

# **PROPOSTA DI UN PROGETTO DI RICERCA ESTIVO SULL'INTELLIGENZA ARTIFICIALE PRESSO IL DARTMOUTH COLLEGE\***

**J. McCarthy, Dartmouth College  
M. L. Minsky, Università di Harvard  
N. Rochester, I.B.M. Corporation  
C.E. Shannon, Bell Telephone Laboratories**

31 agosto 1955

Proponiamo che 10 persone per 2 mesi svolgano una ricerca sull'intelligenza artificiale durante l'estate del 1956 presso il Dartmouth college di Hanover, New Hampshire. Lo studio procederà sulla base della congettura per cui, in linea di principio, ogni aspetto dell'apprendimento o una qualsiasi altra caratteristica dell'intelligenza possano essere descritte così precisamente da poter costruire una macchina che le simuli. Si tenterà di capire come le macchine possano utilizzare il linguaggio, formare astrazioni e concetti, risolvere tipi di problemi riservati per ora solo agli esseri umani e migliorare se stesse. Riteniamo che si possano fare passi avanti significativi in uno o più problemi del genere se un gruppo accuratamente selezionato di scienziati potrà occuparsi di queste tematiche per un'estate.

Quelli che seguono sono alcuni degli aspetti problematici dell'intelligenza artificiale.

## **1. Calcolatori automatici**

Se una macchina può eseguire un compito, allora un calcolatore automatico può essere programmato per simulare quella macchina. La velocità e la capacità di memoria di un calcolatore attuale possono essere insufficienti per simulare molte delle funzioni superiori del cervello umano, ma il problema maggiore non è la mancanza di capacità computazionali quanto la nostra incapacità di scrivere programmi che sfruttino al meglio ciò che abbiamo.

## **2. Come programmare un calcolatore per manipolare un linguaggio**

Si può ipotizzare che una larga parte del pensiero umano consista nell'usare parole in base a regole di ragionamento e a regole di congettura. Da questo punto di vista, la formazione di una generalizzazione consiste nell'accettare una nuova parola e in alcune regole per cui le frasi che contengono la nuova parola implicano o sono implicate da altre frasi. Questa idea non è mai stata formulata con precisione e neanche ne sono mai stati pensati degli esempi.

---

\* Nota del Traduttore: Questo documento si presenta come una bozza di progetto seguita senza soluzione di continuità dalle dichiarazioni dei suoi proponenti. Non è uno scritto indirizzato alla pubblicazione ma ha il carattere di documento interno da distribuire solo tra i partecipanti alla conferenza e pochi altri. Il testo in alcuni punti è lacunoso, lo stile è provvisorio e cambia a seconda dell'autore, il linguaggio tecnico in alcuni casi è primordiale, la formattazione non è curata. Essendo un documento di importanza storica, ho tentato comunque di rimanere il più fedele possibile al testo originale (Gianluca Paronitti).

### **3. Reti di neuroni**

Come si può organizzare una rete di neuroni (ipotetici) per formare dei concetti? Un notevole lavoro teorico e sperimentale è stato condotto su questo problema da Uttley, Rashevsky e il suo gruppo, Farley e Clark, Pitts e McCulloch, Minsky, Rochester e Holland, e altri. Si sono raggiunti risultati parziali e il problema richiede un ulteriore sviluppo teorico.

### **4. Teoria delle dimensioni di un calcolo**

Dato un problema ben definito (cioè un problema per il quale sia possibile controllare in maniera meccanica se una data risposta sia valida o no), un modo per risolverlo è quello di provare ordinatamente tutte le possibili risposte. Questo metodo è inefficiente, ma per scartarlo si deve possedere un qualche criterio per misurare l'efficienza del calcolo. Alcune considerazioni mostreranno che per ottenere questa misura è necessario avere a disposizione un metodo per misurare la complessità dei dispositivi di calcolo; tale metodo, a sua volta, può essere concepito solo disponendo di una teoria della complessità delle funzioni. Alcuni risultati parziali sono stati ottenuti da Shannon e anche da McCarthy.

### **5. Auto-miglioramento**

Probabilmente una macchina davvero intelligente compirà delle attività che possono essere meglio classificate come auto-miglioramento. Per realizzare un tale processo si sono proposti alcuni schemi che, tuttavia, andrebbero studiati ulteriormente. Questa questione può essere affrontata anche in maniera astratta.

### **6. Astrazioni**

Alcuni tipi di "astrazioni" possono essere definiti con chiarezza e altri meno. Sembrerebbe proficuo un tentativo diretto di classificare le astrazioni e di descrivere metodi meccanici che portino alla formazione delle astrazioni a partire da dati sensoriali o di altro tipo.

### **7. Casualità e Creatività**

Una congettura abbastanza attraente e tuttavia ancora chiaramente incompleta è quella per cui la differenza tra pensiero creativo e pensiero competente ma non creativo consista nell'introduzione di un qualche grado di casualità. In altre parole, l'ipotesi plausibile o l'intuizione immettono casualità controllata in un pensiero altrimenti ordinato.

Oltre ai problemi appena indicati, che abbiamo formulato collettivamente, si è anche richiesto che ognuno di coloro che parteciperà al seminario descriva quello di cui si occuperà. Alleghiamo a questa proposta di progetto le dichiarazioni dei quattro proponenti.

Suggeriamo di organizzare il lavoro di gruppo come segue.

Copia di questa proposta sarà inviata ai potenziali partecipanti a cui sarà richiesto se vorranno lavorare in gruppo sulle problematiche dell'intelligenza artificiale e su quali preferiranno concentrarsi. Gli inviti saranno fatti dal comitato organizzatore sulla base della valutazione del possibile contributo individuale al lavoro di gruppo. Prima di questo periodo, i partecipanti faranno circolare il loro lavoro precedente e le loro idee riguardanti i problemi che saranno discussi.

L'incontro sarà scandito da seminari di ricerca e per i partecipanti ci sarà l'opportunità di lavorare sia individualmente sia in piccoli gruppi informali.

I proponenti di questo progetto sono:

1. C. E. Shannon, matematico, Bell Telephone Laboratories. Shannon ha sviluppato la teoria statistica dell'informazione, l'applicazione del calcolo proposizionale ai circuiti di commutazione, ottenendo risultati notevoli sulla sintesi efficiente di questo tipo di circuiti, sulla progettazione di macchine che apprendono, sulla crittografia e sulla teoria delle macchine di Turing. Insieme a J. McCarthy è curatore del volume "Automata Studies" per gli *Annals of Mathematical Studies*.

2. M. L. Minsky, Harvard Junior Fellow in Matematica e Neurologia. Minsky ha costruito una macchina per simulare l'apprendimento attraverso reti nervose e ha scritto una tesi di dottorato in matematica a Princeton dal titolo "Neural Nets and the Brain Model Problem" che include alcuni risultati nella teoria dell'apprendimento e sulla teoria delle reti neurali casuali.

3. N. Rochester, Manager dell'Information Research, IBM Corporation, Poughkeepsie, New York. Rochester si è occupato dello sviluppo dei radar e di macchine calcolatrici per sette anni. Ha partecipato alla progettazione dell'IBM 701 che è un calcolatore automatico di grandi dimensioni attualmente molto diffuso. Rochester ha elaborato alcune delle tecniche di programmazione automatica oggi più diffuse e si è occupato del problema di come far eseguire a una macchina dei compiti che precedentemente potevano essere eseguiti solo da persone. Si è anche occupato di simulazione di reti nervose con particolare attenzione all'utilizzo di calcolatori per controllare teorie neuro-psicologiche.

4. J. McCarthy, Professore associato di Matematica, Dartmouth College. McCarthy ha lavorato su diverse questioni connesse alla natura matematica dei processi di pensiero inclusa la teoria delle macchine di Turing, la velocità dei calcolatori, la relazione di un modello cerebrale al suo ambiente e la manipolazione dei linguaggi da parte delle macchine. Alcuni risultati di questo lavoro sono inclusi nel prossimo volume degli "Annals of Mathematical Studies" curato da Shannon e McCarthy. Altri lavori di McCarthy riguardano il campo delle equazioni differenziali.

È stato richiesto supporto finanziario per il progetto alla Rockefeller Foundation sulle seguenti basi.

1. 1200\$ di stipendio per ogni partecipante di grado universitario che non verrà finanziato dalla sua organizzazione. Ci si aspetta, per esempio, che i partecipanti dei Bell Laboratories e dell'IBM Corporation saranno finanziati da queste organizzazioni

mentre quelli di Dartmouth e di Harvard avranno bisogno del supporto della fondazione.

2. 700\$ di stipendio per non più di due studenti universitari.
3. Rimborso ferroviario per i partecipanti che arrivano da lontano.
4. Rimborso per le spese di affitto a coloro che contemporaneamente pagano l'affitto altrove.
5. 650\$ di spese amministrative, 500\$ per un/a segretario/a e 150\$ per spese di riproduzione.
6. 200\$ di spese organizzative. (Incluse le spese per riprodurre il lavoro preliminare dei partecipanti e i viaggi necessari per motivi organizzativi).
7. Spese per due o tre persone che verranno in visita per brevi periodi.

#### Spese Stimate

6 stipendi da 1200	7200
2 stipendi da 700	1400
8 rimborsi di viaggio e affitto a una media di 300	2400
Spese amministrative e organizzative	850
Spese di viaggio addizionali	600
Imprevisti	550
	----
	13,500\$

Vorrei dedicare la mia ricerca a uno o a entrambi gli argomenti indicati qui sotto. Pur sperando di farcela, è possibile che per motivi personali non riesca a essere presente per tutti e due i mesi. Tuttavia intendo essere presente per tutto il tempo possibile.

1. L'applicazione dei concetti della teoria dell'informazione ai calcolatori e ai modelli del cervello. Uno dei problemi basilari nella teoria dell'informazione è quello della trasmissione affidabile di informazione attraverso un canale disturbato. Un problema analogo nelle macchine calcolatrici è quello dell'affidabilità della computazione utilizzando elementi inaffidabili. Questo problema è stato studiato da von Neumann riguardo agli operatori di Sheffer e da Shannon e Moore per i relais; ma restano ancora da risolvere alcuni problemi. Tra gli argomenti rilevanti ci sono il problema di trattare un elevato numero di elementi, lo sviluppo di concetti come quello della capacità del canale, l'analisi più raffinata dei limiti superiori e inferiori della ridondanza necessaria ecc.. Un'altra questione riguarda la teoria delle reti dove l'informazione viaggia in vari anelli chiusi (in contrasto con il semplice canale unidirezionale che di solito è preso in considerazione dalla teoria della comunicazione). I problemi del ritardo (del segnale) divengono molto importanti nel caso degli anelli chiusi, e un approccio completamente nuovo sembra necessario. Questo potrebbe probabilmente coinvolgere concetti come quello di entropie parziali quando è nota una parte della storia passata di un insieme di messaggi.

2. L'approccio agli automi basato sul modello "cervello-ambiente". In generale una macchina o un animale possono solamente adattarsi o operare in una classe limitata di ambienti. Anche il cervello umano, per quanto complesso, inizialmente si adatta agli aspetti più semplici del suo ambiente e gradualmente ne ricostruisce le caratteristiche più complesse. Propongo di studiare una sintesi dei modelli del cervello sviluppando parallelamente una sequenza di ambienti e di modelli cerebrali che si adattino a questi. L'enfasi è sulla specifica del modello di ambiente, e sul modo di rappresentarlo come una struttura matematica. Spesso nella discussione riguardante l'intelligenza meccanica, noi immaginiamo delle macchine che realizzino le attività più avanzate dell'intelletto umano: dimostrare teoremi, scrivere musica o giocare a scacchi. Quello che propongo è invece di partire da attività più semplici, vale a dire da un ambiente che non sia né ostile (ma solamente indifferente) né complesso e di svilupparlo attraverso una serie di passaggi elementari nella direzione di quelle attività più avanzate.

Non è difficile progettare una macchina che esibisca il seguente tipo di apprendimento. La macchina è fornita di canali di ingresso e di uscita e di un sistema interno per fornire differenti risposte di uscita in base agli ingressi in modo che, attraverso un processo per "tentativi ed errori", possa essere "addestrata" ad acquisire una delle molteplici funzioni di ingresso-uscita. Una macchina del genere, qualora fosse posta in un ambiente appropriato e le fosse fornito un criterio per la valutazione del "successo" o del "fallimento", potrebbe essere addestrata a esibire un comportamento "orientato all'obiettivo". A meno che la macchina non sia fornita o comunque non sia capace di sviluppare un modo per astrarre dal materiale sensoriale, questa sarà in grado di progredire in un ambiente complicato solo attraverso dei passaggi penosamente lenti, e in generale non raggiungerà un comportamento di alto livello.

Ora ammettiamo che il criterio per il successo non sia solamente la comparsa di un certo schema di attività desiderata nel canale di uscita della macchina, ma piuttosto la prestazione di una data manipolazione in un dato ambiente. Allora in un certo senso lo stato motorio sembra essere un doppiato di quello sensoriale, e i progressi possono essere ragionevolmente veloci solo se la macchina è ugualmente in grado di assemblare un insieme di "astrazioni motorie" collegando la sua attività di uscita ai cambiamenti nell'ambiente. Tali "astrazioni motorie" possono essere significative solo se sono collegate con cambiamenti nell'ambiente che possano essere rilevati dalla macchina come cambiamenti nello stato sensoriale, cioè se sono collegati, attraverso la struttura dell'ambiente, alle astrazioni sensoriali che la macchina sta utilizzando.

Ho studiato sistemi di questo genere per lungo tempo e mi pare che se si potesse progettare una macchina nella quale le astrazioni motorie e quelle sensoriali, una volta formate, potessero essere disposte in modo da soddisfare certe relazioni, potrebbe venirne fuori un comportamento di alto livello. Queste relazioni consistono nell'accoppiamento delle astrazioni motorie con quelle sensoriali per produrre un nuovo stato sensoriale. Questo nuovo stato dovrebbe rappresentare i cambiamenti nell'ambiente che ci si aspetta qualora l'atto motorio corrispondente si verificasse.

Il risultato desiderabile sarebbe che la macchina cominciasse a costruirsi internamente un modello astratto dell'ambiente nel quale è posta. Infatti, se gli fosse dato un

problema da risolvere, potrebbe esplorare alcune soluzioni nel modello astratto interno dell'ambiente e poi tentare di sperimentarle all'esterno. A causa di questo studio preliminare interno, questi tentativi esterni apparirebbero come piuttosto intelligenti, e il comportamento potrebbe essere considerato "immaginativo".

Una proposta molto provvisoria di come questo potrebbe essere realizzato è descritta nella mia tesi e intendo continuare a lavorare in questa direzione. Per l'estate del 1956 spero di avere un modello di questo tipo di macchina che sia prossimo allo stadio della programmazione al calcolatore.

### **Originalità nella Prestazione della Macchina**

Normalmente, programmando un calcolatore automatico, si tenta di fornire alla macchina un insieme di regole che gli permettano di affrontare ogni imprevisto in cui si possa imbattersi. Ci si aspetta che la macchina segua pedissequamente questo insieme di regole e che non esibisca originalità o senso comune. Per questo, bisognerebbe irritarsi solo con se stessi quando la macchina si comporta in maniera confusa, perché le regole che gli sono state fornite conducono a contraddizioni anche lievi. A volte, programmando, si arriva alla soluzione dei problemi seguendo percorsi molto laboriosi laddove, se la macchina avesse solo un poco di intuizione o potesse azzardare delle ipotesi ragionevoli potrebbe essere molto più diretta. In questo articolo si propone un'ipotesi sulla costruzione di una macchina che si comporti in maniera più sofisticata riguardo al problema indicato qui sopra. L'articolo discute una questione sulla quale ho lavorato saltuariamente per circa cinque anni e che vorrei indagare ulteriormente durante il prossimo Progetto di Ricerca Estivo sull'Intelligenza Artificiale.

### **Il Processo di Invenzione o Scoperta**

Il nostro ambiente culturale ci fornisce delle procedure per risolvere molti problemi e, tuttavia, proprio il funzionamento di queste procedure non è chiaro. Discuterò questo aspetto della questione nei termini di un modello proposto da Craik<sup>1</sup>. Egli suggerisce che l'azione mentale consista fondamentalmente nella costruzione di piccoli dispositivi all'interno del cervello i quali possano simulare e quindi anticipare le astrazioni, collegandole con l'ambiente. In questo senso, la soluzione di un problema consiste nel processo seguente.

1. L'ambiente fornisce i dati con i quali si formano certe astrazioni.
2. Le astrazioni insieme a certe abitudini interne o impulsi forniscono:
3.
  1. una definizione del problema nei termini di una condizione che si desidera raggiungere nel futuro, cioè un obiettivo.
  2. Una azione che viene suggerita per risolvere il problema.
  3. La stimolazione che avvii nel cervello il dispositivo che corrisponde a quella situazione.
4. A questo punto il dispositivo si mette in funzione per prevedere lo stato futuro dell'ambiente e la reazione proposta.

---

<sup>1</sup> K.J.W. Craik, *The Nature of Explanation*, Cambridge University Press, 1943 (ristampato nel 1952), p. 92.

5. Se la previsione corrisponde all'obiettivo, l'individuo procede all'azione come indicato.

La previsione corrisponderà all'obiettivo se l'ambiente culturale ha fornito all'individuo la soluzione del problema. Per quanto riguarda l'individuo considerato come un calcolatore a programma memorizzato, è il programma che deve contenere le regole per affrontare una determinata evenienza.

Una situazione più complessa esige regole più complicate. Tali regole potrebbero richiedere di sperimentare ogni possibile insieme di azioni per determinare quale di queste possa fornire la soluzione. Un insieme ancora più complesso di regole potrebbe occuparsi dell'ambiguità dell'ambiente, per esempio giocando a tris si deve considerare non solo la propria mossa ma anche le varie possibili mosse dell'ambiente (cioè dell'avversario).

Ora consideriamo un problema per il quale non si dispone di una soluzione nel proprio ambiente culturale e che ha resistito ai vari tentativi di risolverlo. Potrebbe essere un tipico problema scientifico attualmente insoluto. L'individuo potrebbe tentare di venirne a capo e scoprire che ogni azione ragionevole porta al fallimento. In altre parole il programma memorizzato contiene delle regole per la soluzione di questo problema ma queste regole sono quelle sbagliate.

A questo punto, l'individuo dovrà ricorrere a qualcosa di insolito o inaspettato rispetto al patrimonio di conoscenze comuni accumulato dalla cultura. Egli potrebbe ottenere questo risultato facendo vari tentativi a caso, ma, generalmente, un tale approccio risulterebbe oltremodo inefficiente. Solitamente ci sono troppe alternative possibili delle quali solo una piccola porzione è accettabile. Gli individui hanno bisogno di una intuizione, di un qualcosa di inaspettato e non del tutto ragionevole. Alcuni problemi, spesso quelli che sono piuttosto nuovi per i quali non si sono ancora fatti grossi sforzi alla ricerca di una soluzione, hanno bisogno solo di poca casualità. Altri, di solito quelli che hanno lungamente resistito al tentativo di risolverli, hanno bisogno di una deviazione davvero bizzarra dai metodi tradizionali. Un problema la cui soluzione richieda originalità, potrebbe portare a un metodo che coinvolga la casualità.

Nei termini del modello di Craik, il dispositivo che dovrebbe simulare l'ambiente all'inizio fallisce nel simularlo correttamente. Perciò, è necessario tentare varie modifiche del dispositivo finché non se ne trovi una che gli faccia fare ciò che è necessario.

Invece di descrivere il problema nei termini di un individuo e della sua cultura lo si potrebbe descrivere nei termini dell'apprendimento di un individuo immaturo. Così quando un individuo si imbatte in un problema al di là della sua esperienza, deve superarlo in una maniera simile.

Finora l'approccio pratico più simile all'utilizzo di questo metodo nella soluzione automatica dei problemi è una estensione del metodo Monte Carlo. Il tipico problema per cui il metodo di Monte Carlo è appropriato è una situazione che è largamente incompresa, che possiede troppi possibili fattori e in cui non si è in grado di decidere quali fattori ignorare nel calcolare una soluzione analitica. Così il matematico lascia fare alla macchina qualche migliaio di esperimenti casuali. Il risultato di questi

esperimenti fornisce una stima approssimata su come la risposta possa essere. L'estensione del metodo Monte Carlo userà questi risultati come una guida per determinare cosa non considerare al fine di semplificare il problema abbastanza da ottenere una soluzione analitica approssimata.

Ci si potrebbe chiedere perché il metodo dovrebbe includere la casualità. Per quale motivo il metodo non dovrebbe essere quello di tentare ogni possibilità in base alla probabilità di successo data dallo stato attuale della conoscenza? Potrebbe darsi che il singolo scienziato, immerso nel proprio ambiente culturale, non riesca a risolvere il problema nel corso della sua vita e quindi sia richiesto lo sforzo di molti. Utilizzando la casualità, gli scienziati potrebbero lavorare tutti insieme senza duplicare inutilmente i propri sforzi. Altrimenti essi avrebbero bisogno di una comunicazione dettagliata fino all'inverosimile. Per l'individuo che cresce nella competizione con altri individui, la necessità di una strategia mista (per utilizzare la terminologia della teoria dei giochi) sostiene la casualità. Per la macchina, probabilmente, la casualità sarà necessaria per superare i pregiudizi e la miopia del programmatore. Sebbene la necessità di introdurre casualità non sia stata dimostrata, molte sono le evidenze in suo favore.

### **La Macchina fornita di Casualità**

Per scrivere un programma che permetta a un calcolatore automatico di operare in maniera originale, si dovrà introdurre casualità con accortezza. Se, per esempio, si scrivesse un programma nel quale una volta ogni 10000 passi il calcolatore generasse un numero casuale e lo utilizzasse come un'istruzione il risultato sarebbe probabilmente il caos. A questo punto, dopo un certo periodo di caos, la macchina tenterebbe qualcosa di proibito o eseguirebbe un'istruzione di arresto e l'esperimento sarebbe concluso.

Due approcci, tuttavia, sembrano ragionevoli. Uno di questi è capire come il cervello affronta e risolve questo tipo di situazione. L'altro è quello di considerare varie classi di problemi reali che richiedono soluzioni originali e tentare di trovare un modo di scrivere un programma che li risolva su un calcolatore automatico. Probabilmente entrambi gli approcci avranno successo prima o poi. Tuttavia non è chiaro né quale sarà il primo, né quanti anni o generazioni richiederà. Finora mi sono occupato soprattutto del primo perché ho ritenuto che sarebbe stato meglio padroneggiare tutta la conoscenza scientifica più rilevante riguardo a un argomento così problematico e, d'altra parte conosco abbastanza bene lo stato attuale dei calcolatori e l'arte di programmarli.

Il meccanismo di controllo del cervello è chiaramente differente dal meccanismo di controllo dei calcolatori attuali. Un sintomo di questa differenza è nella maniera in cui si presenta un guasto. Tipicamente, un guasto in un calcolatore si manifesta come qualcosa di abbastanza irragionevole. Un errore nella memoria o nella trasmissione dei dati si troverà tanto nell'istruzione più rilevante quanto in quella meno. Un errore nel controllo può fare praticamente di tutto. Potrebbe eseguire l'istruzione sbagliata o far funzionare l'unità di ingresso-uscita errata. Invece gli errori umani nel parlare possono risultare in asserzioni che hanno più o meno senso (si consideri qualcuno quasi addormentato, un po' alticcio o febbricitante). Forse il funzionamento del cervello è tale che un piccolo errore nel ragionamento introduce la casualità proprio

nel modo giusto. Forse il meccanismo che controlla l'ordine sequenziale del comportamento<sup>2</sup> guida il fattore casuale in modo da aumentare l'efficienza dei processi immaginativi al di sopra della pura causalità.

Del lavoro è stato fatto nella simulazione di reti neurali su un calcolatore automatico. Uno degli scopi era proprio quello di vedere se sarebbe stato possibile introdurre la casualità in modo appropriato. Sembra che si possa concludere che ci sono troppe connessioni sconosciute tra l'attività dei neuroni e il problem-solving perché questo approccio abbia successo nell'immediato futuro. I risultati hanno gettato un po' di luce sul comportamento delle reti e dei neuroni, ma non hanno prodotto nessun nuovo modo per risolvere problemi che richiedono originalità.

Un importante aspetto di questo lavoro è stato quello di far sì che la macchina si formi e manipoli concetti, astrazioni, generalizzazioni, e nomi. Si è tentato di controllare una teoria<sup>3</sup> di come il cervello faccia questo. Il primo insieme di esperimenti ha portato a una revisione di alcuni dettagli della teoria. Il secondo insieme di esperimenti è ancora in atto. Per la prossima estate questo lavoro sarà finito e ne sarà proposta una sintesi finale.

Il mio progetto è quello di scrivere successivamente un programma per risolvere problemi che siano elementi di una qualche classe limitata di problemi che richiedano originalità nella loro risoluzione. È troppo presto per prevedere esattamente a quale stato del lavoro arriverò per la prossima estate o semplicemente come definirò allora il problema immediato. Tuttavia, la questione sottostante, che è descritta in questo articolo, è quella che intendo studiare. Per dirla in breve: come posso costruire una macchina che risolva un problema in maniera originale?

Durante il prossimo anno e durante il Progetto di Ricerca Estivo in Intelligenza Artificiale, mi propongo di studiare la relazione tra linguaggio e intelligenza. Sembra chiaro che la diretta applicazione di metodi basati su "tentativi ed errori" alla relazione tra dati sensoriali e attività motoria non condurrà a nessun comportamento molto complicato. Piuttosto è necessario applicare i metodi per "tentativi ed errori" a un più alto livello di astrazione. La mente umana, apparentemente, utilizza il linguaggio come un mezzo per affrontare fenomeni complicati. Frequentemente i processi per tentativi ed errori, a un più alto livello, si configurano come la formulazione e controllo di congetture. La lingua inglese ha un numero di caratteristiche che manca a un qualsiasi linguaggio formale finora descritto.

1. Gli argomenti in inglese supportati da matematica informale possono essere concisi.
2. L'inglese è universale nel senso che può sviluppare qualsiasi altro linguaggio al suo interno e poi utilizzare quel linguaggio dove sia appropriato.
3. Il parlante inglese può riferirsi a se stesso e formulare asserzioni che riguardano il suo progresso nel risolvere il problema di cui si sta occupando.
4. Oltre alle regole di dimostrazione, l'inglese se completamente formulato avrebbe delle regole di congettura.

---

<sup>2</sup> K.S. Lashley, "The Problem of Serial Order in Behavior", in *Cerebral Mechanism in Behavior, The Hixon Symposium*, curato da L.A. Jeffress, John Wiley & Sons, New York, pp. 112-146, 1951.

<sup>3</sup> D. O. Hebb, *The Organization of Behavior*, John Wiley & Sons, New York, 1949.

I linguaggi logici finora formulati sono stati liste di istruzioni per far sì che i calcolatori eseguano calcoli specificati in precedenza, o altrimenti formalizzazioni di parti della matematica. Il secondo caso è stato costruito per:

1. essere facilmente descritto nella matematica informale,
2. consentire la traduzione di asserzioni dalla matematica informale nel linguaggio,
3. semplificare la possibilità di argomentare se certe dimostrazioni di [lacuna nel testo].

Non è stato fatto nessun tentativo per rendere le dimostrazioni nei linguaggi artificiali tanto brevi quanto quelle informali. Perciò sarebbe desiderabile tentare di costruire un linguaggio artificiale – che un calcolatore potrebbe essere programmato a usare – su problemi che richiedono congettura e auto-referenzialità. Dovrebbe corrispondere all'inglese nel senso che le brevi asserzioni inglesi circa un dato argomento dovrebbero avere dei corrispondenti altrettanto brevi nel linguaggio e allo stesso modo ci dovrebbe essere una corrispondenza sia per le argomentazioni sia per gli argomenti congetturali. Spero di riuscire a formulare un linguaggio con queste proprietà e che anche contenga le nozioni di oggetto fisico, evento ecc., con la speranza che utilizzando questo linguaggio sia possibile programmare una macchina che impari a giocare bene e a eseguire molti altri compiti.

Lo scopo della lista è far sì che coloro che ne fanno parte sappiano chi è interessato a ricevere materiale su questi problemi. Le persone sulla lista riceveranno copie del resoconto del Programma Estivo sull'Intelligenza Artificiale del Dartmouth College. [nota del 1996: non c'è stato nessun resoconto.]

La lista consiste delle persone che hanno partecipato o hanno visitato il Programma Estivo sull'Intelligenza Artificiale del Dartmouth College, o che si sa essere interessate all'argomento. La lista sarà inviata a tutti gli iscritti e a pochi altri.

Per gli scopi attuali si ritiene che il problema dell'intelligenza artificiale sia quello di costruire una macchina che si comporti in modi che sarebbero considerati intelligenti nel caso di un essere umano.

Una lista aggiornata sarà compilata al più presto, in modo che chiunque altro sia interessato ad aggiungersi o che voglia modificare il suo indirizzo possa scrivere a:

[nota del 1996: non tutte le persone della lista sono andate alla conferenza di Dartmouth. Si tratta di persone che ritenevamo potessero essere interessate all'intelligenza artificiale].

La lista comprende:

Adelson, Marvin  
Hughes Aircraft Company  
Airport Station, Los Angeles, CA

Ashby, W. R.  
Barnwood House

Gloucester, England

Backus, John  
IBM Corporation  
590 Madison Avenue  
New York, NY

Bernstein, Alex  
IBM Corporation  
590 Madison Avenue  
New York, NY

Bigelow, J. H.  
Institute for Advanced Studies  
Princeton, NJ

Elias, Peter  
R. L. E., MIT  
Cambridge, MA

Duda, W. L.  
IBM Research Laboratory  
Poughkeepsie, NY

Davies, Paul M.  
1317 C. 18th Street  
Los Angeles, CA.

Fano, R. M.  
R. L. E., MIT  
Cambridge, MA

Farley, B. G.  
324 Park Avenue  
Arlington, MA.

Galanter, E. H.  
University of Pennsylvania  
Philadelphia, PA

Gelernter, Herbert  
IBM Research  
Poughkeepsie, NY

Glashow, Harvey A.  
1102 Olivia Street  
Ann Arbor, MI.

Goertzel, Herbert  
330 West 11th Street

New York, New York

Hagelbarger, D.  
Bell Telephone Laboratories  
Murray Hill, NJ

Miller, George A.  
Memorial Hall  
Harvard University  
Cambridge, MA.

Harmon, Leon D.  
Bell Telephone Laboratories  
Murray Hill, NJ

Holland, John H.  
E. R. I.  
University of Michigan  
Ann Arbor, MI

Holt, Anatol  
7358 Rural Lane  
Philadelphia, PA

Kautz, William H.  
Stanford Research Institute  
Menlo Park, CA

Luce, R. D.  
427 West 117th Street  
New York, NY

MacKay, Donald  
Department of Physics  
University of London  
London, WC2, England

McCarthy, John  
Dartmouth College  
Hanover, NH

McCulloch, Warren S.  
R.L.E., M.I.T.  
Cambridge, MA

Melzak, Z. A.  
Mathematics Department  
University of Michigan  
Ann Arbor, MI

Minsky, M. L.  
112 Newbury Street  
Boston, MA

More, Trenchard  
Department of Electrical Engineering  
MIT  
Cambridge, MA

Nash, John  
Institute for Advanced Studies  
Princeton, NJ

Newell, Allen  
Department of Industrial Administration  
Carnegie Institute of Technology  
Pittsburgh, PA

Robinson, Abraham  
Department of Mathematics  
University of Toronto  
Toronto, Ontario, Canada

Rochester, Nathaniel  
Engineering Research Laboratory  
IBM Corporation  
Poughkeepsie, NY

Rogers, Hartley, Jr.  
Department of Mathematics  
MIT  
Cambridge, MA.

Rosenblith, Walter  
R.L.E., M.I.T.  
Cambridge, MA.

Rothstein, Jerome  
21 East Bergen Place  
Red Bank, NJ

Sayre, David  
IBM Corporation  
590 Madison Avenue  
New York, NY

Schorr-Kon, J.J.  
C-380 Lincoln Laboratory, MIT  
Lexington, MA

Shapley, L.  
Rand Corporation  
1700 Main Street  
Santa Monica, CA

Schutzenberger, M.P.  
R.L.E., M.I.T.  
Cambridge, MA

Selfridge, O. G.  
Lincoln Laboratory, M.I.T.  
Lexington, MA

Shannon, C. E.  
R.L.E., M.I.T.  
Cambridge, MA

Shapiro, Norman  
Rand Corporation  
1700 Main Street  
Santa Monica, CA

Simon, Herbert A.  
Department of Industrial Administration  
Carnegie Institute of Technology  
Pittsburgh, PA

Solomonoff, Raymond J.  
Technical Research Group  
17 Union Square West  
New York, NY

Steele, J. E., Capt. USAF  
Area B., Box 8698  
Wright-Patterson AFB  
Ohio

Webster, Frederick  
62 Coolidge Avenue  
Cambridge, MA

Moore, E. F.  
Bell Telephone Laboratory  
Murray Hill, NJ

Kemeny, John G.  
Dartmouth College  
Hanover, NH