICHE DI PROTEZIONE

Oltre alla tutela giuridica, la normativa riconosce all'autore il diritto di proteggere tecnologicamente le proprie opere; in particolare, il d.lgs. 9 aprile 2003 n.68 ha introdotto due importanti disposizioni sul tema:

1. l'art. 102-quater permette all'autore di *“apporre sulle opere [...] misure tecnologiche di protezione efficaci”*; che possono differenziarsi tra misure “antiaccesso” (le quali non consentono ai terzi non autorizzati l'accesso all'opera, per esempio, mediante la richiesta di una *password* o con l'impiego di tecniche di crittografia) e misure “anticopia” (le quali pur consentendo l'accesso all'opera, ne impediscono le utilizzazioni non consentite, per esempio, la riproduzione).

2. l'art. 102-quinquies ammette l'inserimento di *“informazioni elettroniche sul regime dei diritti”*, vale a dire di note informative che dovrebbero indicare all'utilizzatore i diritti dell'autore.

L'utilizzo di dispositivi che permettono di eludere tali misure tecnologiche di protezione costituisce illecito amministrativo punito dall'art. 174 ter con:

a. una sanzione amministrativa del pagamento di € 154;

b. una sanzione accessoria della confisca del materiale;

c. un'eventuale sanzione accessoria di pubblicazione del provvedimento sulla stampa.

Invece la produzione o la commercializzazione

dei dispositivi elusivi di cui sopra sostanzia un reato ai sensi dell'art. 171 ter lett. f bis) e, pertanto, implica le seguenti sanzioni:

- a. una sanzione pecuniaria pari al doppio del prezzo dell'opera moltiplicato per il numero di duplicazioni abusive eseguite;
- b. la sospensione dell'attività commerciale;
- c. e, in caso di recidiva, la revoca della licenza commerciale.

Riguardo a questo argomento non si può sotto-
cendere il famoso "caso Sklyarov", dal nome del programmatore russo arrestato dall'FBI nel 2001 per la violazione del *Digital Millennium Copyright Act*² poiché la sua società commercializzava un programma che permette di superare gli accorgimenti predisposti dalla Adobe a tutela della integrità del contenuto dei file aventi il formato "PDF".

Il programma incriminato, in realtà, non era necessariamente destinato alla riproduzione abusiva, potendo, infatti, permettere agli utenti di riprodursi una copia privata "di sicurezza" di un file legittimamente posseduto, funzione assolutamente lecita secondo il cosiddetto *fair use*.

La giuria, il 17 dicembre 2002, emise verdetto di piena assoluzione.

6. IL DECRETO URBANI

Come ricordato, la maggior parte degli interventi normativi in materia di diritto d'autore prende le mosse dall'ordinamento comunitario, infatti proprio il d.lgs. n.68/03 che disciplina le misure tecnologiche di protezione recepisce l'importante direttiva europea 2001/29/CE.

Tuttavia, nell'anno in corso il legislatore italiano ha approvato un provvedimento (la legge 21 maggio 2004 n.128 con cui è stato convertito il decreto legge 22 marzo n.72, detto "*decreto Urbani*" dal ministro che lo ha proposto) finalizzato a contrastare la diffusione telematica abusiva di opere dell'ingegno.

In estrema sintesi, le disposizioni più significative stabiliscono che:

1. il reato di duplicazione e diffusione abusiva di opere dell'ingegno viene punito qualora sia commesso "*per trarne profitto*";
2. in tale delitto è compreso espressamente l'uso della tecnologia peer-to-peer, quando non sia "*per uso personale*";
3. è necessario versare a favore della SIAE un "compenso" sul prezzo di supporti ottici (CD e DVD) e dei software di masterizzazione;
4. nel caso in cui un'opera dell'ingegno venga messa on line, è obbligatorio fornire un idoneo avviso sull'avvenuto assolvimento degli obblighi derivanti dalla normativa sul diritto d'autore e sui diritti connessi;
5. gli *Internet Service Provider* (ISP) hanno l'obbligo, su ordine dell'autorità giudiziaria, di rimuovere i contenuti dei siti che violano la normativa in esame (ovvero di impedirne l'accesso), a pena di una sanzione amministrativa da € 50.000 a € 250.000 per chi non coopera! Si tratta di norme che hanno suscitato molte polemiche, tanto che - prima ancora di una sua completa attuazione - non sono mancate le proposte (e le promesse) di modifica!

7. CONCLUSIONI

Al di là delle leggi, non sempre di agevole interpretazione e applicazione, la maniera più efficace per tutelare le opere dell'ingegno probabilmente ancora una volta va ricercata nel campo tecnologico.

Negli ultimi anni ha avuto una grande diffusione la tecnica del *watermarking*, ossia la possibilità di inserire in un documento digitale (testo, video, audio ecc.) una sequenza di bit impercettibile dall'uomo che però può essere rilevata dal computer, permettendo anche di riconoscere la provenienza (per esempio, può contenere il nome dell'autore, l'anno di produzione e altre indicazioni utili).

Incidentalmente si ricorda che questa tecnica è stata utilizzata anche dalle forze dell'ordine che "marchiavano", in modo invisibile, le immagini pedo-pornografiche per poterne seguire, nel corso dell'indagine penale, gli scambi fatti in rete.

All'inserimento di un watermark nell'immagine digitale si accompagna frequentemente un servizio di monitoraggio e tracciamento della circolazione abusiva del documento protetto da *copyright* su Internet, con l'individuazione del-

² Il *Digital Millennium Copyright Act* (DMCA), testo legislativo suddiviso in 5 parti, è stato promulgato il 28 ottobre 1998 e costituisce una delle principali fonti normative statunitensi sul diritto d'autore; per approfondimenti si rinvia a <http://www.copyright.gov/legislation/dmca.pdf>

l'originale colpevole dell'illecita diffusione (la cui identità è registrata).

Dopo il "caso Napster", molti hanno ipotizzato la vendita via Internet di canzoni marchiate, onde impedirne l'illecita diffusione; il watermarking, infatti, potrebbe contenere il nome

dell'acquirente, in modo che in caso di distribuzione illecita, sia possibile risalire al responsabile della violazione.

Tutto ciò a patto che i watermark apposti siano particolarmente robusti e resistenti ai tentativi di cancellazione!

ANTONIO PIVA laureato in Scienze dell'Informazione, Presidente, per il Friuli - Venezia Giulia, dell'ALSI (*Associazione Nazionale Laureati in Scienze dell'Informazione ed Informatica*) e direttore responsabile della Rivista di Informatica Giuridica.

Docente a contratto di Informatica giuridica all'Università di Udine.

Consulente sistemi informatici, valutatore di sistemi di qualità ISO9000 e ispettore AICA per ECDL base e advanced.

antonio_piva@libero.it

DAVID D'AGOSTINI avvocato, ha conseguito il master in informatica giuridica e diritto delle nuove tecnologie, fornisce consulenza e assistenza giudiziale e stragiudiziale in materia di *software*, *privacy* e sicurezza, contratti informatici, *e-commerce*, nomi a dominio, computer crime, firma digitale. Ha rapporti di partnership con società del settore ITC nel Triveneto.

Collabora all'attività di ricerca scientifica dell'Università di Udine e di associazioni culturali.

david.dagostini@adriacom.it



DENTRO LA SCATOLA

Rubrica a cura di

Fabio A. Schreiber

Il Consiglio Scientifico della rivista ha pensato di attuare un'iniziativa culturalmente utile presentando in ogni numero di Mondo Digitale un argomento fondante per l'Informatica e le sue applicazioni; in tal modo, anche il lettore curioso, ma frettoloso, potrà rendersi conto di che cosa sta "dentro la scatola". È infatti diffusa la sensazione che lo sviluppo formidabile assunto dal settore e di conseguenza il grande numero di persone di diverse estrazioni culturali che - a vario titolo - si occupano dei calcolatori elettronici e del loro mondo, abbiano nascosto dietro una cortina di nebbia i concetti basilari che lo hanno reso possibile. La realizzazione degli articoli è affidata ad autori che uniscono una grande autorevolezza scientifica e professionale a una notevole capacità divulgativa.

L'aritmetica dei residui e il suo uso per la realizzazione di unità aritmetiche specializzate particolarmente veloci

Renato Stefanelli

1. INTRODUZIONE

Una contadina desiderava sapere quante fossero le uova contenute in un cesto, ma sapeva contare solo fino a sette, e le uova erano sicuramente più di sette; allora dal cesto ne tolse tre, poi altre tre, poi ancora altre tre finché non ne rimasero che due; rimise le uova nel cesto e poi da questo tolse a più riprese gruppi di cinque uova fino a che non ne rimasero che quattro; poi rifecce l'operazione togliendone a gruppi di sette finché non ne rimase che una.

Il matematico cinese Sun Tsu Suan-Ching, circa 1650 anni fa si domandò se tali risultati definissero in modo univoco il numero di uova nel cesto e arrivò alla seguente conclusione:

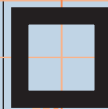
il cesto contiene 29 uova. Infatti, chiamando Q_1 , Q_2 , Q_3 i gruppi di tre, cinque e sette uova, e chiamando N il numero di uova nel cesto, si ha: $N = 3 Q_1 + 2$; $N = 5 Q_2 + 4$; $N = 7 Q_3 + 1$ e il numero 29, diviso per 3, 5, 7 fornisce resti 2, 4, 1. Per ottenere questo risultato egli definì il concetto di "residuo", ne studiò le proprietà ed enunciò e dimostrò il "teorema cinese sui residui", la base per la moderna aritmetica dei resi-

dui. Ovviamente Sun Tsu Suan-Ching non immaginava certo l'applicazione che le sue idee avrebbero avuto 1650 anni dopo. Nell'ultimo mezzo secolo, infatti, si è verificato come tali idee potessero essere applicate con successo alla realizzazione di unità aritmetiche particolarmente veloci.

Questo articolo tratta, in primo luogo, la definizione di residuo e le sue proprietà aritmetiche; si enuncerà il teorema cinese dei residui, già citato in precedenza; poi se ne studieranno le applicazioni; in particolare, si studierà il modo di ottenere strutture aritmetiche particolarmente veloci e si accennerà ai settori ove tale rappresentazione può avere un effettivo interesse (poco per unità aritmetiche di calcolatori, molto in unità aritmetiche specializzate per elaborazione di segnali e immagini).

2. I RESIDUI E LORO USO PER RAPPRESENTARE NUMERI INTERI POSITIVI

Si consideri un numero intero positivo N e lo si divida per una base intera positiva b ; si ottiene



un quoziente Q e un resto R tali per cui $N = Qb + R$ ove il resto R , compreso tra 0 e $b - 1$, viene chiamato “residuo del numero N rispetto alla base b ” e viene indicato come $|N|_b$. Per esempio il residuo del numero 29 rispetto alla base 3 vale $|29|_3 = 2$ in quanto $29 = 9 \cdot 3 + 2$. Ciò definisce un singolo residuo del numero N rispetto a una sola base b . Si considerino ora tre basi $b_1=3, b_2=5$ e $b_3=7$; è allora possibile calcolare i tre residui del numero N rispetto alle tre basi. Si può verificare che, se N è inferiore a un certo limite (105 nell'esempio) e se le tre basi sono state scelte in modo opportuno (come si vedrà in seguito) allora esiste una relazione biunivoca tra il numero e la terna dei suoi residui. Si consideri, infatti, la tabella 1 dove sono riportate le terne di residui per le basi 3, 5, 7 e per numeri tra zero e 106; si verifica intanto che al numero 29 corrisponde la terna 2, 4, 1 e a nessun altro numero corrisponde la stessa terna (in realtà, per questioni di spazio, la tabella riporta solo alcuni casi; pregherei il lettore di credere all'affermazione precedente). Si verifica poi che la relazione biunivoca è presente solo per numeri inferiori a 105; infatti, a partire dal numero 105, la tabella si ripete fino al numero 208 ecc..

La condizione che deve essere verificata nella scelta delle basi è che queste siano “prime tra loro a coppie”; se ciò è verificato allora il limite di rappresentabilità del numero N corrisponde al prodotto delle basi; nell'esempio, $3 \cdot 5 \cdot 7 = 105$. La condizione è sicuramente verificata nell'esempio in quanto si sono scelte basi rappresentate da numeri primi; ma andrebbero bene anche basi come (8, 9, 11) oppure (6, 7, 11, 25) dove alcune delle basi non sono numeri primi ma ogni coppia di basi ha massimo comune divisore uguale ad uno.

Si esamineranno ora le proprietà aritmetiche dei residui, la conversione da binario a residuo e quella inversa da residuo a binario e le applicazioni alle unità aritmetiche veloci.

3. PROPRIETÀ ARITMETICHE DEI RESIDUI

Le proprietà aritmetiche dei residui, per le operazioni di somma e di prodotto, si basano sul seguente teorema:

Si considerino due numeri X e Y rappresentati rispettivamente dai residui $x_1 \dots x_n$ e $y_1 \dots y_n$ nelle basi $b_1 \dots b_n$; sia inoltre Z la somma (il prodotto) di X e Y e Z abbia i seguenti residui $z_1 \dots z_n$; allora, per ogni base b_i , il risultato Z ha residui pari, al residuo della somma (del prodotto) dei residui degli addendi (fattori), ossia $z_i = |x_i + y_i|_{b_i}$ ($z_i = |x_i \cdot y_i|_{b_i}$). Come esempio si consideri il caso $3 + 4 = 7$ per la somma e $2 \cdot 4 = 8$ per il prodotto: esaminando la tabella si ha:

$$|0+1|_3 = 1 = |7|_3; |3+4|_5 = 2 = |7|_5;$$

$$|3+4|_7 = 0 = |7|_7$$

$$|2 \cdot 1|_3 = 2 = |8|_3; |2 \cdot 4|_5 = 3 = |8|_5;$$

$$|2 \cdot 4|_7 = 1 = |8|_7$$

Ne risulta, quindi, che, per fare la somma e il prodotto di due numeri, è sufficiente eseguire la somma e il prodotto, in modulo, dei residui omologhi. Attenzione: se somme e prodotti possono essere eseguiti in “modo facile”, non esistono proprietà analoghe per eseguire il confronto tra due numeri, il rapporto tra due numeri oppure, cosa abbastanza complicata, il controllo se il risultato di un'operazione ha superato o meno i limiti di rappresentabilità. Questi sono ancora argomenti di ricerca e tuttora non si conoscono algoritmi efficienti per la loro implementabilità.

4. LA CONVERSIONE DALLA RAPPRESENTAZIONE BINARIA A QUELLA A RESIDUI E QUELLA CONTRARIA DA RESIDUI A BINARIO

Si suppone di partire da dati in ingresso a una unità aritmetica rappresentati in binario e si desidera

base\N	0	1	2	3	4	5	6	7	8	9	..	28	29	30	..	103	104	105	106
3	0	1	2	0	1	2	0	1	2	0	..	1	2	0	..	1	2	0	1
5	0	1	2	3	4	0	1	2	3	4	..	3	4	0	..	3	4	0	1
7	0	1	2	3	4	5	6	0	1	2	..	0	1	2	..	5	6	0	1

TABELLA 1

Terne di residui per basi 3, 5, 7

che il risultato di un'operazione aritmetica sia ancora in formato binario. La conversione da binario a residuo con n basi può essere ottenuta mediante n circuiti divisori; si può notare però che in tali circuiti il divisore è un numero fisso e relativamente piccolo; la loro complessità è, quindi, inferiore a quella di un circuito divisore di tipo generale. Tali circuiti hanno complessità circuitale (area di silicio) e tempi di ritardo paragonabili a quelli di un normale addizionatore binario.

Più complessa è la conversione opposta, che costituisce il "teorema cinese sui residui", il contributo principale del matematico cinese Sun Tsu Suan-Ching.

Si fa l'esempio, per semplicità, delle tre basi $b_1 = 3$, $b_2 = 5$, $b_3 = 7$ (facile è l'estensione a basi in numero e valore qualsiasi). Si suppone di aver precalcolato tre numeri D_1, D_2, D_3 tali per cui:

$$|D_1 b_2 b_3|_{b_1} = 1 \quad |b_1 D_2 b_3|_{b_2} = 1 \quad |b_1 b_2 D_3|_{b_3} = 1$$

Nel caso delle tre basi 3, 5, 7, per tentativi si trova $D_1 = 2$, $D_2 = 1$, $D_3 = 1$.

Si consideri ora il numero $\underline{A} = a_1 D_1 b_2 b_3 + a_2 b_1 D_2 b_3 + a_3 b_1 b_2 D_3$. Per verificare che tale numero rappresenti il risultato richiesto, se ne calcolano i tre residui onde controllare che questi siano uguali a quelli iniziali a_1, a_2, a_3 e che il numero calcolato sia compreso entro i limiti di rappresentabilità.

Si lascia, per brevità, al lettore la prima parte (calcolo dei tre residui) ricordando che il residuo della somma (prodotto) è il residuo della somma (prodotto) dei residui, che un termine proporzionale a una base b_i ha residuo nullo rispetto alla stessa base e che ogni residuo a_i è inferiore alla propria base b_i e quindi $|a_i|_{b_i} = a_i$. Per verificare se $\underline{A}$ è compreso entro i limiti di rappresentabilità si considera l'esempio dei residui 2, 4, 1 nelle basi 3, 5, 7 che, dalla tabella, dovrebbe corrispondere al numero $A = 29$. Si ha:

$$\underline{A} = 2 \cdot 2 \cdot 5 \cdot 7 + 4 \cdot 1 \cdot 3 \cdot 7 + 1 \cdot 1 \cdot 3 \cdot 5 = 140 + 84 + 15 = 239,$$

numero che esce dai limiti di rappresentabilità; ma il numero 239 ha residuo 29 rispetto al prodotto M delle tre basi, pertanto:

$$A = |\underline{A}|_M \text{ ove } M = b_1 \cdot b_2 \cdot b_3 = 105;$$

infatti, $239 = 2 \cdot 105 + 29$

5. APPLICAZIONE ALLE UNITÀ ARITMETICHE

Si desideri eseguire la somma (o il prodotto) di due numeri rappresentati in binario e si desidera che il risultato sia in forma binaria. Per eseguire tale operazione occorre:

1. Convertire i due numeri in residui: occorrono $2n$ circuiti di conversione per i due numeri e per le n basi (ogni circuito di conversione "lavora" indipendentemente dagli altri, quindi, i $2n$ circuiti operano contemporaneamente). L'operazione è *lenta* (tempi paragonabili a quelli di un addizionatore binario).

2. Eseguire l'operazione aritmetica: occorrono n sommatore (moltiplicatori), ciascuno che elabori due residui omologhi per produrre un residuo del risultato. Si noti che, dato che le operazioni avvengono "in modulo", ogni sommatore (moltiplicatore) è diverso dagli altri. Anche in questo caso, ogni circuito aritmetico opera indipendentemente dagli altri e, quindi, possono "lavorare" contemporaneamente (non si ha riporto da un circuito all'altro) e questo è il motivo principale di un eventuale vantaggio della rappresentazione a residui rispetto a quella tradizionale binaria. L'operazione aritmetica vera e propria è *velocissima* in quanto gli n circuiti aritmetici operano in parallelo su numeri costituiti da pochi bit.

3. Ottenuti i residui del risultato occorre riconvertire questo alla rappresentazione binaria e tale operazione richiede molte moltiplicazioni e somme, su numeri molto grandi, e il calcolo di un residuo rispetto al prodotto delle basi, normalmente un numero grande. Nel complesso la terza operazione è fortemente onerosa sia rispetto alla complessità (area di silicio), sia rispetto ai tempi, per cui la si può indicare come *lentissima*.

Ma allora, se un'operazione con i residui è fortemente svantaggiosa sia come area di silicio che, e soprattutto, per la velocità, a cosa serve e dove può essere applicata?; si ha l'impressione di aver sparato a una mosca con un cannone (e forse senza essere sicuri di aver centrato la mosca!).

Si consideri l'unità aritmetica di un normale calcolatore: questa deve eseguire un'operazione singola per ogni istruzione di tipo aritmetico; in questo caso, l'unità aritmetica risulterebbe estremamente lenta rispetto a una soluzione tradizionale. Inoltre, tale unità



dovrebbe poter eseguire anche divisioni e confronti; nella rappresentazione mediante residui queste due operazioni richiederebbero una conversione da residui a binario per poterle eseguire in modo tradizionale. Il tutto sarebbe fortemente svantaggioso, e questa è la ragione per cui i residui non hanno avuto applicazione nelle unità aritmetiche dei calcolatori.

Si consideri ora un'unità aritmetica per un'applicazione "speciale", come, per esempio, un filtro digitale per segnali, una trasformata di Fourier per immagini bidimensionali ecc.: in questo caso, si hanno relativamente pochi dati in ingresso da convertire da binario a residui, relativamente pochi dati in uscita da convertire da residuo a binario, mentre si hanno molte operazioni di somme e prodotti da eseguire in modo veloce nel passo 2 precedentemente definito (in questo caso le n unità aritmetiche che operano contemporaneamente nella fase 2 corrisponderebbero non al calcolo di una sola somma o moltiplicazione, ma al calcolo completo dell'algoritmo di filtraggio o

della trasformata di Fourier); inoltre, non si hanno operazioni di divisione e di confronto. In questo caso, la fase 2 risulta preponderante rispetto alle fasi 1 e 3; pertanto, si può avere un effettivo vantaggio, in termini di velocità, dall'uso della rappresentazione mediante residui. In questa ultima frase consiste l'effettivo interesse per la rappresentazione dei numeri mediante residui.

Bibliografia

Esistono diversi libri che trattano l'argomento generale dell'aritmetica; alcuni di questi hanno un capitolo dedicato all'aritmetica dei residui; si indica qui un testo che tratta l'argomento in modo succinto ma chiaro:

Koren I.: *Computer arithmetic algorithms*. Prentice Hall, 1993.

Esiste poi una diffusissima bibliografia specifica: cercando su internet con Google.com usando come frase chiave "Chinese Remainder Theorem" si trovano 12.900 siti; se la ricerca avviene sulla frase "Residue Number System" si trovano ben 464.000 siti! Il lettore interessato si armi di tanta pazienza!

RENATO STEFANELLI Professore ordinario di Calcolatori Elettronici presso la Facoltà di Ingegneria del Politecnico di Milano. È autore di circa 160 pubblicazioni principalmente sulla tolleranza ai guasti sia in unità aritmetiche che in sistemi regolari di multiprocessori, sull'elaborazione di immagini, su circuiti aritmetici ad elevata velocità, su circuiti aritmetici basati sull'uso di residui. Ha diretto per diversi anni un Centro di Studio del CNR sull'Informatica presso il Politecnico di Milano.
stefanel@elet.polimi.it