

Marco Somalvico, Francesco Amigoni, Viola Schiaffonati

# Intelligenza artificiale

In questa trattazione viene presentata l'*intelligenza artificiale* (nel seguito indicata con la sigla IA), ponendo particolare attenzione al suo inquadramento nell'ambito dell'informatica e della robotica intelligente.

Nella presente introduzione la trattazione sarà limitata all'illustrazione generale dell'IA, per poi passare nelle sezioni che seguono all'esame di alcuni aspetti storici, epistemologici e scientifici inerenti al problema del rapporto tra la macchina e l'essere umano. Tali considerazioni faranno inoltre riferimento ad alcuni importanti ambiti applicativi dell'IA.

È opportuno in primo luogo fornire una definizione non ambigua dell'IA che, per sua natura, si presta a svariate interpretazioni e a qualche fraintendimento. L'IA è una disciplina recente che negli anni ha fornito un importante contributo al progresso dell'intera informatica. Per questo motivo riveste una indubbia importanza per quanti siano interessati alla valutazione delle potenzialità dell'informatica e degli effetti che il progresso di questa disciplina ha per l'uomo. L'IA è stata inoltre influenzata da numerose discipline fra le quali la filosofia, la matematica, l'economia, le neuroscienze, la psicologia, la cibernetica, le scienze cognitive e la linguistica.

L'IA è quella disciplina, appartenente all'informatica, che studia i fondamenti teorici, le metodologie e le tecniche che consentono di progettare sistemi hardware e sistemi di programmi software capaci di fornire all'elaboratore elettronico prestazioni che, a un osservatore comune, sembrerebbero essere di pertinenza esclusiva dell'intelligenza umana (Somalvico 1987). È questa una delle definizioni, accettate a livello internazionale, di un'area scientifica di notevole importanza strategica che si propone di superare le attuali frontiere dell'informatica, ponendosi l'obiettivo di migliorare ed estendere le prestazioni offerte dall'elaboratore elettronico. Si osservi che l'IA è una disciplina sperimentale, nel senso che un sistema di IA soddisfa gli obiettivi richiesti solamente quando le prestazioni desiderate sono effettivamente misurabili.

Scopo di questa disciplina non è quello di replicare o simulare l'intelligenza umana, obiettivo che per taluni è addirittura non ammissibile, bensì di riprodurre o emulare l'intelligenza umana. Non vi è alcun motivo che impedisca a priori che talune (ma non tutte!) prestazioni dell'intelligenza umana – per es., la capacità di risolvere problemi mediante processi inferenziali – possano anche essere fornite da una macchina. Nel caso dell'emulazione le prestazioni intelligenti sono ottenute utilizzando meccanismi propri della macchina, eventualmente differenti da quelli dell'uomo, ma tali da fornire prestazioni qualitativamente equivalenti e quantitativamente superiori a quelle umane. Considerando attentamente la storia e la natura dei risultati ottenuti dalle ricerche in IA, si può dire che gli studiosi hanno valutato volta per volta diversi modelli della razionalità umana, scegliendo il migliore per concepire una nuova tecnica o un nuovo sistema di IA. In molti casi, infatti, si preferisce parlare di *razionalità*, come concetto ideale di intelligenza, piuttosto che di intelligenza per eliminare i riferimenti antropomorfi di quest'ultimo termine. I modelli di razionalità adottati sono stati sia di tipo antropomorfo (ossia modelli che consentono non soltanto di ottenere prestazioni analoghe a quelle dell'uomo ma anche di adottare gli stessi metodi usati dall'uomo) sia di tipo non antropomorfo, con l'obiettivo di garantire i migliori risultati possibili. In generale, i modelli non antropomorfi hanno dimostrato caratteristiche di maggiore adattamento e integrazione con i modelli, le tecniche e i sistemi di IA già disponibili. Non esiste quindi in IA un'ideologia che porti a preferire a priori l'adozione di modelli antropomorfi, benché questi non siano esclusi programmaticamente; è ormai evidente, invece, la constatazione che, sulla base della valutazione dell'esperienza passata, quelli utilizzati sono stati spesso modelli non antropomorfi.

Un altro punto importante da considerare per caratterizzare l'IA è che essa non si basa su un insieme definito e completo di tematiche ma, al contrario, si occupa di problemi di ricerca aperti e di frontiera, estendendo costantemente i limiti del suo interesse e della sua azione. Pertanto, una prestazione che oggi potrebbe apparire come esclusiva dell'intelligenza umana e quindi, secondo la definizione data, essere di pertinenza dell'IA, in futuro potrebbe essere comunemente fornita da un sistema artificiale e non rientrare così nell'IA.

È bene osservare inoltre che l'IA è al tempo stesso una scienza e un'ingegneria (Nilsson 1998). È una scienza perché, nel momento in cui vengono emulati con determinati sistemi artificiali alcuni dei comportamenti intelligenti (dove, come si è già detto, intelligente va inteso nell'accezione che fa riferimento a comportamenti ritenuti esclusivi dell'uomo), l'uomo consegue l'obiettivo di formulare modelli oggettivi e rigorosi e di ottenere conferme sperimentali; ciò permette di ottenere un indiscutibile progresso nello studio, svolto con metodi scientifici, dell'intelligenza nell'uomo. L'IA, inoltre, è un'ingegneria perché, quando si ottengono dalle macchine prestazioni che emulano taluni comportamenti, erroneamente ritenuti inaccessibili all'ambito artificiale, si fornisce un oggettivo progresso al contributo che l'ingegneria (e le macchine progettate nel suo ambito) offre al miglioramento della vita dell'uomo.

## 1. Storia dell'intelligenza artificiale

L'IA ha una data di nascita ufficiale, il 1956, l'anno del famoso seminario estivo tenutosi presso il Dartmouth College di Hanover nel New Hampshire durante il quale la nuova disciplina viene fondata programmaticamente, a partire dalla raccolta dei contributi sviluppati negli anni precedenti e in direzione delle potenzialità future. Questa data di nascita convenzionale, sulla quale vi è accordo unanime nella comunità scientifica, potrebbe far pensare che la questione delle origini possa considerarsi definitivamente risolta. Tuttavia, quando si parla di storia dell'IA non si comincia descrivendo gli sviluppi successivi al 1956 ma si fa riferimento anche alla cibernetica e all'avvento dei primi calcolatori elettronici. Si citano inoltre anche Charles Babbage (1791-1871) e la sua macchina analitica, Gottfried Wilhelm Leibniz (1646-1716) e il suo progetto di meccanizzare la ragione, risalendo fino alla macchina logica di Raimondo Lullo (1235-1315) e agli automi semoventi di Erone di Alessandria (I sec. d.C.). Sarebbe però fuorviante tentare di risolvere la questione considerando tutto ciò che è avvenuto prima del 1956 come appartenente alla categoria dei precursori dell'IA. In primo luogo perché una parte della ricerca informatica degli anni Cinquanta dello scorso secolo è, a buon diritto, aspetto essenziale e fondante dell'IA. In secondo luogo perché considerare Lullo o Leibniz precursori dell'IA e Babbage padre dell'informatica significherebbe attualizzare impropriamente tali figure.

In questo contesto appare invece utile parlare di tradizione di ricerca (McCorduck 1979) per attenuare l'eterogeneità culturale, concettuale e soprattutto di intenti di coloro che possono essere citati nell'ambito della storia dell'IA. Tale tradizione, che è stata chiamata 'tradizione formalistica' di indagine sulla mente, convoglia un'idea importante, ossia che la prestazione artificiale faccia parte della pratica umana come la prestazione naturale, nella direzione di un continuo tentativo dell'uomo di imitare e di riprodurre se stesso e la natura. In questo senso l'inquadramento storico di una disciplina, che nel caso dell'IA appare particolarmente ricco ma anche problematico, rappresenta il punto di avvio per un'analisi più ampia, nella quale considerare anche questioni legate alla determinazione del suo particolare statuto ontologico.

La tradizione formalistica di indagine sulla mente, la nascita delle prime macchine calcolatrici e la tendenza dell'uomo ad autoimitarsi sono evidenti, per esempio, nel progetto di Babbage della 'macchina alle differenze' e della 'macchina analitica'. L'esigenza primaria in questo caso risiede nella necessità di calcolare automaticamente le tavole dei logaritmi, evitando gli errori umani tipici. Nella macchina alle differenze si possono rappresentare fino a 20 numeri di 18 cifre ciascuno; le cifre sono rappresentate da ruote e le ruote per le cifre di uno stesso numero sono incolonnate in modo tale che ogni colonna possa comprendere fino a 18 ruote. Nella macchina analitica, che non fu effettivamente mai realizzata per problemi di ordine economico, i numeri risultanti come uscita della macchina alle differenze avrebbero dovuto essere utilizzati come dati in ingresso per un calcolo successivo. La parte centrale avrebbe dovuto essere costituita dalla 'macina', una sorta di processore in grado di elaborare i numeri che le vengono inviati per produrre i risultati richiesti. La macchina analitica, inoltre, avrebbe dovuto possedere un magazzino (memoria) contenente sia i numeri non ancora elaborati dalla macina, sia quelli risultanti dalle operazioni della macina stessa. La macchina analitica, infine, avrebbe dovuto essere dotata anche di un sistema di controllo, in grado di eseguire automaticamente una sequenza di operazioni secondo il meccanismo delle schede perforate utilizzato nei telai.

Oltre a considerare questa vasta tradizione di ricerca, per delineare criticamente la storia dell'IA è importante riconoscere come questa abbia ereditato molte idee, punti di vista e tecniche da altre discipline, in particolare dalla filosofia, dalla matematica e dalla psicologia. Più precisamente, derivano dalla filosofia i risultati relativi al dibattito sulla natura dell'intelligenza e della razionalità, dalla matematica l'approccio formale basato sulla logica e dalla psicologia l'analisi delle relazioni fra conoscenza e azione. Tuttavia, è senza dubbio con la cibernetica e l'informatica che queste influenze si fanno più manifeste e aprono la strada alla nascita ufficiale dell'IA.

All'inizio degli anni Quaranta dello scorso secolo (Cordeschi 1996) si comincia a indicare con il termine *cibernetica* lo studio sistematico dei processi riguardanti la comunicazione e il controllo sia negli animali sia nelle macchine. Warren McCulloch e Walter Pitts propongono nel 1943 il primo modello di neuroni artificiali, attingendo alla conoscenza della fisiologia e delle funzioni di base dei neuroni, alla logica proposizionale e alla teoria della computabilità di Alan Turing. L'idea alla base del progetto cibernetico è quella di studiare i meccanismi dell'autoregolazione e del controllo presenti sia negli organismi viventi sia nelle macchine con retroazione, in grado cioè di rispondere in modo adattativo alle sollecitazioni dell'ambiente modificando il proprio comportamento. Uno dei risultati più significativi consiste nel mostrare come ogni funzione calcolabile possa essere elaborata da una qualche rete di neuroni connessi. A partire da ciò nel 1949 Donald Hebb dimostra come una semplice regola di aggiornamento per modificare le forze di connessione fra i neuroni possa dare luogo a processi di apprendimento. Nonostante questi significativi successi iniziali, il destino della cibernetica è segnato: dopo pochi anni (intorno alla metà degli anni Cinquanta dello scorso secolo) le risorse sono ormai quasi completamente convogliate sull'IA, a causa del disinteresse della cibernetica per le crescenti prestazioni dell'informatica e di una limitazione degli obiettivi iniziali. Questa tradizione risorgerà verso la metà degli anni Ottanta dello scorso secolo con il riemergere, in seno all'IA, del paradigma delle reti neurali.

Poiché per avere successo l'IA deve disporre di un sistema artificiale nel quale riprodurre, emulandoli, i fenomeni dell'intelligenza, l'elaboratore è stato considerato fin dall'inizio come il miglior candidato a questo ruolo. Alla base dell'elaboratore vi è il concetto di *macchina di Turing*, una macchina concettuale in grado di

trovarsi in un numero finito di stati diversi e di eseguire un numero limitato di azioni, al fine di poter esprimere qualsiasi tipo di procedura definita. La macchina di Turing è composta da un nastro di lunghezza infinita, suddiviso in celle lette da una testina, che può spostarsi di una cella avanti o indietro, e da un organo di controllo capace di leggere il simbolo che si trova nella cella sotto la testina. A un dato istante, l'azione che la macchina intraprende è determinata dal simbolo letto e dalla configurazione in cui la macchina si trova in quel momento. Dopo avere letto il simbolo stampato sulla cella, la testina può compiere due operazioni alternative: lasciare il simbolo così com'è oppure cancellarlo e stamparne un altro. Il concetto di algoritmo può essere ricondotto alla sequenza di operazioni svolte dalla macchina di Turing. Il contributo di Alan Turing (1912-1954) all'IA deve essere riconosciuto non solo nell'ambito dei fondamenti dell'informatica, ma anche in quello del dibattito filosofico relativo ai limiti e alle potenzialità delle nuove "macchine pensanti". In un noto articolo (Turing 1950), Turing propone il famoso test che porta il suo nome per verificare la presenza o meno di intelligenza in una macchina.

Se nei secc. XVIII, XIX e XX con la formalizzazione delle scienze e della matematica sono state create le condizioni per lo studio dell'intelligenza e delle sue possibili artificializzazioni, è solo con l'avvento dei primi elaboratori elettronici, a cavallo della Seconda guerra mondiale, che tale interesse può imboccare una strada concreta, arrivando alla definizione del programma di ricerca espresso dal seminario di Dartmouth nel 1956. Nell'estate di quell'anno un gruppo di studiosi si riunisce al Dartmouth College con lo scopo di esaminare, come appare nella proposta per il seminario redatta l'anno precedente, la congettura che ogni aspetto dell'intelligenza possa essere, in linea di principio, descritto in modo tanto preciso da poter far sì che una macchina lo simuli. Questa data segna l'inizio ufficiale di una nuova disciplina che il matematico John McCarthy, assistant professor a Dartmouth e uno degli organizzatori del seminario, propose di chiamare intelligenza artificiale. Gli altri organizzatori erano Marvin Minsky, junior fellow di matematica e neurologia a Harvard, Nathaniel Rochester, direttore della ricerca sull'informazione in un centro ricerche dell'IBM, e Claude Elwood Shannon, il matematico già famoso per la teoria dell'informazione, allora ai Bell Telephone Laboratories. Il seminario aveva le caratteristiche del 'brainstorming', ossia di un dibattito aperto e poco strutturato dal quale emerse, attraverso le discussioni comuni, un nuovo approccio teoretico teso a definire la possibilità della riproduzione dell'intelligenza da parte di un elaboratore elettronico. Il seminario si proponeva inoltre di raccogliere e analizzare i programmi caratterizzati da prestazioni definibili come intelligenti – come il Logic Theorist (LT) di Allen Newell, Bernard Shaw e Herbert A. Simon in grado di dimostrare teoremi della logica del primo ordine – e di proporre una serie di obiettivi ambiziosi che avrebbero dovuto essere verificati dieci anni dopo in un nuovo incontro.

Quelli che seguono il seminario di Dartmouth sono gli anni delle grandi aspettative, alimentate anche dai successi dovuti ai miglioramenti vertiginosi dei supporti informatici utilizzati. Indicativamente possono essere rilevate due tendenze: da una parte il gruppo guidato da Newell, Shaw e Simon interessato alla simulazione dei processi cognitivi umani per mezzo dell'elaboratore, che con il GPS (General Problem Solver) del 1958 intendeva estendere l'ambito delle applicazioni del LT oltre quelle puramente logiche (paradigma della simulazione); dall'altra tutti coloro i quali dedicavano le loro forze al raggiungimento della migliore prestazione possibile per i programmi, indipendentemente dal fatto che questa potesse essere realizzata adottando procedure più o meno imitative dei procedimenti seguiti dall'uomo (paradigma della prestazione o dell'emulazione).

In questo periodo si assiste al temporaneo declino dei modelli a reti neurali (risorti, come detto, anni dopo), a causa della critica decisiva di Minsky al Perceptron di Frank Rosenblatt, un sistema neurale incapace di riconoscere stimoli visivi anche molto semplici. Parallelamente, il paradigma della prestazione produceva risultati di successo, come i primi programmi per il gioco della dama e degli scacchi e quelli basati sul tipo di rappresentazione della conoscenza introdotta dall'Advice Taker di McCarthy: un programma (mai effettivamente realizzato) che può essere visto come il primo sistema completo di IA capace di impiegare conoscenza generale sul mondo che può essere facilmente modificata. L'enfasi di questo primo periodo si pone quindi sui meccanismi generali di ricerca e su una concezione limitata della nozione di intelligenza, quale per esempio l'abilità di giocare a scacchi o la capacità di risolvere problemi matematici.

Ben presto però i ricercatori cominciano a incontrare i primi insuccessi: metodi adeguati per casi semplici si rivelano totalmente inadeguati in contesti più complessi e ampi. Le grandi aspettative iniziali devono fare i conti con il fallimento dei progetti di traduzione automatica fra linguaggi naturali; i programmi che agiscono a livello della semplice manipolazione sintattica si rivelano infatti totalmente inadeguati, causando il ritiro delle ingenti sovvenzioni dei governi americano e inglese. A tutto ciò si aggiunge l'incapacità di trattare l'esplosione combinatoria: l'estensione a problemi più ampi non si affronta semplicemente con hardware più veloce e memorie più grandi, esistendo insormontabili limitazioni insite nella natura stessa di alcuni problemi.

Queste difficoltà conducono, a partire dal 1970, a concentrarsi su aree più ristrette di competenza con la realizzazione dei primi *sistemi esperti*, nei quali l'importanza di una conoscenza dettagliata dello specifico dominio in esame gioca un ruolo fondamentale. La conoscenza non è più legata soltanto alla comprensione teorica del problema ma anche a specifiche regole euristiche convalidate dall'esperienza. In un sistema esperto il complesso di queste due conoscenze, presente in un esperto umano di uno specifico settore, è codificato e rappresentato in una forma che l'elaboratore utilizza allo scopo di risolvere problemi simili a quelli affrontati dall'esperto umano. Il primo sistema esperto, DENDRAL, viene programmato per inferire la struttura di molecole organiche in base alle loro formule chimiche; MYCIN, forse il più conosciuto tra i sistemi esperti,

incorpora conoscenza medica specifica che usa per diagnosticare e prescrivere trattamenti per le infezioni batteriche del sangue a partire da informazioni sui sintomi incomplete e incerte. L'importanza della conoscenza del dominio risulta evidente anche nell'area della elaborazione del linguaggio naturale: in questi anni si assiste al passaggio dall'attenzione per la sintassi a quella per la semantica, che porta a risultati interessanti.

A partire da questo ridimensionamento delle aspettative si assiste, all'inizio degli anni Ottanta dello scorso secolo, alla nascita dell'IA come industria. Nel 1982 viene progettato il primo sistema esperto commerciale di successo, che supporta le configurazioni di ordini per nuovi sistemi di elaborazione in un'azienda produttrice. L'IA viene a far parte di un ampio sforzo che include la progettazione di chip e la ricerca relativa alle interfacce uomo-macchina. Parallelamente si assiste al ritorno dell'approccio basato sulle reti neurali: intorno al 1985 quattro differenti gruppi di ricerca inventano nuovamente un algoritmo di apprendimento, già scoperto anni prima, basato sulla retropropagazione dell'errore e lo applicano con successo a molti problemi di apprendimento in informatica e ingegneria. Il ritorno di questo approccio viene promosso anche dalla nascita di una nuova disciplina: le *scienze cognitive* che, nel 1979, si consacrano ufficialmente come disciplina autonoma in cui confluiscono molte delle ambizioni di una parte della psicologia e di quella parte dell'IA, più scientifica e meno ingegneristica, che ha sempre considerato la macchina come strumento privilegiato per lo studio della mente.

Negli ultimi anni l'IA è stata caratterizzata da numerosi cambiamenti sia a livello metodologico sia a livello contenutistico. Attualmente la tendenza comune è quella di partire da teorie già esistenti piuttosto che inventarne di nuove, basando le asserzioni su teoremi rigorosamente dimostrati o su evidenze sperimentali. L'attenzione è stata rivolta a problemi reali chiaramente delimitati, come il riconoscimento del parlato o la pianificazione delle attività di una fabbrica. Ciò che resta della caratterizzazione dell'IA delle origini è la pluralità di approcci: accanto al tradizionale approccio logico della rappresentazione della conoscenza, ha acquistato peso crescente l'approccio *subsimbolico*, concepito per dotare i sistemi di IA di prestazioni intelligenti anche senza una rappresentazione dettagliata della conoscenza. Vasta attenzione è stata inoltre dedicata negli ultimi anni ai metodi probabilistici e fuzzy, utilizzati per consentire un ragionamento efficiente sulla base di evidenze incerte. Si delinea così l'estensione pragmatica dei confini della disciplina, ogniqualvolta si raggiungano nuovi risultati e si progettino nuove applicazioni, nella direzione di una visione dell'intelligenza non limitata al solo 'pensare razionalmente' bensì sempre più estesa all' 'agire razionalmente'.

## 2. L'intelligenza artificiale e l'informatica in una struttura culturale unitaria

Come è stato messo in luce anche nella disamina della sua storia, un corretto inquadramento dell'IA richiede di valutare l'opportuna collocazione di questa disciplina all'interno di una disciplina più ampia, l'informatica, della quale l'IA è una parte e con la quale è in relazione sia a livello dei fondamenti sia a livello delle applicazioni. Nel campo dei fondamenti dell'informatica l'IA tende a ricercare e a sviluppare le teorie di base, sia già esistenti (per es., la logica matematica), sia del tutto nuove (come la teoria dei problemi). Nel campo delle applicazioni dell'informatica, l'IA rappresenta una disciplina sperimentale che si propone di affrontarne di nuove, specialmente di tipo non numerico. Un completo inquadramento dell'IA va affrontato, oltre che nell'ambito dell'informatica, estendendo talune considerazioni anche alla robotica, in modo tale da pervenire ad alcune conclusioni di natura generale.

È interessante osservare come questa contestualizzazione porti alla considerazione di una struttura generale che meglio può aiutare a comprendere tutto il quadro teorico al quale si fa riferimento (Somalvico 1987). Partendo da una visione che potremmo definire 'ontologica', ossia relativa alla struttura della realtà, possiamo distinguere il 'reale' nelle due componenti del 'naturale' e dell' 'artificiale' (fig. 1).

Il 'naturale' è a sua volta costituito da due elementi: l' 'uomo' e il 'mondo' (inteso come il naturale esterno all'uomo), mentre l'artificiale è riferito alla nozione di 'macchina', che per l'informatica consiste nell' 'elaboratore' e per la robotica consiste nel 'robot' (che, più opportunamente, potrebbe essere denominato 'interelaboratore'). Pertanto, sempre da un punto di vista molto astratto, si può considerare l'intero scenario del reale come articolato nelle tre entità di uomo, mondo e macchina, rappresentabili in un diagramma tripolare del reale (fig. 2).

Con l'informatica l'uomo concepisce la realizzazione di una macchina, l'elaboratore, che riproduce rozzamente talune facoltà dell'intelligenza umana che l'uomo impiega quando riflette sul reale e risolve alcuni dei problemi sorti conseguentemente a tale riflessione. Pertanto l'uomo, progettando l'elaboratore, sostituisce se stesso con l'elaboratore nell'intelligere senza coinvolgere l'interagire con il mondo. Con la robotica l'uomo concepisce la realizzazione di una macchina, il robot (l'interelaboratore), che riproduce rozzamente sia talune facoltà dell'intelligenza umana che riflette sul reale, sia talune facoltà dell'interazione dell'uomo, con le quali egli agisce e percepisce (cioè interagisce) nel reale, sia infine talune facoltà che gli permettono di risolvere alcuni dei problemi causati da tale riflessione e da tale azione e percezione. Dunque l'uomo, progettando il robot, sostituisce se stesso con il robot nell'intelligere e nell'interagire con il mondo.

Si noti inoltre che, mentre la problematica dell'interazione della macchina con il mondo è di rilevanza solo per la robotica, quella dell'interazione della macchina con l'uomo è di profonda importanza non soltanto nell'ambito della robotica ma anche nell'ambito dell'informatica. L'informatica è infatti legata alla macchina

elaboratore mentre la robotica è legata alla macchina robot (la macchina che ha la capacità di interagire con il mondo) e ogni macchina (elaboratore e robot) è dotata della capacità di interagire con l'uomo. In questo contesto la robotica viene considerata non come una disciplina che comprende l'informatica come sua sottodisciplina ma come distinta da essa. In linea di principio, tutte le attività di elaborazione dell'informazione, che nell'ambito dell'informatica sono considerate come disponibili all'elaboratore, nell'ambito della robotica possono essere considerate come disponibili al robot in modo integrato con le altre attività di interazione con il mondo peculiari della robotica.

È in questo quadro di riferimento concettuale che è opportuno trattare brevemente gli aspetti dell'IA rispetto alle altre aree dell'informatica dal punto di vista della loro natura e delle loro potenzialità. A tale scopo, è necessario esaminare la natura dell'elaborazione dell'informazione, sia nell'elaboratore sia nel robot. Si osservi anzitutto che per 'informazione', intesa come oggetto dell'attività di elaborazione esplicita dall'elaboratore e dal robot, si intende quella classe di modelli del reale che hanno la proprietà di essere sottoposti, come 'operandi', alle attività di elaborazione e di interelaborazione svolte da quegli 'operatori' che sono le macchine (elaboratori nel caso dell'informatica e robot nel caso della robotica). Quindi elaboratori e robot sono macchine che operano su modelli del reale chiamati informazione.

Quando si esamina l'elaborazione dell'informazione si possono sommariamente identificare tre tipi di informazione: (1) l'algoritmo; (2) il dato; (3) il problema. Si noti che, quando il termine tecnico 'informazione' è adoperato per indicare solamente il modello di un problema, può essere sostituito dal più specifico termine tecnico 'conoscenza'. Di conseguenza si possono identificare sommariamente anche tre tipi di elaborazione dell'informazione: (1) l'esecuzione di algoritmi; (2) la gestione di dati, termine composto che include la scrittura, la memorizzazione, la modifica e la lettura dei dati; (3) la risoluzione di problemi; una dizione equivalente a quest'ultimo tipo di elaborazione dell'informazione è quella di 'inferenza sulla conoscenza'.

Quando si esamina la robotica, cioè l'interelaborazione dell'informazione, ai tre precedenti tipi di informazione (algoritmo, dato, problema) si aggiungono due ulteriori tipi: (1) il modello dei fenomeni prodotti dal robot; (2) il modello dei fenomeni percepiti dal robot. Conseguentemente, vengono identificati due tipi di interazione con il mondo: (1) la produzione di fenomeni conformi a un loro opportuno modello (l'informazione) posseduto dal robot, che esplica tale attività tramite il sistema degli attuatori (un'interfaccia di interazione in uscita dal robot verso il mondo); (2) la percezione di fenomeni descritti da un loro opportuno modello (l'informazione) posseduto dal robot, che esplica tale attività tramite il sistema dei sensori (un'interfaccia di interazione in entrata nel robot dal mondo).

È allora possibile affermare che l'IA studia gli elaboratori e i robot nei quali è svolta elaborazione risolutoria o inferenziale (Somalvico 1991). Va osservato che la parte della robotica d'interesse per l'IA prende il nome di 'robotica intelligente'. Oggigiorno è usuale indicare l'elaboratore o il robot che svolge elaborazione risolutoria con l'espressione 'agente intelligente' o 'sistema intelligente' (tav. I).

### 3. Intelligenza artificiale e risoluzione di problemi

Uno dei contributi fondamentali che l'IA ha fornito all'informatica è stato l'aver posto l'accento sulla nozione di problema, affiancandola a quella di algoritmo e di dato, studiando ed evidenziando le varie modalità per la sua modellizzazione.

L'attività di modellizzazione di un problema prende il nome di 'rappresentazione del problema' o 'rappresentazione della conoscenza'. Sotto questo punto di vista, particolare rilevanza ha acquisito il tema specifico della 'risoluzione automatica dei problemi', ossia l'obiettivo di dotare l'elaboratore di metodi generali e di programmi efficienti che gli consentano di costruire, secondo processi completamente automatici o interattivi, l'algoritmo che risolve un problema (ottenendo quindi, dopo la costruzione e l'esecuzione di tale algoritmo, la soluzione del problema stesso). Si noti come questo approccio all'utilizzo dell'elaboratore richieda che l'uomo sia in grado di fornire alla macchina un algoritmo capace di costruire un altro algoritmo, ossia un meta-algoritmo. In IA tali algoritmi prendono il nome di 'algoritmi inferenziali'. Il loro impiego permette di fornire all'elaboratore, come dato di ingresso, la rappresentazione del problema che la macchina deve risolvere automaticamente. Per la risoluzione automatica dei problemi la macchina deve essere non soltanto abile e veloce nell'esecuzione di algoritmi costruiti dall'uomo ma anche capace di costruire autonomamente gli algoritmi necessari per risolvere problemi (Nilsson 1971). Tali algoritmi costruiti dalla macchina possono essere eseguiti sia dalla macchina stessa sia, in alcune situazioni, dall'uomo.

Per analizzare gli aspetti concettuali e funzionali della risoluzione di un problema e per chiarire l'affermazione secondo la quale l'IA studia gli elaboratori e i robot che svolgono elaborazione risolutoria o inferenziale, è utile partire da un parallelo con il mondo della fisica (fig. 3).

In fisica esiste una differenza tra l'occorrenza di un fenomeno percepito dall'uomo e la formalizzazione di un modello; quest'ultimo descrive taluni parametri fisici che definiscono il fenomeno e ne fornisce pertanto una rappresentazione. Analogamente, in IA esiste una differenza tra un problema intuitivo, cioè l'esigenza avvertita dall'uomo di un problema da risolvere, e la sua formalizzazione, che in IA viene denominata 'problema rappresentato'.

Inoltre, il fisico utilizza il modello di un fenomeno al fine di derivarne, mediante un'attività inferenziale, una legge che consenta di caratterizzare le proprietà del fenomeno e di prevederne di nuovi. Allo stesso modo, in IA il risolutore automatico di problemi è in grado di operare un'elaborazione sul problema rappresentato che permetta di costruire il 'problema risolto', espresso sia dall'algoritmo risolvete il problema rappresentato sia dalla sua esecuzione, che produce effettivamente la soluzione del problema.

Pertanto ai tre concetti in fisica, di fenomeno, di modello e di legge corrispondono, nell'IA, quelli di problema intuitivo, problema rappresentato e problema risolto.

È interessante a questo proposito analizzare con attenzione il ruolo dell'uomo e dell'elaboratore nel complesso processo della risoluzione di un problema. Mentre in fisica le attività sono compiute solamente dall'uomo, nell'IA esse sono svolte tanto dall'uomo quanto dall'elaboratore. In particolare, l'IA si pone nell'ambiziosa prospettiva di fornire alla macchina la capacità di procedere alla risoluzione automatica dei problemi, ossia di passare dal problema rappresentato al problema risolto. Sono invece mantenuti come esclusiva prerogativa dell'attività umana, almeno nell'approccio adottato in questo contributo, tanto il passaggio dal problema intuitivo al problema rappresentato, ossia la capacità di formalizzazione di un problema, quanto il passaggio dal problema risolto al problema intuitivo, ossia la capacità di confronto della soluzione ottenuta con l'esigenza da soddisfare, dalla quale ha preso le mosse l'intero procedimento di risoluzione.

Considerando da un punto di vista epistemologico tanto il caso della fisica quanto il caso dell'IA, appare chiaro che le attività in gioco possono essere articolate nel modo seguente: (1) passaggio (chiamato formalizzazione in fisica e rappresentazione in IA) dal mondo della realtà (fenomeno, problema intuitivo) al mondo della conoscenza della realtà (modello, problema rappresentato); (2) passaggio (chiamato derivazione e formulazione in fisica, risoluzione ed esecuzione in IA, inferenza in filosofia), nell'ambito della conoscenza della realtà, dal modello alla legge, dal problema rappresentato al problema risolto; (3) passaggio (chiamato aspettazione e previsione in fisica e confronto e soddisfacimento in IA) dal mondo della conoscenza della realtà (legge, problema risolto) al mondo della realtà (fenomeno, problema intuitivo).

È relativamente all' 'inferenza' (operante all'interno della conoscenza della realtà) che l'IA si concentra nell'emulazione delle attività intellettuali dell'uomo, con alcune differenze nei diversi casi.

Possiamo infatti affermare che in talune situazioni, per esempio nell'applicazione della risoluzione automatica di problemi ai giochi (come nel caso degli scacchi), la macchina, dopo avere costruito l'algoritmo risolvete, lo utilizza eseguendolo. Più precisamente essa: (1) esegue l'algoritmo inferenziale che nel gioco degli scacchi è quello che esplora, costruendolo fino a una certa profondità, il cosiddetto 'albero del gioco', ossia l'articolata espressione gerarchica delle sequenze di tutte le mosse convenienti, contromosse, controcontromosse, ecc.; (2) costruisce l'algoritmo risolvete il problema, che nel gioco degli scacchi è il cammino lungo l'albero del gioco che partendo dalla radice, caratterizzante la posizione attuale sulla scacchiera, raggiunge la 'foglia' che definisce la posizione di scacco matto, oppure la posizione che viene valutata come la più vantaggiosa tra quelle corrispondenti a tutte le foglie dell'albero; (3) esegue l'algoritmo risolvete il problema, del quale ottiene la soluzione, che, nel gioco degli scacchi, è la mossa corrispondente al primo elemento del cammino lungo l'albero del gioco adottato.

Si noti però che in altre situazioni, per esempio nei sistemi esperti di diagnosi delle malattie nell'uomo (come nel caso del sistema esperto MYCIN (Buchanan, Shortliffe 1984) dedicato alla diagnosi delle malattie del sangue), la macchina costruisce l'algoritmo risolvete, mentre l'uomo lo utilizza eseguendolo. Più precisamente, la macchina: (1) esegue l'algoritmo inferenziale che in MYCIN è quello che esplora, costruendolo fino a una certa profondità, il cosiddetto albero delle concatenazioni (costituite da coppie causa-effetto), le quali arrivano alla diagnosi di una malattia e alle correlate prognosi e terapie partendo dalla valutazione dei sintomi rilevati; (2) costruisce l'algoritmo risolvete il problema, che in MYCIN è il cammino lungo l'albero delle concatenazioni causa-effetto che partendo dalla radice, caratterizzante i sintomi iniziali, si articola in modo consequenziale fino a raggiungere l'identificazione diagnostica della malattia. A questo punto l'uomo, e non più la macchina, come accadeva nella precedente situazione del gioco degli scacchi, (3) esegue l'algoritmo risolvete il problema, del quale ottiene la soluzione, che nel sistema esperto MYCIN riguarda la prima domanda formulata al medico o il primo esame diretto o di laboratorio indicato dall'albero delle concatenazioni. La risposta alla domanda o l'esito dell'esame viene inserito nella macchina, che continua nella costruzione del cammino coerentemente con le risposte e gli esiti ottenuti dal medico. È opportuno quindi notare che nel sistema MYCIN (e nei sistemi esperti in generale) è la macchina che costruisce l'algoritmo risolvete il problema, mentre è l'uomo che è invitato a eseguirlo.

#### **4. Aspetti scientifici all'intelligenza artificiale**

La definizione di IA data all'inizio risulta ora ulteriormente precisata alla luce dell'illustrazione della risoluzione di problemi. Un sistema di IA deve essere in grado di acquisire, rappresentare ed elaborare conoscenza (e non semplicemente dati simbolici separati dal loro significato) relativa al compito da eseguire e di applicarla in meccanismi di elaborazione intelligenti, al fine di fornire le prestazioni richieste. I sistemi di IA sono pertanto caratterizzati non soltanto dalla capacità di fornire prestazioni che si potrebbero considerare esclusive dell'intelligenza dell'uomo, ma anche dalla capacità di saper gestire, elaborare e scambiare

conoscenza mediante meccanismi tipici dell'intelligenza umana, quali (ma non solamente): l'inferenza, la deduzione, il ragionamento con incertezza, il ragionamento analogico, la generalizzazione, la particolarizzazione, la generazione e la valutazione di ipotesi, l'apprendimento.

L'IA può anche essere definita come la scienza che affronta il problema di come rappresentare, manipolare e costruire conoscenza circa fatti, azioni e leggi di causalità. Da un punto di vista metodologico, i ricercatori di IA hanno l'obiettivo di costruire programmi per l'elaboratore che realizzino alcune attività intellettuali proprie dell'uomo, congiuntamente al tentativo di spiegare i principi basilari dell'intelligenza. Da un punto di vista pratico, diverse sono le aree di ricerca che hanno fornito e che forniscono le tecniche e i metodi per realizzare tali programmi; in particolare: (1) 'la ricerca nello spazio degli stati', per programmi che determinano automaticamente una sequenza di azioni per risolvere un problema; (2) 'il soddisfacimento di vincoli', per programmi che individuano automaticamente un insieme di valori che soddisfano un dato insieme di vincoli; (3) 'la formalizzazione e il ragionamento basati sulla logica', per applicazioni che agiscono su rappresentazioni del mondo basate su insiemi di espressioni logiche per derivare proprietà del mondo (implicitamente contenute nella rappresentazione iniziale); (4) 'la formalizzazione e il ragionamento probabilistico e fuzzy', per la progettazione di programmi che agiscono in condizioni di incertezza riguardo all'ambiente esterno al programma; (5) 'la pianificazione', per determinare automaticamente una sequenza di azioni per raggiungere un obiettivo; (6) 'l'apprendimento automatico', per la realizzazione di programmi che migliorano le loro prestazioni in base all'esperienza; (7) 'le reti neurali', per applicazioni basate su complesse funzioni non lineari apprese automaticamente dai dati; (8) 'gli algoritmi genetici', che ottimizzano una funzione basandosi su una metafora evolutiva; (9) 'i sistemi di programmazione e i linguaggi dell'IA', che supportano lo sviluppo di programmi che incorporano le tecniche precedenti.

A titolo di esempio illustrativo del funzionamento di alcuni sistemi di IA, analizziamo ora più in dettaglio il metodo della ricerca nello spazio degli stati, storicamente rilevante nello sviluppo dell'IA. Tale metodo e gli algoritmi inferenziali da esso derivati permettono di situare uno dei fondamenti teorici dell'IA nella branca della matematica costituita dall'algebra astratta.

L'approccio formale alla risoluzione di problemi presentato qui di seguito ha il pregio di illustrare le abilità meta-algoritmiche o inferenziali che l'IA ha fornito alla macchina e riveste notevole importanza, sia dal punto di vista teorico sia dal punto di vista applicativo. Basti ricordare, a tale proposito, i risultati eccellenti che l'IA ha ottenuto nella pianificazione e nel gioco degli scacchi.

Questo approccio formale è denominato 'metodo dello spazio degli stati'. Esso rappresenta un modello molto utile per potere concepire e realizzare non solo la formalizzazione dei metaalgoritmi ma anche la loro programmazione e, in tal modo, costruire quel complesso sistema di programmi denominato 'risolutore automatico di problemi'.

Lo spazio degli stati può essere pensato come un modo di rappresentare un problema  $P$  attraverso una quadrupla  $P=(S,G,i,F)$ , dove  $S$  è un insieme (lo spazio degli stati) non necessariamente finito di elementi detti stati;  $G$  è un insieme finito di funzioni  $g$  definite su  $S$  (gli operatori),  $g:S \rightarrow S$ ;  $i$  è uno stato di  $S$  (lo stato iniziale),  $i \in S$ ;  $F$  è un sottoinsieme di  $S$  (l'insieme degli stati finali),  $F \subseteq S$ .

Esistono due modi per rappresentare l'insieme  $S$ . Uno esplicito, ossia indicando tutti gli stati possibili:  $S=\{s_0, s_1, \dots, s_n\}$ ; l'altro implicito, fornendo un algoritmo  $E$  di enumerazione dei possibili stati candidati ad appartenere a un dato spazio degli stati e un algoritmo  $AS$  di accettazione, in grado di decidere sull'appartenenza o meno di uno stato candidato allo spazio degli stati. Analogamente si può ragionare per l'insieme degli stati finali, fornendo un algoritmo  $AF$ , in grado di decidere sull'appartenenza o meno di uno stato candidato all'insieme degli stati finali.

Applicando un operatore a uno stato si ottiene un nuovo stato e, se l'operatore applicato è legale e lo stato iniziale appartiene a  $S$ , anche lo stato ottenuto appartiene a  $S$ .

Si definisce soluzione del problema  $P$  un cammino  $X[P]$  dallo stato iniziale  $i$  a uno degli stati finali  $f$  appartenenti a  $F$ .  $X[P]$  è pertanto una sequenza finita di operatori legali:  $\langle g_1, g_2, \dots, g_k \rangle$  tale che:

$$i = s_0 \rightarrow g_1 \rightarrow s_1 \rightarrow g_2 \rightarrow s_2 \rightarrow \dots \rightarrow g_k \rightarrow s = f.$$

L'attività di generazione di tutti i nuovi stati ottenibili a partire da un vecchio stato, mediante l'applicazione di tutti gli operatori applicabili a quest'ultimo, è detta 'espansione' del vecchio stato. Il vecchio stato si dice il 'predecessore' del nuovo stato e il nuovo stato si dice il 'successore' del vecchio stato.

Questo metodo di formalizzazione permette di associare alla rappresentazione di un problema l'isomorfa descrizione di un grafo (dello spazio degli stati) orientato ed etichettato, dove i vertici corrispondono agli stati, gli archi agli operatori e la soluzione a un cammino del grafo, che unisce lo stato iniziale  $i$  a uno degli stati finali  $f$  di  $F$ .

È pertanto evidente che i metaalgoritmi necessari per costruire la soluzione di un problema così formalizzato coincidono con gli algoritmi necessari per espandere un grafo al fine di costruire il cammino che unisce un vertice (detto iniziale) con un altro vertice (detto finale), appartenente a un insieme di vertici finali.

Si chiama 'spazio di ricerca' un sottoinsieme dello spazio degli stati che deve essere effettivamente esplicitato al fine di individuare un cammino (solitamente di costo minimo, se si associa a ogni operatore un costo) che costituisca una soluzione (possibilmente la soluzione ottima) del problema.

## 5. Applicazioni dell'intelligenza artificiale

Le applicazioni dei sistemi di IA sono innumerevoli e una loro completa catalogazione risulta molto difficile a causa della quantità ed eterogeneità dei settori interessati. Come emerge dalle applicazioni sommariamente tratteggiate nel seguito, la concezione al momento dominante nello sviluppo dei sistemi di IA è quella dell' 'agire razionalmente', che individua una visione più ampia e più complessa di intelligenza legata non soltanto alle capacità intellettuali ma anche all'abilità di saper interagire con il mondo. In questa direzione l'enfasi è posta sempre più sull'efficienza dei risultati dei sistemi di IA piuttosto che sulla loro generalità. Ciò comporta spesso lo sviluppo di sistemi con competenze molto specifiche in contesti applicativi ben delimitati, invece di sistemi con generiche caratteristiche di 'intelligenza' applicabili a diversi contesti. Un'imprecisa, incompleta, ma utile approssimazione tassonomica prevede le seguenti aree di applicazione per i sistemi di IA.

*La pianificazione autonoma* di attività e operazioni è l'area che maggiormente interessa la produzione industriale e la logistica, anche se sistemi di pianificazione autonoma sono utilizzati in altre applicazioni come, per esempio, a bordo di sonde spaziali. I sistemi che operano in quest'area sono in grado di ricevere un obiettivo di alto livello, di generare un piano composto da una sequenza di operazioni semplici per raggiungere l'obiettivo assegnato e, infine, di monitorare l'esecuzione del piano.

*I giochi* sono una delle aree tradizionali dell'IA da un punto di vista metodologico. Lo studio dei giochi ha determinato lo sviluppo e la messa a punto di numerose tecniche, soprattutto quelle della ricerca nello spazio degli stati. Nel 1997, per la prima volta nella storia, un programma sviluppato dalla IBM (Deep Blue) ha sconfitto il campione del mondo Garry Kasparov con il punteggio di 3,5 a 2,5.

*Il controllo autonomo* è un campo, nell'ambito del controllo di sistemi complessi, nel quale alcuni sistemi di IA si sono dimostrati molto efficaci, come nel caso del controllo di automobili e di sonde spaziali.

*La dimostrazione automatica di teoremi matematici e la programmazione automatica* sono due aree classiche dell'IA da un punto di vista metodologico e comportano spesso ricerche di natura teorica. In passato la ricerca nella prima di queste aree era volta a scoprire e provare nuovi teoremi nell'aritmetica, nella geometria, nella logica e nella meccanica. Di recente, invece, i ricercatori si sono interessati prevalentemente allo sviluppo di tecniche inferenziali generali che, da un dato insieme di fatti, consentano di dedurre fatti nuovi o loro proprietà e relazioni. Ciò influisce sul progetto di diversi sistemi di IA quali, per esempio, i sistemi esperti e i 'sistemi domanda e risposta'. La programmazione automatica ha come obiettivo la costruzione di sistemi i quali, partendo da specifiche ad alto livello delle funzioni che un programma deve eseguire, siano in grado di generare automaticamente il programma stesso, corretto ed efficiente. I risultati ottenuti, sebbene non ancora concretamente utilizzabili nella costruzione di sistemi reali, fanno prevedere un notevole impatto sulle tecniche di produzione di software.

*La robotica e la visione artificiale.* L'area della robotica intelligente (cioè quella di interesse per l'IA) si articola nelle sottoaree della manipolazione e della navigazione. In generale, le aree di ricerca della manipolazione e della navigazione si estendono a un contesto multidisciplinare che va oltre il contributo fornito dall'IA per comprendere anche la meccanica, la sistemistica e l'elettronica. L'obiettivo è la realizzazione di macchine autonome, capaci di sostituirsi all'uomo nell'esecuzione di attività manuali, ripetitive, pesanti e nocive. La manipolazione e la navigazione si occupano dei problemi geometrici e di movimento connessi alla realizzazione di robot intelligenti e hanno raggiunto da tempo risultati concreti e rilevanti anche dal punto di vista applicativo. L'area di ricerca della visione artificiale riguarda il problema dell'elaborazione dell'informazione raccolta dai sensori visivi, per esempio da telecamere, al fine di riconoscere e classificare forme, oggetti e scene bidimensionali e tridimensionali. I risultati teorici e sperimentali della visione artificiale sono assai brillanti e significative applicazioni sono attualmente disponibili sul mercato. Uno fra i principali risultati ottenuti in questi ambiti è la vettura autonoma Stanley, sviluppata dalla Stanford University, che nell'ottobre del 2005 è stata in grado, viaggiando senza alcun intervento umano, di percorrere 132 miglia nel deserto del Mohave (in California) in poco meno di 7 ore nell'ambito della competizione DARPA Grand Challenge.

*L'elaborazione del linguaggio naturale.* I problemi affrontati in quest'area di ricerca sono estremamente complessi, riguardando alcuni dei meccanismi meno noti e più sfuggenti del comportamento umano. La comunicazione, infatti, presenta ancora molti problemi aperti, sia dal punto di vista linguistico sia da quello psicologico; la realizzazione di funzioni per la comprensione e per la generazione di espressioni linguistiche per mezzo dell'elaboratore non ha ancora raggiunto risultati tecnologicamente assestati. Vi sono parecchi sistemi funzionanti che consentono di risolvere in modo soddisfacente solo problemi limitati, quali dialogare con banche dati, impartire ordini a un robot, consultare sistemi esperti, tradurre approssimativamente semplici frasi da una lingua a un'altra. Collegata alla elaborazione del linguaggio naturale è l'area dell'elaborazione del linguaggio vocale: il riconoscimento e la sintesi del linguaggio vocale dell'uomo richiedono l'elaborazione di informazioni complesse di natura fonetica, morfologica, sintattica, semantica e pragmatica e rappresentano un obiettivo molto ambizioso. I risultati delle ricerche teoriche e sperimentali svolte hanno permesso di costruire sistemi dotati di capacità limitate di riconoscimento e comprensione (parole pronunciate in modo staccato, vocabolario limitato), nonché sistemi di sintesi in grado di produrre artificialmente il linguaggio parlato, leggendo un testo scritto. Il potenziale applicativo di tali ricerche è senz'altro molto rilevante.



*I sistemi esperti e ontologie.* Il mondo industriale si riconosce intorno all'ingegneria della conoscenza, termine che individua gli aspetti ingegneristici dell'IA e che ha per anni posto in primo piano la tecnologia dei sistemi esperti impiegata, per esempio, in compiti di diagnosi. I sistemi esperti hanno rivoluzionato l'approccio all'impiego dell'elaboratore, non richiedendo più, al fine di risolvere problemi applicativi, di programmare algoritmi per la macchina algoritmica ma esigendo solamente di rappresentare la conoscenza che la macchina algoritmica utilizza per costruire inferenzialmente la soluzione dei problemi applicativi (purché questi siano inerenti alla conoscenza posseduta dal sistema esperto). L'interesse per i sistemi esperti è stato accresciuto dall'opportunità di modellizzare la conoscenza e non di meccanizzare gli algoritmi risolutivi, rendendo possibile l'interazione con la macchina direttamente a coloro che possiedono la conoscenza stessa. A partire da diverse ricerche precedenti, comprese quelle sulla rappresentazione della conoscenza, negli ultimi anni lo studio delle *ontologie*, ovvero delle descrizioni formali di domini specifici, ha raggiunto un buon grado di maturità e promette di aprire nuovi settori applicativi.

In conclusione, ribadiamo che gli ambiti di applicazione dell'IA non si esauriscono in quelli che abbiamo citato poiché è in continua espansione, muovendosi su un territorio di frontiera della ricerca, e poiché interessa anche altri contesti, quali l'elaborazione dei segnali, l'analisi dei dati, l'intrattenimento e la medicina, che contribuiscono a caratterizzarla come una disciplina dinamica, aperta e ricca di promettenti prospettive.

## Bibliografia

- Buchanan, Shortliffe 1984: Buchanan, Bruce G. - Shortliffe, Edward H. (a cura di), *Rule-Based Expert Systems: The MYCIN Experiments of the Stanford Heuristic Programming Project*, Reading (MA), Addison-Wesley, 1984.
- Cordeschi 1996: Cordeschi, Roberto, *L'intelligenza artificiale*, in: Geymonat, Ludovico, *Storia del pensiero filosofico e scientifico. Il Novecento*, a cura di Enrico Bellone e Corrado Mangione, Milano, Garzanti, 1996, VIII, pp. 145-200.
- McCorduck 1979: McCorduck, Pamela, *Machines who think*, San Francisco, W. H. Freeman, 1979.
- Nilsson 1971: Nilsson, Nils J., *Problem solving methods in artificial intelligence*, New York, McGraw Hill, 1971.
- Nilsson 1998: Nilsson, Nils J., *Artificial intelligence: a new synthesis*, San Mateo (CA), Morgan Kaufmann, 1998.
- Russell, Norvig 2003: Russell, Stuart J. - Norvig, Peter, *Artificial intelligence. A modern approach*, 2 ed., Upper Saddle River, (N.J.), Prentice Hall, 2003 (1. ed.: Englewood Cliffs (N.J.), Prentice Hall, 1994).
- Somalvico 1987: Somalvico, Marco, *L'intelligenza artificiale*, Milano, Rusconi Editore, 1987.
- Somalvico 1991: Somalvico, Marco, *Intelligenza artificiale*, in *Enciclopedia italiana di scienze lettere ed arti*, Roma, Istituto della Enciclopedia Italiana, 1991, Appendice V, pp. 735-738.
- Turing 1950: Turing, Alan M., "Computing machinery and intelligence", *Mind*, Volume 59, 1950, pp. 433-460.
- Wooldridge 2002: Wooldridge, Michael, *An introduction to multiagent systems*, Chichester, Wiley, 2002.

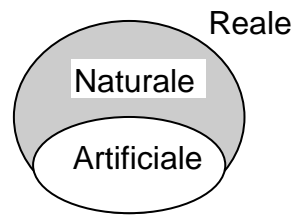
## Tavola I: Agenti intelligenti

Un *agente intelligente* (chiamato anche agente autonomo, agente razionale o, semplicemente, agente) è un'entità in grado di percepire l'ambiente e di agire su di esso (fig. 4). I dispositivi che permettono all'agente di percepire l'ambiente sono i sensori, mentre quelli che permettono all'agente di agire sull'ambiente sono gli attuatori. Un agente intelligente può essere descritto come una sistema che, sulla base delle proprie percezioni passate e presenti, sceglie di eseguire l'azione che si aspetta possa massimizzare la misura di prestazione dell'agente. In questo senso, si può dire che un agente intelligente è tale quando si comporta "in modo razionale", cercando di catturare, con tale dizione, il concetto di "fare la cosa giusta al momento giusto".

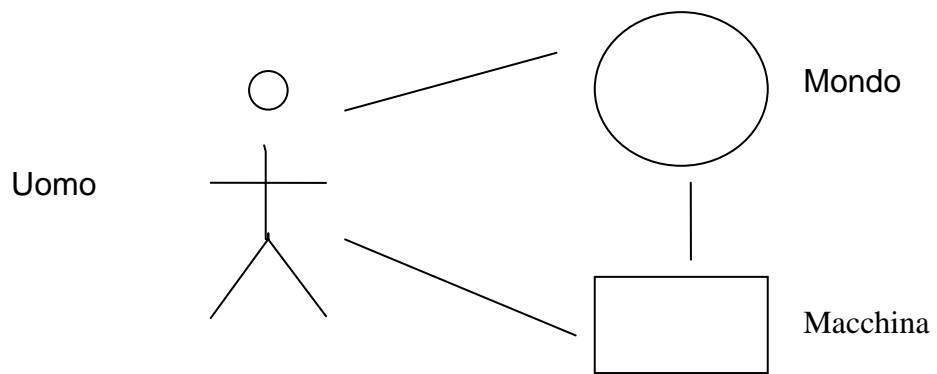
La recente affermazione della nozione di agente intelligente in seno all'IA è conseguenza di almeno due tendenze. Da un lato, tale nozione permette di focalizzare meglio lo scopo dell'IA che, sempre più spesso, è identificato con la realizzazione di sistemi, di agenti intelligenti, che siano in grado di agire razionalmente nell'ambiente nel quale operano (Russell, Norvig 2003). Dall'altro lato, l'agente intelligente rappresenta un riferimento comune intorno al quale riorganizzare le diverse aree di ricerca dell'IA che si sono sviluppate, pur con punti di contatto significativi, sostanzialmente in modo indipendente. Questa tendenza è esemplificata dalla strutturazione didattica dei libri di testo che sono adottati nelle maggiori università (Nilsson, 1998; Russell, Norvig 2003). Operando in un ambiente caratterizzato da incertezza, sia rispetto alle percezioni, che possono non riflettere esattamente lo stato dell'ambiente, sia rispetto alle azioni, che possono non portare agli effetti attesi, un agente deve essere realizzato sulla base di un insieme di tecniche che gli permettano, fra le altre cose, di interpretare i segnali provenienti dai sensori, di effettuare ragionamenti deduttivi in condizioni di incertezza, di esplorare l'ambiente e di apprendere dall'esperienza.

Si ricordi, inoltre, che l'IA si dedica anche allo studio dei sistemi di agenti intelligenti o *sistemi multiagente* (Wooldridge 2002), i quali coniugano, accanto all'elaborazione risolutoria o inferenziale e all'eventuale interazione con il mondo, anche l'interazione fra gli agenti (per esempio la cooperazione), così da riprodurre rozzamente, e in modo astratto, il comportamento delle società umane.

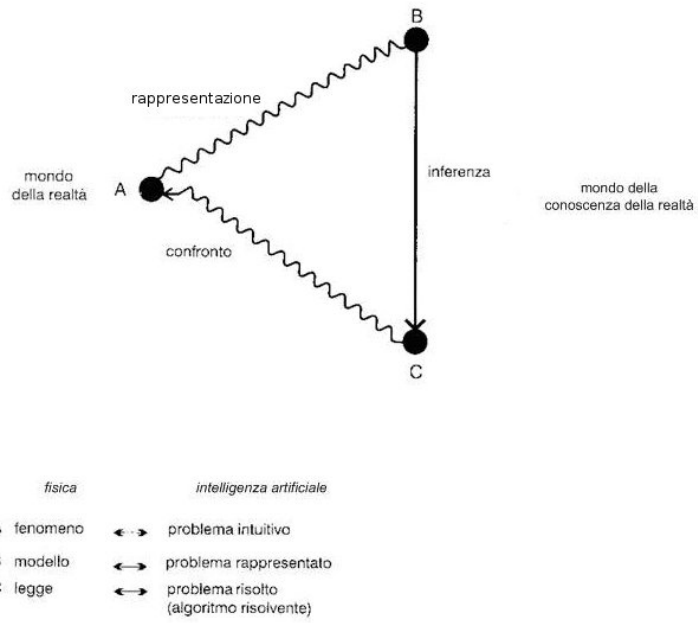
**Figure**



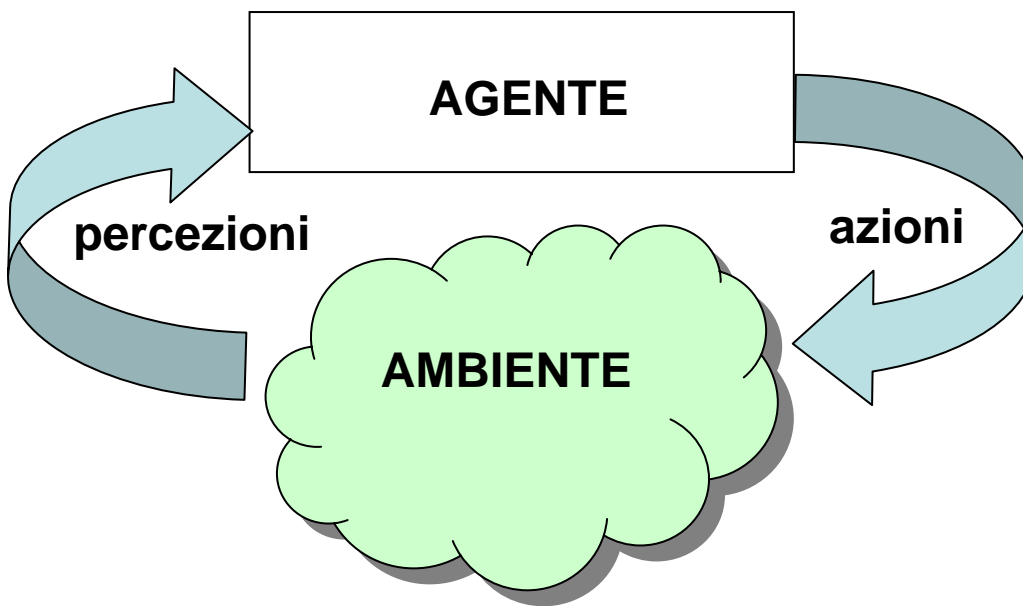
**Figura 1 - Articolazione del reale in naturale e artificiale**



**Figura 2 - Diagramma tripolare del reale**



**Figura 3 - Risoluzione dei problemi**



**Figura 4 - Agente**