

La natura della cultura.

**Modelli simulativi dell'apprendimento
sociale.**

Alberto Acerbi

Prefazione	p. 1
Introduzione	p. 5
1. Le simulazioni come strumento di conoscenza	p. 13
Ontologia naturalistica	p. 15
Dati empirici, teorie quantitative, laboratori	p. 18
Cos'è una simulazione	p. 21
Sistemi complessi e interdisciplinarietà	p. 26
Cosa ci dicono le simulazioni sulla realtà?	p. 29
2. Quali simulazioni?	p. 35
Alle origini: macchine di Turing e meccanismi di retroazione	p. 36
Realismo biologico: reti neurali (1)	p. 40
Ontogenesi: reti neurali (2)	p. 43
Filogenesi: algoritmi genetici	p. 45
<i>Embodiment</i> : istanze teoriche	p. 48
<i>Embodiment</i> : reti neurali, organismi simulati, robots	p. 51
3. Trasmissione culturale intergenerazionale e intragenerazionale	p. 55
Un modello di trasmissione culturale	p. 57
Risultati	p. 62
Un possibile ruolo per la trasmissione intragenerazionale	p. 63
Trasmissione intragenerazionale in ambiente dinamico	p. 68
La probabilità che un individuo apprenda dai coetanei viene ereditata	p. 71
Conclusioni	p. 73
Evoluzione biologica ed evoluzione culturale	p. 75
Appendice 3.1: La regola Delta	p. 81
4. Apprendimento sociale e apprendimento individuale	p. 83
Set – up sperimentale	p. 85
L'algoritmo di apprendimento	p. 87
Risultati (apprendimento individuale)	p. 88

Apprendimento sociale	p. 90
Risultati (apprendimento sociale)	p. 92
Trasmissione culturale iterata	p. 95
Cos'è la "trasmissione culturale"?	p. 100
Jukebox, Camino del Diablo e attrattori: metafore per la cultura	p. 102
Trasmettere e copiare	p. 105
Appendice 4.1: L'algoritmo di <i>simulated annealing</i>	p. 111
5. Meccanismi indiretti di trasmissione sociale	p. 115
Set – up robotico	p. 119
La simulazione e le condizioni sperimentali	p. 121
Risultati	p. 124
Il modello analitico	p. 129
Simulazioni e modelli matematici	p. 132
Bibliografia	p. 135

Portare a termine un dottorato di ricerca, senza borsa di studio, in antropologia, in Italia, richiede, oltre a perseveranza e ad una qualche attitudine (spero), una buona dose di fortuna. Nel mio caso, la fortuna è stata quella di incontrare Domenico Parisi, iniziando a collaborare con il LARAL (Laboratory of Autonomous Robotics and Artificial Life) all'Istituto di Scienze e Tecnologie della Cognizione (ISTC) del CNR di Roma. Senza l'aiuto di Domenico Parisi e di Stefano Nolfi, che tramite una borsa di studio di un progetto ISFOL prima e un assegno di ricerca all'interno del progetto europeo ECAGents poi, mi hanno materialmente permesso di fare ricerca, semplicemente questa tesi non sarebbe esistita. Da ultimo, un finanziamento del network europeo euCognition ha facilitato la mia attuale visita al Max Planck Institute for Evolutionary Anthropology di Leipzig.

Tuttavia, oltre al necessario sostegno economico, la vera fortuna è stata quella di trovare nell'ISTC un ambiente perfetto dove studiare, imparare, lavorare. Pur con tutte le difficoltà di un centro di ricerca italiano, poco finanziato (*tutti* i fondi per i “giovani” ricercatori provengono da progetti esterni, supportati in genere dalla Comunità Europea) e con qualche problema logistico, l'ISTC è stato ed è una quotidiana fonte di stimoli intellettuali, di discussione, di confronto, che mi ha permesso, come non credo altre esperienze avrebbero potuto, non solo di crescere intellettualmente ma anche di sperimentare la pratica quotidiana del lavoro di ricerca, fatta anche di scadenze di conferenze, organizzazione di workshop, review da finire di notte e caffè che diventano ore a discutere animatamente con qualche collega.

In particolare, Domenico Parisi e Stefano Nolfi hanno provato a trasmettermi la loro passione e la loro curiosità e sono stati coloro che, principalmente, mi hanno insegnato tutto ciò che conosco sulla robotica e sulla vita artificiale. La mia vita, forse non solo dal punto di vista scientifico, sarebbe stata completamente diversa senza averli incontrati.

Il mio gruppo, il LARAL, oltre ad essere un gruppo internazionalmente riconosciuto e ad avere un programma ed una metodologia (uno *stile*, verrebbe da dire) che non hanno nulla da invidiare a più noti e più finanziati centri di ricerca, è, anche e soprattutto, un gruppo di amici che non avrebbero potuto non influire su questo lavoro. Dagli “esperti” Orazio Miglino, Gianluca Baldassarre e Federico Cecconi (che per primo mi ha introdotto alla programmazione in C++), ai miei compagni d’ufficio Diana Giorgini, Vito Trianni e Onofrio Gigliotta (che, oltre a condividere con me la scrivania, è la voce dei FF) a Marco Mirolli, altro profugo senese, a Davide Marocco (il mandolinista dei FF, che ha la pazienza di discutere con me per ore ed ore di embodiment, apprendimento, epistemologia e quant’altro ci passa per la mente) ad Andrea di Ferdinando, Massimiliano Caretti, Giovanna Mazzapioda, Gianluca Massera (un “vero” informatico, a cui continuo sempre a fare domande che probabilmente giudica estremamente stupide), a Simone Giansante, a Francesco Mannella, Stefano Zappacosta, Dimitri Ognibene, Massimiliano Schembri (ideatore dello “zanzarone”, logo del Terzo Workshop Italiano di Vita Artificiale), la lista delle persone da ringraziare è davvero lunga.

All’ISTC ho approfittato anche del sostegno e dell’amicizia di molte altre persone: Paolo Turrini (con cui ho condiviso una terribile casa a Rebibbia), Gennaro di Tosto, Giulia Andrighetto, Francesca Giardini, Francesca Marzo, Luca Tummolini, Fabio Paglieri, Giovanni Pezzulo e, non ultimo, il direttore Cristiano Castelfranchi.

Dopo la laurea, quando non sapevo se le ricerche di antropologia cognitiva che intendevo portare avanti sarebbero interessate a qualcuno, Massimo Marraffa mi ha offerto di scriverci sopra un articolo per Sistemi Intelligenti e un capitolo per un libro (ora entrambi pubblicati) e mi ha incoraggiato a proseguire. Dan Sperber è passato dall’essere un’entità mitizzata ad una persona reale a cui scrivo mail e con cui parlo dei progetti futuri. A Siena, Pier Giorgio Solinas ha seguito a distanza il mio lavoro per questo dottorato e Massimo Squillacciotti ha addirittura inserito due miei scritti nel programma d’esame di antropologia cognitiva (ogni tanto un po’ di superflua autostima...). Tommaso Sbriccoli e Carlo Mathieu sono sempre stati disponibili ad

ascoltare e commentare le mie idee oltre che a condividere piacevoli cene; Leonardo Gatti, forse più critico, mi ha aiutato indirettamente, impersonificando la difficile doppia figura di amico e barista.

Qui a Leipzig sto approfittando della meravigliosa ospitalità che mi viene concessa al Max Planck Institute for Evolutionary Anthropology, un luogo che, agli occhi di un dottorando italiano, appare quasi irreali per l'efficienza, la ricchezza e la quantità di stimoli che può fornire. Oltre a Charlie Nunn, il mio supervisor, molte altre persone dell'istituto mi sono state (e mi sono) d'aiuto nella – non difficile, a dire il vero – impresa di ambientarmi in Germania: in ordine un po' sparso sono Mathias Franz, Brian Preston, Damien Caillaud, Dieter Lukas, Daniela Hedwig, Claudia Barelli, Thomas Breuer, Shelly Masi, Roger Mundry, Laura Muniz, Claudio Tennie ed Heinz Gretsche.

Michela, anche se ora vive a Brooklyn, mi è stata vicina per buona parte del periodo occupato dal dottorato, anche nei momenti più difficili; e sembra che i miei genitori, da un po' di tempo a questa parte, siano un po' più convinti che io non sia *solo* un simpatico eccentrico. Grazie a tutti.

Leipzig, 7 giugno 2007

Introduzione

La natura della cultura può apparire un titolo paradossale. Se si aggiunge che le ricerche sperimentali che compongono la parte più sostanziale di questo lavoro sono state condotte all'interno di un laboratorio che si situa programmaticamente nell'ambito di una disciplina che prende il nome di *vita artificiale*, tutto ciò potrebbe scoraggiare anche il lettore più benintenzionato. Così, la mia intenzione, in questa breve introduzione, oltre ad essere quella di presentare una panoramica generale della struttura della tesi, sarà anche quella di cercare di persuadere, almeno tale lettore benintenzionato, del fatto che questi ossimori non siano del tutto privi di significato.

In antropologia è probabilmente poco o per nulla diffusa l'idea di un dualismo natura / cultura, ossia l'idea che si possano distinguere due classi distinte di fenomeni, sia che ciò avvenga sul piano ontologico (i fenomeni sono differenti nella loro *essenza*) sia su quello epistemologico (i fenomeni devono essere conosciuti e analizzati con metodologie differenti). Già Clifford Geertz (Geertz 1988) aveva messo in guardia gli studiosi di scienze sociali dall'accettare in modo più o meno implicito quella che definiva una concezione stratigrafica – o, più prosaicamente, “a cipolla” – del comportamento umano (*Ibidem*, p. 95):

L'emergere sincronico di un proencefalo espanso, di forme elaborate di organizzazione sociale e di modelli di cultura istituzionalizzati indica che la procedura standard di trattare in serie i parametri biologici, sociali e culturali – il primo considerato primario rispetto al secondo, e il secondo al terzo – è

errata. Al contrario, questi cosiddetti livelli dovrebbero intendersi come reciprocamente collegati e quindi da esaminare congiuntamente.

Giustissimo. Tuttavia, la proposta alternativa di Geertz, che pare divenuta senso comune nell'antropologia, era quella di considerare, detto in modo un po' semplicistico, i fatti naturali come *prodotto* dei fatti culturali, finendo per riproporre una concezione stratigrafica invertita, in cui "il cuore della cipolla" era la cultura e il resto derivava, almeno sul piano epistemologico, da essa (una trattazione più approfondita di questa lettura di Geertz è in Acerbi 2003). Probabilmente nessun antropologo contemporaneo accetterebbe di sottoscrivere una posizione così estrema (e sicuramente non accetterebbe di sottoscriverla come posizione di Geertz) e, in effetti, essa appare più come una caricatura di un'opzione epistemologica che come una posizione epistemologica concreta. Questo solo per poter dire che assumere una posizione che potrebbe apparire come opposta, ossia una concezione naturalistica della cultura, non significa ne' contrapporsi ad una presunta concezione comune dell'antropologia, in cui la cultura ha una priorità ontologica ed epistemologica sulla natura (perché questa è una caricatura), ne' riproporre lo stesso stereotipo – la natura ha una priorità ontologica ed epistemologica sulla cultura – che Geertz aveva correttamente giudicato sterile.

Ma, allora, cosa significa sottoscrivere una concezione naturalistica della cultura? In questa tesi non c'è spazio per una trattazione sistematica, teorica, di questa questione, tuttavia spero che gli esperimenti presentati contribuiscano a chiarirlo. Per iniziare, comunque, è doveroso esplicitare due vincoli fondamentali che l'accettazione di una concezione naturalistica della cultura dovrebbe imporre:

1) la cultura o, meglio, i fenomeni che vengono etichettati come "culturali", appartengono allo stesso dominio ontologico di tutti gli altri fenomeni. Questo significa solo, con le parole di Dan Sperber, che "tutto ciò che ha un potere causale lo deve esclusivamente alle sue proprietà fisiche" (Sperber 1999, p. 15). Questo impegno ontologico deve, o dovrebbe, implicare un corrispondente impegno epistemologico, per cui la spiegazione dei fenomeni culturali, pur non dovendo essere in alcun modo ricondotta, o *ridotta*, a queste remote "proprietà fisiche", dovrebbe utilizzare concetti che, quantomeno, non siano in contrasto con ciò che conosciamo a proposito delle scienze naturali e, se possibile, utilizzare dati e ipotesi provenienti da esse per chiarire e sviluppare i concetti stessi.

2) le metodologie sviluppate nelle scienze naturali possono essere usate per le indagini tipiche dell'antropologia. Questo tema è approfondito nel primo capitolo, ma per ora è possibile anticipare l'idea fondamentale: alcune metodologie che rappresentano, probabilmente, la ragione del successo delle scienze naturali, come l'importanza accordata all'osservazione diretta, lo stretto rapporto tra teorie e dati empirici, l'espressione delle teorie in termini quantitativi, o l'uso degli esperimenti di laboratorio, possono essere utilizzate, una volta apportate opportune modifiche, anche nelle scienze sociali.

In generale, la maggior parte degli antropologi sono decisamente contrari a qualsiasi pretesa di scientificità nel loro ambito di studio. Ritengo che, purtroppo, almeno in alcuni casi, questo atteggiamento sia un po' pregiudiziale e dovuto ad una mancata riflessione sulla pratica scientifica¹, ma, naturalmente, alcune critiche colgono nel segno: gli strumenti e le metodologie sviluppati nelle scienze naturali non possono essere trasportati così come sono da un ambito all'altro e riutilizzati immediatamente. Essi devono essere modificati e riadattati al differente oggetto di studio e alle differenti esigenze dei ricercatori. Secondo alcuni, lo strumento che potrebbe rendere possibile questo "trasferimento" metodologico sarebbe proprio quello della simulazione (Parisi 2001): un'indagine critica di questa posizione è l'oggetto del primo capitolo di questa tesi.

Nel secondo capitolo si passa ad esaminare più in dettaglio i presupposti

¹ Un esempio, fra tanti possibili, è quello dell'*oggettività* della ricerca. In realtà il tema richiederebbe esso stesso una tesi di dottorato, ma mi interessa almeno usarlo per far capire al lettore cosa intendo dire. Come è noto, in antropologia la prassi comune è quella di sottolineare, al contrario, la *soggettività* della propria esperienza di ricerca e della propria analisi, lasciando intendere che, ad ogni modo, anche la ricerca nelle scienze naturali sarebbe influenzata da inclinazioni personali, contesti sociali e culturali, interessi politici ed economici – e tralasciamo per clemenza il principio di indeterminazione di Heisenberg. Naturalmente tutto ciò è vero! Ma è vero anche che le scienze naturali – al contrario del rifiuto contemporaneo dell'antropologia – sono ben consapevoli del problema e cercano, quantomeno, di attuare una serie di pratiche per *ridurre* gli effetti di questi fattori e per stimolare un confronto imparziale. Tra queste pratiche si possono includere l'importanza che viene data al carattere collettivo della ricerca (quasi mai una ricerca è frutto del lavoro di un singolo ricercatore; per pubblicare esiste un sistema anonimo di *peer review* – fallibile quanto si vuole, ma esiste!), l'importanza di formulare le teorie in termini quantitativi in modo da renderle confrontabili e l'importanza di basare le proprie ipotesi su esperimenti controllati, quindi replicabili e falsificabili (vedi anche Parisi 2001).

metodologici ed epistemologici delle simulazioni che saranno poi presentate nella seconda parte della tesi. L'ambito della ricerca che fa uso di simulazioni, anche limitandosi alle scienze umane, è decisamente vasto e frammentato: con il termine "simulazione" possono essere etichettati lavori che, pur condividendo alcuni principi fondamentali – trattati, come detto, nel primo capitolo –, si differenziano molto l'uno dall'altro.

Gli esperimenti qui presentati possono essere fatti rientrare nel sotto – ambito della *vita artificiale* (Langton 1989) o della simulazione di sistemi *complessi e adattivi* (Holland 1975, 1995). Come viene illustrato più in dettaglio nel secondo capitolo, le simulazioni presentate possiedono caratteristiche peculiari, che le differenziano da altri approcci simulativi, in particolare:

- 1) l'uso di architetture di controllo basate su *reti neurali*.
- 2) l'uso di *algoritmi evolutivi e di apprendimento* per determinare le caratteristiche del comportamento degli agenti (il comportamento non viene quindi programmato a priori dallo sperimentatore).
- 3) l'attenzione nei confronti di una simulazione realistica, oltre che dei sistemi di controllo, delle proprietà fisiche degli organismi e dell'ambiente (*embodiement*).
- 4) l'importanza che viene accordata all'interazione dell'organismo con l'ambiente simulato (che può comprendere risorse, pericoli, altri organismi, ecc.).

Lo scopo di questa prima parte della tesi è, quindi, quello di fornire una "cassetta degli attrezzi" metodologica, necessaria per poter affrontare la parte sperimentale di questo lavoro. Non ha la pretesa di essere né una *review* dello stato dell'arte dell'uso delle simulazioni in antropologia o nelle scienze sociali, né di discutere in modo sistematico *tutte* le opzioni teoriche presentate².

² Mi limito a segnalare alcuni riferimenti bibliografici per il lettore interessato: Kohler e Gumerman (2000) hanno curato una raccolta di lavori simulativi particolarmente rilevanti da una prospettiva antropologica. Un "classico" che ha avuto grande influenza, incentrato su modelli ad agenti per esplorare l'evoluzione dei comportamenti cooperativi, è Axelrod (1984); un altro classico sull'uso delle simulazioni con agenti "a bassa complessità" che producono dinamiche sociali complesse è Epstein e Axtell (1996). La rivista *JASSS: Journal of Artificial Societies and Social Simulations* (tutte le uscite sono disponibili all'indirizzo: <http://jasss.soc.surrey.ac.uk/>) è interamente dedicata, come si evince dal titolo, all'uso delle simulazioni nelle scienze sociali.

Più specifico rispetto al particolare approccio utilizzato in questa tesi è, in italiano, Parisi

La seconda parte della tesi, composta dai capitoli tre, quattro e cinque, presenta gli esperimenti simulativi realizzati. Tutti riguardano tematiche inerenti alla trasmissione e all'evoluzione culturale. Con il terzo capitolo si incomincia a presentare un primo modello di come sia possibile, all'interno di questo orizzonte metodologico, realizzare una simulazione del processo di trasmissione culturale. L'esperimento, in particolare, riguarda le conseguenze della trasmissione intragenerazionale (o orizzontale) nelle dinamiche di evoluzione culturale. Il quarto capitolo presenta uno scenario simile, ma con l'aggiunta dell'apprendimento individuale e di una simulazione più realistica dei comportamenti sensomotori degli agenti e mostra come questi possano avere effetto sul processo di trasmissione culturale. Infine, nell'ultimo capitolo, viene analizzato un altro modello, in cui il processo di trasferimento "sociale" dell'informazione viene realizzato in modo indiretto, attraverso l'influenza dell'interazione sociale sull'apprendimento individuale degli agenti. Per questo modello, inoltre, viene effettuato un confronto dei risultati della simulazione con un'analisi matematica della stessa dinamica.

Nei capitoli della seconda parte della tesi ho cercato di mantenere il più possibile lineare il resoconto degli esperimenti. Le note integrano il testo con riferimenti alla letteratura più specificatamente antropologica o psicologica e, all'inizio o alla fine dei capitoli, sono presenti sezioni che, traendo spunto dai risultati delle simulazioni, affrontano temi più generali: in particolare, queste sezioni riguardano le differenze tra trasmissione ed evoluzione culturale e trasmissione ed evoluzione biologica (capitolo terzo), l'importanza di tenere in considerazione i processi individuali di trasmissione per rendere conto delle dinamiche culturali a livello popolazionale (capitolo quarto), i meccanismi indiretti di trasferimento sociale dell'informazione (capitolo quinto) e una breve riflessione sul rapporto tra simulazioni ad agenti e modelli formali matematici (ancora capitolo quinto). Nelle

(2001), dedicato al contributo che le simulazioni possono dare, in particolare, alle scienze umane. Nolfi e Floreano (2000) ha un orientamento esplicitamente robotico, e dà grande rilievo all'uso di algoritmi di ricerca evolutivi e all'importanza dell'*embodiement*. Clark (1997) è un'utile – ed equilibrata – introduzione alla critica, in scienza cognitiva, agli approcci "classici" (intelligenza artificiale, funzionalismo, ecc.) che proviene, appunto, dalla vita artificiale, dalla robotica e dalla simulazione dei sistemi dinamici. La rivista *Artificial Life* ha dedicato recentemente, nel 2003, un numero speciale (9/4) al tema: "Collective Effects on Human Behavior" e, nel 2008, uscirà un numero speciale, a cura di Acerbi, Marocco e Vogt, della rivista *Connection Science*, esplicitamente dedicato alla simulazione dei processi di apprendimento sociale ed evoluzione culturale.

appendici sono contenuti dettagli tecnici o implementativi che ritengo interessanti anche se non essenziali per la comprensione degli esperimenti.

~

La maggior parte di questa tesi (soprattutto la parte sperimentale) si basa su materiale che ho avuto la fortuna di poter presentare in conferenze o seminari o di pubblicare su riviste, atti di conferenze o capitoli di libri. In particolare:

- parte del capitolo primo (su come le simulazioni possano essere strumenti adatti per integrare nelle scienze sociali alcune delle metodologie delle scienze naturali) è stata presentata oralmente, in maniera decisamente informale, al Collegio Santa Chiara dell'Università di Siena, il 12 – 9 – 2006, in forma di un tutorial (tenuto insieme a Gennaro di Tosto) dal titolo “Modelli computazionali di fenomeni sociali: apprendimento e cooperazione” in occasione del “Terzo Workshop Italiano di Vita Artificiale”. Allo stesso modo, la discussione sull'ontologia naturalistica dei fatti culturali e' ripresa da un articolo di review sull'antropologia cognitiva che ho pubblicato sulla rivista “Sistemi Intelligenti” (Acerbi 2005).

- la simulazione descritta nel capitolo terzo è stata presentata oralmente al LabSEC (Laboratorio per lo studio dei sistemi socio-economici complessi) del Centro dei Sistemi Complessi dell'Università di Siena il 13 – 1 – 2005; all'ISTC di Roma il 4 – 3 – 2005 al “Secondo Workshop di Vita Artificiale”; come poster il 16 – 9 – 2005 all'University College di Londra, in occasione della conferenza “Pattern and Process in Cultural Evolution” ed infine pubblicata sulla rivista *JASSS - Journal of Artificial Societies and Social Simulations* (Acerbi & Parisi 2006)

- una versione estesa dell'ultima parte del capitolo terzo (“Evoluzione biologica ed evoluzione culturale”) è stata presentata oralmente all'Università di Tor Vergata, Roma, il 7 – 10 – 2006, in occasione della conferenza “Epistemic Complexity, Scientific Invariants and Cultural Evolution”.

- la simulazione presentata nel capitolo quarto è stata presentata oralmente al Collegio Santa Chiara dell'Università di Siena, il 13 – 9 – 2006, in occasione del

“Terzo Workshop Italiano di Vita Artificiale”; il 28 – 10 – 2006 a Genova, al “Terzo Convegno Nazionale di Scienze Cognitive”; il 5 – 4 – 2007 a Honolulu (USA), nell’ambito della conferenza “IEEE Symposium on Artificial Life 2007”. Due versioni differenti sono state pubblicate negli atti del convegno nazionale di scienze cognitive (Acerbi & Nolfi 2006) e in quelli della conferenza IEEE – Alife (Acerbi & Nolfi 2007).

- l’ultima parte del capitolo quarto (“Cos’è la trasmissione culturale?”, “JukeBox, Camino del Diablo e attrattori: metafore della cultura” e “Trasmettere e copiare”) è basata sul capitolo “Le trasmissioni della cultura” inserito nel volume a cura di Alessandro Lutri *Modelli della mente e processi di pensiero: il dibattito antropologico contemporaneo* (Acerbi 2007).

- la prima parte del capitolo quinto è ancora basata sul capitolo inserito nel volume a cura di Alessandro Lutri (Acerbi 2007).

- la simulazione presentata nel capitolo quinto è stata accettata per la pubblicazione negli atti della 9th European Conference of Artificial Life (Acerbi, Marocco & Nolfi, *in press*) e, insieme al modello matematico, sarà presentata al workshop – associato alla conferenza – dal titolo “Social Learning in Embodied Agents” (Acerbi, Cecconi, Marocco & Zappacosta, 2007).

Colgo l’occasione per ringraziare qui tutti i *reviewer* anonimi delle pubblicazioni e tutti coloro che sono intervenuti in occasione delle presentazioni orali con domande di chiarimento, commenti o critiche. La forma attuale di questi capitoli risente in grande misura delle loro suggestioni. Inoltre, ringrazio tutti i coautori di questi lavori (Domenico Parisi, Stefano Nolfi, Davide Marocco, Federico Cecconi, Stefano Zappacosta) senza il contributo dei quali questi tesi sarebbe stata sicuramente differente.

1. Le simulazioni come strumento di conoscenza

L'antropologia, fin dall'inizio – o da quello che viene considerato il suo inizio “moderno”, nella seconda metà dell'ottocento – si è trovata a dover affrontare un dilemma interno, probabilmente in modo più profondo rispetto ad altre discipline, sulla natura del sapere prodotto dai suoi praticanti. Da una parte ci sono coloro che ritengono che lo scopo dell'antropologia sia più che altro descrittivo, o evocativo; che le descrizioni prodotte dagli etnologi debbano riguardare singole società o culture senza pretese di generalizzazione e che l'antropologia si debba porre al fianco delle altre discipline umanistiche (la storia, ma anche la retorica e la critica letteraria) per quanto riguarda gli obiettivi e il tipo di strumenti teorici utilizzati. Dall'altra sono schierati coloro che ritengono che l'antropologia sia una scienza nomotetica (che, cioè, possa e debba individuare le leggi generali che governano i fenomeni) e che si debba porre al fianco delle scienze naturali (come la biologia e la psicologia, anche se non molti accetterebbero di chiamare quest'ultima scienza naturale).

Questa netta separazione ha fatto sì che l'antropologia si dividesse in numerose sotto – discipline con grosse difficoltà di integrazione (con l'antropologia fisica o biologica schierate decisamente sul lato “scienze naturali” e l'antropologia culturale e sociale più su quello umanistico), ma che anche paradigmi di ricerca che vengono identificati con la stessa etichetta – per esempio nel caso dell’“antropologia culturale” – siano estremamente differenti tra loro e oscillino tra i due poli scienze

naturali / discipline umanistiche.

Una spesso citata osservazione di Eric Wolf coglie probabilmente la sensazione di molti di coloro che praticano l'antropologia culturale: "l'antropologia è allo stesso tempo la più scientifica delle discipline umanistiche e la più umanistica delle scienze" (Wolf 1964, traduzione mia). L'allusione lascia intendere che si tratterebbe, insomma, di un falso problema, o al massimo di una questione identitaria dei singoli antropologi, di qualche dubbio nella collocazione della disciplina rispetto alle tradizionali cesure accademiche. Tuttavia, come hanno notato gli psicologi evolucionistici John Tooby e Leda Cosmides (Tooby & Cosmides 1992), in un vibrante attacco all'epistemologia che sembra guidare la maggior parte della ricerca antropologica contemporanea, proprio questa collocazione incerta potrebbe essere una delle ragioni principali della "crisi" delle discipline antropologiche (se, come Tooby e Cosmides, si è dell'opinione che ci sia stata, o ci sia, una "crisi" delle discipline antropologiche). Tooby e Cosmides iniziano il loro ragionamento analizzando quella che chiamano *integrazione concettuale* delle scienze naturali, ossia il fatto che queste "hanno sviluppato una robusta combinazione di coerenza logica, descrizione causale, potere esplicativo, ripetibilità dei risultati, e sono divenute esempio di come la conoscenza umana possa essere profondamente soddisfacente e affidabile" (*ibidem*, traduzione mia). Ciò, al contrario, non sarebbe avvenuto nelle scienze sociali: anzi, ancora secondo Tooby e Cosmides, la recente fioritura di posizioni decisamente anti-scientifiche in antropologia sarebbe dovuta alla medesima sensazione di insoddisfazione, solo presa, per così dire, dal punto di vista completamente opposto. Per usare le parole di Marvin Harris: "il problema non è che abbiamo avuto troppa scienza sociale positivista, è che ne abbiamo avuta troppo poca" (Harris 1991, traduzione mia).

Eppure, anche se la sensazione è che, al momento, l'antropologia, o almeno l'antropologia culturale, sia abbastanza schiacciata su posizioni che sono decisamente contrarie a quella che, con un'ironia voluta, Harris chiama scienza sociale "positivista", esistono molti esempi che vanno in questa direzione, che essi siano etichettati con il termine antropologia (si veda per esempio Atran 1998, Boyer 1994, Boyd & Richerson 2006, Hirschfeld 1996, Sperber 1996) oppure no (per esempio Aunger 2000, Axelrod 1984, Barkow *et al.* 1992, Epstein & Axtell 1996, Tomasello 2005).

Il lavoro che presenterò in questa tesi è, anch'esso, orientato in questa

direzione. La motivazione che lo guida è che sostenere un approccio *naturalistico* allo studio dei fenomeni sociali e culturali implichi una serie di impegni epistemologici e metodologici ben precisi, tra i quali, come vedremo più avanti, utilizzare un'ontologia fisicalista, riconoscere l'importanza di alcuni strumenti utilizzati dalle scienze naturali (per esempio, gli esperimenti di laboratorio, l'attenzione ai dati empirici, le teorie espresse in modo quantitativo, ecc.) e l'importanza dell'integrazione con le altre discipline che studiano il comportamento degli esseri umani da un punto di vista naturalistico. Tuttavia, come cercherò di spiegare nel corso di questo capitolo, non mi pare nemmeno sensato un atteggiamento di cieca reverenza nei confronti delle scienze naturali attuali, che implicherebbe un "trasferimento" diretto delle metodologie utilizzate da queste alle scienze sociali.

Ad oggi non c'è alcun accordo su quale sia il modo migliore per realizzare questo progetto e con questa tesi non pretendo certo di fornire una risposta ad una domanda così ambiziosa. Questo capitolo, e il prossimo, non hanno nessuna pretesa di completezza, non vogliono difendere in modo rigoroso un paradigma teorico, ma, come anticipato nell'Introduzione, solo presentare alcuni strumenti (e qualcuna delle loro tracce nel dibattito all'interno dell'antropologia e delle scienze cognitive) che verranno poi utilizzati nei capitoli sperimentali della tesi.

Ontologia naturalistica

In un capitolo significativamente intitolato "Come essere un vero materialista in antropologia" (Sperber 1999), Sperber invita gli antropologi ad interrogarsi sulla natura ontologica dei fatti da loro studiati (si veda anche Acerbi 2003, 2005). Di che natura sono i fatti culturali? Quando gli antropologi parlano di "tabù", "matrimonio", "etnia", a cosa si stanno riferendo? La letteratura antropologica non abbonda, in effetti, di riflessioni articolate sullo statuto ontologico dei propri oggetti di studio.

A proposito del dibattito sullo statuto scientifico o meno dell'antropologia, con cui abbiamo aperto questo capitolo, Sperber ha, se non altro, il merito di sostenere una tesi ben precisa: le scienze propriamente dette formano un corpus unitario perché utilizzano la medesima ontologia per gli oggetti studiati, che si può definire materialista o, più propriamente, fisicalista. Con ciò si vuole semplicemente affermare

che “tutto ciò che ha un potere causale lo deve esclusivamente alle sue proprietà fisiche” (Sperber 1999); tuttavia, non è immediato capire in cosa consistono precisamente le proprietà fisiche di un tabù o di un matrimonio. Fino a pochi anni fa, gli antropologi erano portati a scegliere fra due opzioni opposte. Da una parte, coloro che accettavano di mettere in rilievo la natura rappresentazionale o simbolica dei fatti culturali negavano che una loro spiegazione dovesse essere collegata, in modo interessante, a qualsiasi causa di natura fisica (Geertz 1973). Dal lato opposto, la tendenza era di sottolineare primariamente l'importanza dei vincoli economici, tecnologici o ecologici, ritenuti genuinamente “fisici”, tacciando gli altri approcci di “intellettualismo” o “idealismo” (Harris 1968).

La strategia proposta da Sperber, invece, consiste nell'ammettere che i fatti culturali siano, fondamentalmente, rappresentazioni condivise, ma, nello stesso tempo, di situare le rappresentazioni nel mondo materiale. Nella maggior parte dei casi, nelle scienze sociali, le rappresentazioni sono state considerate a un livello puramente astratto e sono stati analizzati i rapporti formali che intercorrono tra di esse a questo livello: tuttavia, “a provocare nel bambino un'eccitante situazione di paura non è la favola di *Cappuccetto Rosso* in astratto, ma la comprensione delle parole della madre” (Sperber 1996) ossia, non si vede in che modo si possa essere in grado di attribuire proprietà causali alle rappresentazioni considerate come entità astratte. Come possiamo studiarle, al contrario, come entità concrete? Secondo Sperber, la rappresentazione astratta “Cappuccetto Rosso”, può essere meglio considerata come un'etichetta che contrassegna una serie di rappresentazioni concrete, causalmente collegate tra loro, che sono sia le rappresentazioni pubbliche di Cappuccetto Rosso (un testo, un'effettiva narrazione) sia le rappresentazioni mentali che hanno come oggetto Cappuccetto Rosso. Se la natura materiale delle rappresentazioni pubbliche è evidente, e può essere accettata di buon grado da qualsiasi antropologo, quella delle rappresentazioni mentali è al centro delle riflessioni della scienza cognitiva e, sebbene con sfumature diverse, è ammesso che possano essere considerate come entità puramente materiali: è così possibile esplicitare le relazioni causali che queste rappresentazioni intrattengono tra di esse all'interno di un'ontologia puramente fisicalista.

Ora, cosa significa ciò dal punto di vista degli impegni metodologici? Naturalmente non si tratta in alcun modo di *ridurre* i fatti culturali a “forze” fisiche, qualsiasi cosa questo voglia dire. Secondo il filosofo Daniel Dennett, quando si parla

di riduzionismo, si possono intendere due cose profondamente differenti (Dennett 1995). L'idea generale è che i concetti di una scienza si possano "ridurre" a quelli di un'altra, quelli della chimica a quelli della fisica, quelli della biologia a quelli della chimica e via dicendo, proseguendo in questo modo fino ad arrivare ai concetti delle scienze sociali. Secondo l'interpretazione riduzionista più radicale, si vogliono "abbandonare i principi, le teorie, il vocabolario e le leggi delle scienze di livello superiore in favore dei termini di quelle di livello inferiore" (*ibidem*), vale a dire che i fatti culturali andrebbero descritti con un vocabolario esclusivamente psicologico, il quale a sua volta dovrebbe essere sostituito da uno biologico, con la conseguenza che alla fine dovremmo arrivare a spiegarli con principi fisici, il che è naturalmente assurdo.

E' assurdo per due ragioni: da una parte, una spiegazione totalmente in termini fisici di un fatto culturale non ci sarebbe di grande aiuto nella comprensione di questo (dovremmo parlare dei movimenti delle melocole presenti nello spazio dove si svolge un rituale? O delle attivazione neurali dei partecipanti? Forse di entrambe?), dall'altra sarebbe, soprattutto, semplicemente *sbagliato*. Fra tutte le suggestioni – a volte purtroppo abbastanza vaghe – che provengono dalla cosiddetta scienza dei sistemi complessi, una sembra abbastanza chiara: i sistemi complessi "tendono ad essere inseriti in gerarchie di sistemi, dove un sistema ad un certo livello costituisce un elemento di un sistema più grande al livello immediatamente successivo della gerarchia e ci sono influenze reciproche tra tutti i livelli" (Parisi 2001), di conseguenza, conoscere le leggi che governano un sistema ad un certo livello è utile, ma non ci serve per predire o spiegare il comportamento del sistema ad un livello gerarchico superiore.

Un'interpretazione riduzionista più ragionevole consiste nell'utilizzare concetti che, quantomeno, non siano in contrasto con ciò che conosciamo a proposito delle leggi delle scienze (o dei sistemi) di livello inferiore e, eventualmente, utilizzare tali leggi per chiarire e sviluppare i concetti stessi.

Ancora due punti a proposito dell'invito di Sperber. L'opzione teorica che Sperber preferisce, nelle scienze cognitive, è la teoria rappresentazionale della mente, associata, in un senso abbastanza forte, all'intelligenza artificiale "classica". Come avremo modo di approfondire in dettaglio nel capitolo successivo, i lavori presentati in questa tesi si discostano da questa opzione teorica per molti aspetti. Tuttavia, l'idea generale di un'ontologia naturalistica per i fatti culturali può essere considerata

indipendente dalle specifiche opzioni teoriche che la sostengono ai livelli “inferiori”, psicologico o biologico, almeno nella misura in cui esse si ispirano agli stessi principi fondamentali.

Infine, una nota cautelativa: anche se qui, e in seguito, l’approccio naturalistico ai fatti culturali è sostenuto con decisione, questo non significa che lo ritenga l’unico approccio possibile e, soprattutto, l’unico approccio *corretto*. Una pluralità di metodologie, a condizione che siano in grado di confrontarsi e di dialogare tra esse, è, probabilmente, la migliore possibilità che l’antropologia e le scienze sociali in generale hanno a disposizione allo stato attuale. La mia impressione è che si tratti soprattutto di una questione pragmatica: a seconda degli ambiti che si vogliono analizzare è necessario, o utile, muoversi tra i vari livelli di spiegazione, ma non abbiamo purtroppo a disposizione un metodo rigoroso per compiere queste scelte. L’unico presupposto che ritengo essenziale è, ripeto, un comune accordo su alcuni principi generali, o quantomeno una disponibilità a valutarne in modo non pregiudiziale le possibilità.

Dati empirici, teorie quantitative, laboratori

Si è detto che l’ontologia naturalistica adottata dalle scienze naturali dovrebbe, almeno in linea di principio, permettere una maggiore integrazione fra le varie discipline, che a sua volta consentirebbe di produrre quell’integrazione concettuale che gli psicologi evoluzionistici ritengono assente nelle scienze sociali. Tuttavia, questa è solo una parte della storia. Le scienze naturali adottano anche una serie di pratiche, di metodologie di ricerca, di procedure condivise senza le quali questa integrazione sarebbe impossibile. Naturalmente questa è una visione idealizzata dell’attività scientifica: come ho anticipato introducendo questa tesi, nessuno si aspetta qualcosa come una conoscenza puramente oggettiva, o come un rapporto completamente trasparente tra dati empirici ed ipotesi esplicative. Le pratiche messe in atto dalle scienze naturali, però, *cercano* di minimizzare i rischi derivanti dalla soggettività, dagli interessi personali dei singoli ricercatori o dei gruppi e dagli errori sempre possibili in qualunque attività umana.

Innanzitutto, la scienza attribuisce un’importanza fondamentale ai dati

empirici (su questo e sul resto del paragrafo si veda anche Parisi 2001). Da un certo punto di vista, tutte le discipline attribuiscono grande importanza all'osservazione empirica; del resto, ogni spiegazione in antropologia è, o almeno così si suppone, la spiegazione di un qualche fatto che viene osservato dall'etnografo. Quello che cambia nella pratica delle scienze naturali è il rapporto che intercorre tra i dati empirici e le ipotesi esplicative. Nel caso delle scienze naturali, c'è, o ci dovrebbe essere, un rapporto molto stretto tra dati empirici e ipotesi esplicative, ossia un'ipotesi esplicativa dovrebbe, almeno in linea di principio, potere sempre essere falsificata da ulteriori osservazioni empiriche.

Questo, mi sembra, non avviene per molte delle spiegazioni in antropologia. Anche se per qualche antropologo contemporaneo l'idea stessa di fornire spiegazioni di qualcosa è in sé sbagliata, l'antropologia ha fornito nel corso della sua storia, e continua a fornire, spiegazioni di fatti osservati. Tuttavia, nella maggior parte dei casi, queste spiegazioni sono interpretazioni di fatti che non possono essere verificate o falsificate da ulteriori osservazioni empiriche. Per quanto riguarda le scienze naturali, invece, i dati empirici vengono raccolti attraverso procedure standardizzate e le teorie vengono espresse in termini quantitativi: così, le teorie sono formulate in modo tale da rendere possibile ricavare da esse delle predizioni ed è esattamente questo che permette la loro falsificazione. Una teoria scientifica dovrebbe incorporare tutto il necessario perché un altro scienziato possa riprodurre le condizioni sperimentali (per questo nella scienza gli esperimenti in laboratorio sono così importanti) e verificare che le predizioni che si ricavano da questa teoria siano corrette per un certo insieme di osservazioni. Gli esperimenti di laboratorio sono importanti per un'altra ragione ancora: le condizioni di laboratorio permettono agli sperimentatori di variare in modo controllato almeno alcuni dei parametri sensibili per l'ipotesi da testare, in modo da poter valutare in modo esatto, almeno idealmente, l'influenza che questi parametri hanno sui risultati nel loro complesso.

Il fatto che nelle scienze sociali, o nella maggior parte di esse, non esista questa stretta integrazione tra predizioni e dati empirici, che abbiamo caratterizzato come legata alla presenza di procedure standardizzate di osservazione (in particolare in relazione agli esperimenti di laboratorio, che permettono di riprodurre esattamente le stesse condizioni e di manipolare le variabili sensibili) e alla formulazione in termini quantitativi delle teorie (che permette di ricavare in modo univoco delle predizioni dalle teorie) non può essere semplicemente considerato come una

“mancanza” da parte degli scienziati sociali, ma è legato alla natura dei fenomeni che essi studiano. La stragrande maggioranza dei fenomeni studiati, per esempio, da un antropologo, non può essere “portata in laboratorio” per una serie di motivi, tutti validissimi. Innanzitutto, i fenomeni studiati dalle scienze sociali sono dipendenti da un contesto ampio, spesso *molto* ampio, dal punto di vista spaziale e temporale, e questo contesto non può essere riprodotto in laboratorio. In quest’ottica, non ha molto senso invitare gli antropologi a studiare i loro fenomeni in laboratorio, semplicemente perché questo non sarebbe, per ragioni pratiche, possibile. Inoltre, la maggior parte dei fenomeni studiati dalle scienze sociali non sono ripetibili: non è possibile, se non in un senso molto vago, ripetere l’osservazione di un fenomeno sociale nelle *stesse* condizioni, se non altro per l’intrinseca complessità delle variabili che intervengono in questa classe di fenomeni. Ancora, nel caso dei fenomeni sociali, non è sempre possibile (in realtà non lo è quasi mai) manipolare le variabili sperimentali, sia per ragioni pratiche che per ragioni etiche e abbiamo visto come questo sia di importanza fondamentale per il metodo delle scienze naturali.

Infine, le scienze sociali non producono, di solito, teorie quantitative e anche questo rende più difficile ricavare da esse predizioni falsificabili. Come per gli esperimenti di laboratorio, questa non è una “mancanza” delle scienze sociali. E’ ragionevole pensare che il compito delle scienze sociali sia più difficile di quello delle scienze naturali, ancora, probabilmente, per un problema di complessità. Nessuno, chiaramente, si aspetta che si arrivi un giorno a predizioni per il comportamento degli esseri umani simili alle predizioni, per esempio, della meccanica classica. Tuttavia, i biologi, per fare un esempio di un'altra scienza naturale, si trovano ad affrontare problemi molto simili: la complessità degli ecosistemi o dei comportamenti animali, anche di quelli che ci possono apparire come estremamente semplici, tende ad impedire loro di formulare predizioni in termini quantitativi. In biologia esistono (1) leggi generali estremamente potenti ma da cui si possono ricavare, a seconda dei contesti, predizioni empiriche molto differenti (il meccanismo darwiniano di mutazione e selezione naturale è una di queste), (2) modelli limitati e astratti, basati su queste leggi generali, che producono predizioni rigorose solo per casi specifici e in condizioni ideali e (3) qualche tentativo di verifica empirica in condizioni controllate (laboratorio) o semi – controllate (osservazioni empiriche con rigidi protocolli sperimentali, ecc.). Da questo punto di vista, il caso delle scienze sociali potrebbe essere, almeno in parte, analogo.

Come discuteremo nella parte successiva di questo capitolo, è stato proposto che le simulazioni possano offrire uno strumento per quelle scienze, come le scienze sociali, che non possono utilizzare così com'è il metodo delle scienze naturali (Parisi 2001).

Cos'è una simulazione

In linea di principio, una simulazione è semplicemente un programma che gira in un computer. In fisica, per esempio, o in altre scienze le cui teorie sono espresse attraverso sistemi di equazioni (un caso che è stato storicamente importante per le simulazioni è la meteorologia), non è possibile, a volte, ricavare “a mano” una soluzione analitica di un dato sistema di equazioni. Per questo motivo, da quando i computer hanno incominciato ad essere diffusi tra i ricercatori, i fisici, per esempio (o i meteorologi...), li hanno utilizzati per risolvere, almeno parzialmente, questi sistemi non trattabili analiticamente. Per fare questo è necessario delimitare un certo spazio dei parametri e lasciare che il computer risolva il sistema di equazioni per ogni punto di questo spazio. In questo modo è possibile osservare determinati andamenti del sistema e ricavare delle predizioni che, almeno in quello spazio dei parametri, sono testabili empiricamente.

Le simulazioni che verranno qui presentate sono però differenti da queste. Invece che descrivere con delle equazioni il comportamento del sistema a livello globale, il comportamento viene descritto al livello degli agenti che operano nel sistema. Una breve descrizione di una celebre “simulazione ad agenti” può essere utile per chiarire il funzionamento di questo tipo di simulazioni. Il modello di segregazione spaziale di Schelling (Schelling 1971) è un semplice modello ad agenti, in cui gli agenti sono disposti in una griglia bidimensionale. Gli agenti possono essere di due “colori” differenti, diciamo rosso e verde, e possono spostarsi di una casella per volta in questo mondo a griglia. Gli agenti hanno una semplice regola comportamentale: essi possono “osservare” i loro vicini – quindi, al massimo, altri otto agenti in tutto – e decidere se spostarsi o no da quella posizione. La decisione viene presa in base al “colore” dei vicini: nel caso del modello di Schelling, se almeno il 37,5 % dei vicini – ossia, nel caso in cui ce ne siano otto, almeno tre (meno della metà) – sono dello stesso colore, l'agente rimane fermo, altrimenti si sposta. Partendo

da una disposizione casuale, ossia da una disposizione in cui gli agenti dei due colori sono mescolati tra loro, dopo un certo numero di interazioni, il modello di Schelling mostra come si giunga ad uno stato stazionario in cui si creano pattern spaziali omogenei, ossia emerge una forma di segregazione spaziale: zone della griglia “abitate” da agenti rossi e zone della griglia “abitate” da agenti verdi, le une separate dalle altre (vedi Fig. 1.1).

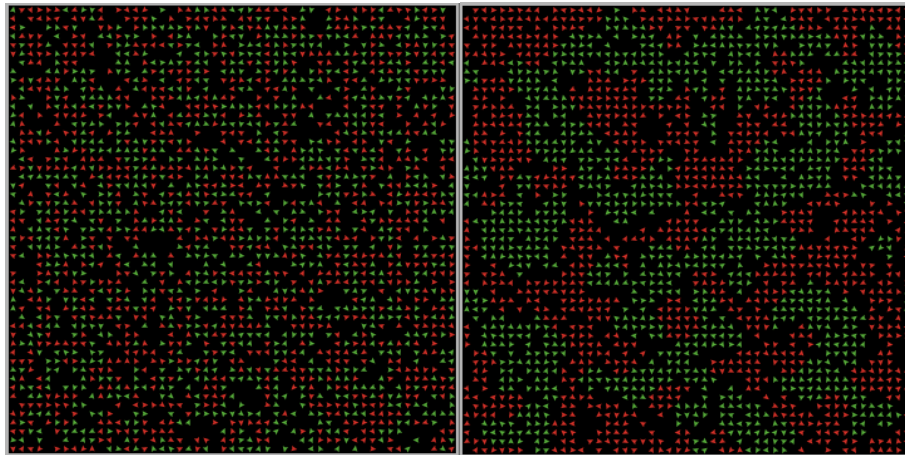


Fig. 1.1: Il modello di segregazione di Schelling, realizzato tramite il software NetLogo (Wilensky 1988). A sinistra, la situazione iniziale, con gli agenti disposti in modo casuale; a destra, la situazione finale, con l’emergenza di aree di segregazione spaziale.

Il modello di Schelling è, ai nostri scopi, interessante per diversi motivi. Innanzitutto mostra un caso estremamente semplice di simulazione ad agenti: il comportamento del sistema, come detto, non è descritto al livello delle proprietà globali del sistema (come potrebbe essere se, per esempio, il modello si limitasse a calcolare un indice globale di segregazione) ma a livello del comportamento degli agenti individuali. In più, il modello mostra una proprietà emergente del sistema, difficilmente prevedibile concentrando la propria analisi al livello dei comportamenti individuali. Gli agenti di Schelling non solo non hanno nessuna tendenza intrinseca a formare aree segregate (né, naturalmente, esiste una qualche nozione di segregazione che possa essere afferrata a livello dei singoli agenti), ma anche hanno un comportamento che può solo debolmente essere definito “segregazionista”, ossia, come detto in precedenza, essi “accettano” la presenza della *maggioranza* di vicini di

colore diverso. Tuttavia, a livello globale, queste regole comportamentali producono un pattern emergente di segregazioni spaziali. Potremmo, conoscendo solo le “attitudini” dei singoli agenti prevedere quello che accade a livello globale, senza simularlo? Probabilmente no. Il concetto di emergenza, che incontreremo ancora nel resto della tesi, anche se abbastanza difficile da definire in modo univoco, è una delle proprietà più interessanti delle simulazioni in generale e le distingue in modo marcato dagli strumenti tradizionali delle scienze naturali.

Come le simulazioni possono permettere un approccio genuinamente naturalistico allo studio dei fatti socio-culturali.

Nel caso del modello di Schelling, le capacità “cognitive” degli agenti e le loro regole di interazione con l’ambiente sono estremamente semplici ed idealizzate. I modelli ispirati alla vita artificiale, che presenteremo nel secondo capitolo, ambiscono invece a simulare in modo più dettagliato alcune caratteristiche reali degli esseri viventi e dell’ambiente in cui essi vivono. Innanzitutto, come si dirà più avanti, gli agenti delle simulazioni che presenteremo sono controllati da reti neurali, che sono modelli che si ispirano al funzionamento biologico dei sistemi nervosi reali. Inoltre, le simulazioni possono incorporare dinamiche di tipo evolutivo, utilizzando algoritmi appositi (denominati, in generale, algoritmi genetici). Oltre a questi due principi fondamentali è, in larga misura, una scelta lasciata agli sperimentatori quanto spingersi nel realismo della simulazione dei processi cognitivi o delle interazioni tra organismi, della complessità dell’ambiente sociale, o della simulazione delle caratteristiche fisiche e corporee degli organismi e via dicendo.

Il punto che mi preme sottolineare è solo che, se, analizzando il modello di Schelling, non si può certo dire che si tratti di un passo in direzione di un approccio naturalistico alle dinamiche sociali (nessuna realistica caratteristica cognitiva o biologica viene, in effetti, simulata), lo strumento delle simulazioni può permettere di realizzare questa aspirazione. Secondo alcuni (Parisi 2001), inoltre, non solo le simulazioni possono permettere questo, ma sono, ad oggi, l’*unico* strumento che lo può permettere. Si potrebbe pensare, infatti, che basti attendere lo sviluppo, per esempio, della biologia o delle neuroscienze per avere ipotesi completamente naturalistiche rispetto al comportamento degli esseri umani. Tuttavia, per gli stessi motivi esposti sopra a proposito della natura dei sistemi complessi, non è realistico

aspettarsi che le discipline che studiano i livelli gerarchici “inferiori” possano fornire spiegazioni *complete* per ciò che avviene ai livelli superiori. L’unica possibilità sembra quella dell’integrazione e dell’analisi contemporanea dei livelli e le simulazioni sembrano essere lo strumento adatto per fare questo. Come è scorretto affermare che il comportamento *non* si trovi nelle menti degli essere umani, ma nelle interazioni socio-culturali, è pure scorretto affermare che il comportamento si trovi *solo* nelle menti degli esseri umani: il comportamento è all’intersezione di tutti questi livelli e le metodologie simulative cercano di dare un qualche contributo iniziale a questo tipo di analisi.

Come le simulazioni possono fornire un analogo degli esperimenti di laboratorio per le scienze sociali

Si è detto di come esistano una serie di motivi validi per cui le scienze sociali non possono utilizzare gli strumenti metodologici utilizzati dalle scienze naturali. Come si ricorderà, non ha senso affermare che gli antropologi dovrebbero studiare in laboratorio i fenomeni che essi di solito studiano e, tuttavia, gli esperimenti di laboratorio permettono, nelle scienze naturali, di riprodurre le stesse osservazioni nelle medesime condizioni, consentendone la falsificazione, e di manipolare le variabili sensibili.

Parisi (*ibidem*, p.36) ha definito le simulazioni “laboratori sperimentali virtuali”:

Così come avviene nel laboratorio reale, i risultati delle simulazione [...] si verificano in condizioni che sono controllate dallo scienziato. Inoltre, come nel laboratorio reale, lo scienziato può manipolare le condizioni che determinano o influenzano il verificarsi di questi fenomeni virtuali, cioè può aggiungere, togliere e modificare parti della simulazione, può intervenire sui fattori e le variabili incorporate nella simulazione, può modificare il valore quantitativo dei parametri, e osservare ogni volta quali sono i risultati di questi suoi interventi. Questo è esattamente quello che fa lo scienziato nel laboratorio sperimentale reale.

In questi “laboratori sperimentali virtuali”, le limitazioni che abbiamo delineato in precedenza, ossia l’impossibilità di “portare il contesto” nel laboratorio e

di manipolare liberamente le variabili, vengono a cadere e, allo stesso modo, è possibile specificare tutte le condizioni necessarie perché altri ricercatori riproducano lo stesso esperimento. Naturalmente, esiste un abisso tra quello che un ricercatore di scienze sociali intende di solito per “contesto” di un fenomeno e quello che può essere effettivamente realizzato attraverso una simulazione. Ciò non toglie che, almeno in linea di principio, in una simulazione è possibile fare rientrare un numero di fattori molto grande, nella misura in cui essi vengono ritenuti necessari dallo sperimentatore per riprodurre o analizzare un fenomeno. Dal punto di vista opposto, se con un numero di fattori più piccolo di quello che avremmo ipotizzato possiamo riuscire a riprodurre qualcosa di analogo al fenomeno che ci interessa studiare, questo può farci “tornare indietro” al fenomeno reale e valutare se, effettivamente, il contesto da tenere in considerazione per analizzarlo deve essere necessariamente così ampio. Questo è un po’ quello che avviene con il modello di Schelling, dove una struttura complessa, come quella della segregazione spaziale, che un ricercatore potrebbe fare risalire a cause egualmente complesse e che necessitano un ampio contesto per essere studiate, viene riprodotta con un numero estremamente esiguo di fattori. Ciò non significa che ogni fenomeno reale di segregazione spaziale possa essere spiegato con il modello di Schelling, ma, quantomeno, ci invita alla parsimonia nelle spiegazioni dei fatti sociali, costringendoci a “ritornare” al fenomeno con una prospettiva differente.

Come le simulazioni possono permettere alle scienze sociali di formulare le proprie teorie in modo più rigoroso

Ancora secondo Parisi, le simulazioni possono essere definite come “un nuovo modo per esprimere le teorie scientifiche” (*ibidem*, p. 28). Una differenza che abbiamo individuato tra le scienze sociali e le scienze naturali riguarda, come si ricorderà, il fatto che, mentre le seconde, in genere, esprimono le proprie teorie in termini quantitativi, le prime esprimono le proprie teorie (quando lo fanno) in termini qualitativi, che sono sostanzialmente quelli del linguaggio comune. Questo rende probabilmente più difficile ricavare predizioni empiriche dalle teorie delle scienze sociali e impedisce, o quasi, la loro falsificazione.

Certamente, anche nelle scienze sociali esistono teorie che vengono sostituite da altre o vengono rifiutate dopo essere state ritenute corrette per un certo periodo. Ma, quando questo avviene, l’impressione è che si sia più perché le “nuove” teorie sono più evocative di quelle precedenti, o perché il linguaggio con cui sono espresse

risponde meglio alle esigenze accademiche o culturali in tempi mutati. Di nuovo, dopo Kuhn, Lakatos e Feyerabend, sarebbe idealistico (o ingenuo) pensare che questo non accada per le scienze naturali, ma l'argomento è lo stesso avanzato in precedenza: le scienze naturali si sforzano di trovare strumenti che mitighino queste tendenze.

Quali conseguenze potrebbero avere le simulazioni, da questo punto di vista, nell'ambito delle scienze sociali? In che senso una simulazione è un nuovo modo per esprimere una teoria? Quando si scrive una simulazione e la si vuole "far girare" su un computer non è possibile essere vaghi nei dettagli implementativi: semplicemente non funzionerebbe. Per questo motivo, il codice in cui è scritta una simulazione, o, per scopi comunicativi, una descrizione abbastanza dettagliata del suo funzionamento, costituisce un analogo di quello che sono le teorie espresse in termini quantitativi delle scienze naturali. Non solo non ci possono essere diverse interpretazioni di un codice, ma le predizioni empiriche che se ne possono ricavare sono, se possibile, ancora più dettagliate di quelle che forniscono le teorie espresse in modo quantitativo: le predizioni empiriche non sono nient'altro che i risultati della simulazione.

Naturalmente, come questi risultati debbano essere collegati ai dati reali è un ulteriore problema (del resto, lo stesso problema esiste nelle scienze naturali) e lo affronteremo alla fine di questo capitolo.

Sistemi complessi e interdisciplinarietà

L'approccio naturalistico qui difeso, con la sua enfasi sull'adozione di un'ontologia fisicalista per i fatti culturali e sull'utilizzo di metodologie ispirate alle scienze naturali, non propone, come spero sia chiaro, nessun tipo di riduzione dei fatti culturali ai fatti psicologici o biologici o, addirittura, chimici o fisici. Al contrario, l'attenzione è posta sull'integrazione dei livelli di analisi e sulla loro interdipendenza reciproca. Un fatto culturale può essere difficilmente compreso, a mio avviso, senza tenere in considerazione, per esempio, la biologia del sistema nervoso degli esseri umani, ma, allo stesso modo, nessuna analisi, per quanto completa, della biologia del sistema nervoso degli esseri umani, o dei loro meccanismi cognitivi, può costituire una spiegazione esaustiva di un fatto culturale.

Per chiarire meglio questo importante punto, a cui sono già stati fatti accenni

nel corso del capitolo, e a cui si ritornerà più volte in questa tesi, è possibile fare riferimento alla distinzione fra sistemi semplici e sistemi complessi¹.

Gli esseri umani tendono a pensare ai fenomeni che osservano come parte di sistemi semplici e allo stesso modo ha fatto la scienza nel corso della storia. E' molto probabile che non esista una distinzione ontologica, reale, tra sistemi semplici e sistemi complessi, ma che dipenda da quanto vogliamo andare in profondità nella nostra osservazione di questi sistemi. Per alcuni fenomeni, per esempio misurare la velocità di una sfera che rotola su un piano inclinato, è di solito sufficiente pensare in termini di sistemi semplici, ma per altri, tra i quali le società e le culture umane, è forse necessario pensare in termini di sistemi complessi.

Per capire cosa si intende con queste etichette è possibile fare un confronto tra alcune delle proprietà che caratterizzano le due classi di sistemi: (1) nei sistemi semplici i rapporti di causazione sono lineari, ossia, tornando al nostro caso della sfera e del piano inclinato, esiste una relazione lineare tra, per esempio, la forza con cui spingo la sfera all'inizio e la velocità di rotazione. Questo vuol dire che conoscendo il rapporto tra forza e velocità in un caso specifico potrò calcolare, per tutte le forze applicate, quale sarà la velocità di rotazione. Nei sistemi complessi, invece, i rapporti di causazione non sono lineari. Una perturbazione di un sistema complesso non è correlata in modo lineare all'effetto che avrà sul sistema stesso. Tipicamente, piccole perturbazioni possono dare luogo, a seconda dei diversi stati del sistema, ad effetti piccoli (lineari), ma anche ad effetti sproporzionatamente grandi oppure possono non avere per nulla effetti. Ancora, sono tipici dei sistemi complessi fenomeni di soglia, per cui esistono dei valori di soglia sotto i quali una perturbazione non ha nessun effetto e, poi, un trascurabile aumento della perturbazione produce grandi effetti. (2) Per i motivi esposti sopra, gli stati futuri dei sistemi semplici sono prevedibili usando, per esempio, dei modelli matematici, mentre gli stati futuri dei sistemi complessi sono un po' più difficili da prevedere. Non che non siano deterministici, ma le relazioni causali sono difficili da tracciare in un modello analitico. (3) Gli elementi che compongono i sistemi semplici tendono ad avere interazioni stabili tra loro, mentre gli elementi che compongono i sistemi complessi tendono a cambiare il tipo di interazioni a seconda dello stato del sistema e quindi

¹ Sulla scienza della complessità si possono vedere buoni volumi divulgativi come (Waldrop 1995, Gandolfi 1999) o le raccolte più specialistiche pubblicate dal Santa Fe Institute (per esempio Nadel & Stein 1991). Sconsigliabile, invece, almeno dal mio punto di vista, è (Taylor 2005).

della *storia* del sistema. Per questo, per valutare le “regole” che determinano il comportamento di un sistema complesso è spesso necessario non solo averne una comprensione in un determinato stato temporale, ma anche essere in grado di seguire (o di riprodurre) la sua evoluzione nel tempo.

Il fatto che, forse, per studiare i fatti sociali e culturali, sia necessario adottare la prospettiva dei sistemi complessi è un altro motivo che spiega la relativa diffidenza degli scienziati sociali nei confronti delle scienze naturali. I metodi tradizionali delle scienze naturali non sono adatti per studiare i sistemi complessi. E’ possibile che un approccio genuinamente interdisciplinare necessiti di strumenti metodologici differenti, come le simulazioni, per trattare questo tipo di sistemi. Le simulazioni possono essere utili per sviluppare una scienza interdisciplinare o non – disciplinare proprio perché permettono di studiare le interazioni tra differenti livelli di analisi in modo più efficace di quanto possono fare i modelli quantitativi “tradizionali”.

Un altro problema che ostacola lo sviluppo di una scienza interdisciplinare è il fatto che, ad oggi, le diverse discipline non solo si differenziano rispetto agli oggetti di studio ma anche, e forse in modo ancora più evidente, per i metodi con cui gli oggetti vengono studiati. Alcune discipline utilizzano osservazioni controllate, altre esperimenti di laboratorio, altre modelli astratti quantitativi, altre descrizioni verbali e via dicendo. Le simulazioni possono essere utili in questa direzione poiché, come abbiamo detto, integrano in un solo strumento alcune delle caratteristiche delle diverse metodologie tradizionali e potrebbero rappresentare una sorta di “lingua franca metodologica” attraverso cui ricercatori di varie discipline potrebbero comunicare tra loro.

L’interdisciplinarietà, tuttavia, si scontra con grandi problemi pratici, che hanno radici nella storia della scienza occidentale che ha, generalmente, enfatizzato la specializzazione dei saperi e favorito una rigida divisione interdisciplinare. Eppure, sembra proprio che nel caso dell’antropologia, il cui oggetto di studio pare il più ampio (o, a voler essere pessimisti, il peggio definito) fra le discipline che studiano il comportamento umano, l’adozione di una prospettiva interdisciplinare, o, almeno, un po’ più interdisciplinare, potrebbe essere decisamente rilevante.

Cosa ci dicono le simulazioni sulla realtà?

Sono, almeno in parte, fiducioso sul fatto che il lettore di questa tesi – il lettore benintenzionato a cui accennavo nell’Introduzione – possa trovare questa presentazione delle simulazioni come strumento metodologico, se non convincente, quanto meno plausibile. Ma, allo stesso modo, sono sicuro che, una volta arrivato alla parte sperimentale, e alle descrizioni delle simulazioni, lo stesso lettore sarà convinto di trovarsi di fronte ad un’incompatibilità insormontabile tra le pretese epistemologiche che sono state qui avanzate e la disarmante semplicità dei modelli descritti. Per questo motivo, voglio concludere questo capitolo esaminando un paio di critiche che possono essere rivolte alla metodologia simulativa e cercando di esporre in modo un po’ più dettagliato quale possa essere il rapporto tra i dati delle simulazioni e i fenomeni reali (su tutto ciò che segue mi sono ispirato a Parisi 2001 e a Di Paolo, Noble & Bullock 2000).

La prima critica è, appunto, quella dell’eccessiva semplificazione rispetto alla realtà. A mio avviso, questa critica, se interpretata letteralmente, non coglie nel segno. Se, come abbiamo detto, le simulazioni possono essere definite come un modo di esprimere le teorie alternativo a quelli tradizionali, allora, nella misura in cui riusciamo a rendere conto di un fenomeno complicato con una teoria semplice, tanto meglio. La semplicità di una teoria, nelle scienze naturali, viene considerata di solito una caratteristica positiva e non negativa, anche se, naturalmente, non è sulla base di questo che possiamo decidere l’accettazione o il rifiuto della stessa. Tuttavia, se due teorie spiegano lo stesso numero di fatti, ed una è più semplice dell’altra, allora si tenderà, in genere, a preferire quella più semplice. Inoltre, come è stato sottolineato più volte nel corso di questo capitolo, le teorie espresse tramite simulazioni sono, o possono essere, più ricche di particolari dei tradizionali modelli scientifici, per cui, semmai, da questo punto di vista, la critica dovrebbe essere opposta, ossia qualcosa come “le simulazioni sono troppo *complicate* per spiegare chiaramente un fenomeno”!

Assumo, quindi, che quando si afferma che le simulazioni sono modelli troppo semplificati si stia in realtà dubitando che in uno specifico modello si siano operate le corrette semplificazioni. Questa è, invece, una critica legittima, ed una critica che tocca quello che è, a mio avviso, uno dei maggiori problemi delle simulazioni allo stato attuale, ossia la mancanza di procedure condivise per valutare come i risultati ottenuti debbano essere rapportati ai fenomeni reali. In generale, chi fa simulazioni tende ad essere più interessato alla verifica interna delle proprie teorie che alla

verifica esterna (Parisi 2001), ossia utilizza le simulazioni per scoprire, partendo da un certo numero di assunzioni essenziali (il comportamento degli agenti, il tipo di ambiente, ecc.), quali risultati la simulazione produce. Si pensi ancora al modello di Schelling: nella descrizione che ne abbiamo fatto esso ci dice solo che pattern di segregazione spaziale possono emergere anche se i singoli agenti non hanno una particolare preferenza per stare vicini ad agenti dello stesso colore. La verifica interna di questa teoria consiste nell'utilizzare il "laboratorio sperimentale" per valutare la sua sensibilità ai parametri (sotto a quale soglia i pattern di segregazione spaziale non emergeranno più?), per modificare i parametri e vedere cosa accade (cosa accadrà se al posto di due "colori" ne abbiamo tre?) e per, eventualmente, modificare lo scenario (per esempio, si potrebbe immaginare di variare nel corso del tempo le preferenze degli agenti).

Valutare in modo rigoroso tutte queste possibilità è una parte fondamentale del lavoro di chi fa simulazioni, ma questo non basta. Fino a questo punto non sappiamo se il modello di Schelling ci sta dicendo qualcosa di corretto su un fenomeno reale o se nel modello sono state operate le semplificazioni corrette. Se noi fossimo sicuri che le semplificazioni operate sono corrette, e riuscissimo a dare un senso rigoroso a questo termine, allora non avrebbe molto senso, a mio avviso, dire che il modello è troppo semplice.

A proposito del rapporto fra risultati delle simulazioni e fenomeni empirici sono state avanzate, come è immaginabile, diverse ipotesi nel corso degli ultimi anni. Partirò dalle due ipotesi più "estremiste", che mi paiono entrambe poco plausibili e poi presenterò quelle più moderate, che mi sembrano più difendibili.

Le due ipotesi estremiste sono di segno opposto: da una parte, i sostenitori di quella che viene a volte definita *strong artificial life* (si veda Di Paolo *et al.* 2000), che si occupano per lo più di modellizzare sistemi biologici "a basso livello" (per esempio le cellule artificiali), ritengono che i risultati di una simulazione non siano un *modello* di un fenomeno reale, ma siano un'ulteriore *istanza* di quel fenomeno, che va ad aggiungersi ai dati empirici reali e come essi va trattata (in questi termini, non è quindi possibile parlare, propriamente, nemmeno di simulazione). Dall'altro lato vi sono coloro che ritengono che i risultati di una simulazione non abbiano alcun valore scientifico: dopo tutto, si tratta solo di simboli che un computer ha prodotto a partire da premesse che il ricercatore già conosceva e utilizzando regole anch'esse conosciute a priori.

Queste ipotesi, come dicevo, mi sembrano entrambe indifendibili: ritorniamo, ancora una volta, al modello di Schelling. Anche accettando che il modello di Schelling sia un'ulteriore istanza di un fenomeno reale di segregazione spaziale non è chiaro, per me, che tipo di aumento porti alle nostre conoscenze. Il punto da indagare è se il meccanismo con cui la segregazione spaziale viene realizzata – nel mio modello – possa rendere conto di *altre* istanze di quel fenomeno, che avvengono plausibilmente nel mondo reale. Per me non è molto importante stabilire se quello che ho realizzato in una simulazione sia un *vero* fenomeno o no, quello che mi importa è stabilire le relazioni tra quello e altri fenomeni, solo perché quello che ho realizzato nella simulazione riesco a conoscerlo e a controllarlo meglio di quelli che osservo nel mondo reale.

Dovrebbe essere chiaro, dopo quello che è stato detto nel corso del capitolo, perché la soluzione opposta non mi sembra neppure convincente. Innanzitutto simulazioni anche molto semplici producono, a volte, come avviene nel modello di Schelling, risultati abbastanza sorprendenti: è vero che il ricercatore conosce le premesse e le regole con cui la simulazione lavora, ma questo non significa che conosca a priori anche i risultati. Inoltre, e questo è il punto più importante, se teniamo ancora presente l'idea delle simulazioni come mezzo per esprimere una teoria, quello che importa è il *modo* in cui i meccanismi sono specificati e quali sono i risultati, non se il ricercatore li conosceva o no a priori. Altrimenti, sarebbe come affermare che qualsiasi legge scientifica non ha valore in quanto deterministica, ma è proprio questo che in genere si chiede ad una legge scientifica. La legge di gravitazione universale, una volta formulata, non ci dice nulla di nuovo sul fatto che gli oggetti cadono verso la terra oppure no, ma ci propone un'ipotesi sul *perché* essi cadono.

A questo punto si può passare alla parte costruttiva: quanto realismo dobbiamo chiedere ad una simulazione, quanta complessità? A questa domanda purtroppo non c'è una risposta precisa, ma il suggerimento che si può avanzare è che sia necessario un continuo “movimento” tra la simulazione e il fenomeno empirico che con essa vogliamo catturare e un continuo riaggiustamento della simulazione (e quindi, della teoria che spiega il fenomeno) conseguente a questo movimento.

Di Paolo *et al.* (2000) hanno definito i modelli simulativi “esperimenti di pensiero opachi”. Questa analogia mi sembra interessante: una simulazione è, almeno all'inizio, qualcosa di analogo ad un esperimento di pensiero, ossia è un artificio che

ci aiuta, tramite la creazione di scenari ipotetici, a comprendere i fatti reali. Tuttavia, una simulazione diventa, man mano che procede il lavoro, qualcosa di più. Innanzitutto, quando scriviamo la simulazione, siamo costretti ad essere estremamente espliciti, cosa che non è per forza necessaria nel caso degli esperimenti di pensiero. Inoltre, come lascia intendere il termine “opachi”, mentre nel caso degli esperimenti di pensiero i postulati iniziali, i meccanismi e, soprattutto, i risultati devono essere nella mente di chi li formula, nelle simulazioni, come abbiamo visto nel caso del modello di Schelling, possiamo attenderci risultati inaspettati.

Ma, a questo punto, cosa che non sempre viene fatta, è necessario tornare al fenomeno che vogliamo studiare. Come minimo possiamo chiederci se la nostra simulazione ci fornisce nuove intuizioni sui fenomeni empirici che, semplicemente osservando la realtà, non avremmo colto. Ma possiamo farci domande più precise: variando i parametri della simulazione osserviamo delle modificazioni nei nostri risultati che hanno qualche corrispondenza con dinamiche reali? Possiamo pensare ad esperimenti (reali) che ci permettano di verificare se le semplificazioni operate nei nostri modelli sono quelle corrette? Al di là del realismo della simulazione, e della sua complessità, l'utilizzo di questi modelli ci dovrebbe permettere, almeno, di guardare ai fenomeni reali con occhi diversi e, guardando i fenomeni reali con occhi diversi, tornare alla simulazione aggiungendo o togliendo particolari, modificando parametri o pensare a nuove simulazioni per, poi, tornare di nuovo alla realtà, in un circolo potenzialmente senza un termine ultimo.

Possiamo pensare ai modelli simulativi come ad uno strumento finalizzato primariamente a rispondere a specifiche questioni, e a fare questo rispettando certi vincoli: questi vincoli sono soprattutto legati alla quantità di informazione che possediamo rispetto alla questione che vogliamo risolvere o, detto in altro modo, a quanta energia e tempo siamo disposti ad utilizzare per raccogliere informazioni. Grimm e Railsback (2005), introducendo le simulazioni *individual – based* in campo ecologico, invitano il lettore a pensare ad un esempio quotidiano quale la scelta della fila alle casse di un supermercato. Il “modello” più veloce e meno costoso dal punto di vista informativo sarebbe semplicemente quello di scegliere a caso, tuttavia questo non si rivela, in generale, molto efficace. Un modello alternativo è quello di considerare la lunghezza delle code: possiamo ipotizzare che a code più corte corrispondano tempi di attesa minori. Questo modello è anch'esso poco costoso, ma più efficace del modello casuale. Naturalmente, con più informazioni anche questo

modello potrebbe rivelarsi troppo semplicistico: potremmo volere conoscere l'esperienza dei cassieri, o la loro abilità, o controllare se a file lunghe corrisponda in effetti più merce da pagare, ecc. Il punto da sottolineare è il seguente: come possiamo dire quale tra questi è il modello *corretto*? La risposta non è univoca: dipende.

Quello che si può ricavare da questa analogia è, comunque, che il mero realismo, o la complessità, di una simulazione non sono un metro valido, o almeno non sempre, per valutarne la bontà. Più che vicini al sistema che si vuole simulare bisognerebbe avere chiare le questioni che si vogliono affrontare ed utilizzare queste come guida per decidere cosa inserire o no nel proprio modello.

2. Quali simulazioni?

Nel capitolo precedente ho presentato, in termini molto generali e personali, l'idea della simulazione come metodologia scientifica, soffermandomi in particolare sulle differenze che intercorrono tra simulazioni e metodologie scientifiche tradizionali e sulle implicazioni che l'uso delle simulazioni può avere per quegli oggetti di analisi che mi sembrano riguardare più da vicino l'antropologia, o quantomeno le scienze sociali. E' stata evidenziata la distinzione tra simulazioni in senso generale e simulazioni ad agenti: in questo capitolo cercherò di entrare nel dettaglio e di illustrare una particolare "sottoclasse" delle simulazioni ad agenti, a cui appartengono quelle presentate nella parte sperimentale di questa tesi.

Per comprendere le motivazioni che sottostanno alla scelta di alcune caratteristiche distintive di queste simulazioni sarà necessario ripercorrere alcune tappe fondamentali del cammino storico e concettuale della scienza cognitiva. La scienza cognitiva, come programma di ricerca unitario e definito, nasce negli Stati Uniti intorno alla metà del secolo scorso, in contrapposizione al programma di ricerca allora dominante nella psicologia d'oltreoceano, il comportamentismo. In uno slogan, mentre per i comportamentisti una psicologia scientifica doveva fare a meno di qualsiasi riferimento ai processi mentali, in quanto entità non osservabili e non misurabili, per i sostenitori della nascente scienza cognitiva la descrizione del comportamento umano non poteva evitare di confrontarsi con essi. Naturalmente i comportamentisti non erano così ingenui da sostenere che i processi mentali non fossero realmente importanti per comprendere il comportamento umano: più

pragmaticamente essi avevano constatato l'enorme differenza tra lo sviluppo e i progressi delle discipline scientifiche come la fisica, la biologia, la chimica (che utilizzavano un rigoroso paradigma sperimentale ed esprimevano le proprie teorie in termini quantitativi) e la psicologia, che invece sembrava, ai loro occhi, continuare a riproporre gli stessi problemi e a porsi le stesse domande senza dare l'impressione di un'effettiva accumulazione di conoscenze. Il loro intento era quello di fornire alla psicologia strumenti simili a quelli delle altre discipline scientifiche: da questo punto di vista, la scienza cognitiva può essere vista non tanto in contrapposizione, quanto in continuità rispetto al comportamentismo (si veda Dennett 1991, in particolare la seconda parte del volume, intitolata "Le caratteristiche della teoria in psicologia"). Il problema, per i comportamentisti, era che, fino agli anni '50 del secolo scorso, nessuno aveva nemmeno una vaga idea di come ciò avrebbe potuto essere fatto per i processi di pensiero.

Alle origini: macchine di Turing e meccanismi di retroazione

In maniera molto interessante, dal nostro punto di vista, quando le idee arrivarono, non fu dalle altre scienze già "affermate", ma da un modo radicalmente nuovo di pensare i processi mentali, che vedeva nella *modellizzazione* in altri artefatti di alcune loro caratteristiche, che venivano ritenute definitorie, la soluzione al problema della trattabilità scientifica del pensiero umano. In particolare, ci soffermeremo su due caratteristiche definitorie, che ritengo importanti per lo sviluppo di questo capitolo. Una è l'idea che il pensare sia, fondamentalmente, compiere operazioni su simboli astratti (idea che costituisce il paradigma ancora oggi dominante della scienza cognitiva) e che venne operazionalizzata dal logico britannico Alan Turing, che descrisse un meccanismo astratto (appunto, la macchina di Turing) che poteva compiere queste operazioni seguendo le leggi note della fisica (Turing 1936). L'altra è l'idea che un organismo "intelligente" sia caratterizzato dalla possibilità di autoregolarsi e debba perciò essere fornito di un sistema di controllo che gli permetta un comportamento modificabile in relazione alle mutazioni dell'ambiente in cui vive. Questa idea fu sviluppata in particolare da un gruppo di studiosi (di cui faceva parte anche Gregory Bateson) e operazionalizzata nel meccanismo di *retroazione*. Questi

studiosi fondarono una disciplina chiamata cibernetica, che ebbe, all'inizio, meno influenza dell'opzione di Turing, benchè le due non siano necessariamente in contraddizione: oggi le istanze proposte dalla cibernetica sono accolte con molto più favore e questo lavoro, come si vedrà nel corso del capitolo, risente probabilmente più di queste che di quelle proposte della scienza cognitiva "classica".

Il fatto che queste idee possano risultare oggi, per qualcuno, non particolarmente sorprendenti, mi sembra un'ulteriore prova del loro successo: meccanismi di retroazione (per esempio i termostati) e macchine di Turing (naturalmente, i computer) fanno parte della nostra vita quotidiana e hanno cambiato radicalmente il nostro modo (o almeno, il modo di alcuni di noi) di pensare ai processi cognitivi, rendono tutto sommato normale l'idea che, se essi possono essere istanziati in artefatti, allora sono solo un tipo diverso, forse estremamente complicato, di processo fisico, tanto che non è facile avere la consapevolezza di cosa queste idee potessero significare allora.

Iniziamo con la cibernetica. Il termine, nel suo senso moderno, fu coniato nel 1947 da Norbert Wiener (Wiener 1982). Wiener aveva lavorato, durante la seconda guerra mondiale, alla progettazione di un meccanismo automatico di puntamento per artiglieria antiaerea, rendendosi conto che, perché questo fosse efficace, era necessario che modificasse continuamente i propri parametri in relazione al comportamento dell'aereo da colpire. Ossia, non era sufficiente che il dispositivo di puntamento fosse *preciso*, ma anche che si adattasse alle condizioni mutevoli dell'ambiente: nel caso specifico, alla velocità e alla traiettoria dell'obiettivo. Wiener, insieme ad altri studiosi, ipotizzò che questo tipo di sistemi, caratterizzati dalla loro flessibilità e dal fatto che il loro comportamento non fosse comprensibile analizzando il congegno *in sé*, ma in relazione al compito per cui era progettato e alle modificazioni dell'ambiente, fossero diffusi nel mondo naturale e, in particolare, potessero rendere conto di alcune caratteristiche dei processi cognitivi umani.

Il meccanismo di base con cui lavorano questi sistemi viene detto *retroazione negativa*: l'idea è che ci sia un rapporto inversamente proporzionale tra gli effetti che il dispositivo produce nel sistema e la sua attivazione, di modo che il sistema tende a stabilizzarsi. Si pensi ad un termostato che deve scaldare un ambiente freddo. Il termostato deve mantenere una determinata temperatura nell'ambiente; per fare questo monitora la temperatura effettiva e, più questa è distante da quella desiderata, più attiva il sistema di riscaldamento. Tuttavia, quando la temperatura si avvicina a

quella desiderata, il termostato tenderà a *diminuire* la propria attivazione, mantenendo stabile, nei limiti del possibile, il sistema.

Un esempio celebre e ancora più semplice di meccanismo autoregolante (anche se un po' più difficile da spiegare) è quello del regolatore di Watt (lo cito poiché il regolatore di Watt è diventato recentemente l'analogia preferita dei processi cognitivi in quella che viene definita *ipotesi dinamicista* nelle scienze cognitive, si veda van Gelder 1998). Il regolatore di Watt ha lo stesso effetto di un termostato ma, a differenza del termostato, non fa uso di rappresentazioni esplicite (la temperatura) per mantenere il sistema stabile. Eccone una descrizione schematica, ma utile per i nostri scopi: un regolatore di Watt è composto da un fuso centrale, che ruota su se stesso, alimentato da una macchina a vapore: lo scopo del sistema è mantenere costante la velocità di rotazione di questo fuso. Per fare questo, all'estremità superiore del fuso è collegata, attraverso un braccio mobile, una sfera; la distanza della sfera dall'asse di rotazione del fuso è determinata dalla forza centrifuga, ossia dalla velocità di rotazione: più il fuso ruota velocemente, più la sfera tenderà ad allontanarsi dal fuso. Se, come avviene nel regolatore di Watt, il movimento del braccio controlla l'apertura della valvola della macchina a vapore, siamo ancora in presenza di un meccanismo di retroazione negativa: più il fuso ruota velocemente, più la sfera tenderà ad allontanarsi e a chiudere l'apertura della valvola. In questo modo, il fuso ruoterà più lentamente, di modo che la sfera tenderà a riavvicinarsi aprendo così di più la valvola e via dicendo. Così, il sistema si mantiene stabile nel tempo.

Molte delle idee che sono implicite in questi meccanismi autoregolanti (l'adattabilità, lo stretto rapporto con le condizioni ambientali, l'assenza di un controllo esplicito – il regolatore di Watt non ha bisogno di una regola che gli “dica” di mantenere la velocità del fuso costante) saranno, come vedremo, fondamentali per comprendere le caratteristiche delle simulazioni che verranno presentate in seguito.

Ma, come si diceva, il modello che all'inizio ebbe maggiore influenza per gli sviluppi della scienza cognitiva fu la macchina di Turing (d'ora in poi MT). Pur essendo, nella formulazione originale di Turing, un congegno meramente astratto e caratterizzato puramente da proprietà logiche e formali (inoltre, Turing, almeno all'inizio, propose la sua idea di MT senza alcun rapporto con la formalizzazione dei processi cognitivi, ma in risposta ad un problema matematico sollevato da David Hilbert), se ne fa di solito una descrizione più prosaica e materiale che, in effetti, rende semplice capire di cosa si tratta, per cui mi atterro volentieri all'uso comune.

Una MT può essere rappresentata da un sistema composto da un nastro e da un dispositivo di lettura/scrittura. Sul nastro ci sono una serie di caselle nelle quali sono presenti dei simboli, che fanno parte di un alfabeto ben definito (e conosciuto a priori dalla macchina). Il dispositivo di lettura/scrittura si trova, ad un tempo t , in corrispondenza di una casella: può decodificare il simbolo presente su quella casella e, a seconda di un insieme di regole predefinito, valutare il proprio stato s e produrre eventualmente come risposta una delle seguenti azioni: muoversi a destra o a sinistra sul nastro o scrivere un nuovo simbolo nella casella in cui si trova attualmente. Dati un insieme di regole finito ed un insieme di stati possibili, anch'esso finito, una MT è in grado di risolvere qualsiasi problema computabile (tesi di Church-Turing).

Dal punto di vista pratico, la MT portò in breve alla creazione dei calcolatori digitali (che sono, tutti, istanze concrete della MT astratta) ma, per quello che ci interessa, dal punto di vista concettuale, la MT contribuì in maniera determinante a fare sì che si iniziasse a prendere sul serio l'idea che il pensiero stesso fosse una forma di computazione su simboli. Questa idea è senza dubbio l'idea per eccellenza intorno alla quale si costituì la scienza cognitiva. Il problema del comportamentismo era superato: era possibile finalmente, secondo l'ottica dei primi scienziati cognitivi, trattare il pensiero stesso come un fatto fisico, quantificabile, soggetto alle leggi e alle regole che valevano per tutti gli altri fenomeni studiati dalle scienze. Sulla base di questa considerazione nacque l'intelligenza artificiale (oggi definita intelligenza artificiale "classica", o *GOFAI*, *good old-fashioned artificial intelligence*) e, parallelamente, si sviluppò una corrente di filosofia della mente estremamente influente per gli sviluppi della scienza cognitiva, detta funzionalismo, secondo cui gli stati mentali sono caratterizzati dalla loro *funzione* rispetto ad altri stati mentali, indipendentemente dal substrato fisico in cui sono realizzati.

Così, in estrema sintesi, il programma di ricerca che caratterizzò la scienza cognitiva degli inizi, influenzato dall'idea di MT e dal funzionalismo, in filosofia, si basava su alcune idee fondamentali: (1) che il pensiero fosse una forma di computazione, (2) che questa computazione venisse effettuata con regole esplicite su simboli espliciti e (3) che l'implementazione fisica dei processi cognitivi non avesse particolare importanza. Come vedremo, tutte queste idee sono state messe in discussione nel corso degli anni successivi (anche se in modo differente: tutti, o quasi, oggi sarebbero d'accordo nel sostenere l'importanza dell'implementazione fisica dei processi cognitivi; l'approccio simbolico è al centro di un dibattito che divide in modo

abbastanza equo la comunità degli scienziati cognitivi, mentre l'idea che il pensiero possa essere produttivamente considerato, almeno ad un certo livello di analisi, come una forma di computazione è accettata da quasi tutti – tranne che dai sostenitori del già citato approccio dinamicista alla cognizione) e passare in rassegna le critiche che sono state rivolte a queste posizioni ci servirà per presentare, gradualmente, gli “strumenti di lavoro” che verranno poi utilizzati nelle simulazioni.

Realismo biologico: reti neurali (1)

Secondo filosofi come Paul Churchland, l'idea che i processi cognitivi possano essere considerati computazioni con regole esplicite su simboli espliciti è semplicemente una trasposizione dell'idea di senso comune che stati mentali espliciti, quali credenze, desideri, e via dicendo, influiscano sulle azioni degli esseri razionali. In un articolo significativamente intitolato *Eliminative Materialism* (Churchland 1981) Churchland mise l'accento soprattutto sulla scarsa, o assente, integrazione che il paradigma classico della scienza cognitiva aveva con i crescenti dati che si possedevano, a livello neurofisiologico, sul funzionamento del cervello.

In effetti, i progressi della neuroscienza (e i paralleli insuccessi dell'intelligenza artificiale classica che, dopo un periodo di grande entusiasmo iniziale, sembrava non essere in grado di riprodurre la maggior parte dei comportamenti caratteristici degli esseri umani, anche se alcuni lavori erano stati considerati notevoli successi, per esempio Marr 1982) iniziarono, a partire dagli anni '80 del secolo scorso, a fare dubitare molti almeno dell'assioma riguardante la non importanza dell'implementazione fisica dei processi cognitivi.

Uno strumento alternativo ai sistemi computazionali simbolici fu identificato nelle reti neurali artificiali. La maggior parte dei concetti di base su cui le reti neurali si fondano erano già stati sviluppati da alcuni decenni (vedi per esempio McCulloch & Pitts 1943), ma si dovette attendere la metà degli anni '80 perché iniziarono a diffondersi fra gli studiosi di scienze cognitive, in relazione alla pubblicazione di *PDP: Parallel Distributed Processing* (Rumelhart & McClelland 1991), che, oltre a risolvere un fondamentale problema di implementazione (veniva perfezionato l'algoritmo di apprendimento supervisionato di *backpropagation*, che era possibile

applicare a reti con strati interni, mentre prima di allora non si era ancora trovato il modo di fare “apprendere” una rete con unità interne – sull’importanza degli algoritmi di apprendimento per le reti neurali vedi il paragrafo successivo), presentava una serie di lavori in cui i modelli a reti neurali venivano applicati a specifici problemi affrontati dalla psicologia cognitiva (linguaggio, percezione, ecc.).

Una rete neurale è un modello estremamente semplificato del cervello che si basa su alcune analogie con il funzionamento di neuroni e sinapsi. Per capire come lavora possiamo considerare una rete molto semplice ed analizzare il funzionamento delle sue componenti. Una rete neurale è composta da unità (analoghe ai neuroni) e connessioni fra le unità (analoghe alle sinapsi).

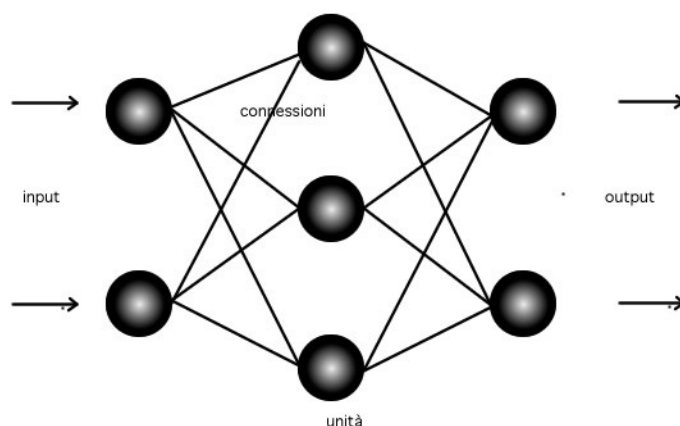


Fig. 2.1: Rappresentazione schematica di una semplice rete neurale.

Le unità di una rete neurale possono essere distinte in unità di input, unità interne e unità di output. Seguendo le frecce della figura 2.1, le due unità di input (a sinistra) ricevono dei valori numerici che provengono dall’ “esterno” della rete. Questi valori possono essere pensati come particolari eventi o particolari condizioni chimico – fisiche dell’ambiente (o, eventualmente, se pensiamo ad una struttura complessa – e probabilmente massivamente modulare – come il cervello, come output di altre reti neurali). Prendiamo come esempio la percezione visiva: la retina riveste la parte posteriore del globo oculare ed è composta da circa un milione di fotorecettori, ossia neuroni sensibili alla differente lunghezza d’onda della luce. Essi possono essere realmente considerati come unità di input di una – estremamente complessa – rete neurale: ogni fotorecettore ha un’attivazione (che può essere rappresentata come una

quantità numerica continua) e questa attivazione si propaga in modo parallelo nella rete, alle unità interne. Nella propagazione, l'attivazione delle unità di input viene modificata dai pesi delle connessioni e risulta nell'attivazione delle unità interne, che, a loro volta, la trasmettono, modificandola ancora, alle unità di output (naturalmente in un sistema biologico realistico gli "strati" di unità interne sono molti). Gli output, ancora, producono dei valori numerici, che questa volta potranno essere interpretati come segnali chimici (per esempio attivazioni muscolari) o, ancora, come input per altre reti neurali.

Quindi, in sostanza, una rete neurale si limita anch'essa ad effettuare una *computazione*, ma ad un livello logico più basso di quello dei simboli formali e, soprattutto, distribuendola tra unità semi-autonome. Oltre a dimostrarsi particolarmente adatte per riprodurre alcune caratteristiche tipiche dei processi cognitivi (per esempio: la robustezza rispetto a stimoli rumorosi o a danneggiamenti locali; la buona capacità di effettuare generalizzazioni; la possibilità di trattare le categorie in modo "fluidico", come può avvenire negli esseri umani, ecc.), le reti neurali avevano un'altra caratteristica di fondamentale importanza per il nostro discorso.

Mentre i sistemi computazionali dell'intelligenza artificiale classica venivano programmati dallo sperimentatore, che cercava di riprodurre, attraverso il suo codice, un comportamento cognitivo analogo a quello che era interessato a studiare, le reti neurali, semplicemente, non possono essere programmate dallo sperimentatore. La strategia tipica dell'intelligenza artificiale classica consisteva nell'analizzare un problema, per come si credeva fosse risolto dai soggetti umani, scomporlo in unità minime, che potessero essere implementate in un linguaggio formale e vedere se il programma in questione si comportava effettivamente come un agente cognitivo razionale. Nelle reti neurali, si può eventualmente usare una strategia simile per decidere quale architettura utilizzare (quante unità, che tipo di connessioni, che funzioni di attivazione delle unità utilizzare) e per trovare il modo migliore di codificare gli input e di decodificare gli output (anche se vedremo nei paragrafi successivi che anche queste scelte possono essere "lasciate al sistema"), ma il comportamento risultante è determinato dai pesi delle connessioni e questi, anche volendo, non si possono decidere a priori. Non è possibile, cioè, a partire da un comportamento desiderato (un *mapping* input – output) stabilire quali pesi lo possano realizzare.

Questa strategia (generalmente definita *bottom – up* in contrapposizione alla

strategia *top – down* tipica dell'intelligenza artificiale classica) non solo si rivelava vicina all'attenzione, sorta in concomitanza con lo studio dei sistemi complessi (e dei sistemi che si autoregolano, come quelli della cibernetica), verso i fenomeni emergenti, verso i sistemi che, pur mostrando una finalità “intelligente” a livello macro, sono composti, a livello micro, da unità che, prese singolarmente, non possono rendere conto di quella finalità, ma anche permetteva di reintrodurre, per così dire, dalla porta principale, una caratteristica del comportamento umano che era stata almeno in parte trascurata nella *GOFAI*, ossia il fatto che esso fosse determinato da processi dinamici di ontogenesi (apprendimento) e filogenesi (evoluzione).

Ontogenesi: reti neurali (2)

Molte importanti correnti della psicologia, come quella piagetiana e quella vygotskijana, per non parlare dell'antropologia culturale e cognitiva *mainstream*, rimasero fredde, se non decisamente ostili, nei confronti del messaggio che la scienza cognitiva stava iniziando a diffondere con sempre più convinzione. Nel caso dell'antropologia, che è quello che conosco meglio, le ragioni furono sicuramente molteplici, ma non poca importanza ebbe l'impressione che la scienza cognitiva relegasse un ruolo secondario all'apprendimento e, anzi, avesse riportato in vita il concetto di innatismo, che gli antropologi pensavano di avere definitivamente allontanato dal “salotto buono” delle scienze umane. La mia opinione, a questo proposito, è che molti degli antropologi commisero – e tuttora continuano a commettere – un errore duplice.

Da una parte, è probabile che gli scienziati cognitivi abbiano, in effetti, ragione nel sostenere che la plasticità del comportamento umano debba essere spiegata con la complessità dei meccanismi cognitivi evoluti per selezione naturale e che almeno alcuni di questi meccanismi siano uguali, o analoghi, in tutti gli esseri umani. Questa ipotesi, inoltre, giusta o sbagliata che sia (intendo da un punto di vista metodologico – ossia: *quanto* dobbiamo tenere conto di questi meccanismi nella spiegazione dei fatti culturali? – perché ho davvero pochi dubbi, per usare un eufemismo, che sia sbagliata nel senso che l'evoluzione non abbia influito sulla struttura del cervello degli esseri umani), ha generato una serie di lavori che trattano

problemi “classici” dell’antropologia, dalla religione, alla folkbiology, alla categorizzazione sociale, alle relazioni interetniche e via dicendo, il cui interesse, che si sia d’accordo o meno con i loro presupposti, è difficilmente negabile (si veda anche Acerbi 2003, 2005, 2006).

Dall’altra, la scienza cognitiva, sicuramente, non è solo innatismo. La scienza cognitiva, nelle sue diverse e a volte confuse forme, ha contribuito allo studio del comportamento umano con un approccio genuinamente interdisciplinare (per quanto possibile), ha formulato teorie dettagliate e spesso rivoluzionarie sui processi cognitivi, ha ripreso temi della biologia evoluzionistica, della scienza della complessità, dalla neurobiologia: che gli antropologi, o la maggior parte di essi, studino gli esseri umani come se nulla fosse successo, mi sembra un impoverimento particolarmente gravoso per la nostra disciplina. Come si diceva, poi, non tutta la scienza cognitiva è innatista (e, comunque, gli “innatisti”, come per esempio gli psicologi evoluzionistici, non propongono quelle caricature di teorie che spesso gli vengono affibiate): i ricercatori che, all’inizio, utilizzavano le reti neurali, muovevano una critica all’intelligenza artificiale classica del tutto analoga a quella che l’antropologia muove – o muoveva – alla scienza cognitiva. La vita artificiale, come vedremo meglio nel prossimo paragrafo, ha un approccio diverso: si ritiene che per studiare il comportamento sia necessario tenere conto *sia* dell’apprendimento, individuale e sociale, nel corso della vita, *sia* delle forze dell’evoluzione naturale (Nolfi, Elman & Parisi 1994).

Chi non sarebbe d’accordo, del resto, con questa opzione? Quello che conta è *come* questo viene fatto. Nella vita artificiale entrambi i processi (apprendimento ed evoluzione) sono visti come complementari e da una prospettiva scevra di tutte quelle sovrastrutture concettuali che rendono il dualismo *nature/nurture* così caro alla scienza divulgativa, ma anche al dibattito accademico. Per capirci, ecco come presentano il problema in un articolo di review Nolfi e Floreano (Nolfi & Floreano 1999, p. 89):

Evolution and learning are two forms of biological adaptation that differ in space and time. Evolution is a process of selective reproduction and substitution based on the existence of a geographically distributed population of individuals displaying some variability. Learning, instead, is a set of modifications taking place within each single individual during its own lifetime. Evolution and

learning operate at different time scales. Evolution is a form of adaptation capable of capturing relatively slow environmental changes that might encompass several generation [...] Learning, instead, allows an individual to adapt to environmental changes that are unpredictable at the generational level. Learning might include a variety of mechanisms that produce adaptive changes in an individual during its lifetime, such as physical development, neural maturation, and synaptic plasticity. Finally, whereas evolution operates on the genotype, learning affect only the phenotype and phenotypic changes cannot directly modify the genotype.

Ma, cosa significa imparare, per una reti neurale? Come si è detto, il comportamento di una rete deriva, in prima approssimazione, dal valore dei pesi delle connessioni. Cambiando il valore dei pesi delle connessioni, cambia il *mapping* tra input ed output effettuato dalla rete, ossia la rete risponde in modo diverso agli stessi stimoli. Quindi, per una rete neurale, apprendere significa modificare il valore dei pesi delle connessioni. Poiché, come anticipato, i valori non possono essere decisi a priori dallo sperimentatore, è parte integrante dell'indagine connessionista¹ la ricerca di algoritmi di apprendimento che permettano alla rete di “scoprire” il corretto insieme di pesi per risolvere un dato problema. Nei capitoli successivi, dedicati alle simulazioni, vedremo quali sono, e come operano, alcuni di questi algoritmi.

Filogenesi: algoritmi genetici

Come possiamo simulare un processo analogo all'evoluzione naturale nel contesto di una simulazione con reti neurali? Tornando alla citazione riportata in precedenza, possiamo affermare che, perché un processo evolutivo possa avere luogo, sono necessarie alcune condizioni, in particolare, la presenza di una *popolazione* che esibisca una qualche forma di variabilità, una dimensione temporale ampia, che possa esprimersi in *generazioni* di individui, l'esistenza di un *genotipo*, attraverso il quale il

¹ Sono riuscito a parlare per queste pagine di reti neurali senza usare il termine “connessionismo”? Eccone la definizione – dalla Stanford Encyclopedia of Philosophy (<http://plato.stanford.edu/entries/connectionism>): “connectionism is a movement in cognitive science which *hopes* to explain human intellectual abilities using artificial neural networks” (corsivo mio).

meccanismo evolutivo possa operare. Nel nostro caso la popolazione è una popolazione composta da individui che sono singole reti neurali e il genotipo di una rete neurale può essere immaginato come una stringa che contiene i valori dei pesi delle connessioni.

Il meccanismo di selezione ed evoluzione naturale viene realizzato per mezzo dei cosiddetti algoritmi genetici (Holland 1975, Mitchell 1996). Un algoritmo genetico opera, in sostanza, con gli stessi principi essenziali del reale processo evolutivo (vedi Figura 2.2). Partiamo da una popolazione di reti neurali con i pesi inizializzati in modo casuale, che devono risolvere un certo compito e testiamo le reti. Con i pesi casuali è molto probabile che nessuna delle reti avrà una prestazione soddisfacente (in analogia con l'evoluzione biologica, la misura della prestazione di un individuo viene detta, generalmente, *fitness*), tuttavia alcune reti saranno un po' meglio di altre. Dopo avere testato tutte le reti, si passa alla seconda generazione. Le reti della prima generazione vengono ordinate secondo la loro prestazione e le migliori vengono fatte riprodurre. La seconda generazione sarà quindi composta da duplicati dei migliori individui (questo corrisponde alla selezione differenziale del reale processo evolutivo), tuttavia, come nella realtà, il genoma delle reti "figlie" non sarà esattamente uguale a quello delle migliori reti della prima generazione: il loro genoma (i pesi delle loro connessioni) verranno mutati attraverso un qualche processo stocastico (il modo più semplice è di avere una bassa probabilità di sostituire ogni peso con un nuovo valore casuale). Iterando questo processo, la popolazione di reti neurali si potrà evolvere in direzione di individui che hanno buone prestazioni rispetto ad un compito specificato.

Naturalmente, un realistico ed efficace algoritmo genetico è a volte molto più complesso: i metodi di selezione degli individui da fare riprodurre possono essere più raffinati (per esempio è spesso utile fare riprodurre, con una certa probabilità, anche individui che non hanno una *fitness* relativamente alta, in modo da assicurare una maggiore variabilità alla popolazione e quindi uno spazio di ricerca più ampio per l'algoritmo), la riproduzione può essere sessuata (permettendo così anche il *crossover* genetico), le mutazioni possono agire in molti modi differenti e via dicendo. Ciò che importa qui è comunque presentare uno schema generale del modo in cui un algoritmo genetico funziona.

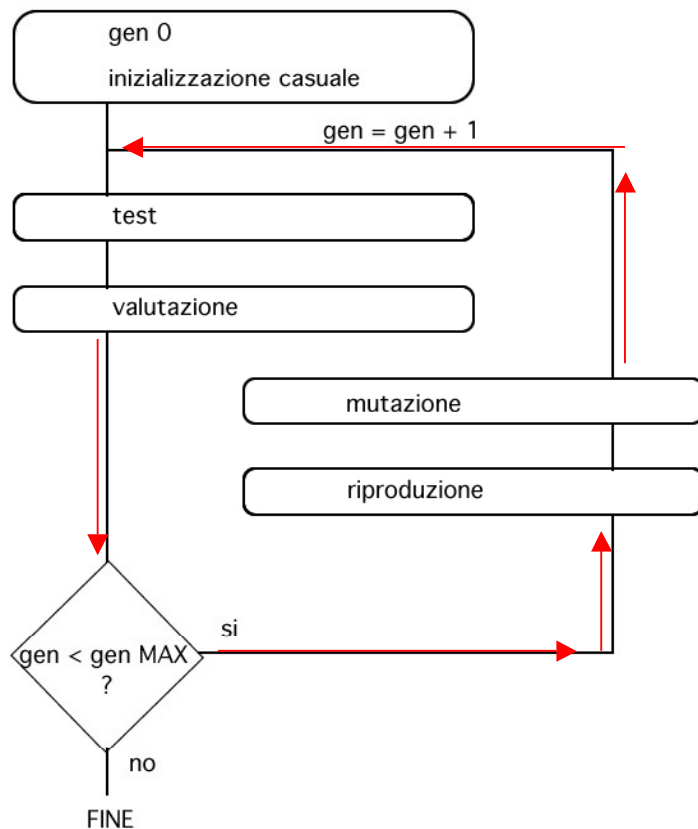


Fig. 2.2: Modello del funzionamento di base di un algoritmo genetico.

Da un punto di vista ingegneristico, gli algoritmi genetici rendono possibile esplorare un insieme più grande di potenziali soluzioni ad un dato problema rispetto a quanto possa fare un algoritmo convenzionale (ed un programmatore – umano – convenzionale) e, quindi, anche se fosse solo per questo motivo, sarebbero uno strumento conveniente da usare per programmare reti neurali per risolvere determinati compiti. Tuttavia, quello che qui interessa di più sottolineare sono altri due punti concettuali. Gli algoritmi genetici rappresentano, come si è già detto per le reti neurali, un ulteriore passo nella direzione della realizzazione di processi e meccanismi complessi, teleologici (che possiedono una finalità), attraverso processi e meccanismi semplici e privi di teleologia. Da un altro punto di vista, inoltre, pur nella loro astrattezza e a prescindere dall'ovvia necessità di semplificazione, essi si sono dimostrati un buon modello del *reale* meccanismo darwiniano di selezione ed evoluzione naturale. In particolare, solo l'uso degli algoritmi genetici e delle

simulazioni al computer ha permesso, negli ultimi decenni, di apprezzare la complessità delle dinamiche evolutive. L'evoluzione naturale è a volte considerata come un processo semplice, lineare, in cui gli organismi tendono in modo intrinseco verso le soluzioni più adattive. Gli algoritmi genetici mostrano come, invece, a partire dalla disarmante semplicità dei meccanismi darwiniani (variabilità, selezione, riproduzione) sia possibile dare origine a dinamiche estremamente complesse, in cui, per esempio, il caso ha un ruolo determinante, in cui minimi locali e soluzioni subottimali sono importanti quanto le soluzioni più adattive e così via.

Una critica immediata che può essere mossa all'uso degli algoritmi genetici con popolazioni di reti neurali è che, comunque, il compito da effettuare, e quindi la *fitness* degli individui siano decisi a priori dallo sperimentatore, dando una sensazione di circolarità. Questo è vero, ma, come vedremo nei prossimi paragrafi, aggiungendo ancora un po' di ingredienti per giungere a qualcosa di simile ad una simulazione di vita artificiale, all'aumentare della complessità degli individui simulati e dell'ambiente, è sempre più difficile per lo sperimentatore sapere, prima di effettuare gli esperimenti, in che modo un particolare compito verrà risolto e se verrà risolto. Uno dei parametri con cui valutare una simulazione, anche se molto difficile da formalizzare, è proprio quanto i risultati siano sorprendenti rispetto alle nostre aspettative (si vedano ancora, a questo proposito, i brevi accenni sul concetto di emergenza presentati nel primo capitolo).

Embodiment: istanze teoriche

Come questo capitolo cerca di rendere manifesto, l'utilizzo di reti neurali ed algoritmi genetici è parte di un più ampio programma di ricerca che cerca di superare alcune delle limitazioni che, nel corso degli anni, sono emerse nella modellistica della scienza cognitiva "classica". Naturalmente una contrapposizione rigida fra i due approcci è solo un artificio che permette di chiarire alcuni dei presupposti teorici che stanno alla base dell'approccio qui utilizzato, che si definisce solitamente con etichette quali *embodied cognition* o *situated cognition*. Sebbene per alcuni ricercatori (si veda per esempio van Gelder 1998) la contrapposizione fra i due approcci possa essere considerata come una vera e propria "rivoluzione" che ha cambiato, o è

destinata a cambiare in breve, l'idea che le scienze cognitive sono andate sviluppando sulla cognizione umana, altri (per esempio Clark 1997), li considerano in una prospettiva "continuista", per cui l'approccio *embodied* completa l'approccio classico, deve tenere conto dei risultati raggiunti da questo, e non è esso stesso privo di problemi (uno, per esempio, riconosciuto da quasi tutti, è la difficoltà di rendere conto di abilità cognitive di alto livello, tipicamente umane). La mia preferenza va a questa seconda interpretazione, ma, tuttavia, come si diceva, è utile, a scopo illustrativo, considerare i due approcci come contrapposti.

L'approccio *embodied* si è sviluppato intorno alla fine degli anni '80 (si veda Brooks 1990, 1991), in parallelo alle sempre maggiori difficoltà dell'intelligenza artificiale classica. Schematizzando (si veda anche la tavola 2.1) l'approccio classico tende a considerare la cognizione come individuale, nel senso che il luogo privilegiato in cui si svolgono i processi cognitivi è il singolo individuo, mentre l'approccio *embodied* tende a localizzare i processi cognitivi in un contesto più vasto che comprende certamente i singoli individui (e le loro *menti*), ma anche le loro interazioni con l'ambiente, che, soprattutto nel caso degli esseri umani, è un ambiente socialmente costruito². Ancora, l'approccio classico tende a considerare la cognizione come razionale, nel senso che il pensiero concettuale degli esseri umani (quello, per

² Solo due commenti. I sostenitori dell'approccio classico non erano – o non sono –, naturalmente, così ingenui da credere che l'ambiente o il contesto sociale non influenzino i processi cognitivi, in particolare quelli degli esseri umani. Si tratta più che altro di una questione di priorità: per rendere trattabile un tema così complesso come quello della cognizione è necessario astrarre da alcuni dettagli e semplificare alcune assunzioni. E' probabile che le astrazioni e le semplificazioni operate dai sostenitori dell'approccio classico abbiano, in effetti, occultato alcuni aspetti fondamentali di quella che usiamo chiamare intelligenza. Tuttavia, è fuor di dubbio che alcuni importanti progressi siano stati realizzati.

Il secondo commento è invece più legato alle discipline antropologiche, quantomeno all'antropologia cognitiva. Per un antropologo, sentirsi dire che "l'approccio *embodied* tende a localizzare i processi cognitivi in un contesto più vasto che comprende certamente i singoli individui (e le loro *menti*), ma anche le loro interazioni con l'ambiente, che è un ambiente socialmente costruito" può sembrare una banalità o semmai un tardivo riconoscimento di quello che l'antropologia ha sempre sostenuto. Qui, il rischio è lo stesso che ho sottolineato discutendo di apprendimento ed evoluzione: la mia opinione è che a volte, in antropologia, si tenda a rifiutare totalmente l'approccio classico delle scienze cognitive, ma senza in effetti conoscere nel dettaglio risultati e alternative possibili e come queste alternative abbiamo incorporato molti dei risultati ottenuti dalle scienze cognitive nel corso dei decenni.

capirci, che si può formalizzare con le regole della logica) viene considerato come la forma più genuina di attività cognitiva, mentre l'approccio *embodied* vede la cognizione come una serie di euristiche, la cui caratteristica fondamentale è la loro economicità, che possono anche essere “scorrette” dal punto di vista della logica, ma che, in determinati contesti, e, probabilisticamente, nella maggior parte dei casi, sono efficaci per risolvere problemi adattivi. Un'ulteriore differenza, già sottolineata in precedenza in questo capitolo, riguarda il fatto che l'approccio classico considera la cognizione come qualcosa di astratto, ossia indipendente dall'implementazione fisica (nel duplice senso di implementazione in un cervello e in un corpo e di implementazione dell'agente in un ambiente fisico). I sostenitori dell'approccio *embodied*, al contrario, ritengono di primaria importanza per la modellizzazione dei processi cognitivi i vincoli fisici imposti dal corpo e dell'ambiente esterno in cui l'agente agisce. Ancora, mentre i sostenitori dell'approccio *embodied* tendono a vedere i processi cognitivi come indissolubilmente legati alla percezione e all'azione, l'approccio classico li considera come separati (*detached*), cioè si ritiene plausibile trattare autonomamente, e in modo seriale, i sistemi cognitivi per la percezione, l'elaborazione concettuale ed infine l'azione (i classici diagrammi di flusso dell'intelligenza artificiale classica). Infine, per l'approccio classico, i processi cognitivi che vengono realizzati nei modelli sono istanze di principi generali, che valgono per qualsiasi agente cognitivo e in qualsiasi circostanza, mentre per l'approccio *embodied* i processi cognitivi sono specifici sia rispetto al tipo di agente, che rispetto al contesto, nonché rispetto alla particolare attività che viene svolta dall'agente.

Intelligenza artificiale classica (GOFAI)	Approccio <i>embodied</i>
Individual	Social
Rational	Heuristic
Abstract	Concrete / Embodied
Detached	Located
General	Specific

Tav 2.1: Come la cognizione viene considerata nell'approccio “tradizionale” dell'intelligenza artificiale e nell'approccio *embodied* (ripresa, con modifiche, da Cantwell Smith 1999).

Embodiement: reti neurali, organismi simulati, robots

La questione che ora si pone è come questi principi teorici astratti possano essere implementati in un modello. Si è visto come l'utilizzo di reti neurali e di algoritmi evolutivi e/o di apprendimento vada in questa direzione, ma per giungere a qualcosa di simile alle simulazioni che verranno presentate nei prossimi capitoli bisogna aggiungere ancora alcuni strumenti.

Innanzitutto, come si è anticipato, possiamo immaginare che gli input e gli output delle reti neurali non siano semplicemente simboli astratti, ma rappresentino degli stimoli chimico – fisici che un agente può ricevere in un ambiente. Per esempio, gli input possono rappresentare degli stimoli percettivi e gli output dei comandi motori. A questo punto le nostre reti neurali possono essere poste in un ambiente simulato ed interagire con questo ambiente. Stimoli percettivi differenti producono input differenti per la rete che si comporterà in modo diverso a seconda delle caratteristiche dell'ambiente in cui si trova, ossia si può “muovere” più velocemente, più lentamente, fermarsi, eccetera. Con opportuni algoritmi i pesi della rete possono essere modificati per rinforzare certe classi di comportamenti.

Gli ambienti in cui gli organismi simulati “vivono” possono essere resi più complessi, aggiungendo oggetti che possono rappresentare risorse, pericoli o quant'altro e, soprattutto, per quanto riguarda i nostri scopi, altri organismi simulati che possono interagire tra loro.

Le simulazioni con reti neurali possono anche essere implementate in artefatti fisici come i robot (vedi figura 2.3). Per esempio, gli input della rete neurale possono essere realmente collegati a dei sensori che si attivano in presenza di oggetti fisici nelle loro vicinanze o che cambiano attivazione a seconda del colore del terreno su cui si muovono, mentre gli output possono attivare dei motori reali che controllano le ruote dei robot. Allo stesso modo, i robot possono essere dotati di microfoni e altoparlanti che permettono di instaurare semplici forme di comunicazione tra di essi.



Fig 2.3: Un robot *e-puck*, realizzato dall'Ecole Polytechnique Fédérale de Lausanne. Questo semplice modello di robot verrà utilizzato nella simulazione descritta nel capitolo 5.

Quali sono i vantaggi di utilizzare robot reali invece che organismi simulati nei modelli? Se per coloro che sono interessati più direttamente agli aspetti ingegneristici è ovvio valutare i risultati del proprio lavoro nei termini della riuscita dell'implementazione fisica e se per coloro che studiano processi cognitivi di “basso-livello” (mi vengono in mente per esempio la locomozione o, più in generale, il controllo dei movimenti corporei) è sicuramente utile impiegare degli artefatti fisici nei propri esperimenti, nel nostro caso non è immediato rispondere a questa domanda. Di certo non è possibile affermare che un'implementazione robotica sia più *realistica* di una simulazione ad agenti, almeno per quanto riguarda il tipo di dinamiche sociali che siamo interessati a studiare. I sensori e i motori di un robot sono tanto poco realistici quanto lo sono le etichette utilizzate come input ed output delle reti neurali. Tuttavia, mentre nel caso degli organismi simulati lo sperimentatore può decidere liberamente e in modo arbitrario come codificare gli input e gli output, nel caso dei robot è necessario rispettare ulteriori vincoli fisici: gli output e gli input devono potere essere realizzati da dei dispositivi reali, anche se questi dispositivi non hanno alcun corrispondente in natura. Utilizzando questo tipo di artefatti è sempre più difficile per lo sperimentatore prevedere come un compito verrà risolto: come vedremo nella simulazione descritta nel capitolo quattro, questi vincoli fisici possono determinare

alcuni risultati non sempre intuitivi che, anche se non possono essere direttamente collegati a fenomeni esistenti in natura, potrebbero fare luce su alcuni principi generali in un modo che sarebbe difficile cogliere usando codificazioni arbitrarie.

Da ora in avanti si cercherà di mostrare come gli strumenti qui descritti possano essere impiegati per costruire modelli di fenomeni reali: in particolare, nel prossimo capitolo, verrà presentata una simulazione che cerca di analizzare gli effetti che la *direzionalità* della trasmissione dei comportamenti – nel nostro caso trasmissione intragenerazionale (o orizzontale) e intergenerazionale – ha sul processo di evoluzione culturale.

3. Trasmissione culturale intergenerazionale e intragenerazionale

Mentre la trasmissione genetica procede necessariamente dai genitori ai figli, la trasmissione culturale può coinvolgere ogni individuo, indipendentemente dalla relazione genetica che intercorre tra coloro che sono coinvolti nella trasmissione. Tradizionalmente vengono distinte tre “direzioni” della trasmissione culturale rispetto a quella genetica (vedi Fig. 3.1): trasmissione verticale (avviene tra individui di generazioni differenti geneticamente correlati – tipicamente genitori e figli), obliqua (avviene tra individui di generazioni differenti *non* geneticamente correlati) ed orizzontale (avviene tra individui della stessa generazione, indipendentemente dalla loro correlazione genetica) (Cavalli – Sforza & Feldman 1981, Boyd & Richerson 1985).

In linea di principio, un individuo può acquisire comportamenti, informazioni, valori sulla base di ogni arbitraria proporzione di trasmissione culturale verticale, obliqua o orizzontale. In pratica, però, la relativa importanza dei tre canali di trasmissione culturale è spesso correlata con alcune caratteristiche dei sistemi sociali: per esempio, la trasmissione verticale appare specialmente importante, rispetto a quella obliqua e orizzontale, nelle società tradizionali, di piccola scala (Hewlett & Cavalli – Sforza 1986). D’altro canto, la trasmissione orizzontale sembra avere acquisito un ruolo sempre maggiore nella società occidentale contemporanea, probabilmente in correlazione con l’aumento della velocità del cambiamento

culturale: per esempio, in un recente studio di psicologia sociale (Knafo & Schwartz 2001) è stata comparata l'importanza della trasmissione orizzontale rispetto a quella verticale e obliqua in Israele, tra locali e tra immigranti degli stati ex-sovietici. I figli degli immigrati sono statisticamente meno disposti ad accettare i valori e i comportamenti dei genitori rispetto ai giovani nativi israeliani. Gli autori spiegano questo risultato attraverso il rapido cambiamento dell'ambiente culturale che gli immigrati devono affrontare: tra gli immigrati, i comportamenti e i valori della generazione adulta, che avrebbero potuto essere efficaci in un ambiente differente, avevano perso parte della loro efficacia nella nuova situazione.

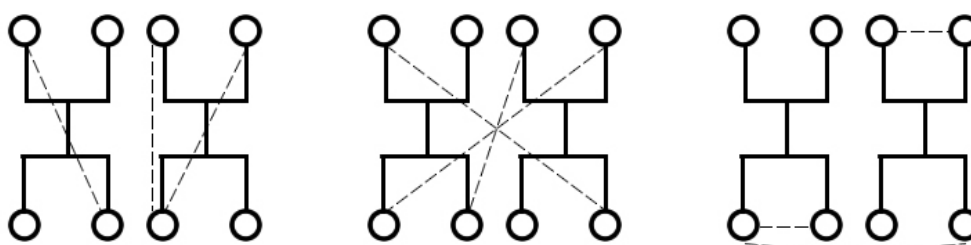


Fig. 3.1: Sinistra: trasmissione culturale verticale. Centro: trasmissione culturale obliqua. Destra: trasmissione culturale orizzontale. Le linee continue rappresentano le relazioni genetiche, le linee tratteggiate la “direzione” della trasmissione culturale.

La scelta di testare un modello simulativo di trasmissione ed evoluzione culturale rispetto a queste tematiche è motivata sia dalla relativa diffusione di modelli analitici con i quali potere, almeno qualitativamente, confrontare i risultati ottenuti (Cavalli – Sforza & Feldman 1981, Boyd & Richerson 1985), sia dal fatto che sembrano essere, almeno intuitivamente, adatti da simulare con le metodologie a nostra disposizione.

Così, l'obiettivo nella realizzazione di questa simulazione è stato duplice: da una parte, un obiettivo di tipo implementativo (*come* realizzare un modello in cui la trasmissione dei comportamenti in una popolazione di agenti controllati da reti neurali avvenga tramite un meccanismo almeno in parte analogo alla trasmissione culturale? *Come* modellizzare le differenze nella direzionalità della trasmissione culturale – verticale, obliqua, orizzontale – in un modo che producesse dinamiche compatibili

con quelle che già conosciamo?) e, dall'alta, un obiettivo più specificatamente conoscitivo (testare empiricamente l'ipotesi del legame tra maggiore adattività della trasmissione orizzontale e aumento della velocità del cambiamento dell'ambiente culturale).

Un modello di trasmissione culturale

L'ambiente simulato è costituito da un quadrato bidimensionale che contiene, distribuiti in modo casuale, 50 oggetti che gli agenti devono raggiungere (funghi commestibili) e 50 oggetti che gli agenti devono evitare (funghi velenosi), anch'essi distribuiti in modo casuale. Il comportamento di ogni agente è controllato da un perceptrone (McCulloch & Pitts 1943), cioè una rete neurale a due strati, con dieci unità di input e tre unità di output (motorie) e nessuna unità interna (Fig. 3.2), le funzioni di attivazione delle unità sono funzioni logistiche. Gli input sono codificati in questo modo (vedi anche Baldassarre 2001): i funghi commestibili e i funghi velenosi hanno due differenti colori, diciamo A e B. Gli agenti sono forniti di una retina monodimensionale che copre un campo visivo di 180° e che è recettiva ai due colori A e B. Il campo visivo è diviso in cinque settori e quello che appare in ognuno dei settori è codificato in due input. Un input viene attivato in corrispondenza della presenza, nel settore corrispondente, di un oggetto di colore A, mentre l'altro è attivato dalla presenza di un oggetto del colore B (Fig. 3.2) Le tre unità di output codificano il modo in cui gli agenti si muovono nell'ambiente: due unità determinano la direzione del movimento (l'angolo), mentre la terza codifica la velocità del movimento, ossia la distanza percorsa in ogni ciclo dall'individuo.

Ogni agente vive nella propria copia dell'ambiente per 10 epoche di 50 cicli input / output, per un totale di 500 cicli, che corrispondono a 500 movimenti. All'inizio di ogni epoca, l'agente è posizionato in un punto dell'ambiente scelto casualmente. L'energia (*fitness*) dell'agente aumenta di un'unità ogni volta che "mangia" un fungo commestibile e decresce di un'unità ogni volta che "mangia" un fungo velenoso. "Mangiare", qui, significa avvicinarsi al fungo al di sotto di una distanza di soglia. Quando un fungo, commestibile o velenoso, viene mangiato, il fungo scompare ed un nuovo fungo, dello stesso tipo, appare in un punto casuale

dell'ambiente.

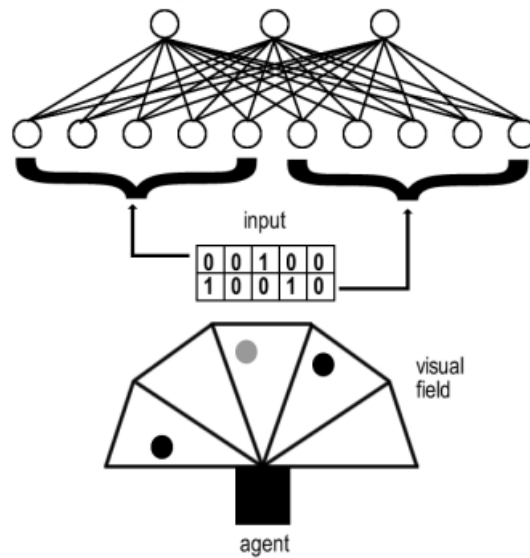


Fig. 3.2: Struttura del campo visivo dell'agente e modo in cui gli input codificano il contenuto del campo visivo. Sopra: architettura di controllo dell'agente.

Ogni generazione è composta da 100 agenti. Alla nascita, ad ogni agente è assegnato un set di valori casuali, compresi nel range $[-5.0; +5.0]$, che vengono attribuiti ai pesi delle connessioni della sua rete neurale. Alla fine della vita di tutti gli agenti, ossia dopo 500 cicli per ognuno, una nuova generazione di 100 agenti sostituisce la precedente. La simulazione continua per 500 generazioni. Poiché i pesi delle connessioni delle reti neurali sono, ad ogni generazione, assegnati in modo casuale alla nascita degli agenti, non esiste eredità genetica (nel senso descritto nel capitolo precedente) e tutte le modifiche comportamentali osservabili nel corso della simulazione sono dovute ai processi che chiameremo di trasmissione ed evoluzione culturale. Per lo stesso motivo, non esistendo eredità genetica, non è possibile distinguere tra trasmissione culturale verticale e obliqua, per cui mi riferirò, d'ora in poi, a trasmissione *intergenerazionale* per la trasmissione culturale tra i membri di una generazione e quelli della successiva, e a trasmissione *intragenerazionale* per la trasmissione culturale tra i membri della medesima generazione.

Il modello di trasmissione culturale che è stato implementato utilizza un

algoritmo di apprendimento supervisionato, per cui gli output generati da una rete neurale, in relazione ad un set di input, sono comparati con gli output desiderati (il set di output desiderati prende il nome di *teaching input*) e la differenza – quantitativa – tra i due set di output (errore) viene usata per modificare in modo graduale i pesi delle connessioni, finchè gli output generati dalla rete sono sufficientemente simili agli output desiderati e l'errore tende a zero. Nel nostro caso, i *teaching input* sono gli output generati da un'altra rete – la rete del modello – in risposta agli stessi input presentati all'osservatore. Il tipo di apprendimento reale che ispira questo modello di trasmissione culturale è l'apprendimento per imitazione: l'individuo che apprende osserva come il modello si comporta in relazione ad alcune specifiche situazioni e, in modo progressivo, impara a rispondere alle stesse situazioni in un modo uguale, o simile, al modello. Anche se all'inizio dell'apprendimento il comportamento dell'osservatore è diverso dal comportamento del modello, poiché i pesi delle sue connessioni sono assegnati in modo casuale, dopo un numero sufficiente di ripetizioni della procedura di apprendimento i pesi sono stati modificati in modo tale che il comportamento dell'osservatore assomigli al comportamento del modello (implementazioni analoghe sono: Hutchins & Hazlehurst 1995, Denaro & Parisi 1996, si veda anche, per una breve review: Parisi 1997).

Usando un'implementazione analoga, Denaro e Parisi (1996) hanno mostrato come, in una popolazione di reti neurali, possono emergere comportamenti inizialmente non esistenti, a patto che due vincoli siano rispettati: (1) i modelli vengono selezionati tra gli agenti della generazione precedente che hanno ottenuto i migliori risultati, in termini di prestazioni, e (2) del rumore viene aggiunto al *teaching input*, in modo che il comportamento dell'osservatore sia simile, ma non perfettamente uguale, a quello del modello. Entrambi i vincoli sembrano essere abbastanza realistici, poiché nella trasmissione culturale esistono statisticamente diverse tendenze che indirizzano la scelta degli individui da cui apprendere (Boyd & Richerson 1985) e ciò che un individuo apprende non è quasi mai una copia esatta del comportamento originale¹. Nel modello qui presentato, il rumore rappresenta fattori

¹ La scarsa fedeltà del processo di trasmissione individuale e la relativa stabilità della cultura a livello popolazionale è un fenomeno attualmente non perfettamente compreso. Dan Sperber (1999) ha avanzato l'ipotesi che tale scarsa fedeltà sarebbe un serio problema per gli approcci evuzionistici alla cultura che si rifanno ad un'analogia stretta tra le unità fondamentali di trasmissione nel caso culturale e biologico, poiché i geni, al contrario dei tratti culturali, si riproducono con un'affidabilità quasi

come gli errori casuali nella trasmissione, le modificazioni intenzionali del comportamento, la ricombinazione di aspetti di comportamenti diversi appresi da diversi modelli e così via, e rende possibile, in pratica, l'introduzione di nuove varianti dei comportamenti culturalmente trasmessi².

Qui, a differenza del modello di Denaro e Parisi, il processo di trasmissione culturale si svolge in diversi momenti nel corso della vita degli agenti, avendo quindi anche una connotazione temporale. L'apprendimento culturale ha luogo all'inizio di ogni epoca, per le prime sei delle dieci di cui è costituita la vita di ciascun agente. Per ogni "sessione" di apprendimento, ad ogni agente viene assegnato un modello, scelto ogni volta tra i cinque individui con l'energia più alta della generazione precedente, nel caso della trasmissione intergenerazionale, o della stessa generazione, nel caso della trasmissione intragenerazionale. Si noti che, tra l'altro, ciò implica che, in entrambi i casi, un individuo ha di norma più di un modello nel corso della propria vita. Duecento pattern di input vengono presentati, uno alla volta, sia al modello che all'osservatore; i due agenti producono un output in risposta a ciascun pattern di input e, infine, i pesi delle connessioni dell'osservatore vengono modificati sulla base delle differenze tra l'output prodotto e quello prodotto dal modello. Ogni pattern di input è un vettore di 10 bit (0 e 1) generati casualmente, anche se, per riflettere le condizioni ecologiche, gli 1 (che rappresentano la presenza di un fungo, commestibile o velenoso, in un particolare settore del capo visivo) sono solo il 16% del totale dei bit perfetta. Sperber ha argomentato che la stabilità a livello popolazionale possa essere spiegata da strutture cognitive universali – moduli – che vincolano in modo sistematico il trasferimento delle informazioni, facendo convergere, in pratica, su "soluzioni" comuni, processi che di per sé porterebbero ad esiti differenti.

A proposito dello stesso argomento, Hernrich e Boyd (2002) ritengono che l'assenza di replicatori a livello micro non invalidi la costruzione di modelli analitici ispirati alla genetica delle popolazioni per rendere conto delle dinamiche culturali a livello macro. Allo stesso modo, pur riconoscendo la scarsa affidabilità dei processi di trasmissione a livello individuale, affermano che per spiegare la stabilità a livello popolazionale non è necessario fare riferimento a strutture cognitive comuni, ma che particolari tendenze (*bias*) nella scelta dei modelli da cui apprendere siano sufficienti (ma si veda la risposta in Cladière & Sperber in stampa). Probabilmente le due spiegazioni non sono in disaccordo (Acerbi 2007).

² Non sarà sfuggito che alcuni dei fattori citati, come la modificazione intenzionale dei comportamenti culturalmente appresi, o la ricombinazione siano, o possano essere, esplicitamente orientati dall'individuo che li attua, tramite apprendimento individuale, differenziandosi così dal rumore casuale. Si ritornerà su questo importante fatto nel capitolo successivo.

(questa percentuale è stata calcolata considerando il numero totale dei funghi presenti nell'ambiente – sempre 100 – e la grandezza del campo visivo degli agenti). Gli output del modello sono utilizzati come *teaching input* dopo avere introdotto, come anticipato sopra, del rumore casuale: ognuno dei tre output che fanno parte di ogni pattern ha una probabilità del 5% di essere sostituito da un valore casuale scelto nel range [0.0;1.0]. L'apprendimento è implementato tramite l'algoritmo di Widrow – Hoff (Widrow & Hoff 1960), più comunemente conosciuto come regola delta, con un tasso di apprendimento (*learning rate*) di 0.3: il tasso di apprendimento definisce di quanto i valori dei pesi sono modificati in funzione della grandezza dell'errore (vedi appendice 3.1). Questo processo, come detto, viene ripetuto all'inizio di ogni epoca per le prime sei epoche della vita dell'agente, per un totale di 1200 pattern presentati.

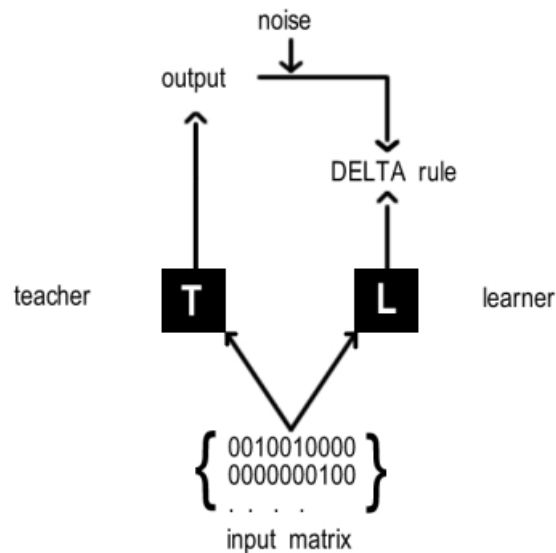


Fig. 3.3: Come viene implementata la trasmissione culturale.

Risultati

Se la trasmissione culturale è puramente intergenerazionale, ossia tutti gli agenti delle successive generazioni apprendono dai migliori individui della generazione precedente, l'abilità di distinguere i funghi commestibili da quelli velenosi e di raggiungere i primi evitando i secondi, che non è, all'inizio, presente, emerge gradualmente e diventa una proprietà stabile della popolazione (Fig. 3.4).

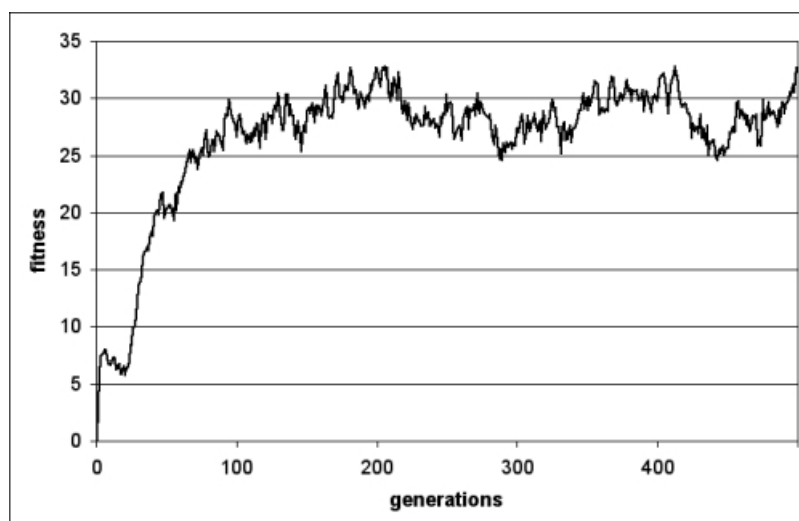


Fig. 3.4: Prestazione media della popolazione di 100 agenti nel caso di trasmissione culturale puramente intergenerazionale [dati medi su 10 repliche].

Al contrario, se la trasmissione culturale è puramente intragenerazionale, ossia tutti gli agenti apprendono dai migliori individui della propria generazione, non si assiste all'emergere di un comportamento appropriato (Fig. 3.5). Nel caso della trasmissione intergenerazionale, i modelli sono individui che hanno avuto l'opportunità di apprendere per tutte le sei epoche della fase di apprendimento ed hanno appreso da individui che, a loro volta, avevano compiuto lo stesso percorso.

Nel caso della trasmissione intragenerazionale, invece, poiché ogni individuo ha, alla nascita, pesi casuali, la probabilità che uno di essi esibisca un comportamento corretto che possa essere appreso dai “coetanei” è estremamente bassa. Inoltre, con una trasmissione culturale puramente intragenerazionale, il processo di trasmissione culturale non può essere cumulativo: i comportamenti appropriati che, per ragioni casuali, dovessero emergere alla fine del processo di apprendimento non possono essere trasmessi alla successiva generazione, con il risultato che ogni nuova generazione deve ricominciare da zero³.

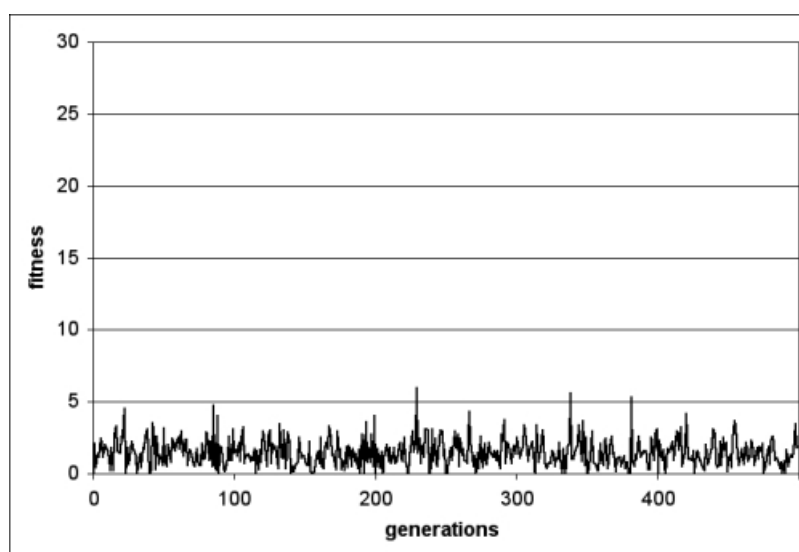


Fig. 3.5: Prestazione media della popolazione di 100 agenti nel caso di trasmissione culturale puramente intergenerazionale [dati medi su 10 repliche].

Un possibile ruolo per la trasmissione intragenerazionale

Se le popolazioni che usano esclusivamente la trasmissione culturale intragenerazionale non sono in grado di evolvere un comportamento adattivo,

³ Michael Tommasello (2005) ritiene che la cumulatività della trasmissione culturale (che egli chiama *ratchet effect*) sia una caratteristica peculiare della specie umana. Gli altri animali, in particolare i primati non umani, oggetto di studi estensivi da parte di Tommasello e dei suoi collaboratori, pur possedendo una qualche capacità di apprendimento sociale, mancherebbero della capacità di mantenere nel tempo le modificazioni apportate ai tratti culturali, non potendo così sviluppare un sistema culturale paragonabile a quello di *homo sapiens*.

dobbiamo concludere che la trasmissione culturale intragenerazionale non possa avere alcun ruolo nel processo di evoluzione culturale? Per rispondere a questa domanda dobbiamo chiederci, prima di tutto, quale possano essere, in generale, gli effetti della trasmissione culturale intragenerazionale. L'ipotesi qui presentata è che la trasmissione intragenerazionale abbia l'effetto di introdurre rumore nel processo di trasmissione culturale⁴.

Come già detto, una certa quantità di rumore è necessaria per l'emergere di comportamenti adattivi, sia nel caso della trasmissione genetica, sia in quello della trasmissione culturale. Nel modello qui presentato, una certa quantità di rumore viene aggiunta agli output del modello (si veda ancora Fig. 3.3), così che, alla fine dell'apprendimento, il comportamento dell'osservatore finisce per essere simile, ma non uguale, a quello del modello. Siccome le differenze tra i comportamenti sono casuali, nella maggior parte dei casi l'osservatore tenderà ad esibire un comportamento meno efficiente se comparato con quello del modello, ma, almeno in qualche occasione, esibirà un comportamento migliore. Gli agenti a cui capiterà ciò saranno, a loro volta, scelti come modelli, e in questo modo i miglioramenti dovuti al rumore avranno la tendenza a consolidarsi ed a diffondersi nella popolazione. Il rumore aggiunto all'output del modello ha lo stesso ruolo, nella nostra simulazione, che le mutazioni genetiche casuali hanno nella trasmissione biologica: entrambe rendono possibile la ricerca non-orientata di nuove e migliori soluzioni.

A conferma di questa analisi, se rimuoviamo il rumore dal processo di trasmissione culturale, così che l'osservatore riceve come *teaching input* gli output, non modificati, del modello, il risultato è che la qualità del comportamento raggiunge il livello del migliore individuo della prima generazione ma non va oltre (Fig. 3.6). In altre parole, la trasmissione culturale è potenzialmente cumulativa, ma non ci sono modificazioni comportamentali che permettono di modificare – e selezionare – nuovi

⁴ E' evidente che questo modello, non prevedendo la possibilità di apprendimento individuale, esclude il fatto che i coetanei abbiano modificato in modo autonomo il loro comportamento in direzione di una maggiore adattività. E' certo che uno dei motivi principali per cui la trasmissione orizzontale rappresenta una forza importante dell'evoluzione culturale è da ascrivere a questo fatto. Inizialmente, questa mi sembrava una debolezza che potenzialmente poteva invalidare i risultati del modello. Tuttavia, più recentemente ho cambiato opinione e credo che questo possa essere, al contrario, un punto di forza della simulazione, una volta esplicitato: il dato nuovo e probabilmente controintuitivo è che, anche *senza apprendimento individuale*, la trasmissione orizzontale può avere un ruolo adattivo nell'evoluzione culturale.

tratti da trasmettere.

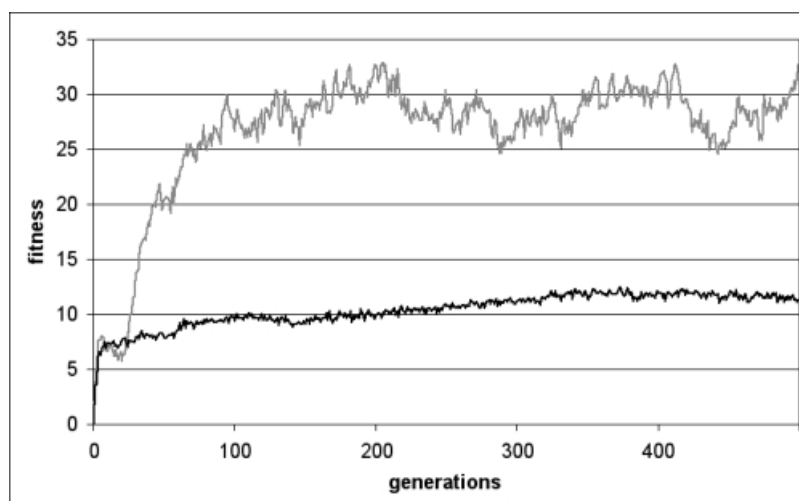


Fig. 3.6: Confronto tra la prestazione media della popolazione di 100 agenti nel caso di trasmissione culturale puramente intergenerazionale con rumore (linea grigia) e senza rumore (linea nera) [dati medi su 10 repliche].

Infatti, è stato osservato (Denaro & Parisi 1996) che se viene rimosso il rumore dal *teaching input*, non solo la popolazione non è in grado di andare oltre il livello del migliore degli individui della prima generazione, ma anche che la prestazione media tende progressivamente a deteriorarsi. Questo avviene perché, ad ogni generazione, il comportamento non viene mai appreso in modo perfetto, poiché l'errore, ossia la differenza tra l'output prodotto dal modello e l'output prodotto dall'osservatore, non scende mai a zero, per le caratteristiche intrinseche dell'algoritmo utilizzato e della struttura delle reti neurali.

In queste simulazioni non solo abbiamo replicato il risultato ottenuto da Denaro e Parisi, ma abbiamo anche mostrato come tale risultato sia influenzato dal tasso di apprendimento applicato al processo di trasmissione culturale. Tassi di apprendimento alti – come il nostro 0.3 – nella simulazione senza rumore possono innescare piccoli miglioramenti, al contrario di tassi di apprendimento più bassi, come nel caso di Denaro e Parisi (1996) (Fig. 3.7). Questo effetto può essere spiegato dal fatto che a tassi di apprendimento più elevati corrispondono maggiori deviazioni rispetto al comportamento da apprendere, il che implica un esplorazione più ampia dello spazio dei comportamenti possibili. Questa esplorazione più ampia, anche senza rumore nel processo di trasmissione, può innescare piccoli miglioramenti (o limitare il deterioramento della prestazione) sebbene la prestazione media rimanga sensibilmente

peggiore rispetto allo scenario in cui il rumore è presente nella trasmissione culturale.

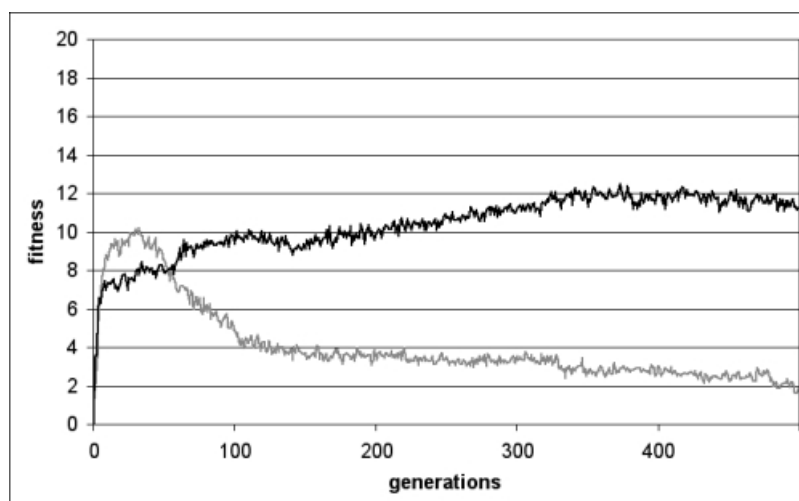


Fig. 3.7: Confronto tra la prestazione media della popolazione di 100 agenti nel caso di trasmissione culturale puramente intergenerazionale senza rumore con tasso di apprendimento = 0.1 (linea grigia) e tasso di apprendimento = 0.3 (linea nera) [dati medi su 10 repliche].

Dopo avere chiarito il ruolo del rumore nel nostro modello di trasmissione culturale, è possibile cercare di capire quale può essere la funzione della trasmissione intragenerazionale, se viene interpretata come un meccanismo in grado di generare rumore.

Immaginiamo che, senza rumore nella trasmissione culturale, in ciascuna generazione, alcuni individui imparino dai migliori fra i loro coetanei – ossia, con trasmissione culturale intragenerazionale – e altri imparino dai migliori individui della generazione precedente – ossia, con trasmissione intergenerazionale. Questa situazione potrebbe avere due vantaggi. Innanzitutto, ci potrebbero essere alcuni individui che, già alla nascita, per ragioni casuali (i pesi delle connessioni), mostrano comportamenti che sono migliori dei comportamenti dei loro coetanei. Con una certa quantità di trasmissione intergenerazionale, la possibilità che essi trasmettano il loro comportamento agli individui della generazione successiva non è completamente esclusa, come nel caso in cui era permessa solo la trasmissione intragenerazionale. Inoltre, apprendere da coetanei che non esibiscono un comportamento “corretto” (secondo gli standard della generazione precedente) può essere, come nel caso dell’aumento del tasso di apprendimento, una sorgente di rumore nel sistema. E’ interessante notare come, in questo caso, il rumore sia “orientato”. Mentre nel caso

dell'aumento del tasso di apprendimento o del rumore intrinseco al processo di trasmissione culturale (la quantità di rumore aggiunta all'output del modello) l'esplorazione dello spazio dei comportamenti possibili procede in un modo completamente casuale, ora solo i migliori comportamenti, tra quelli non necessariamente in linea con la generazione precedente, vengono replicati dagli agenti coinvolti nel processo di trasmissione culturale.

Per testare questa ipotesi, è stato creato uno scenario in cui non c'è rumore intrinseco nella trasmissione culturale e una variabile α specifica la probabilità che un agente utilizzi la trasmissione intragenerazionale piuttosto che quella intergenerazionale. Per bassi livelli di α (0.1) i risultati sono più o meno gli stessi di quelli dello scenario con trasmissione intergenerazionale "pura". Tuttavia, con valori più alti di α , la prestazione della popolazione è notevolmente migliore (Fig. 3.8).

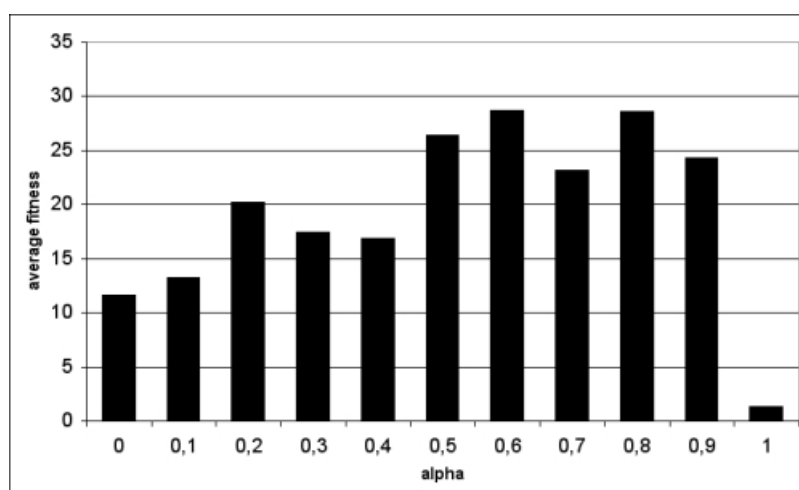


Fig. 3.8: Prestazione media (nelle ultime 100 generazioni) in funzione del livello di α ($\alpha = 0$: trasmissione puramente intergenerazionale; $\alpha = 1$: trasmissione puramente intragenerazionale) [dati medi su 10 repliche].

La conclusione che possiamo trarre è che, in assenza di rumore, un certo grado di trasmissione culturale intragenerazionale (in questo scenario il risultato migliore si ottiene con $\alpha = 0.6$, si veda Fig. 3.8 e Fig. 3.9) rende possibile un progressivo miglioramento della prestazione che, nel caso di trasmissione puramente intergenerazionale senza rumore è impossibile. Per $\alpha = 1$ (trasmissione puramente intragenerazionale) la prestazione è vicina a zero (vedi quanto detto sopra). Ciò

nondimeno, anche una piccola quantità di trasmissione intergenerazionale ($\alpha = 0.9$) è sufficiente per produrre una buona prestazione nelle ultime 100 generazioni, in virtù del processo cumulativo innescato dalla trasmissione intergenerazionale.

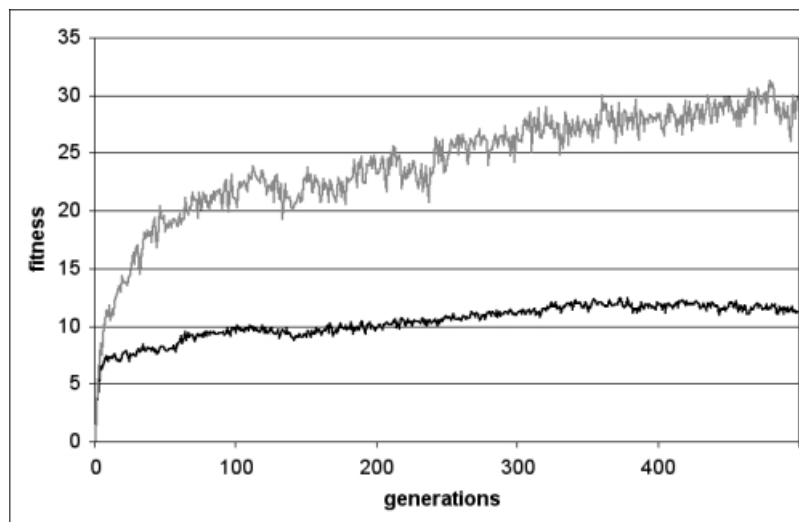


Fig. 3.9: Confronto tra la prestazione media della popolazione di 100 agenti nel caso di trasmissione culturale senza rumore con $\alpha = 0.6$ (linea grigia) e con $\alpha = 0$ (linea nera) [dati medi su 10 repliche].

Trasmissione intragenerazionale in ambiente dinamico

Fino ad ora, lo scenario preso in considerazione era costituito da un ambiente statico, ossia da un ambiente che non cambiava durante il tempo. In questo ambiente il problema era di distinguere i funghi commestibili da quelli velenosi, raggiungere i primi ed evitare i secondi e questo problema rimaneva lo stesso durante tutta la simulazione. Ma, in generale, l'ambiente, e i problemi da affrontare in un ambiente, cambiano nel tempo e, quando questo avviene, i comportamenti sviluppati in un certo ambiente per risolvere certi problemi potrebbero non essere più appropriati nel nuovo ambiente e per i nuovi problemi. E' in questa situazione, cioè in ambienti dinamici che cambiano rapidamente, che i limiti della trasmissione intergenerazionale e, al contrario, il ruolo positivo della trasmissione intragenerazionale emergono con chiarezza⁵.

⁵ "Ambiente", in questo caso, va inteso in senso ampio, come l'insieme dei fattori (naturalisti,

Per testare questa ipotesi è stato sviluppato uno scenario in cui il rumore è presente nella trasmissione culturale (e, quindi, da questo punto di vista, la trasmissione intergenerazionale sarebbe efficace), ma in cui l'ambiente cambia improvvisamente nel corso della simulazione, dopo che la popolazione ha avuto il tempo necessario per adattarsi alle condizioni iniziali. All'inizio la situazione è identica alle simulazioni precedenti, ma dopo 100 generazioni l'ambiente cambia: le caratteristiche percettive dei funghi vengono invertite o, in altre parole, i funghi che erano commestibili diventano improvvisamente velenosi e viceversa. Chiaramente, nel nuovo ambiente, i comportamenti che erano in precedenza appropriati non lo sono più: la popolazione ha a disposizione altre 100 generazioni per evolvere i nuovi comportamenti (la simulazione ha quindi una durata totale di 200 generazioni, a differenza delle precedenti, che continuavano per 500 generazioni).

In questa simulazione, durante le prime 100 generazioni, la trasmissione culturale è puramente intergenerazionale e, come sappiamo, ciò è sufficiente, in queste condizioni, per sviluppare un comportamento corretto. Tuttavia, dopo che l'ambiente è cambiato, alcuni individui apprendono anche dai propri coetanei. Come sopra, la variabile α specifica la probabilità che un individuo scelga come modello uno tra i cinque migliori individui della propria generazione al posto che uno tra i cinque migliori individui della generazione precedente.

I risultati mostrano che quando l'ambiente cambia, la prestazione della popolazione, in tutte le condizioni, crolla inevitabilmente e, in seguito, i nuovi comportamenti emergono gradualmente. Se la trasmissione culturale è puramente intragenerazionale ($\alpha = 0$) tuttavia questo processo è significativamente più lento che nel caso in cui la trasmissione è sia intergenerazionale che intragenerazionale (Fig. 3.10 e Fig. 3.11). Le differenze tra le due condizioni possono essere spiegate negli stessi termini di cui sopra: la possibilità di sfruttare e trasmettere i rari miglioramenti comportamentali dovuti al fatto che, alla nascita, i pesi delle connessioni degli agenti sono settati in modo casuale e la più ampia – e orientata – ricerca nello spazio dei comportamenti possibili che la trasmissione intragenerazionale rende possibile.

sociali, culturali) con cui un individuo interagisce nel corso della propria vita.

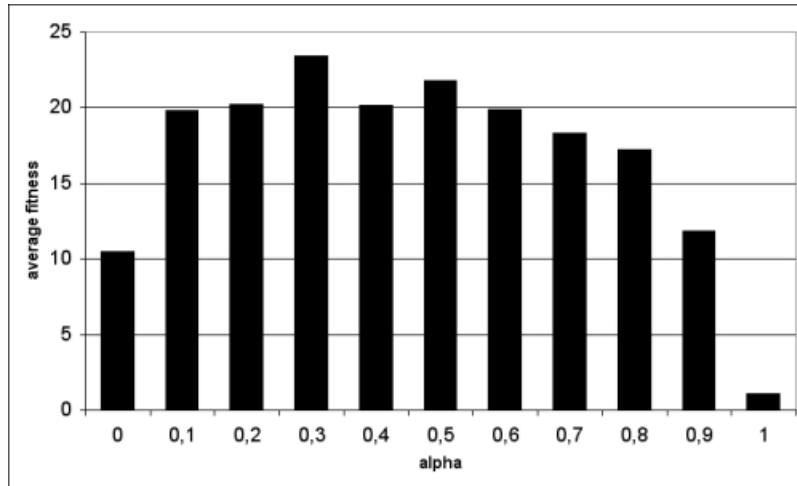


Fig. 3.10: Prestazione media (nelle ultime 100 generazioni) in funzione del livello di α ($\alpha = 0$: trasmissione puramente intergenerazionale; $\alpha = 1$: trasmissione puramente intragenerazionale) [dati medi su 10 repliche].

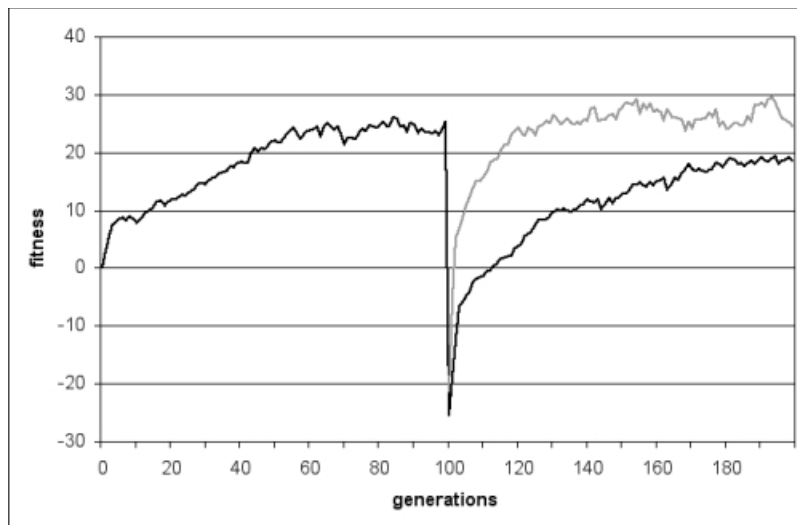


Fig. 3.11: Confronto tra la prestazione media della popolazione di 100 agenti con $\alpha = 0.3$ (linea grigia) e con $\alpha = 0$ (linea nera) [dati medi su 10 repliche].

La probabilità che un individuo apprenda dai coetanei viene ereditata

Nelle simulazioni precedenti la probabilità di utilizzare la trasmissione intragenerazionale al posto che quella intergenerazionale (α) veniva manipolata direttamente: in questa simulazione il valore di α è controllato da un “gene” ed evolve usando un algoritmo genetico (vedi capitolo 2).

Lo scenario è esattamente uguale a quello delle simulazioni precedenti – in ambiente dinamico – con la sola eccezione che ora ogni agente ha un gene con un valore compreso nel range [0.0;1.0] che determina la probabilità di scegliere come modello un individuo della propria generazione. All’inizio della simulazione i valori dei geni sono assegnati a caso ai 100 individui della popolazione iniziale ma, ad ogni generazione, i 20 individui con la migliore prestazione trasmettono il valore del proprio gene a 5 “figli” ciascuno [5 X 20 = 100 individui della generazione successiva], con una probabilità dell’ 1% che questo valore sia casualmente riassegnato (mutazione genetica). Come nelle simulazioni precedenti, i pesi delle connessioni dell’architettura di controllo vengono inizializzati in modo casuale alla nascita degli agenti, cioè lo sviluppo della capacità di distinguere i funghi e di muoversi correttamente nell’ambiente rimane interamente dovuto alla trasmissione e all’evoluzione culturale.

Basandosi sui risultati precedenti ci si può aspettare che, per le prime 100 generazioni, quando l’ambiente rimane stabile, il valore medio del gene tenderà a scendere verso 0, ossia gli individui saranno geneticamente predisposti ad imparare dagli individui della generazione precedente. Tuttavia, possiamo aspettarci che, al cambiamento dell’ambiente, il valore medio del gene venga spinto velocemente verso l’alto, in quanto la selezione tenderà a favorire i mutanti che apprendono *anche* dai propri coetanei, in quanto si può ipotizzare che tali mutanti abbiano, in media, prestazioni migliori degli agenti che utilizzano la trasmissione intergenerazionale in modo esclusivo.

Infatti, questo è ciò che accade. Le figure 3.12 e 3.13 mostrano, rispettivamente, il valore medio del gene che controlla la probabilità che un agente usi la trasmissione culturale intragenerazionale al posto di quella intergenerazionale e la

prestazione media della popolazione.

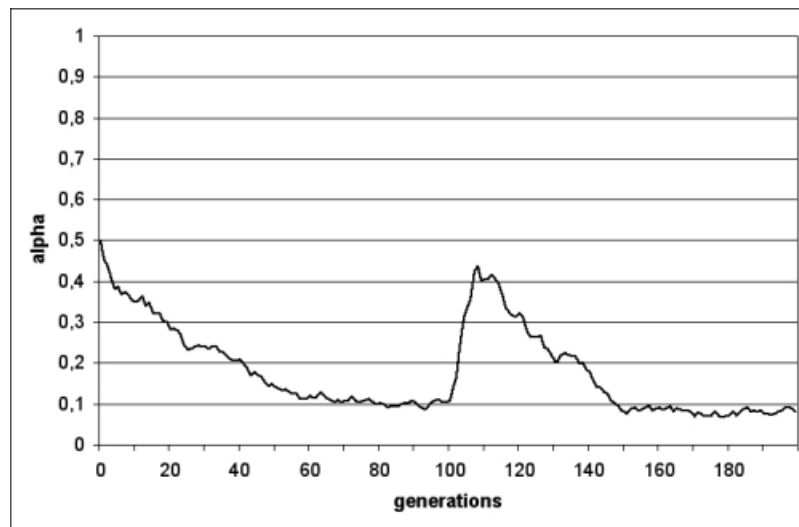


Fig. 3.12: Valore del gene che controlla la probabilità α [dati medi su 10 repliche].

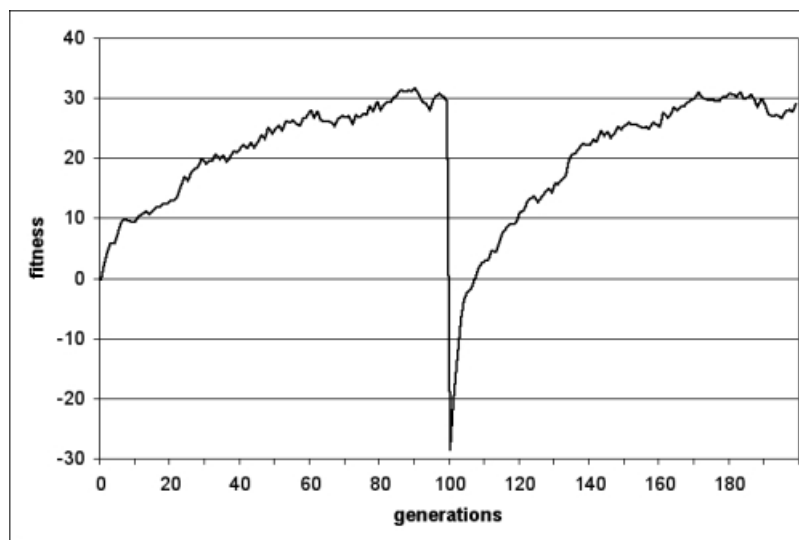


Fig. 3.13: Prestazione media della popolazione con il valore α controllato dall'algoritmo genetico [dati medi su 10 repliche].

Inoltre, come nella simulazione precedente, la prestazione della popolazione è significativamente migliore rispetto alla scenario in cui è possibile solo la trasmissione culturale intergenerazione (Fig. 3.14).

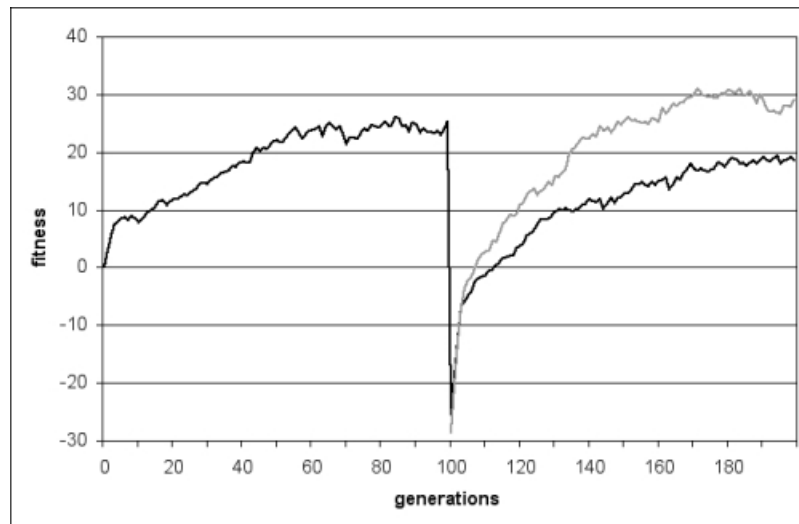


Fig. 3.14: Confronto tra la prestazione media della popolazione di 100 agenti con α controllato da algoritmo genetico (linea grigia) e con $\alpha = 0$ (linea nera) [dati medi su 10 repliche].

Conclusioni

Malgrado lo scenario molto semplice (l'ambiente di un agente consiste solo di funghi commestibili e velenosi e l'interazione con i conspecifici è limitata alla procedura standardizzata determinata dall'algoritmo di trasmissione culturale che è stato implementato) e le scelte di utilizzare generazioni che non si sovrappongono tra loro e di ignorare il ruolo dell'apprendimento individuale, queste simulazioni possono fornire alcuni suggerimenti rispetto all'importanza della trasmissione intragenerazionale nell'evoluzione culturale.

In questa simulazione il modello di trasmissione ed evoluzione culturale è stato considerato in stretta analogia con il modello di trasmissione ed evoluzione genetica basato sulla selezione naturale. Entrambi, come ogni processo di evoluzione basato su meccanismi analoghi alla selezione naturale (Dennett 1995), necessitano di due meccanismi per operare: un meccanismo di riproduzione selettiva (nel nostro caso, determinato dal fatto che i modelli venivano scelti fra i cinque individui con la prestazione migliore) ed un meccanismo in grado di aggiungere variabilità (nel nostro caso, il rumore intrinseco al processo di trasmissione culturale). In questa

simulazione, per esempio, la trasmissione intragenerazionale è stata interpretata come una forza specifica dell'evoluzione culturale e il suo ruolo – in assenza di apprendimento individuale – è stato interpretato come un meccanismo per aggiungere ulteriore variabilità al processo evolutivo. E' stato mostrato come, se la trasmissione culturale è accompagnata da un rumore intrinseco, in modo che l'osservatore acquisisca un comportamento simile ma non uguale a quello del modello, ha luogo evoluzione culturale anche se la trasmissione è puramente intergenerazionale e tutti gli agenti delle nuove generazioni apprendono esclusivamente dai migliori agenti delle generazioni precedenti. Se, tuttavia, la trasmissione culturale non è accompagnata da rumore, con una trasmissione puramente intergenerazionale, il progresso evolutivo si ferma al livello raggiunto dai migliori individui della prima generazione: manca, in altri termini, un meccanismo di generazione di variabilità. In queste circostanze, una certa quantità di trasmissione intragenerazionale si rivela essere utile. La trasmissione intragenerazionale funziona come meccanismo di aggiunta di rumore al processo evolutivo perché i coetanei sono modelli "imperfetti", poiché non hanno essi stessi completato il loro ciclo di apprendimento e perché, a loro volta, hanno imparato da loro coetanei. In questo modo, la trasmissione intragenerazionale rende possibile la scoperta di nuovi comportamenti che possono risultare più efficienti di quelli già presenti nel *pool* culturale della popolazione.

Questa analisi viene confermata dalle simulazioni in cui l'ambiente cambia: anche se il rumore intrinseco è presente nella trasmissione culturale, la necessità di adattarsi a nuove condizioni rende necessario aumentare in qualche modo questo rumore, per potere, nello stesso tempo, eliminare i comportamenti non più adattivi, ed esplorare lo spazio dei comportamenti possibili alla ricerca di comportamenti che si rivelino efficaci nelle nuove condizioni ambientali.

Allo stesso modo, le simulazioni mostrano che, anche in caso di rapidi cambiamenti ambientali, le due componenti della trasmissione culturale devono essere entrambe presenti: se la trasmissione diventa puramente intragenerazionale l'evoluzione culturale non è possibile. Senza apprendere dalle generazioni precedenti, non c'è modo di conservare i tratti culturali nel tempo e la trasmissione culturale è, per principio, non cumulativa. Se, al contrario, la componente di trasmissione intragenerazionale è completamente assente, l'adattamento ad ambienti nuovi risulta essere lento.

Evoluzione biologica ed evoluzione culturale

Se le analogie tra trasmissione ed evoluzione culturale e trasmissione ed evoluzione genetica sono sicuramente uno strumento estremamente potente che potrebbe permettere agli studiosi della cultura di utilizzare le metodologie già sviluppate in biologia e di giungere a conclusioni potenzialmente innovative sulle dinamiche culturali, oltre che fornire un quadro teorico in grado di unificare i disparati ambiti in cui sono frammentate le discipline che si occupano della cultura umana (Mesoudi, Whiten & Laland 2006), è anche vero che è necessario essere molto cauti nel sottoscrivere queste analogie e, comunque, tenere in considerazione anche le differenze (vedi Tav. 3.1).

Evoluzione	Biologica	Culturale
Mutazioni intenzionali	Poche (selezione artificiale?)	Molte (apprendimento individuale)
Replicatori	Si	Solo alcuni
Tratti discreti	Si	Solo alcuni
Scelta intenzionale dei modelli	Poche (selezione sessuale?)	Molte (<i>bias</i>)
Direzionalità	Una (ingegneria genetica?)	Molte

Tav 3.1: Alcune differenze tra evoluzione biologica e culturale.

Per cominciare, è ragionevole pensare che almeno una parte delle mutazioni culturali siano intenzionali, ossia dovute ad apprendimento individuale direzionato (si ritonerà diffusamente su questo tema nel prossimo capitolo). Ciò non significa affermare che nell'evoluzione culturale l'apprendimento individuale sia così

essenziale da direzionare intenzionalmente *tutta* l'evoluzione culturale, o, al contrario, che nell'evoluzione biologica sia del tutto impossibile dare direzionalità alle mutazioni (basti pensare a forme di selezione artificiale, anche indiretta). Probabilmente, è corretto affermare che in entrambe le evoluzioni, biologica e culturale, sia possibile parlare di un continuum che va da una maggiore ad una minore, o nulla, intenzionalità e che, in questo continuum, l'evoluzione culturale si ponga, rispetto a quella biologica, più vicina, in media, all'estremità intenzionale (vedi anche Dennett 2005, Dennett & McKay 2007).

Inoltre, non tutte le varianti culturali sono replicatori (vedi anche nota 2 in questo capitolo). Alcuni modellisti dell'evoluzione culturale, i *memeticisti* (vedi Blackmore 2002, per una valutazione complessiva della memetica vedi Aunger 2000), tendono ad assumere che credenze, abilità e valori possono essere considerati come replicatori in senso pieno del termine. Tuttavia è quantomeno dubbio che i tratti culturali possano essere considerati replicatori, almeno nello stesso senso in cui i geni lo sono. Il punto è stato espresso con chiarezza da Dan Sperber (1999): la trasmissione – nella cultura – è più una questione di replicazione che di copia. Insomma, non c'è nessuna garanzia che due individui, che osservano il medesimo pattern comportamentale (l'effetto fenotipico per stare all'analogia memetica), acquisiscano la stessa variante culturale (l'ipotetico meme). Boyd e Richerson (2006, p. 114) affermano:

L'informazione sarà riprodotta quando viene trasmessa da una mente all'altra solo se la maggioranza delle persone deduce un'unica regola da un dato risultato fenotipico. Anche se spesso può essere così, le differenze genetiche, culturali o legate allo sviluppo possono fare dedurre varianti culturali diverse dalla stessa osservazione.

Una questione differente è se il fatto che non tutti i tratti culturali possono essere considerati replicatori invalidi o meno l'utilizzo di modelli formali ispirati all'evoluzione genetica – che considerano, in effetti, i tratti come replicatori – per studiare l'evoluzione culturale. Al momento, la maggior parte di coloro che si occupano di queste tematiche tendono a pensare che con opportune modifiche si possa tenere conto di questa differenza (Henrich & Boyd 2002).

Un altro problema dell'analogia tra unità biologiche – geni – e tratti culturali

risiede nel fatto che i secondi non sono, in generale, entità discrete. Non solo la stessa variante culturale può essere acquisita da modelli differenti, ma, come hanno spesso sottolineato gli antropologi contrari all'impiego dei modelli evuzionistici, le culture sono insiemi unitari fortemente integrati che non possono essere scomposti in tratti indipendenti ed atomistici. Anche in questo caso, credo che il punto non sia che i modelli formali e computazionali di evoluzione culturale ispirati ai modelli di evoluzione biologica siano di per sé sbagliati, ma che debbano essere costruiti tenendo conto delle analogie e delle differenze fra i due processi. Qui, probabilmente, le domande chiave sono se, in effetti, *tutti* i tratti culturali non sono discreti (quali lo sono? Quali no?) e, soprattutto, in quali casi considerare i tratti culturali *come se* fossero discreti – al fine da rendere i modelli trattabili – possa privarci di una corretta comprensione delle dinamiche culturali reali.

Un'importante differenza che rende le dinamiche di evoluzione culturale potenzialmente differenti da quelle dell'evoluzione biologica risiede nel fatto che è possibile direzionare la prima in modo intenzionale, non solo mutando individualmente i tratti, come già detto, ma anche scegliendo attivamente quali tratti culturali assumere, per esempio a seconda delle caratteristiche dei tratti stessi (*direct bias*), della loro diffusione (*frequency bias*) o delle caratteristiche dei portatori dei tratti (*indirect bias*) (vedi Boyd e Richerson, 1985, 2006)⁶.

Boyd e Richerson si riferiscono, con il termine *direct bias*, alle scelte individuali, tra varianti culturali pre-esistenti, che si basano direttamente sul loro contenuto. Queste scelte possono essere effettivamente rappresentate da decisioni cosce, che prevedono una valutazione diretta dei costi e dei benefici delle alternative culturali, ma possono anche essere determinate in modo indiretto: l'importante è che ciò avvenga sulla base delle caratteristiche intrinseche dei tratti culturali. Nell'ottica di Boyd e Richerson, l'idea, portata avanti dagli psicologi evuzionistici (Barkow, Tooby & Cosmides 1992) e da un gruppo di antropologi cognitivi, come Dan Sperber (1999), Scott Atran (1998), Pascal Boyer (1994) e Lawrence Hirschfeld (1996) (per

⁶ Anche in questo caso vale quanto detto per le mutazioni intenzionali determinate dall'apprendimento individuale. Una visione tutto / niente è probabilmente poco fruttuosa: molte volte, la scelta dei tratti culturali è guidata da forze inconsce e non intenzionali che agiscono come potrebbe agire una selezione "cieca". Al contrario, nell'evoluzione biologica, dinamiche come quelle legate alla selezione sessuale (che peraltro produce effetti molto simili a quelli determinati, nell'evoluzione culturale, dagli *indirect bias*) possono direzionare in modo "non-cieco" l'evoluzione.

una valutazione si veda Acerbi 2003, 2005, 2006), che alcuni tratti culturali siano più adatti di altri a diffondersi e a rimanere stabili poiché le loro caratteristiche sono tali da renderli facilmente assimilabili, data una struttura cognitiva come quella degli esseri umani, non sarebbe altro che un caso di *direct bias*.

Intuitivamente, almeno quando il *direct bias* consiste in una valutazione conscia, diretta, esso, date le dimensioni e la complessità del repertorio culturale degli esseri umani, è relativamente costoso da mettere in pratica: sembra plausibile, quindi, che la selezione naturale abbia favorito altre euristiche più frugali. Le tendenze basate sulla diffusione dei tratti culturali (*frequency bias*) risponderebbero a questa esigenza. La più comune di queste è, ancora secondo Boyd e Richerson, una tendenza conformista (*conformist bias*), ossia una tendenza ad assumere, in una data popolazione, il tipo di variante più diffuso⁷. Questa strategia ha senso, in quanto un certo numero di processi, incluso il *direct bias* e, ovviamente, la selezione naturale, tendono a fare sì che i comportamenti adattivi si diffondano maggiormente rispetto a quelli non adattivi. Il *conformist bias* produce, nel caso di una specifica variante culturale, indipendentemente dal suo contenuto, una caratteristica curva di diffusione (Fig. 3.15). All'inizio, quando la variante è poco comune, il processo di diffusione è lento, finché si raggiunge una soglia critica di individui che hanno adottato la variante – simile ad una transizione di fase – a partire dalla quale la diffusione diventa rapida.

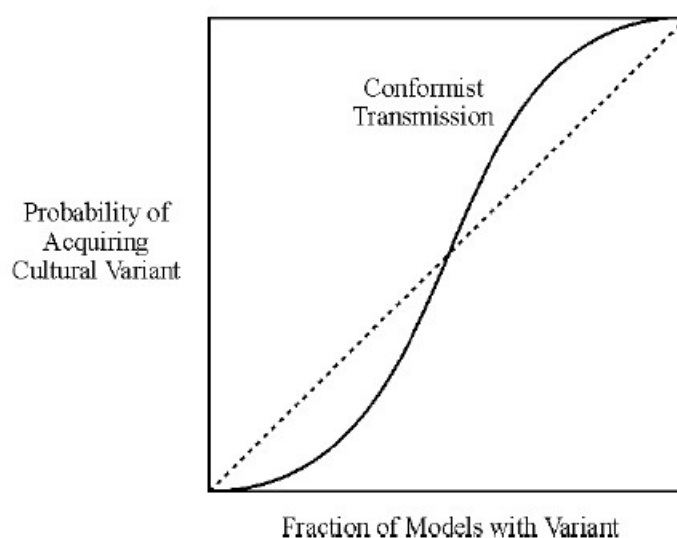


Fig. 3.15: curva di diffusione di una variante culturale in presenza di *conformist*

⁷ In alcune situazioni, tuttavia, potrebbe essere più utile scegliere un tratto basandosi sulla sua rarità: Boyd e Richerson suggeriscono l'esempio della scelta di una professione nella società occidentale contemporanea.

bias (da Boyd & Richerson manoscritto)

Un'altra euristica possibile è l'*indirect bias*: in questo caso l'idea è che il possesso di tratti culturali adattivi sia correlato con alcuni indicatori, per cui, per esempio, imitare i tratti degli individui che ci appaiono di successo (*prestige bias*) è spesso una strategia adattiva, poiché, in generale, è più facile determinare *chi* ha successo, piuttosto che *come* avere successo. Il *prestige bias* rappresenta, come le altre tendenze descritte da Boyd e Richerson, una compensazione tra l'esigenza di avere a disposizione euristiche rapide e facili da utilizzare e la possibilità che tali euristiche siano erranee. Per esempio, nel caso del *prestige bias* è evidente che non è possibile sapere, tra tutti i tratti culturali adottati da un individuo di successo, quali sono i tratti correlati con l'effettivo raggiungimento del successo, e quali no. Le dinamiche di evoluzione culturale che risultano dalla trasmissione influenzata dal *prestige bias* tendono ad essere più rapide – e imprevedibili – di quelle generate dalla trasmissione influenzata dal *conformist bias*.

Per concludere, l'ultima differenza tra evoluzione culturale ed evoluzione biologica, che mi limiterò a citare, in quanto è stata esemplificata dalla simulazione presentata in questo capitolo, risiede nella direzionalità. Come già detto, mentre la trasmissione biologica può procedere solo in una direzione (tralascio il caso dell'ingegneria genetica che, comunque, sarebbe curioso analizzare da questo punto di vista, ossia come una forma di cambiamento di direzionalità determinata da tendenze intenzionali simili a quelle che avvengono nelle dinamiche culturali), quella culturale può procedere, e generalmente procede, in più direzioni.

Appendice 3.1

La regola Delta

L'algoritmo di Widrow – Hoff (Widrow & Hoff 1960), comunemente conosciuto come regola Delta, è un tipo algoritmo di apprendimento supervisionato. L'apprendimento supervisionato consiste nell'addestrare una rete neurale a produrre, in risposta a specifici pattern di input, dei pattern di output che già si conoscono, che vengono denominati *teaching input*. La differenza tra gli output prodotti e i *teaching input* viene chiamata errore. L'algoritmo di apprendimento agisce modificando la matrice dei pesi delle connessioni della rete neurale, al fine di minimizzare l'errore.

La regola Delta, in particolare, minimizza l'errore usando una tecnica di *discesa del gradiente*, ossia correggendo i pesi in misura proporzionale alla grandezza dell'errore.

$$\Delta W_{i,j} = \eta(t_i - y_i)x_j \quad (1)$$

L'equazione (1) indica il funzionamento della regola delta: il peso (W) tra l'unità di output i e l'unità di input j viene modificato di un valore uguale al tasso di apprendimento (η) moltiplicato per la differenza tra la risposta desiderata (t) e l'output effettivo (y), moltiplicati per l'attivazione dell'input (x).

Il tasso di apprendimento (η) è un numero compreso tra [0.0;1.0] che determina di quanto verrà modificato il peso. A tassi di apprendimento alti corrisponde una modifica ingente del peso in questione che, se da una parte può velocizzare il processo di apprendimento, dall'altra può portare a situazioni

oscillatorie, ossia situazioni in cui l'output prodotto si avvicina al *teaching input*, ma rimanendo lontano da una riduzione ottimale dell'errore.

4. Apprendimento sociale e apprendimento individuale

La simulazione che verrà presentata in questo capitolo è, per molti aspetti, simile a quella precedente. Come prima, si tratta di un semplice compito di riconoscimento di pattern (“funghi commestibili” e “funghi velenosi”, per continuare l’analogia) e di spostamento in un ambiente bidimensionale per raggiungere i primi ed evitare i secondi. In certi riguardi, il set-up è ancora più semplice: nell’ambiente è presente, alternativamente, solo uno degli oggetti e le popolazioni sono più piccole, composte da soli dieci agenti.

Tuttavia esistono, tra le altre, almeno due differenze importanti. La prima riguarda il fatto che, ora, la simulazione degli individui è improntata ad un maggiore realismo fisico e può essere implementata in un reale esperimento robotico (si veda il capitolo 2): come verrà mostrato più avanti, questo maggiore realismo avrà conseguenze decisive sull’efficacia del processo di apprendimento. La seconda è invece legata al particolare algoritmo di apprendimento, che in questo caso è ispirato al *simulated annealing* (vedi appendice 4.1). La scelta del *simulated annealing* è dovuta al fatto che questo algoritmo è abbastanza astratto e generalizzabile da poter essere utilizzato per simulare sia l’apprendimento sociale che l’apprendimento individuale, nonché la combinazione tra i due.

Il focus di questo esperimento riguarda i rapporti tra apprendimento individuale e sociale. La presenza di dinamiche di apprendimento sociale e la loro

efficacia adattiva è stata chiaramente dimostrata tra gli esseri umani e nel caso di molti altri vertebrati, sia empiricamente che con modelli teorici (Galef & Laland 2005, Boyd & Richerson 1985). Tuttavia, le condizioni specifiche in cui l'apprendimento sociale è effettivamente vantaggioso rispetto all'apprendimento individuale variano da caso e dipendono da numerosi fattori, quali le caratteristiche specifiche del comportamento da apprendere, la velocità in cui variano le condizioni ambientali, l'esistenza di necessari prerequisiti cognitivi e comportamentali, la quantità di strategie alternative presenti in una popolazione e così via. Per esempio, come dimostrato da un modello analitico dovuto all'antropologo Alan Rogers (1989) e ulteriormente sviluppato da Boyd e Richerson (1995), imitare il comportamento degli altri è adattivo solo quando l'imitazione è relativamente poco diffusa e molti individui acquisiscono i propri tratti comportamentali tramite apprendimento individuale. La logica dell'argomento è abbastanza semplice: si immagina un ambiente che varia nel tempo tra due stati (per esempio caldo / freddo) ai quali corrispondono due comportamenti adattivi differenti. L'apprendimento individuale è relativamente più costoso dell'apprendimento sociale, ma permette di tenere traccia dei cambiamenti ambientali; l'apprendimento sociale è meno costoso, ma si corre il rischio di copiare un individuo che non si è ancora adattato al nuovo cambiamento ambientale. Gli imitatori sono, in questo modello, dei *free-rider* che approfittano degli individui che hanno imparato individualmente. Se sono pochi, la *fitness* della popolazione rimane buona, ma, all'aumentare del numero degli imitatori, la *fitness* complessiva tenderà a scendere, finché possiamo immaginare una situazione in cui una popolazione composta interamente da imitatori non avrà nessuna possibilità di adattarsi alle condizioni ambientali mutevoli.

Una possibile soluzione a questo paradosso risiede nella possibilità per gli imitatori di non essere semplicemente dei “copiatori di comportamenti” passivi, ma di utilizzare in modo critico i vantaggi legati alla relativa convenienza dell'apprendimento sociale aggiustandolo, se necessario, con l'apprendimento individuale. Per esempio, essi potrebbero utilizzare all'inizio l'apprendimento sociale, in quanto meno costoso, ma poi valutare se il comportamento acquisito è adattivo nella particolare condizione ambientale in cui si trovano. Nel caso in cui non lo sia, essi potrebbero ricorrere all'apprendimento individuale (Enquist, Eriksson & Ghirlanda, *working paper*). Nella simulazione qui presentata viene mostrato come la stessa necessità – far seguire all'apprendimento sociale una fase di apprendimento

individuale – sia legata anche alle caratteristiche intrinseche del comportamento da apprendere. In particolare, almeno nel caso in cui l'apprendimento sociale consiste nel tentativo di copiare “a livello fine” il comportamento mostrato da un modello, una simulazione realistica di questa dinamica mostra come differenze minime tra il *mapping* sensomotorio prodotto dal modello e quello prodotto dall'osservatore possono condurre, nel caso di comportamenti continui e sequenziali, a comportamenti differenti, che influiscono in modo significativo sulle prestazioni individuali.

Come sarà chiarito in seguito, in questa simulazione si intende per apprendimento sociale e apprendimento individuale un processo di apprendimento in cui le variazioni dei parametri sono mantenute o scartate, rispettivamente, sulla base del fatto che riducono le differenze tra il *mapping* sensomotorio dei due individui coinvolti nella trasmissione culturale (in modo analogo a quanto avveniva nella simulazione descritta nel capitolo precedente) o sulla base del fatto che migliorano la prestazione dell'individuo.

Set – up sperimentale

Si consideri il caso di un agente che vive in un ambiente che contiene funghi commestibili e velenosi (oggetti α e oggetti β) e che debba essere in grado di raggiungere gli oggetti α e di evitare gli oggetti β . Esistono due varietà diverse di oggetti α e due varietà diverse di oggetti β e ognuna di queste varietà ha differenti proprietà percettive.

Gli agenti “vivono” per 200 epoche, ed ogni epoca dura 50 cicli. All'inizio di ogni epoca, l'agente è posizionato in un'arena quadrata (il lato dell'arena è considerato l'unità di misura di riferimento: lato = 1.0 u.), che contiene nel suo centro un singolo oggetto, scelto a caso fra i quattro oggetti possibili. L'agente è posto in una posizione casuale, ma ad una distanza minima di 0.5 u. dall'oggetto (Fig. 4.1). La prestazione dell'agente è determinata assegnando un punteggio di +0.01 e -0.01 per ogni epoca in cui raggiunge, rispettivamente, un oggetto α o un oggetto β . Poiché durante la vita ogni agente trova 100 oggetti per ciascun tipo, la sua performance può variare tra [-1.0;+1.0].

Il comportamento degli agenti è controllato da una rete neurale provvista di 8

unità di input, 5 unità interne e 2 unità di output (Fig. 4.2). L'architettura della rete neurale è composta da due moduli: (1) un modulo A che riceve come input le proprietà percettive degli oggetti ("11" o "00" nel caso degli oggetti α ; "10" o "01" nel caso degli oggetti β) e che produce come output la categorizzazione dell'oggetto e (2) un modulo B che riceve come input l'output del modulo A e i dati di una videocamera e che produce come output l'azione motoria.

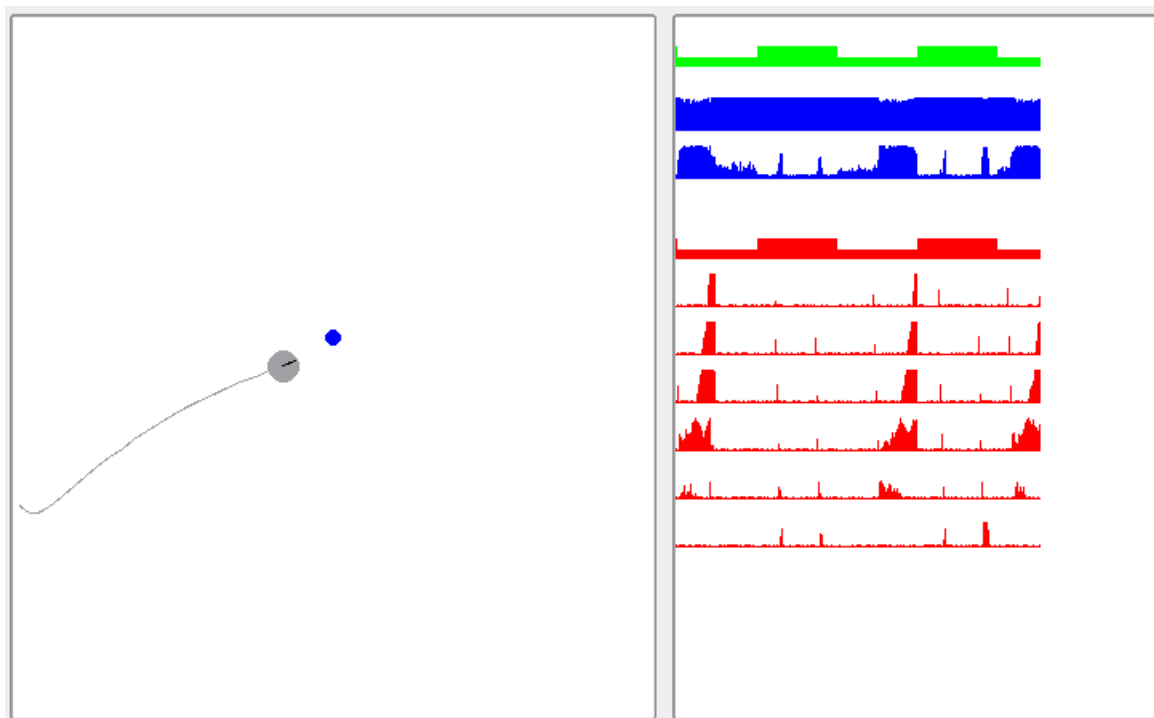


Fig. 4.1: L'arena e un comportamento tipico esibito da un agente in grado di risolvere il compito nel caso di un oggetto α . Sulla destra sono tracciate alcune attivazioni della rete nel corso del tempo: dall'alto, la prima rappresenta l'output del modulo A, la seconda e la terza gli output motori (modulo B) e le ultime sette gli input del modulo B (la prima corrisponde all'output del modulo A e le ultime sei alla videocamera).

La videocamera è una videocamera lineare con un'ampiezza di 60° , fornita di 60 fotorecettori con un angolo di 1° ciascuno. Ogni fotorecettore è attivato in modo binario (1/0) a seconda della presenza o dell'assenza di un oggetto nel suo campo visivo. Ognuno dei sei input del modulo B codifica l'attivazione media dei 10 fotorecettori corrispondenti. Il rumore è simulato aggiungendo, per ciascun input, un

valore compreso tra $[0.0;0.05]$ al valore dell'attivazione. I due output del modulo B controllano due ruote simulate: l'attivazione dei due output è normalizzata tra i valori $[-0.03;+0.03]$, che corrispondono alla distanza massima coperta dalle ruote in ogni passo della simulazione.

Si noti che la struttura modulare dell'architettura di controllo corrisponde al fatto che il compito coinvolge due abilità qualitativamente differenti: l'abilità di categorizzare gli oggetti (modulo A) e quella di muoversi nell'ambiente (modulo B). La prima abilità rappresenta un compito astratto e discreto (si tratta, in pratica, del classico problema dello XOR), mentre la seconda richiede la capacità di produrre un comportamento continuo e sequenziale.

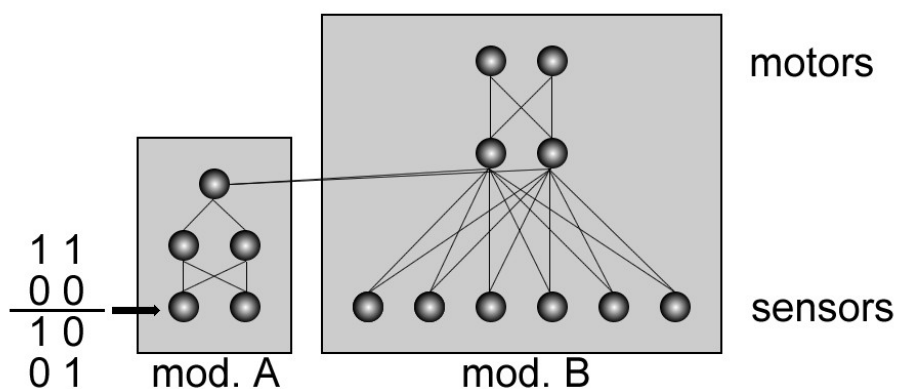


Fig. 4.2: L'architettura neurale di controllo degli agenti.

L'algoritmo di apprendimento

Come anticipato, l'algoritmo di apprendimento utilizzato è basato sull'algoritmo di *simulated annealing*. La procedura, in dettaglio, è la seguente: dopo avere inizializzato i parametri dell'architettura neurale di controllo degli agenti (ossia i pesi delle connessioni) con dei valori casuali compresi tra $[-5.0;+5.0]$,

1. i parametri vengono sottoposti ad un processo di mutazione: ogni parametro ha una probabilità del 5 % di venire sostituito da un valore casuale compreso tra $[-5.0;+5.0]$.
2. la prestazione dell'agente viene valutata sul periodo della vita (200

epoche) e una funzione di valutazione specifica se la mutazione viene mantenuta o scartata. Per ridurre i costi computazionali, le variazioni non adattive vengono scartate dopo le prime 20 epoche.

3. i passi (1) e (2) vengono ripetuti per 100.000 cicli di apprendimento.

Le mutazioni che portano ad un risultato positivo o uguale al precedente (secondo la funzione di valutazione) vengono sempre mantenute. Le mutazioni negative, invece, vengono mantenute con una probabilità che è inversamente proporzionale ai loro effetti negativi e che viene progressivamente ridotta durante il processo di apprendimento. Il fatto di accettare *alcune* mutazioni negative, soprattutto nella prima fase dell'apprendimento, è una caratteristica fondamentale dell'algoritmo di *simulated annealing* ed è motivato dalla necessità di ampliare lo spazio di ricerca e di evitare di rimanere bloccati in minimi locali, trovando soluzioni che sono facili da raggiungere, per una data configurazione di pesi, ma subottimali (vedi appendice 4.1).

Nel caso dell'apprendimento individuale, la funzione di valutazione misura la prestazione dell'agente: le mutazioni sono mantenute o scartate sulla base del loro impatto sulla capacità dell'agente di riconoscere gli oggetti α e β e di sviluppare un comportamento corretto rispetto ad essi (avvicinarsi agli oggetti α ed evitare gli oggetti β).

Risultati (apprendimento individuale)

L'esperimento è stato replicato per 10 individui: la prestazione varia in modo considerevole nelle diverse repliche (vedi Fig. 4.3), mostrando che il compito è difficile da imparare con un algoritmo di apprendimento individuale. Le analisi della prestazione durante il processo di apprendimento mostrano come le due abilità (categorizzazione e sensomotoria) siano, se prese singolarmente, relativamente facili da apprendere, ma che proprio la loro combinazione, come ipotizzato nella costruzione dell'esperimento, rende il compito complesso. Infatti, le mutazioni del sistema di controllo che migliorano la categorizzazione non sono di per sè adattive se non sono collegate a mutazioni che migliorano le abilità sensomotorie e viceversa.

I dati ottenuti mostrano come, per esempio, una categorizzazione corretta degli oggetti (cioè una categorizzazione in cui gli oggetti α , pur avendo caratteristiche

percettive diverse, vengono categorizzati in modo simile nello spazio rappresentazionale e in modo diverso dagli oggetti β , che a loro volta devono essere simili tra loro) venga ottenuta, in pratica, solo in due casi, nelle repliche 7 e 10 (Fig. 4.4) che corrispondono, in effetti, alle migliori prestazioni (Fig. 4.3).

A questo punto, è possibile verificare se un algoritmo di apprendimento sociale, usando come modello il migliore degli individui prodotti tramite l'apprendimento individuale, sia in grado di migliorare queste prestazioni.

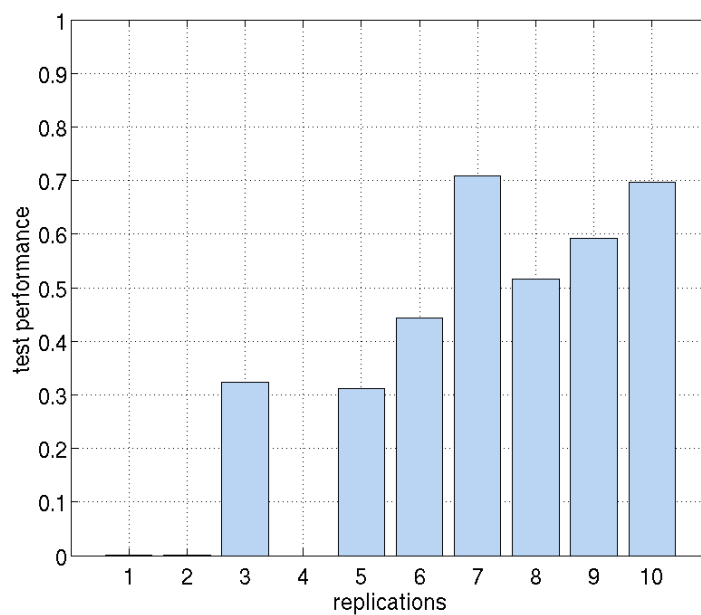


Fig. 4.3: Prestazione per 10 repliche dell'esperimento usando un algoritmo di apprendimento individuale. Le prestazioni sono ottenute testando l'agente alla fine del procedimento di *simulated annealing* testandoli per 1000 epoche.

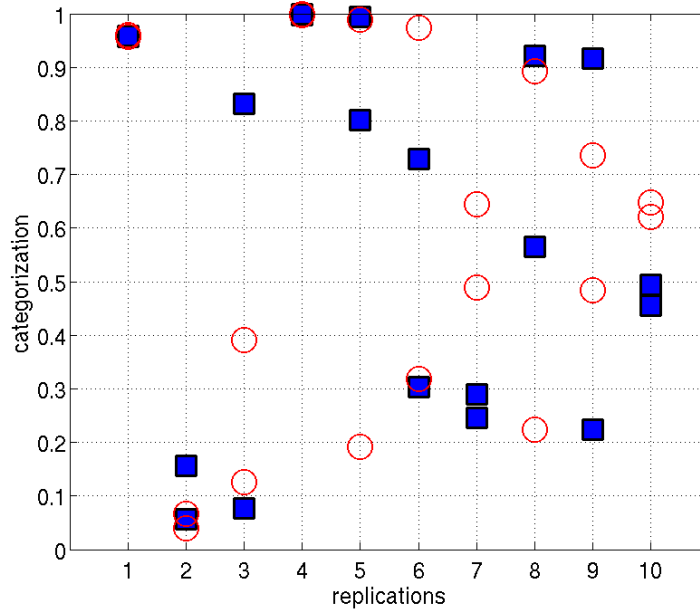


Fig. 4.4. Output del modulo di categorizzazione per i quattro oggetti. Ogni colonna rappresenta la categorizzazione prodotta da un individuo per ogni replica dell’esperimento alla fine del ciclo di apprendimento. I cerchi e i quadrati indicano rispettivamente gli oggetti α e gli oggetti β .

Apprendimento sociale

Per simulare l’apprendimento sociale è stato utilizzato il medesimo algoritmo di apprendimento. Questa scelta porta a due vantaggi: da una parte, i risultati delle due condizioni (apprendimento sociale e apprendimento individuale) sono confrontabili in modo più evidente e dall’altra, come verrà mostrato in seguito, questo permette di realizzare un algoritmo “misto”, in cui, ad una prima parte in cui gli agenti apprendono socialmente, segue una fase in cui il comportamento appreso viene modificato tramite apprendimento individuale.

Nel caso dell’apprendimento sociale, però, la funzione di valutazione non riguarda più la prestazione dell’agente, ma la somiglianza del suo comportamento rispetto al comportamento di un altro agente, che serve come modello. In particolare, la funzione di valutazione misura, ad ogni ciclo, la differenza tra gli output prodotti dal modello e quelli prodotti dall’osservatore (e quindi, come si ricorderà, è, da

questo punto di vista, del tutto analoga a quella descritta nel capitolo precedente). Mentre nella simulazione del capitolo precedente, tuttavia, al modello e all'osservatore venivano presentati dei pattern di input generati casualmente, in questo caso il modello e l'osservatore si muovono realmente nell'ambiente e, ad ogni ciclo, l'osservatore, a prescindere da come si muova, viene spostato nella posizione in cui si trova il modello, in modo da ricevere gli stessi input¹.

Poiché la struttura di controllo neurale degli agenti è composta da due moduli sono possibili tre diverse condizioni di confronto tra gli output: (1) l'apprendimento sociale tende a minimizzare la differenza tra gli output motori (gli output del modulo B), (2) l'apprendimento sociale tende a minimizzare sia la differenza tra gli output motori sia la differenza tra gli output di categorizzazione (gli output del modulo A) e, infine, (3) l'apprendimento tende a minimizzare la differenza tra gli output di categorizzazione (Fig. 4.5).

Il processo di apprendimento "misto" sociale / individuale funziona invece in questo modo: durante la prima fase, che dura metà dell'algoritmo di apprendimento (ossia 50.000 cicli), l'apprendimento è esclusivamente sociale, ossia agisce per minimizzare le differenze tra gli output del modello e gli output prodotti dall'osservatore. Durante la seconda fase, invece, le mutazioni sono mantenute o scartate sulla base della somma pesata dei due tipi di apprendimento, cioè sulla differenza tra gli output dell'osservatore e gli output del modello (apprendimento sociale) e sulla prestazione dell'agente (apprendimento individuale). Questa somma pesata (Δw) è calcolata in accordo con le seguenti equazioni:

¹ E' necessario notare che in una situazione davvero realistica gli output prodotti dal modello non potrebbero essere accessibili direttamente all'osservatore, come avviene in questo caso, ma andrebbero ricostruiti sulla base degli stati sensoriali dell'osservatore, i quali fornirebbero un'informazione solo indiretta e spesso incompleta. Questo implica che, nel nostro caso, gli agenti non devono risolvere il cosiddetto "problema della corrispondenza" (Dautenhahn & Nehaniv 2002), la cui soluzione consiste nell'essere in grado di traslare queste informazioni indirette e incomplete su quello che il modello sta facendo negli opportuni output motori. Tuttavia, ciò che è interessante, nello scenario qui presentato, è che anche *senza* dovere risolvere il problema della corrispondenza un algoritmo di apprendimento sociale che si basi semplicemente sulla copia del comportamento di un modello possa non essere adattivo. Si ritornerà diffusamente su questi aspetti alla fine del capitolo.

$$\Delta w = f(i)W + f(S)(1 - W) \quad (1)$$

$$W = C/C_{max} \quad (2)$$

dove $f(I)$ indica la funzione di valutazione dell'apprendimento individuale, $f(S)$ indica la funzione di valutazione dell'apprendimento sociale, C è il ciclo di apprendimento durante la seconda fase (che varia da 0 a 50.000) e C_{max} è una costante uguale a 50.000.

In altre parole, dopo una prima fase in cui l'agente si limita a “copiare” il comportamento del modello, durante la seconda metà dell'apprendimento le sue mutazioni vengono considerate positive se, eventualmente discostandosi in modo graduale dal comportamento del modello, ne migliorano comunque la prestazione.

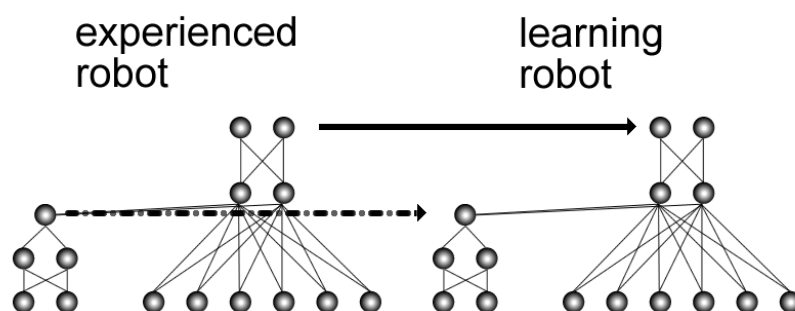


Fig. 4.5: Un modello schematico del processo di apprendimento sociale. L'osservatore - *learning robot* - approssima gli output motori del modello - *experienced robot* - (freccia piena) e/o la sua categorizzazione (freccia punteggiata).

In conclusione, esistono quindi sei condizioni sperimentali in cui l'apprendimento sociale viene testato:

1. SL1: apprendimento sociale (approssimazione output motori)
2. SL2: apprendimento sociale (approssimazione output motori e categorizzazione)
3. SL3: apprendimento sociale (approssimazione categorizzazione)
4. SIL1: apprendimento “misto” (approssimazione output motori)
5. SIL2: apprendimento “misto” (approssimazione output motori e

categorizzazione)

6. SIL3: apprendimento “misto” (approssimazione categorizzazione)²

Risultati (apprendimento sociale)

I risultati mostrano che, in effetti, la possibilità di apprendere socialmente da un modello migliora sensibilmente la prestazione media della popolazione rispetto al caso in cui lo sviluppo del comportamento era lasciato al solo apprendimento individuale.

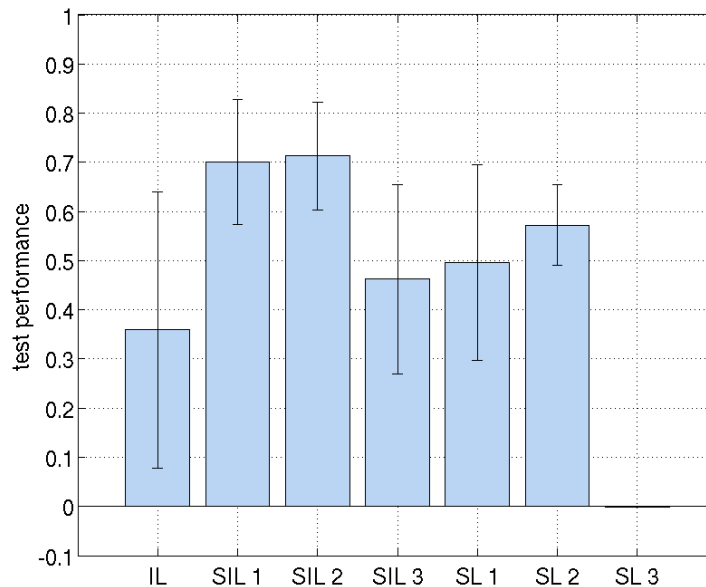


Fig. 4.6: Prestazione media e deviazione standard per 10 replicazioni dell’esperimento nelle condizioni descritte nel testo. La prima condizione (IL) si riferisce all’apprendimento individuale.

Le prestazioni medie di tutte le condizioni sperimentali in cui è presente una qualche forma di apprendimento sociale (tranne SL3³) sono infatti superiori rispetto

² SL sta per *social learning*; SIL per *social - individual learning*.

³ Si noti che la condizione SL3, in cui gli agenti approssimano solo la categorizzazione del modello e non imparano individualmente non può essere mai adattiva, in quanto gli agenti non hanno, per definizione, possibilità di apprendere il compito rispetto al quale vengono valutati, ossia raggiungere i funghi commestibili ed evitare quelli velenosi.

alla condizione in cui è presente solo l'apprendimento individuale (IL). Tuttavia è possibile notare una differenza significativa fra le condizioni in cui gli agenti apprendono prima socialmente e poi individualmente (SIL1 e SIL2) e le condizioni in cui gli agenti apprendono solo socialmente (SL1 e SL2): nei primi casi, infatti, le prestazioni sono migliori, sia dal punto di vista della media delle differenti repliche, sia da quello del migliore individuo ottenuto. A cosa è dovuta questa differenza tra apprendimento “misto” e apprendimento solo sociale?

Il fatto che l'apprendimento sociale da solo conduca a prestazioni peggiori rispetto alla combinazione di apprendimento sociale ed individuale può essere spiegato considerando che la riduzione della differenza tra gli output prodotti dal modello e gli output prodotti dall'osservatore non produce necessariamente un miglioramento della prestazione di quest'ultimo (Fig. 4.7).

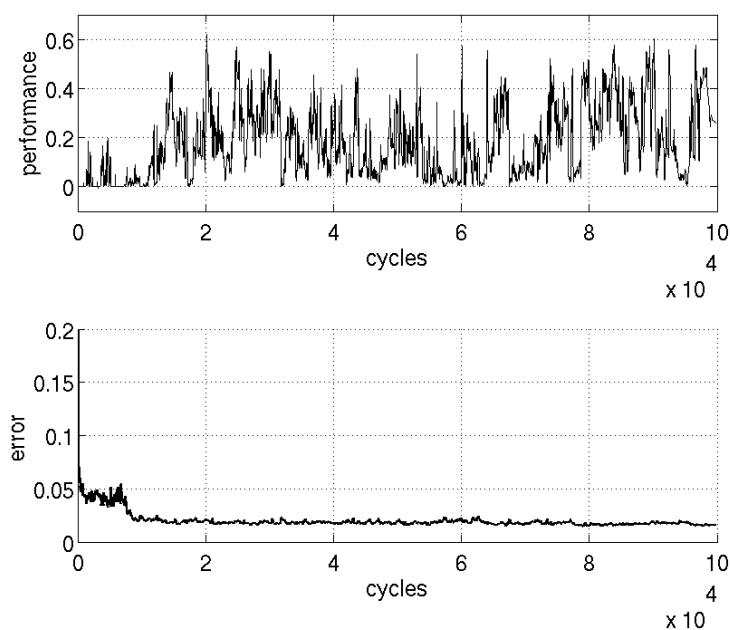


Fig. 4.7: Andamento della differenza tra gli output del modello e gli output dell'osservatore (errore) e prestazione dell'osservatore durante il processo di apprendimento sociale.

Infatti, piccole differenze tra gli output delle due reti neurali (errore < 0.05) possono corrispondere a notevoli differenze di prestazione. Questo fatto, a sua volta, può essere spiegato tenendo in considerazione che abilità comportamentali sequenziali, come quella richiesta nell'esperimento qui presentato, sono proprietà dinamiche che risultano da una sequenza di interazioni non lineari tra l'agente e

l'ambiente e in cui piccole differenze al livello di queste interazioni possono determinare, al livello del comportamento risultante, significative differenze (Chiel & Beer 1997, Nolfi 2006).

In altri termini, come approfondiremo alla fine del capitolo, anche fornendo agli osservatori la conoscenza diretta degli output del modello e un algoritmo di copia efficiente, la trasmissione sociale, almeno nel caso di comportamenti continui, non può essere considerata come un semplice processo passivo di copia, ma deve essere considerata come un processo attivo, in cui le informazioni ricavate dal comportamento del modello vengono utilizzate per *ricostruire* il comportamento corretto.

Trasmissione culturale iterata

La differenza fra le due condizioni (apprendimento sociale ed individuale – apprendimento sociale) diviene ancora più evidente iterando il processo di apprendimento. In sostanza, si tratta di applicare un meccanismo simile a quello descritto nel capitolo precedente a questo nuovo scenario: il processo di apprendimento viene ripetuto per un certo numero di generazioni, selezionando come modello per la generazione $g + 1$ il migliore individuo della generazione g ed utilizzando, ad ogni generazione, agenti con pesi inizializzati in modo casuale. Quindi, come nel capitolo precedente, non esiste trasmissione genetica dei comportamenti, e, ora, nemmeno orizzontale (intragenerazionale): semplicemente, si tratta di verificare se il processo di apprendimento sociale è robusto o meno, ossia, se ripetendolo per più generazioni esso tende a mantenere una buona prestazione della popolazione (o a migliorarla) oppure no.

Ogni generazione è composta da 10 individui, e il processo di iterazione viene ripetuto 10 volte. Le condizioni sperimentali prese in considerazione sono due: apprendimento sociale ed individuale con approssimazione degli output motori e della categorizzazione (SIL2) ed apprendimento solo sociale con approssimazione degli output motori e della categorizzazione (SL2), cioè, in entrambi i casi, sono state iterate le dinamiche che nel singolo processo di apprendimento avevano prodotto le prestazioni migliori. Tutti gli altri parametri sono stati mantenuti come descritti nelle

sezioni precedenti.

I risultati mostrano che, nella condizione in cui gli agenti apprendono socialmente ed individualmente, le abilità sviluppate dal migliore individuo della generazione precedente si trasmettono con successo agli individui delle generazioni successive (Fig. 4.8), mentre, come anticipato, nella condizione in cui gli agenti apprendono solo socialmente, la prestazione tende a deteriorarsi nel corso delle generazioni (Fig. 4.9)⁴.

⁴ I risultati sono apparentemente simili a quelli descritti nel capitolo precedente nel caso della trasmissione culturale senza rumore con *learning rate* bassi, per cui ci si potrebbe legittimamente domandare se il decadimento della prestazione nel caso SL2 non sia semplicemente dovuto all'assenza di rumore. E' questa una domanda interessante, ma è necessario tenere conto del fatto che in questo scenario esiste una forma di rumore intrinseco: come si ricorderà, ai sensori collegati alla camera viene sempre aggiunto un valore casuale compreso tra [0.0;0.05], per cui anche se modello e osservatore sono nella stessa posizione non riceveranno mai esattamente gli stessi input. Tuttavia, credo che la questione sia più profonda. L'idea è che una forma di rumore statistico possa funzionare come "motore" per la variabilità culturale nel caso di compiti *discreti* – per esempio un compito di categorizzazione, o di apprendimento linguistico – ossia compiti in cui una piccola differenza rispetto al modello dà in genere lo stesso tipo di risultato, fino a quando, superata una soglia critica, dà un risultato completamente differente. Nel caso qui descritto, per esempio, il modulo di categorizzazione svolge un compito di questo genere. Tuttavia, nel caso di compiti *continui*, come quello svolto dal modulo sensomotorio, come abbiamo visto, piccole differenze al livello micro delle interazione organismo - ambiente, possono portare a notevoli differenze nel comportamento risultante. Per questo motivo, in casi simili credo sia possibile affermare che, *in generale*, un tipo di apprendimento basato esclusivamente sulla copia esplicita del comportamento di un modello debba non essere efficace. Acerbi e Nolfi (2006) interpretavano, in effetti, i risultati in questa direzione, ma ritengo siano necessarie ulteriori verifiche per sostenere questa ipotesi.

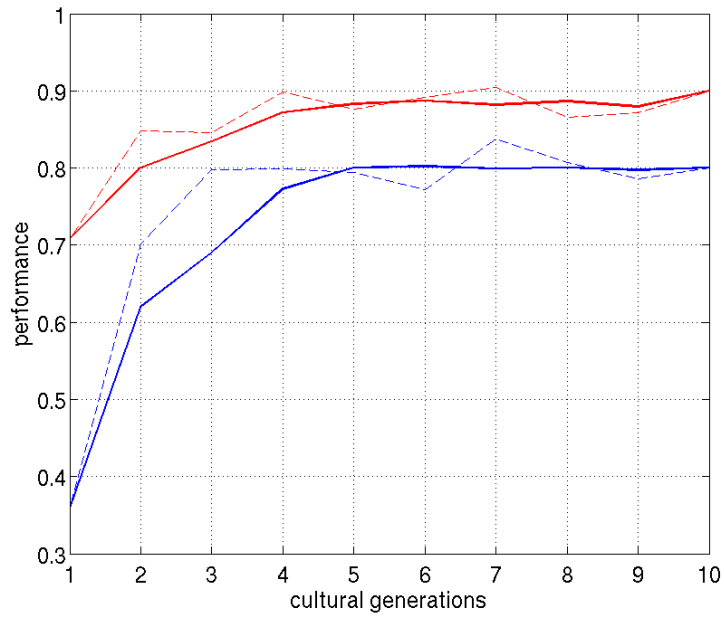


Fig. 4.8: Prestazione massima e media nel processo di trasmissione culturale iterata nel caso in cui gli agenti apprendono tramite una combinazione di apprendimento sociale ed individuale (la linea piana rappresenta la media mobile dei dati con periodo = 3).

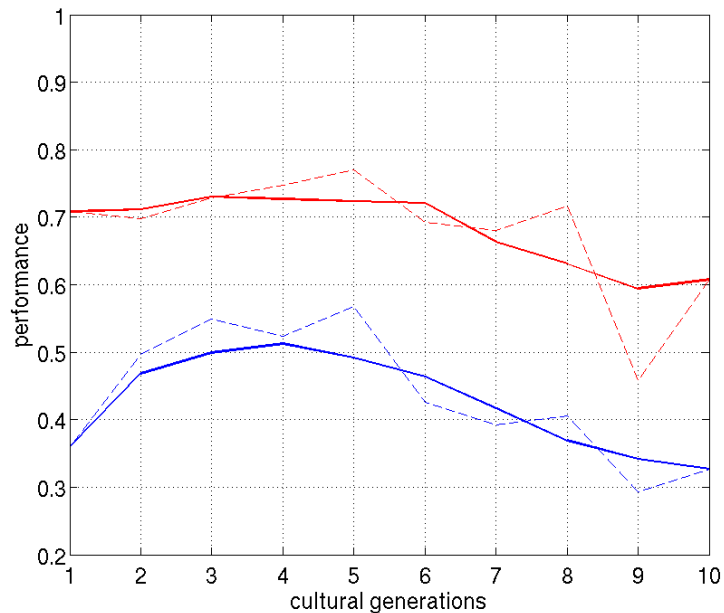


Fig. 4.9: Prestazione massima e media nel processo di trasmissione culturale iterata nel caso in cui gli agenti apprendono tramite apprendimento sociale (la linea piana rappresenta la media mobile dei dati con periodo = 3).

Nel primo caso, inoltre, la crescita della prestazione media e della prestazione del migliore agente indica che non solo il comportamento sviluppato individualmente dagli agenti della prima generazione si è trasmesso e diffuso con successo, ma anche che l'iterazione della trasmissione culturale ha generato una forma di evoluzione culturale *cumulativa*, ossa in cui i tratti o alcune cratteristiche dei tratti culturalmente trasmessi si modificano nel corso delle generazioni e tendono ad essere trasmessi e conservati nel tempo.

Queste conclusioni sono ulteriormente dimostrate dalla comparazione della prestazione di 100 agenti che sono stati sottoposti a: (1) un processo di apprendimento individuale, (2) un processo di apprendimento sociale ed individuale non iterato (cioè gli agenti della generazione 2) e (3) un processo di apprendimento individuale iterato (cioè gli agenti della generazione 10).

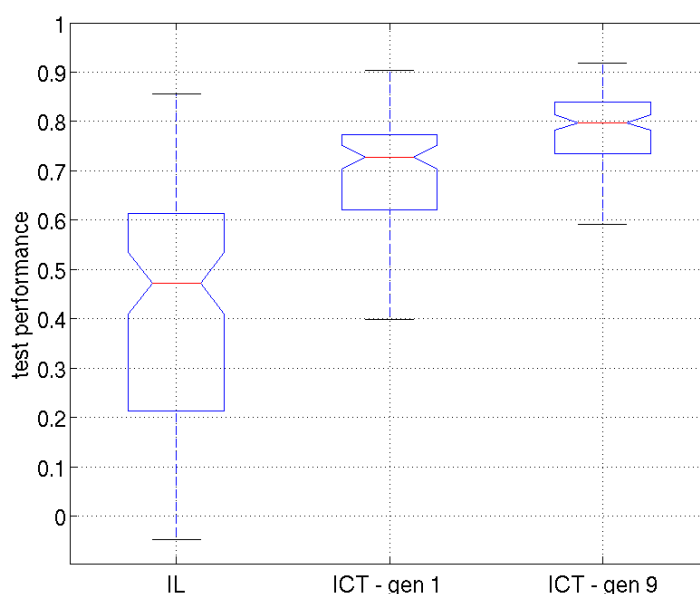


Fig. 4.10: Prestazione di 100 agenti nelle seguenti condizioni sperimentali: (IL) apprendimento individuale, (ICT – gen1) apprendimento sociale e individuale non iterato, (ICT – gen 10) apprendimento sociale ed individuale iterato. Il grafico (detto *box-plot*) va letto in questo modo: gli estremi dei *box* sono il primo e il terzo quartile delle distribuzioni, mentre la linea centrale è la mediana. Le linee orizzontali, esterne ai *box*, rrpresentano il valore massimo e il valore minimo.

Come mostra la figura 4.10, infatti, gli agenti sottoposti ad apprendimento iterato ottengono migliori prestazioni non solo degli agenti sottoposti ad apprendimento individuale, ma anche degli agenti sottoposti ad un processo di trasmissione culturale non iterata. Le differenze tra le tre condizioni sono significative rispetto ad un'analisi non parametrica della varianza (test di Kruskal – Wallis, $p < 0.001$).

Per identificare la natura delle innovazioni che hanno avuto luogo durante il processo di trasmissione culturale iterata, sono state misurate le percentuali degli elementi α e β raggiunti (Fig. 4.11) ed il tempo medio necessario per raggiungere gli oggetti α (Fig. 4.12) per gli agenti della generazione 2 e per gli agenti della generazione 10 (ossia per agenti che non sono o sono stati sottoposti all'iterazione del processo di trasmissione culturale). I risultati mostrano che la prestazione si differenzia sia rispetto all'abilità di produrre un comportamento corretto (raggiungere gli oggetti α ed evitare gli oggetti β) sia rispetto alla velocità con cui si raggiungono gli oggetti α , in modo da ridurre il rischio di non raggiungere l'oggetto nel tempo di 50 cicli, che è a disposizione degli agenti. Il fatto che gli agenti della generazione 10 raggiungono gli oggetti α più velocemente di quanto riescano a fare gli agenti della generazione 2 (Fig. 12) dimostra che, nel corso delle generazioni, ci sono stati progressi per quanto riguarda le abilità motorie. Il fatto che gli agenti della generazione 2 tendano a raggiungere, erroneamente, anche l'oggetto β_2 (ossia uno dei due "cibi velenosi") in circa il 10% dei casi, mentre questo non avviene nella generazione 10, dimostra come il miglioramento sia avvenuto anche rispetto alle capacità di categorizzazione (in particolare, in questo caso, rispetto alla capacità di trattare nello stesso modo gli oggetti β_1 e β_2).

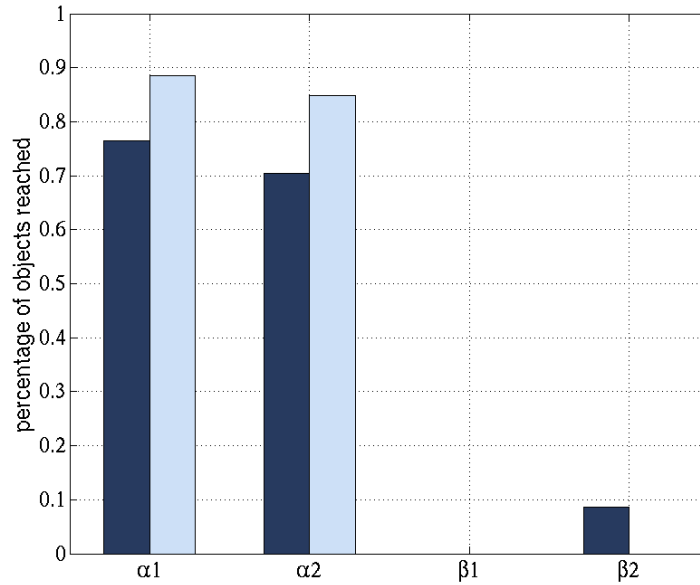


Fig. 4.11: Percentuale di oggetti raggiunti in 1000 repliche dell'esperimento. L'istogramma scuro rappresenta la percentuale per gli agenti della generazione 2, quello chiaro la percentuale per gli agenti della generazione 10.

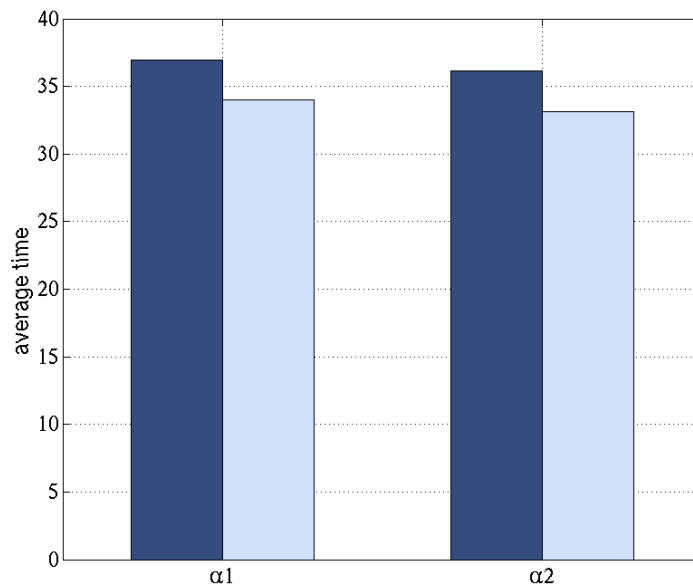


Fig. 4.12: Tempo medio (in passi della simulazione) in cui gli agenti raggiungono i due oggetti α . L'istogramma scuro rappresenta la percentuale per gli agenti della generazione 2, quello chiaro la percentuale per gli agenti della generazione 10.

Cos'è la “trasmissione culturale”?

I risultati presentati in questo capitolo suggeriscono che i reali processi di apprendimento sociale e di trasmissione culturale non possano essere considerati processi in cui un comportamento viene semplicemente *copiato* da un modello e riprodotto da un osservatore.

Nella maggior parte dei casi, sembra che, in antropologia, si tenda a pensare i processi di trasmissione culturale e di apprendimento sociale come processi *trasparenti*. Con questo intendo dire che, una volta osservato un pattern comportamentale, la *spiegazione* di questo viene ricercata nei rapporti che questo pattern intrattiene con altri fenomeni culturali, senza tenere in particolare considerazione i meccanismi a livello delle interazioni individuali. Semmai, viene fatto riferimento ad un concetto di imitazione abbastanza vago, ma entrando raramente nei dettagli relativi a come questo processo venga in pratica istanziato.

Negli ultimi decenni, tuttavia, si sono sviluppati importanti settori di ricerca che si situano ai confini tra antropologia, psicologia, etologia e biologia evoluzionistica, che hanno seriamente messo in questione questo tipo di impostazione. Psicologi evoluzionistici e antropologi cognitivi hanno proposto che la trasmissione culturale sia vincolata in modo importante da strutture cognitive comuni a tutti gli esseri umani (Sperber 1999) o, in modo più estremo, che molte di quelle forme comportamentali, che vengono definite come culturali, siano in realtà frutto della variazione ambientale che evoca risposte differenti, a partire da questa struttura cognitiva comune; risposte comunque iscritte nel patrimonio genetico (Tooby & Cosmides 1992). Oppure, l'imitazione, lungi da essere considerata un processo trasparente, è oggi al centro dell'interesse di ricercatori provenienti da numerose discipline che cercano di definirne le caratteristiche peculiari, la diffusione in specie diverse dagli esseri umani o i meccanismi neurali che la rendono possibile (una rassegna recente di queste problematiche è: Hurley & Chater 2005). Ancora, gli etologi hanno mostrato come molti fenomeni che possono essere descritti, a livello popolazionele, come esempi di trasmissione culturale nel regno animale e che erano stati interpretati come forme di imitazione, sono determinati da diverse forme di interazione più semplici, che possono essere descritte come apprendimento individuale socialmente guidato (Visalberghi & Frigaszy 1990; Galef & Laland

2005): non è irragionevole pensare, se non altro facendo riferimento al principio di parsimonia, che queste possano avere un ruolo importante anche nella nostra specie.

E' possibile, anche se non è la mia opinione, che la migliore strategia si rivelerà un giorno quella di astrarre da questi particolari, continuando a focalizzare le analisi antropologiche al livello macro delle dinamiche culturali. Personalmente lo ritengo improbabile: il dibattito che ha caratterizzato molte scienze sociali (si pensi all'economia) rispetto all'opportunità di studiare i fenomeni a livello macro o micro - forse in antropologia più conosciuto come contrasto tra olismo e individualismo metodologico (Sperber 1997; si veda anche: Boyd & Richerson 2005) - pare spesso sterile. I “fenomeni naturali complessi”, come la cultura, oltre ad essere difficili da catturare con definizioni semplici, hanno la caratteristica - spiacevole, forse, per chi cerca di comprenderli dal punto di vista scientifico - di presentare legami complicati tra ciò che avviene ad un livello (nel nostro caso, le interazioni individuali) e ciò che avviene ad un altro (le dinamiche culturali): forse dobbiamo mettere in conto di occuparci di entrambi e di sviluppare strumenti metodologici che ci permettano di farlo.

Jukebox, Camino del Diablo e attrattori: metafore per la cultura

Gli psicologi evuzionistici John Tooby e Leda Cosmides propongono ai propri lettori un esperimento di pensiero (Tooby & Cosmides 1992, pp. 115-116): chiedono di immaginare una “popolazione” di jukebox identici tra loro, forniti di un ampio repertorio di pezzi musicali e di un meccanismo che permette di stabilire la latitudine e la longitudine. I pezzi musicali che vengono suonati sono determinati dall'output di questo meccanismo. Distribuendo questi jukebox sulla superficie terrestre osserveremmo delle variazioni geografiche rispetto alle melodie suonate: tutti i jukebox posizionati, per ipotesi, a Rio suonerebbero una melodia uguale tra loro ma differente rispetto ai jukebox posizionati, per esempio, a Pechino. Allo stesso modo, se un jukebox venisse spostato da Rio a Pechino cambierebbe la propria melodia, adeguandosi, per così dire, a quelle prodotte dai jukebox locali, poiché il suo meccanismo di rilevazione geografica gli fornirebbe il nuovo output.

Nell'ottica di molti psicologi evuzionistici, gli antropologi hanno, in genere, sovrastimato l'importanza dei processi di trasmissione e apprendimento sociale nella spiegazione della variabilità comportamentale umana. Se, come essi sostengono, ogni essere umano è fornito di una ricca struttura cognitiva innata e sensibile alle variazioni ambientali, evoluta per selezione naturale, molte delle differenze etichettate come "culturali" sarebbero semplicemente dovute alle complesse interazioni tra questa struttura e le condizioni ambientali: Tooby e Cosmides definiscono questi pattern comportamentali come "cultura evocata" (*evoked culture*). Anche se ammettono che lo scenario dei jukebox è un caso irrealisticamente estremo e che viene proposto più che altro come un'immagine speculare dello scenario proposto dall'antropologia *mainstream*, in cui la diffusione culturale è sostanzialmente libera da vincoli (che ritengono irrealistico nella stessa misura), l'esperimento di pensiero di Tooby e Cosmides fornisce spunti interessanti.

I comportamenti etichettabili come cultura evocata non richiedono, in realtà, nessuna forma di apprendimento o di trasmissione, usando i termini nel loro significato di senso comune; di più, non richiedono nemmeno nessuna forma di interazione tra gli individui. Anche assumendo una definizione intuitiva, rimane il fatto che perché un dato comportamento possa essere considerato come "culturale", l'emergere di quel comportamento in un individuo deve essere preceduto da un'interazione con l'ambiente in un senso un po' più ampio rispetto ai jukebox degli psicologi evuzionistici e, in particolare, da un'interazione con gli altri individui presenti in questo ambiente.

Se fornissimo i jukebox di un microfono e di un meccanismo che, al posto che rilevare la posizione geografica, potesse rilevare le frequenze delle prime battute delle canzoni che vengono suonate nelle vicinanze e, in base a queste, cercare nel proprio repertorio le melodie più simili, saremmo più disposti a chiamare culturale il comportamento della nostra popolazione di jukebox? Forse, ma probabilmente mancherebbe ancora qualcosa. Una dinamica di questo genere viene definita come *stimulus priming*: un esempio, nel caso degli esseri umani, noto a tutti, è quello dello sbadiglio. La presenza di altri individui che sbadigliano provoca in breve la replicazione dello stesso comportamento. Un etologo un po' ingenuo, o semplicemente estremamente rigoroso, basandosi su di una definizione secondo cui le condizioni sufficienti e necessarie per la trasmissione culturale sono l'assenza di un dato comportamento, l'interazione e poi la sua replicazione, dovrebbe ammettere di

trovarsi di fronte ad un caso paradigmatico di questo fenomeno. Allora, qualcosa sembra essere andato storto. Intuitivamente vorremmo che il comportamento replicato sia *nuovo*: questa condizione, che sembra banale, è in realtà alquanto difficile da definire con precisione, ma racchiude probabilmente un'idea importante.

Un nuovo esperimento di pensiero, proposto da Robert Boyd e Peter Richerson, ci aiuterà a chiarirla (Boyd & Richerson 2005, pp. 64-67). Questa volta dobbiamo immaginare di trovarci in un ambiente arido, ma nemmeno troppo estremo, con delle provviste per alcuni giorni e di dovercela poi cavare da soli. Boyd e Richerson hanno in mente una zona desertica al confine tra il Messico e gli Stati Uniti, attraversata da parte di un percorso detto Camino del Diablo, che veniva regolarmente usato dai pionieri fino all'arrivo della ferrovia. Allo stesso modo, la zona era regolarmente abitata dagli indiani Tohono O'odam: per restare all'argomento classico della psicologia evoluzionistica, l'ambiente non dovrebbe essere troppo diverso da quelli abitati nel Pleistocene, ai quali, secondo gli psicologi evoluzionistici, le strutture cognitive della nostra specie sono adattati. Insomma, come ce la caveremmo?

Estremizzando la logica dei jukebox non dovrebbe essere troppo complicato, ma, come Boyd e Richerson, sono certo che quasi nessuno è di questa opinione. Ciò che si vuole sottolineare con questo esempio artificioso è che il repertorio comportamentale degli esseri umani è in larga misura costituito da informazioni che non sono trascritte nel codice genetico, istanziate nelle menti umane e pronte ad essere attivate in relazione a determinati stimoli ambientali o sociali. Al contrario, la cultura è adattiva perché permette di conservare, non nei genotipi, ma nelle menti e nei comportamenti, come negli artefatti e nelle modificazioni ambientali in genere, i risultati dell'apprendimento individuale e dei processi di trasmissione sociale che coinvolgono potenzialmente tutti gli esseri umani che interagiscono tra loro.

L'insegnamento che possiamo ricavare dall'esempio dei jukebox è duplice. Da una parte, è necessario essere cauti nel considerare qualsiasi variazione comportamentale che possieda determinate caratteristiche (somiglianza intra-gruppo, differenza tra gruppi, cambiamento nel tempo, adattamento all'ambiente specifico) come dovuta a dinamiche culturali. Dall'altra, è necessario tenere presente che un'architettura cognitiva complessa, come quella comune a tutti gli esseri umani, può influire sui processi di apprendimento sociale e trasmissione culturale, pur senza vincolarli in modo completo.

Rispetto a quest'ultimo punto, una versione più "moderata" dell'argomento dei jukebox proviene dall'antropologo Dan Sperber (Sperber 1999). Anch'egli ritiene che la cognizione umana sia frutto di meccanismi evoluti per selezione naturale e comuni in tutta la specie. Come gli psicologi evoluzionistici, Sperber ritiene che l'architettura della mente sia modulare, ossia che la mente sia composta da meccanismi specializzati e distinti, evoluti per effettuare compiti specifici (si veda anche: Hirschfeld & Gelman 1994). Sperber è convinto che tale architettura influenzi i processi di trasmissione culturale e che, per spiegare la diffusione, la stabilità o la variabilità delle varianti culturali, sia necessario fare riferimento ad essa. La metafora, qui, è quella dell'attrattore: la trasmissione culturale è un processo attivo, in cui chi apprende non *copia* un comportamento da un altro individuo, ma lo *ricostruisce* sulla base delle informazioni che ha a disposizione. Poiché queste informazioni sono, di norma, non sufficienti per una ricostruzione completa del comportamento, la stabilità dei tratti culturali, nel tempo e nello spazio, è molto più robusta di quanto dovremmo aspettarci. Secondo Sperber, ciò è dovuto, come si diceva, all'esistenza di struttura cognitive comuni che vincolano le traiettorie nello spazio dei comportamenti possibili funzionando, in questo spazio astratto, da attrattori. Altri antropologi hanno difeso la plausibilità di ipotesi simili rispetto a domini culturali quali la religione (Boyer 1994), la *folkbiology* (Atran 1998) o la categorizzazione razziale e sociale (Hirschfeld 1996)

Per distinguere queste posizioni ritorniamo all'immagine dello spazio astratto dei comportamenti possibili, in cui muoversi equivale ad apprendere socialmente. Se, nell'immagine dell'ipotetico antropologo ingenuo, questo spazio è completamente piano ed ogni direzione è equivalente e, nell'immagine estrema invocata della psicologia evoluzionistica, è composto da una serie di cavità, dalle quali è impossibile spostarsi, nell'immagine di Sperber è meglio pensarlo come ad un paesaggio collinoso, con qualche valle da cui è difficile muoversi e con vie preferenziali rispetto ad altre. Una delle possibilità per gli antropologi è cercare di definire una cartografia di questo spazio o, quantomeno, utilizzare le mappe che vengono fornite dagli psicologi per verificare le proprie ipotesi sulle spiegazioni delle dinamiche culturali⁵.

⁵ Naturalmente, questo non significa che altri fattori non siano importanti nel determinare le traiettorie in questo spazio astratto. Mi riferisco a fattori ambientali o, in particolare, a fattori più propriamente "culturali" (politici, storici, economici, ecc.). Semplicemente, non sono oggetto di questa analisi e, comunque, ritengo che, almeno rivolgendosi a chi si occupa di antropologia, non sia troppo necessario rimarcare l'importanza.

Tuttavia, in qualche modo, come ci suggerisce l'esempio del Camino del Diablo, il valore adattivo della cultura deriva dall'opportunità di muoversi in questo spazio e muoversi in questo spazio significa modificare il proprio comportamento attraverso l'interazione con altri individui e in direzioni che non sono state completamente specificate dalla genetica. Come questo possa essere fatto è ciò di cui ci occuperemo ora.

Trasmettere e copiare

Come anticipato, come esseri umani siamo portati a pensare che l'apprendimento sociale sia un processo trasparente e che un generico meccanismo che permette l'imitazione possa spiegarlo. Ci muoviamo nello spazio dei comportamenti possibili osservando gli altri, *copiando* i loro comportamenti. Ho già fatto riferimento allo scetticismo di antropologi come Dan Sperber rispetto a questa visione, ma ora è possibile entrare nei dettagli. Ricerche di etologia (che mostrano come l'imitazione non sia affatto un processo diffuso negli animali), di robotica (che mostrano come sia estremamente difficile costruire un artefatto con capacità imitative) e di psicopatologia (che mostrano in che casi per gli esseri umani è difficile, o impossibile, imitare) ci aiutano ad evidenziare alcuni difetti dell'equazione apprendimento sociale come copia del comportamento.

Mettiamoci nei panni di chi debba costruire un robot capace di imitare e che cerchi di implementare questa abilità con un algoritmo che copi esplicitamente il comportamento di un modello, o dimostratore. Il robot è fornito delle migliori tecnologie possibili dal punto di vista di sistemi di percezione del mondo esterno (telecamere, microfoni, ecc.), sistemi di controllo delle parti fisiche, memorie, e via dicendo. Per iniziare, cominceremo ad addestrare il nostro robot ad imitare comportamenti molto semplici: per esempio, fare un nodo. Il dimostratore, probabilmente un essere umano, si metterà davanti alla telecamera del robot e incomincerà a fare movimenti con una corda. Il sistema visivo del robot effettuerà una dettagliata scansione della scena, che verrà conservata senza problemi nell'abbondante memoria di cui lo abbiamo premurosamente fornito.

E ora? Come è possibile ricavare da questa registrazione dei movimenti altrui

delle informazioni su come “muovere se stessi”? Questo problema, straordinariamente difficile da trattare per chi si occupa di robotica, è stato definito il “problema della corrispondenza” (Dautenhahn & Nehaniv 2002). E’ evidente che l’algoritmo che abbiamo utilizzato non sarebbe sufficiente per risolverlo: dovremmo aggiungere un altro meccanismo che sia in grado di traslare i movimenti osservati in opportuni comandi motori. In questo caso, osservando le proprie mani, il robot dovrebbe procedere per prove ed errori (quindi apprendendo, in parte, individualmente) e cercare di effettuare i movimenti che meglio approssimano quelli del modello, che ha registrato, dovendo così, per poterli confrontare, effettuare anche una sorta di inversione della prospettiva visuale. Il rischio è che i costi computazionali di un algoritmo così complesso superino i vantaggi di potere apprendere un comportamento da un modello. Inoltre, il problema sarebbe ancora meno trattabile per comportamenti ancora più semplici, come per esempio le espressioni facciali. Sappiamo che, tra gli esseri umani, le espressioni facciali sono uno dei primi comportamenti che vengono imitati: Meltzoff e Moore (Meltzoff & Moore 1983) hanno scoperto che i neonati imitano le espressioni facciali degli adulti ad un’età media di 32 ore. All’interno del loro campione, il neonato più precoce nell’imitare le espressioni facciali lo fece a 42 minuti dalla nascita! Se nel caso del nodo, il nostro robot avrebbe potuto confrontare i movimenti delle proprie mani con quelli delle mani del modello, cosa avrebbe potuto fare in questo caso, non avendo informazioni percettive a proposito delle proprie espressioni facciali?

Vorrei continuare ancora con la metafora della costruzione del robot. Spero sia chiaro quale sia il significato di questo modo di procedere: da una parte permette di pensare ai meccanismi implicati nel processo di trasmissione culturale nel modo più distaccato possibile, cercando di non essere influenzati da come lo valutiamo in quanto esseri umani, profondamente coinvolti in questo processo (tanto da ritenerlo “naturale”). Dall’altra, permette di testare la plausibilità delle nostre ipotesi sui meccanismi e sui loro funzionamenti, fornendo un “grado zero” di plausibilità: se in un robot non funzionerebbe è *probabile* che non funzioni così nemmeno negli esseri umani. Nulla di più.

Comunque, disperando di riuscire a risolvere il problema della corrispondenza, ma testardamente convinti che il nostro algoritmo di copia si efficace, optiamo per un trucco: non forniamo più alla percezione visiva del robot (la telecamera) i movimenti del modello, ma gli forniamo direttamente le informazioni

sulle attivazioni motorie e muscolari del modello. Il problema della corrispondenza è risolto d'ufficio. A questo punto il robot comincerà a trafficare a caso con la corda, terrà traccia delle proprie attivazioni motorie e le confronterà con quelle del modello. Un algoritmo adatto modificherà gradualmente le attivazioni fino a renderle simili così che, dopo un po' di tentativi, il robot associerà determinati input a determinati output nella stessa maniera del dimostratore. In altre parole, imparerà a fare il nodo.

Ancora, tuttavia, ci troviamo di fronte a dei problemi. Il primo, che è un altro aspetto del problema della corrispondenza, riguarda le differenze nelle strutture e nelle dinamiche corporee (Dautenhahn & Nehaniv 2002). Tipicamente, tra gli esseri umani, i bambini imparano dagli adulti e, banalmente, le strutture e le dinamiche corporee dei bambini sono differenti da quelle degli adulti. Inoltre, gli esseri umani, almeno in linea di principio, possono imitare anche comportamenti prodotti non da esseri umani, ma da animali o da oggetti meccanici che producono un qualsiasi movimento. Un secondo problema, messo in luce da Rodney Brooks (Adams *et al.* 2000), un ricercatore nel campo della robotica e dell'intelligenza artificiale, riguarda *quali* attivazioni motorie imitare. Il nostro ben disposto dimostratore di nodi, oltre a compiere i movimenti necessari per legare la corda, sbatterà sicuramente gli occhi con un certo ritmo, assumerà espressioni facciali, sposterà il proprio peso da un piede all'altro, o, addirittura, potrebbe starnutire o grattarsi il naso. Pensando da esseri umani, sembra banale isolare i movimenti da replicare da quelli non essenziali, ma come dovrebbe riuscirci il nostro robot⁶?

Da ultimo, anche se il robot avesse una struttura corporea esattamente uguale a quella del dimostratore, fosse riuscito in qualche modo ad isolare le attivazioni da replicare, e incominciasse a cercare di modificare le proprie attivazioni per rendere simili a quelle del modello, è ancora possibile, come mostrato nella simulazione

⁶ In realtà, ho in mente anche un corrispettivo "umano" del problema sollevato da Brooks. Si pensi ad attività che richiedono una complessa coordinazione senso-motoria come quelle sportive o alcune di quelle artistiche: per esempio, imparare a suonare uno strumento musicale. Dall'osservazione spesso non è chiaro quali siano le attività motorie da tenere in considerazione per *imitare* il comportamento del maestro. Da dilettante suonatore di pianoforte, ho sperimentato più volte come una particolare postura delle braccia, o della schiena, che può non venire nemmeno notata, abbia profonda influenza sul suono prodotto o produca una modificazione a cascata di altri comportamenti motori. Tuttavia, non mi è chiaro se questo rappresenti una caratteristica generale dell'apprendimento sociale umano o, appunto, sia limitato a particolari forme di espressione corporea, quantitativamente trascurabili tra i comportamenti appresi.

presentata in questo capitolo, che qualcosa vada storto. Anche utilizzando un algoritmo di copia ragionevolmente efficace, in cui gli output del robot “allievo” sono *quasi* uguali a quelli del robot “maestro”, nel caso di comportamenti continui (come muoversi in un ambiente bidimensionale, con delle ruote simulate in modo realistico, o, per continuare l’analogia, fare un nodo), piccole differenze al livello delle attivazioni si accumulano nel tempo, producendo, a livello del comportamento risultante, differenze rilevanti. Tutto ciò, per fare un nodo.

Alla luce di queste considerazioni, trovo giustificato lo scetticismo di antropologi come Dan Sperber rispetto alla riduzione della trasmissione culturale a copia. Per risolvere questo problema, sono state adottate diverse strategie. Come si è detto, Sperber e altri antropologi affermano che problemi come quelli a cui abbiamo accennato vengano risolti dal fatto che le strutture cognitive comuni alla specie umana ci facciano convergere, per default, verso soluzioni comuni. Altri ricercatori (Tomasello 1996) hanno iniziato a trattare l’imitazione come un fenomeno cognitivo complesso, caratteristico degli esseri umani. Michael Tomasello (vedi anche: Tomasello, Kruger & Ratner 1993) ritiene che per imitare, nel senso proprio della parola, non sia sufficiente copiare delle sequenze comportamentali, ma sia necessario comprendere che queste sequenze sono finalizzate ad un obiettivo, realizzando così un *mapping* complesso tra azioni e risultati delle azioni. I primati non umani, sempre secondo Tomasello, avrebbero una tendenza a riprodurre solo i risultati delle azioni (emulazione), ma non il modo in cui questi risultati sono stati ottenuti⁷.

⁷ Un’opinione differente proviene da altri studiosi dell’apprendimento sociale nei primati (per esempio Whiten *et al.* 2005) che hanno effettuato studi su scimpanzè con cui mostrerebbero che essi sono in grado di copiare anche le azioni dei modelli. Nei loro esperimenti utilizzano compiti che possono essere risolti in due modi differenti ottenendo lo stesso risultato, per esempio liberare del cibo contenuto in tubo inclinato, spingendolo con un bastone verso l’estremità rialzata oppure utilizzando lo stesso bastone per rimuovere un ostacolo che impedisce al cibo di cadere dalla parte ribassata del tubo. Dal punto di vista dei risultati, essi sono completamente equivalenti, per cui, se gli scimpanzè emulassero e non imitassero, essi dovrebbero essere del tutto indifferenti rispetto all’azione compiuta dal dimostratore. Tuttavia, Whiten *et al.* hanno mostrato che i soggetti sono più inclini a ripetere anche l’azione del modello e sostengono che le differenze tra le capacità imitative degli esseri umani e degli altri primati sarebbe una questione quantitativa e non qualitativa. Secondo il gruppo di Tomasello si tratterebbe però di un fenomeno spurio (Tennie *et al.* 2006) dovuto ad un effetto di *local enhancement* (vedi capitolo successivo).

Appendice 4.1

L'algoritmo di *simulated annealing*

L'algoritmo di *simulated annealing* (Kirkpatrick, Gelatt & De Vecchi 1983) trae ispirazione dal procedimento reale di tempra (*annealing*) dei solidi. L'analogia fondamentale con il procedimento di tempra è quella di una discesa graduale di un parametro detto temperatura (T), che è direttamente proporzionale alla probabilità di accettare nuovi stati del sistema che sono "peggiori" rispetto a quello in cui già si trova. Nel nostro caso, la temperatura decresce in modo lineare all'aumentare del numero di passi dell'algoritmo, partendo da un valore di $T = 0.02$ ed arrivando a $T = 0.0$. Oltre che dalla temperatura, come detto in precedenza, la probabilità di accettare mutazioni maladattive è determinata dall'entità del peggioramento del sistema, secondo la funzione di valutazione. In termini quantitativi, la probabilità di accettare una mutazione è data da:

$$p = \exp\left(\frac{f(t) - f(t-1)}{T}\right) \quad (3)$$

dove f è la funzione di valutazione, t è il ciclo di apprendimento e T è una variabile chiamata temperatura. Si può notare come, se:

$$f(t) - f(t-1) \geq 0 \quad (4)$$

ossia se la nuova funzione di valutazione dà un risultato migliore o uguale a quella precedente, tutto il termine compreso tra parentesi tonde nell'equazione (3) è maggiore di 0, per cui p è sempre maggiore di 1, ossia le mutazioni sono sempre accettate. Se invece il termine tra parentesi è minore di zero, p varia tra 0 e 1 (vedi Fig. 4.13).

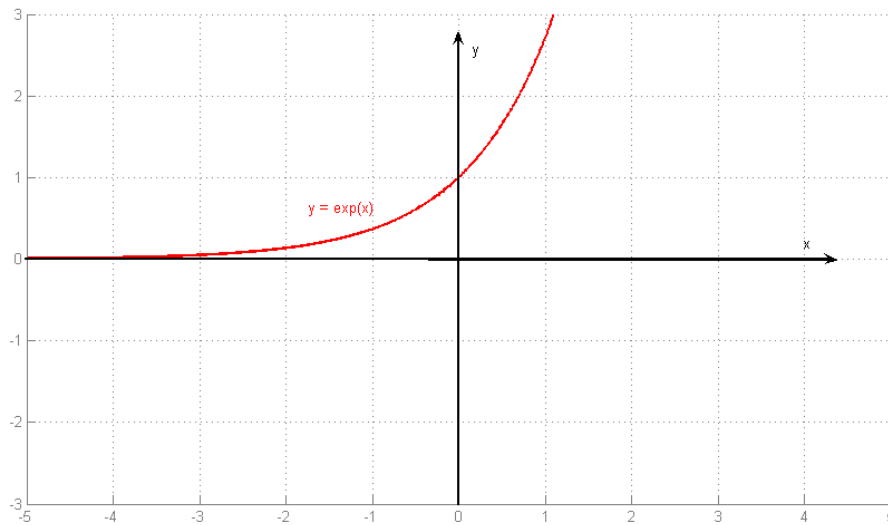


Fig. 4.13: andamento della funzione $y = \exp(x)$

Accettare mutazioni negative serve per evitare di rimanere intrappolati in minimi locali. Vale forse la pena di spendere qualche parola in più a proposito di questo tema, poiché è di fondamentale importanza e anche perché, probabilmente, permette di entrare un po' nel modo di ragionare necessario per costruire simulazioni come quelle qui descritte. Un algoritmo di ottimizzazione è, fondamentalmente, un processo di ricerca degli stati di un sistema che soddisfano determinate condizioni (si veda anche Mitchell 1996). Si può immaginare questa ricerca come un movimento all'interno di uno spazio che ha tante dimensioni quanti sono i parametri su cui l'algoritmo può agire (nel nostro caso si tratterebbe di uno spazio che ha tante dimensioni quanti sono i pesi e i bias della rete neurale). Attraverso le mutazioni, l'algoritmo muove il sistema in una posizione univocamente definita in questo spazio. Un'altra dimensione di questo spazio rappresenta la prestazione del sistema.

Consideriamo una situazione molto semplice, in cui l'algoritmo agisce su due soli parametri: si tratterebbe qui di uno spazio tridimensionale in cui x e y sono i valori dei due parametri e z è la prestazione del sistema (vedi Fig. 4.14). Questo spazio prende generalmente il nome di *fitness landscape*. L'algoritmo muove il sistema in

questo paesaggio astratto cercando di portarlo nei punti in cui la prestazione è maggiore (le vette del paesaggio).

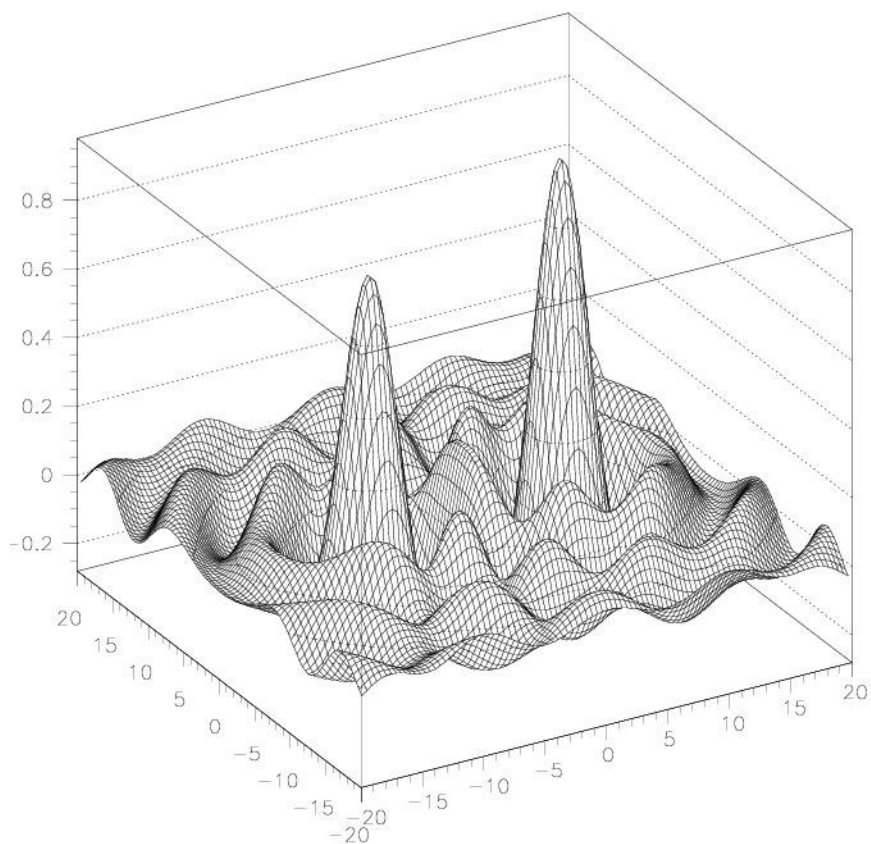


Fig. 4.14: Un ipotetico *fitness landscape*.

Ora, nella maggior parte delle situazioni, questi paesaggi presentano una struttura irregolare che determina più stati in cui il sistema si può bloccare. Possiamo immaginare che un sistema sottoposto ad una procedura analoga al *simulated annealing*, ma in cui non si accettassero mutazioni negative, potrebbe finire in tempi relativamente brevi su una delle “colline” laterali ma non raggiungere mai una delle due “vette” centrali, poiché le mutazioni lo farebbero spostare di poco e verrebbero sempre rifiutate. Accettando mutazioni negative, invece, al sistema è permesso di fare, almeno all’inizio, anche tratti di “discesa”, aumentando in questo modo le probabilità che il risultato finale sia effettivamente quello ottimale.

5. Meccanismi indiretti di trasmissione sociale.

Nel capitolo precedente è stato osservato come appaia riduttivo considerare le dinamiche di trasmissione culturale come processi di copia esplicita di un comportamento e come, per alcuni studiosi, sia necessario considerare il fenomeno che definiamo in modo intuitivo come “imitazione” in modo più approfondito.

Un'altra possibilità ancora è quella di considerare forme differenti di trasmissione sociale dell'informazione. Gli studi etologici, attualmente, convergono su due risultati: da una parte, gli animali non sembrano in grado di “imitare” nel senso proposto, per esempio, da Tomasello (si veda ancora il capitolo precedente) ma, dall'altra, l'apprendimento sociale sembra potersi realizzare in altre forme, diverse da questa, e avere un ruolo decisivo nella formazione di pattern comportamentali complessi in molte specie (Galef & Laland 2005).

Prima, vorrei offrire qualche motivazione per valutare questa possibilità. Non credo sia un pregiudizio antropocentrico affermare che, tra gli esseri umani, la cultura svolge un ruolo che non è paragonabile a quello che ha in altre specie viventi. Tuttavia, in una prospettiva naturalistica, dobbiamo accettare il fatto che questa apparente discontinuità sia dovuta ai consueti meccanismi evolutivisti e che, quantomeno, studiare forme semplici di trasmissione sociale possa servire a riconoscerne l'eventuale presenza nelle dinamiche culturali umane (presenza di cui faccio fatica a dubitare) e, contemporaneamente, a capirne meglio le origini

evoluzionistiche.

Altre due motivazioni mi derivano da principi che ho imparato dai lavori di vita artificiale. Molto spesso accade che fenomeni che interpretiamo come complicati e dovuti a meccanismi egualmente complicati possano essere riprodotti artificialmente come risultato di interazioni e di meccanismi molto più semplici (Noble & Todd 2002). Quando questo accade è assai plausibile che la selezione naturale, in genere parsimoniosa, abbia realizzato i fenomeni in questione con i meccanismi più semplici. Allo stesso modo, i pattern comportamentali che interpretiamo di solito come risultato di dispendiose elaborazioni cognitive “solipsistiche” possono spesso venire realizzati efficacemente da dispositivi più semplici, che sfruttano le interazioni tra sistema nervoso, corpo, ambiente e, eventualmente, conspecifici (Nolfi 2006). Anche in questo caso, probabilmente, il processo evoluzionistico tenderebbe preferenzialmente verso le soluzioni del secondo tipo (si veda, a questo proposito, il secondo capitolo della tesi).

Immaginiamo, per iniziare, una specie che vive in un ambiente con due tipi di piante che presentano caratteristiche percettive differenti. Uno dei due tipi è nutriente e porta energia all’animale, mentre l’altro è velenoso e lo fa ammalare. Sfortunatamente gli animali nascono senza avere nessuna facoltà innata di distinguere una pianta dall’altra: i loro jukebox, per continuare con la metafora proposta dalla psicologia evoluzionistica, rimangono tristemente silenziosi. Se questi animali sono “socialmente ciechi” l’unica possibilità che hanno è di provare una pianta, vedere che effetto ne ricavano e poi regolarsi di conseguenza, sperando di imparare velocemente e che le piante velenose non lo siano troppo.

Immaginiamo invece che questi animali possano anche seguire una regola comportamentale molto semplice, che suonerebbe più o meno così: “quando vedi un adulto seguilo”. Ora, gli adulti, in questa storiella, hanno una caratteristica molto importante: sono vivi e, se sono vivi, significa che hanno imparato a distinguere le piante commestibili da quelle velenose. Quindi, se sono un giovane, stando tra i piedi di un adulto avrò buone probabilità, quando dovrò iniziare ad assegiare le piante, di incappare al primo colpo in una pianta commestibile, poiché è probabile che gli adulti si aggireranno di più nei pressi delle piante commestibili che nei pressi di quelle velenose. Così, avrò imparato individualmente quali sono le piante commestibili, ma, in un altro senso, è chiaro che il mio apprendimento è stato socialmente guidato. Di più, questo comportamento è in un senso pieno un comportamento nuovo e tuttavia

non ho dovuto esplicitamente copiare alcunchè.

Questa piccola storia mostra un comportamento che è noto in etologia come “rinforzo locale” (local enhancement) o “rinforzo per stimolazione” (stimulus enhancement): un processo del genere potrebbe rendere conto, per esempio, della capacità, trasmessa socialmente negli scimpanzè, di rovesciare ceppi di legno per cibarsi delle larve di insetto sottostanti. Anche in questo caso, ai giovani scimpanzè basterebbe imparare che i ceppi sono “oggetti interessanti”, senza copiare in dettaglio le azioni di altri individui: non è necessario, quindi, imitare il comportamento degli adulti nel senso pieno della parola, ma il seguirli rende saliente una determinata classe di stimoli, rispetto alla quale si interagisce in modo individuale. L’etologo Bennett Galef ha spiegato in questi termini la diffusione, tra le cinciarelle (*Cyanistes caeruleus*), della capacità di bucare i tappi di alluminio delle bottiglie di latte che, fino alla fine degli anni ’70, venivano lasciate davanti alle porte delle case inglesi. Anche in questo caso la lettura prevalente era quella di apprendimento per imitazione, ma utilizzando situazioni di controllo sperimentali è stato mostrato come le cinciarelle, pur non imparando mai, o quasi, da sole, questo tipo di comportamento, riuscivano ad impararlo nella stessa misura sia quando il comportamento di un dimostratore (un’altra cinciarella) gli veniva esplicitamente fornito, sia quando gli veniva semplicemente fornito un esempio di bottiglia con il tappo già bucatto. (Sherry & Galef 1984).

Un fenomeno solo leggermente diverso avviene in una popolazione di ratti (*Rattus norvegicus*), studiata ancora da Galef (Galef 1996), che sviluppano preferenze per determinati cibi sulla base delle interazioni con altri conspecifici. In particolare, è stato scoperto che se un ratto interagisce con un conspecifico che ha da poco mangiato del cibo, svilupperà una preferenza per il tipo di cibo mangiato da quell’individuo. Se, tuttavia, in condizioni sperimentali, viene presentato al ratto un batuffolo di cotone impregnato dall’odore di quel cibo, il soggetto non sviluppa alcuna preferenza per quel cibo, a conferma del fatto che, in questo caso, l’interazione sociale è fondamentale perché il comportamento venga messo in atto.

Abbiamo già detto, discutendo l’esempio dei jukebox modificati, di come la dinamica definibile come stimulus priming (lo sbadiglio tra gli esseri umani) non possa essere considerata genuinamente culturale, in quanto manca la caratteristica della novità del comportamento. Tuttavia se immaginiamo che, insieme al priming, sia anche presente una forma rudimentale di apprendimento per associazione, l’unione

delle due può dare origine a fenomeni che assomigliano da vicino a ciò che si intende di solito come trasmissione culturale (Noble & Todd 2002). Se, per esempio, in una specie esiste l'istinto a scappare quando gli altri scappano (il che sembra abbastanza ragionevole) e, per un certo numero di volte, questo avviene in presenza di un dato predatore, un individuo giovane potrebbe facilmente associare l'istinto di fuga a quel predatore e, in seguito, scappare alla sua presenza, anche senza il bisogno che questa reazione sia causata socialmente.

L'ultimo esempio che riporto riguarda uno studio realizzato su dei resi – una specie di scimmie – allevati in cattività (Mineka & Cook 1988). I resi nati e cresciuti in laboratorio, scoprirono i ricercatori, non reagivano, o reagivano poco, alla presenza dei serpenti, al contrario dei resi nati in libertà che mostravano, nella stessa situazione, reazioni molto pronunciate. Se, tuttavia, ad un reso di laboratorio venivano fatti vedere dei conspecifici che si spaventavano alla presenza di un serpente (o dei video che mettevano in scena questa situazione), anche questo sviluppava, e manteneva a lungo, lo stesso tipo di reazione. In modo interessante, inoltre, i resi non sviluppano una paura indiscriminata per ogni stimolo che viene associato, in una situazione sociale, ad una reazione di paura da parte di un conspecifico. Nello stesso studio, vennero manipolati dei video per far sembrare che i resi nati in libertà si spaventassero alla presenza di fiori, al posto che alla presenza dei serpenti, ma nessun reso di laboratorio imparò ad associare la reazione di paura ai fiori. Questo caso mette in luce una dinamica interessante tra evoluzione genetica, apprendimento ed interazione sociale che potrebbe suonare più o meno in questo modo: l'evoluzione genetica fornisce una reazione istintiva (la paura) e circoscrive, senza definirla nei particolari, una classe di stimoli a cui associare questa reazione (di cui fanno parte i serpenti) e una a cui non associarla (di cui fanno parte i fiori). Tuttavia, la delimitazione precisa di queste classi viene lasciata all'apprendimento, il che ha, evolutivamente, un senso, se pensiamo a quanti sono i pericoli a cui una specie animale può essere vulnerabile e a come essi varino nel tempo. Ma, spesso, nel caso dei pericoli, non è possibile apprendere individualmente, perché quando si incontra un pericolo, anche per la prima volta, è necessario sapere già cosa fare, perché potrebbe non esserci tempo per imparare per prove ed errori, da cui la necessità di sfruttare l'interazione sociale.

In questo capitolo verrà presentata una simulazione che mostra come un meccanismo indiretto di trasferimento sociale dell'informazione, in cui

l'apprendimento individuale degli agenti viene modulato dalla presenza di altri individui nelle vicinanze, possa produrre un comportamento efficace in un semplice compito di discriminazione di differenti aree dell'ambiente in cui gli agenti vivono. In particolare, verrà mostrato come, in questo scenario, tale meccanismo produca risultati migliori di altri algoritmi adattivi basati sull'evoluzione genetica o sull'apprendimento individuale.

Set – up robotico

Un team di 10 robots e-puck (vedi Figura 2.2 nel Capitolo Secondo) “vive” in un ambiente che consiste di un'arena quadrata di 200 cm. X 200 cm. circondata da muri. Il fondo dell'arena è di colore grigio e contiene due aree circolari di 60 cm. di diametro colorate una di bianco e una di grigio (*target areas*) (si veda la Figura 5.1).

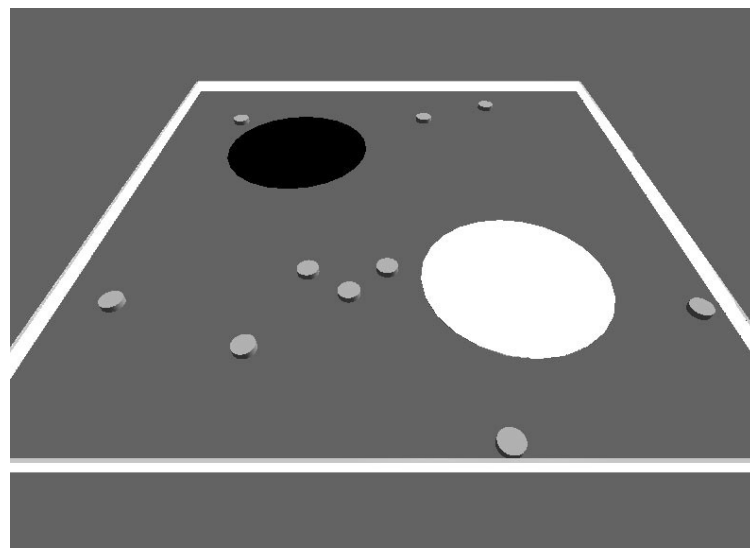


Fig. 5.1: L'ambiente dell'esperimento. I cerchi di colore bianco e di colore nero rappresentano le due *target areas*. I cerchi grigi più piccoli rappresentano i 10 robots di cui è composto il team.

I robots sono dotati di semplici capacità senso – motorie che gli permettono di muoversi, raccogliere informazioni sull'ambiente e produrre dei segnali sonori con un'intensità fissa. In dettaglio, i robots sono composti da un “corpo” circolare di raggio 37 mm. con due ruote e sono forniti di 8 sensori ad infrarossi disposti sul lato

esterno del corpo, 1 sensore del terreno (*ground sensor*), 1 microfono ed infine 1 altoparlante. I robots sono controllati da una rete neurale con 10 unità di input e 3 unità di output (motorie) (si veda la Figura 5.2).

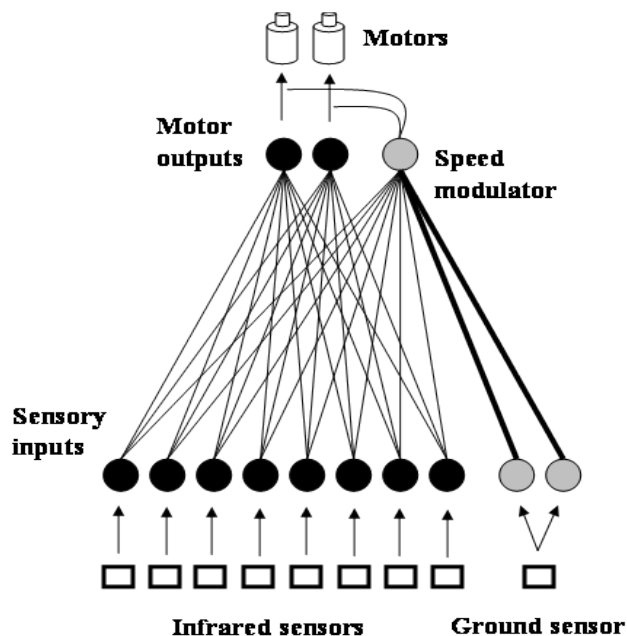


Fig. 5.2: La rete neurale che controlla il comportamento dei robots. Le linee marcate rappresentano le connessioni i cui pesi vengono modificati nella seconda parte della simulazione. Come mostra l'immagine, il modulatore della velocità non opera sugli output, ma direttamente sull'attivazione dei due motori che controllano le ruote dei robots.

Otto delle unità di input codificano l'attivazione degli otto sensori ad infrarossi (che si attivano in modo binario quando rilevano un ostacolo – ossia, in questo caso, un muro od un altro robot – alla distanza massima di circa 4 cm.) e le altre due codificano l'attivazione del *ground sensor*, ossia il colore del terreno sotto al robot, con questi valori: [0 0] grigio, [1 0] bianco, [0 1] nero. Le prime due unità motorie codificano la velocità delle due ruote dei robots, mentre la terza agisce come un modulatore che regola questa velocità, nel suo complesso, inibendola (ossia, più l'unità è attiva più la velocità del robot decrementa: se il valore dell'unità è uguale ad uno, il robot, indipendentemente dalle attivazioni delle altre due unità motorie, starà fermo). Le otto unità di input che codificano l'attivazione dei sensori a infrarossi sono

connesse a tutte le tre unità di output. Le due unità di input che codificano il colore del terreno dell'arena (ground sensor), invece, sono connesse solo alla terza unità di output (ossia a quella che agisce come modulatore).

I robots emettono un segnale sonoro prefissato quando si trovano in una o nell'altra area circolare, che può essere rilevato dagli altri robots attraverso il microfono di cui sono dotati.

La simulazione e le condizioni sperimentali

L'addestramento dei robots è diviso in due fasi. Nella prima fase vengono addestrati a sviluppare le abilità di evitare ostacoli e di esplorare l'ambiente, ossia devono riuscire a non collidere con i muri dell'arena e con gli altri robots del team e devono, nello stesso tempo, muoversi il più possibile nell'arco del periodo in cui vengono testati. Durante la seconda fase, invece, i robots devono imparare a discriminare tra le due aree circolari e, in particolare, ad evitare quella nera.

Nella prima fase, i pesi delle connessioni tra le otto unità di input che codificano l'attivazione dei sensori ad infrarossi e le tre unità motorie vengono evoluti attraverso un algoritmo genetico. I robots vengono posizionati in un ambiente privo di *target areas*. I pesi delle connessioni delle reti neurali che controllano i robots vengono inizializzati con dei valori casuali compresi tra [-10.0; +10.0] e i robots vengono valutati, come detto, rispetto alla loro capacità di muoversi nell'ambiente evitando gli ostacoli. I migliori 20 individui, dopo ogni test, vengono selezionati per la riproduzione e generano 5 "figli" ciascuno, che ereditano i valori dei pesi delle connessioni, con una probabilità del 3% che ogni valore sia mutato, ossia venga reinizializzato con un valore casuale. Il processo di valutazione, selezione e riproduzione viene ripetuto per 50 generazioni.

Nel corso della seconda fase, i pesi delle connessioni che collegano le due unità che codificano l'attivazione del *ground sensor* (che, come si ricorderà, rileva il colore del terreno) e la terza unità motoria (che regola l'attività del modulatore), vengono modificati attraverso differenti procedure, che verranno illustrate in dettaglio in seguito, e che rappresentano le cinque condizioni sperimentali che sono confrontate

in questa simulazione. Gli altri pesi della rete neurale, che consentono al robot di muoversi nell'ambiente evitando gli ostacoli, non vengono modificati. In tutte le condizioni sperimentali, i robots vengono valutati per un'epoca di 10000 passi. All'inizio di ogni epoca, i robots sono posizionati in maniera casuale nell'ambiente ed hanno un'energia di 200 unità. Per ogni passo in cui il robot si trova nell'area circolare nera, l'energia viene diminuita di un'unità. Nulla invece accade nell'altra area. Quando un robot finisce la propria energia "muore" e viene sostituito da un nuovo robot, ancora una volta posizionato in modo casuale nell'ambiente.

Si noti che i robots non hanno nessuna informazione diretta ne' sulla "qualità" delle *target area* ne' sul proprio stato energetico: essi hanno informazioni solo sul colore di un area, ma non possono sapere come lo stare in una o nell'altra *target area* abbia impatto sulla loro energia. In pratica, dato un comportamento generale che consente di esplorare l'ambiente (acquisito in modo uguale per tutti durante la prima fase dell'addestramento), ciò che i robots devono acquisire in questa seconda fase è la capacità di allontanarsi velocemente dalla *target area* nera ed, eventualmente, di fermarsi nell'area circolare bianca, minimizzando, in questo modo, la possibilità di finire, per ragioni casuali, nell'area nera, mentre si muovono nell'arena.

Qui di seguito sono descritte le differenti procedure di modifica dei pesi:

- (1) *Apprendimento Individuale*: Possiamo dire che, in questa condizione, l'apprendimento individuale agisce rinforzando la preferenza per le esperienze già compiute. In altre parole, all'inizio dell'epoca, i robot tenderanno ad uscire da entrambe le aree nello stesso modo, ma, col passare del tempo, acquisiranno una "preferenza" per l'area in cui per ragioni causali sono stati più tempo, e questa preferenza agirà come un meccanismo di feedback positivo (ossia, più acquisiranno una preferenza per una determinata area, più tenderanno a stare in quest'area e più tenderanno a stare nell'area, più acquisiranno preferenza e via dicendo). Ancora, si noti come questo meccanismo è indipendente dal colore dell'area e di conseguenza dall'effetto che le aree hanno sull'energia dei robots. Nel dettaglio, questo meccanismo di feedback positivo viene realizzato inizializzando i pesi tra *ground sensor* e modulatore con un valore nullo (in questo modo, quando all'inizio il *ground sensor* viene attivato per il passaggio del robot in una delle due aree, il modulatore non viene attivato e il robot continua la sua esplorazione dell'ambiente) e

modificandoli di un valore casuale compreso tra [-learning rate; +learning rate] (con learning rate = 0.1) ogni volta che il robot si trova in un'area. Le variazioni vengono conservate solo se hanno l'effetto di rallentare la velocità del robot.

- (2) *Apprendimento Sociale*: L'apprendimento sociale agisce nello stesso modo dell'apprendimento individuale, ma il learning rate è modulato dalla presenza di altri robots nell'area in cui il robot che apprende si trova. Più precisamente, il learning rate viene determinato dall'intensità del segnale sonoro che viene ricevuto dal robot (si ricordi che i robots emettono in modo automatico un segnale di intensità fissa quando si trovano nelle aree). Ogni segnale percepito aumenta di 0.1 il learning rate, ossia, se il robot si trova da solo nell'area percepirà solo il proprio segnale ed avrà un learning rate di 0.1, mentre se, per esempio, tutti i 10 robots si trovano nella stessa area, il learning rate avrà il valore massimo di 1.0. In altre parole, comparando questa situazione a quella descritta in precedenza (apprendimento individuale) è come se, nella prima delle due, i robots fossero “socialmente ciechi” e avvertissero solo la propria presenza, mentre in questo caso possono sfruttare, per calibrare la forza del proprio apprendimento, un'informazione sociale, ossia la presenza o meno di altri robots nella *target area* in cui si trovano. Si noti che, comunque, anche in questo caso, i processi di apprendimento non sono direzionati, ossia, ancora una volta i robots non hanno nessuna informazione sull'effetto che le differenti aree hanno sulla loro energia. In termini generali, questi meccanismi di apprendimento sono plausibili da un punto di vista biologico: molti animali hanno, per esempio, dal punto vista alimentare, la tendenza ad essere neofobici – ossia non sono inclini ad assaggiare cibi che non hanno mai provato prima – ma la socialità tende a ridurre questa neofobia, ossia diventano più inclini a cibarsi di alimenti che i loro conspecifici mangiano (si veda anche Adessi & Visalberghi 2006).
- (3) *Evoluzione Genetica*: In questa condizione, all'inizio di ogni epoca, i pesi delle connessioni sono inizializzati in maniera casuale con un valore compreso tra [-10.0; +10.0]. Quando un robot “muore”, però, viene rimpiazzato da un nuovo robot che eredita i pesi dal robot del team che, in quel momento, ha la più alta energia e vive da più tempo. I pesi vengono

mutati aggiungendo al valore ereditato una quantità casuale compresa tra [-2.0;+2.0] (questo valore è stato scelto dopo diversi test, in quanto è il valore che produce i migliori risultati in questa condizione).

- (4) *Random Search (ricerca casuale)*: Questa è una condizione di controllo, identica all'evoluzione genetica, ma con la differenza che i pesi dei nuovi robots non vengono ereditati, ma vengono ogni volta inizializzati, come all'inizio dell'epoca, con un valore casuale compreso tra [-10.0; +10.0].
- (5) *Apprendimento Individuale (1.0)*: Quest'ultima condizione è uguale alla condizione (1), con la differenza che il learning rate è fissato al valore massimo (ossia 1.0). Questo è utile per valutare se eventuali differenze tra la condizione (1) e la condizione (2) siano effettivamente dovute al fatto che nella condizione (2) (apprendimento sociale) il valore del learning rate cambia in modo dinamico a seconda della presenza o dell'assenza degli altri robots nell'area oppure se esse siano semplicemente legate alla possibilità di avere un learning rate più alto, e quindi un apprendimento più veloce, indipendentemente dall'influenza sociale.

Risultati

Analizzando i risultati ottenuti è possibile osservare che la prestazione varia in modo significativo nelle cinque differenti condizioni sperimentali e che la prestazione nella condizione (2), ossia l'apprendimento sociale, è migliore delle altre.

La Tavola 1 mostra i valori medi di alcuni indici che caratterizzano il comportamento degli agenti, nelle differenti condizioni, alla fine dei 10000 passi di cui è composta ogni epoca. Gli indici denominati *Target W* e *Target B* indicano la percentuale di robots del team che si trovano, rispettivamente, nell'area bianca o nell'area nera; l'indice denominato *Mortality* indica il numero di robots che sono stati sostituiti nel corso del test, poiché hanno terminato la loro energia. La prestazione complessiva (*Overall Performance*) viene calcolata attraverso la seguente equazione:

$$\frac{W_{perc} + \left(\frac{M_{MAX} - M_n}{M_{MAX}} \right)}{2} \quad (0)$$

dove W_{perc} rappresenta la percentuale di robots che si trovano, alla fine dell'epoca, nell'area bianca, M_{max} è il numero massimo di robots sostituiti in tutte le cinque condizioni sperimentali e M_n è il numero dei robots sostituiti nella condizione sperimentale presa in esame. In altre parole, questo indice misura nello stesso tempo l'abilità di discriminare tra le due aree, “preferendo” l'area bianca a quella nera (W_{perc}) e in modo più indiretto, il tempo necessario per sviluppare questa “preferenza”, che è correlato in modo inversamente proporzionale al numero di robots sostituiti nel corso dell'epoca (M_n).

Infine, le ultime due colonne della Tavola 1 mostrano il valore medio dei due pesi collegati al modulatore (ossia gli unici a venire modificati durante la seconda fase dell'esperimento), da cui dipende il comportamento degli agenti nei confronti delle due aree. In particolare, *Weight 1* determina il comportamento dei robots nei confronti dell'area bianca (come si ricorderà, l'area bianca attiva il ground sensor con i valori [1 0], per cui, quando un robot si trova sull'area bianca, è solo questo peso a determinare il funzionamento del modulatore): più il valore del peso è alto (cioè positivo), più il modulatore viene attivato, inibendo i motori che controllano le ruote e di conseguenza facendo fermare i robots in quell'area. *Weight 2* è invece il peso che determina il comportamento nei confronti dell'area nera: in questo caso i valori ottimali sono vicini a zero o negativi.

E' interessante notare come, nelle varie condizioni sperimentali, non ci sia una spinta particolare ad ottimizzare il valore di questo peso, anzi, come il valore “migliore” venga ottenuto nella condizione con la peggiore prestazione complessiva (*Random Search*). Nell'esperimento – almeno per quanto riguarda le condizioni in cui c'è apprendimento – ciò avviene per due ragioni: da una parte, in generale, se l'algoritmo è efficace, i robots passano più tempo, in media, nell'area bianca che in quella nera, per cui l'apprendimento agisce in modo più efficace per favorire la propensione nei confronti dell'area bianca, che per favorire l'avversione nei confronti di quella nera (questo spiega perché nella condizione *Random Search*, in cui l'algoritmo *non* è efficace, si sviluppi anche una lieve avversione per l'area nera). Dall'altra, poiché i robots iniziano con i pesi settati a 0.0, essi non hanno bisogno di apprendere ad evitare un'area, poiché questo è già il loro comportamento all'inizio della seconda parte della simulazione, ma solo devono “scegliere” in quale delle due aree fermarsi.

In particolare, osservando la prestazione complessiva, si può notare come le migliori prestazioni vengano ottenute, nell'ordine, utilizzando un algoritmo di apprendimento sociale, un algoritmo di apprendimento individuale con tasso di apprendimento uguale a 1.0 e un algoritmo di apprendimento individuale con tasso di apprendimento uguale a 0.1.

	Target W	Target B	Mortality	Overall Performance	Weight 1	Weight 2
Apprendimento individuale 0.1	0.78	0.018	14.7	0.63	8.13	1.95
Apprendimento individuale 1.0	0.97	0.003	12.75	0.76	9.71	0.09
Apprendimento sociale	0.98	0	6	0.88	9.86	1.45
Evoluzione genetica	0.56	0.037	26.04	0.32	5.63	0.74
Random Search	0.37	0.052	28.75	0.18	1.77	-1.34

Tav. 5.1: Valori medi degli indici comportamentali per le 5 condizioni sperimentali (si veda il testo per i dettagli). Ogni dato rappresenta la media di 100 repliche della simulazione.

Analizzando come la percentuale di robots presenti nell'area bianca (Fig. 5.3) e la mortalità (il numero di robots sostituiti, Fig. 5.4) variano nel corso del tempo è possibile notare come le curve che si riferiscono alle condizioni *Evoluzione Genetica* e *Random Search* si sovrappongono durante la prima fase e incomincino a divergere solo dopo 2000 e 6000 passi, rispettivamente.

La prestazione negativa dell'algoritmo di ottimizzazione basato sull'evoluzione genetica è, per certi versi, sorprendente, tenendo conto dell'importanza che gli algoritmi genetici hanno nella maggior parte dei lavori di vita artificiale o robotica evolutiva. Nel nostro scenario, questo risultato può essere compreso tenendo conto del fatto che l'algoritmo opera su popolazioni piccole (10 individui), motivo per cui la prestazione è fortemente dipendente dalle condizioni iniziali.

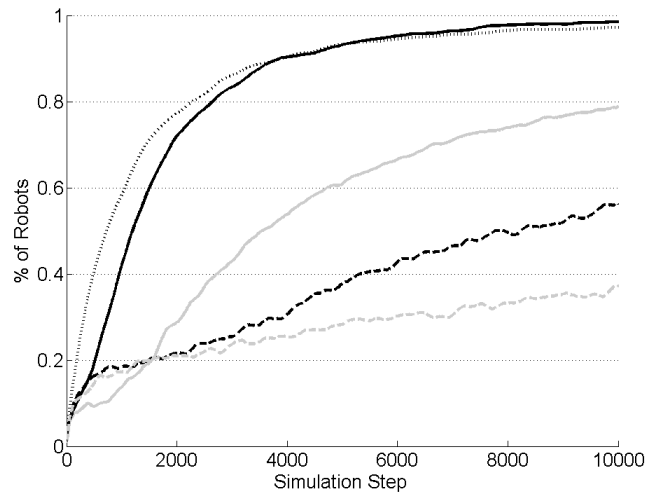


Fig. 5.3: Percentuale di robots localizzati nell'area bianca durante il processo di addestramento nelle cinque condizioni sperimentali (media su 100 repliche degli esperimenti). Legenda: linea nera continua = apprendimento sociale; linea nera puntata = apprendimento individuale (1.0); linea grigia continua = apprendimento individuale (0.1); linea nera tratteggiata = algoritmo genetico; linea grigia tratteggiata = random search.

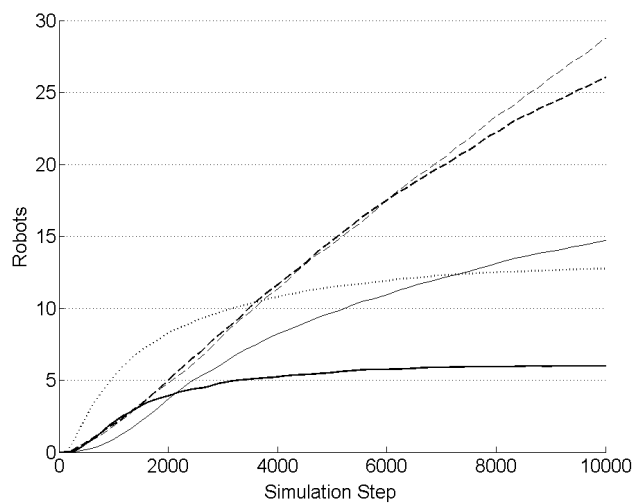


Fig. 5.4: Mortalità (numero di robots sostituiti) durante il processo di addestramento nelle cinque condizioni sperimentali (media su 100 repliche degli esperimenti). Per la legenda vedi Fig. 5.3.

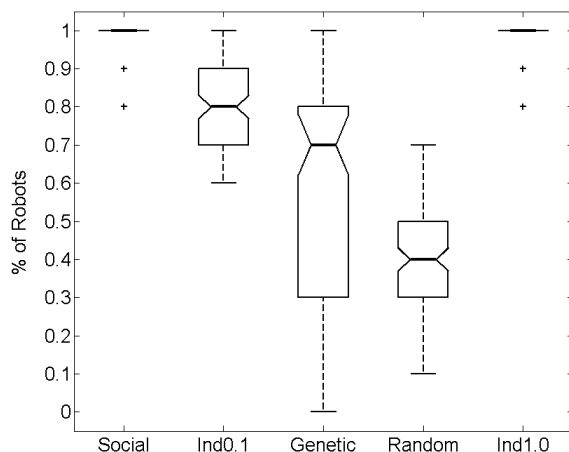


Fig. 5.5: Distribuzione della percentuale finale dei robots nell'area bianca per le 100 repliche delle cinque condizioni sperimentali.

La dipendenza dalle condizioni iniziali è dimostrata dall'analisi della variabilità delle prestazioni osservate, rispetto alla percentuale di robots localizzati nell'area bianca, nelle differenti repliche (100) delle medesime condizioni sperimentali: come si può osservare nella Figura 5.5, la variabilità della condizione *Evoluzione Genetica* è la maggiore, anche rispetto alla condizione *Random Search*.

Allo stesso modo, osservando la Figura 5.4, si può notare come l'indice di mortalità dell'apprendimento sociale sia, all'inizio, simile a quello dell'evoluzione genetica e della ricerca casuale, piuttosto che a quelli delle altre condizioni di apprendimento (individuale), e che incominci a divergere in modo significativo solo intorno ai 1500 passi. Inoltre, l'indice relativo all'apprendimento individuale con tasso di apprendimento 0.1 è, fino a circa 2000 passi, più basso di quello dell'apprendimento sociale. Questo significa che, almeno all'inizio del processo, la possibilità di interazione sociale produce, in media, un *aumento* dei comportamenti non adattivi all'interno del gruppo (si veda la Figura 5.6 che mostra la percentuale di robots presenti nell'area nera, ossia l'area da evitare, nel caso dell'apprendimento sociale e nel caso dell'apprendimento individuale con learning rate uguale a 0.1: nella prima condizione sono presenti *più* robots nell'area "sbagliata" che nella seconda). Infatti, come è già stato osservato, i meccanismi che regolano l'apprendimento individuale e sociale non sono di per sé adattivi, ragione per cui è necessaria, nel caso dell'apprendimento sociale, la *combinazione* di questi meccanismi "ciechi" per giungere a risultati ottimali.

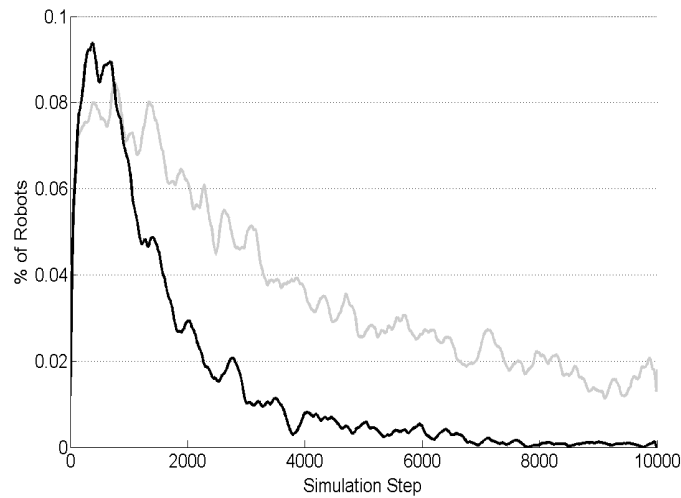


Fig. 5.6: Percentuale di robots nell'area nera nel corso del tempo (media su 100 repliche degli esperimenti). Legenda: linea grigia = apprendimento individuale (0.1); linea nera = apprendimento sociale.

Nel complesso, i risultati mostrano come la combinazione di meccanismi semplici, come, nel nostro caso, un algoritmo di apprendimento individuale non direzionato, un meccanismo di influenza sociale che modula la forza dell'apprendimento individuale, anch'esso non direzionato, e un processo selettivo che opera al livello della popolazione possono dare luogo ad una dinamica adattiva estremamente efficace.

Il modello analitico

Dati i risultati degli esperimenti robotici, nelle sezioni seguenti verrà descritta la dinamica di un modello matematico che replica gli esperimenti, anche se solo rispetto alle condizioni sperimentali *Apprendimento Individuale 1.0* e *Apprendimento Sociale* che sono, del resto, quelle sulle quali si è maggiormente concentrata l'analisi precedente. Qui di seguito verrà solo spiegata la logica generale del modello, senza entrare nei dettagli matematici (la descrizione completa del modello si trova in Acerbi *et al.* 2007).

Il modello è composto da un sistema di equazioni differenziali. Per descrivere

il comportamento del sistema, si denotano con $O(t)$, $W(t)$ e $B(t)$, le densità, al tempo t rispettivamente dell'area grigia, dell'area bianca e dell'area nera. Queste densità possono essere definite in termini di contributi di segno positivo (robots che entrano in un'area) e di segno negativo (robots che escono da un'area). Per un robot che si trova nell'area grigia, entrare in una delle due aree (bianca o nera) è semplicemente una questione casuale, per cui si possono denotare semplicemente con pw e pb le probabilità di entrare, a partire dall'area grigia, nell'area nera o nell'area bianca. Al contrario, se un robot si trova in una delle due aree, bianca o nera, il fatto che rimanga o meno all'interno di questa dipende dall'apprendimento del robot. In particolare, nel caso dell'apprendimento individuale, la probabilità di rimanere in una delle due aree, per ogni robot (Πw e Πb) varia aumentando di un valore costante (σ) ad ogni step (l'equivalente del *learning rate* nel caso dell'esperimento robotico). Nel caso dell'apprendimento sociale, invece, l'aumento della probabilità è direttamente correlato alle densità delle due aree ($W(t)$ e $B(t)$).

La figura 5.7 mostra il confronto tra i risultati del modello analitico e quelli dell'esperimento robotico per quanto riguarda la densità dei robots nell'area bianca, mentre la figura 5.8 mostra il confronto tra le mortalità.

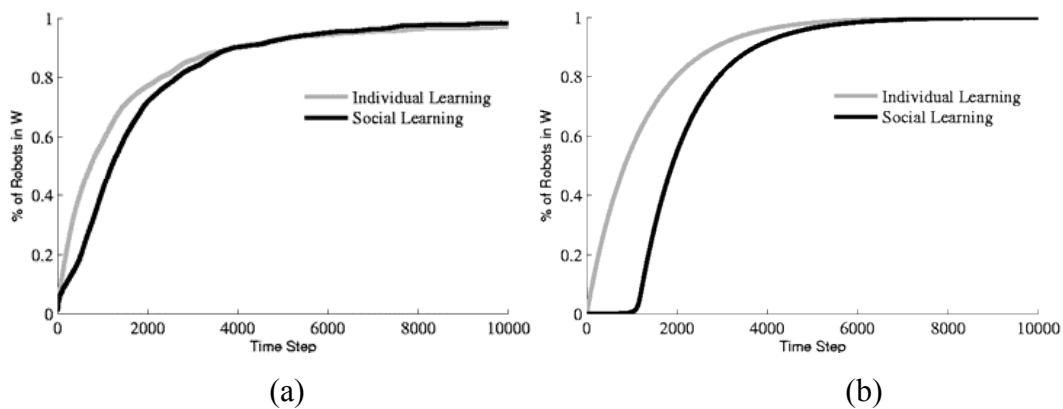


Fig. 5.7: (a) Percentuale di robots nell'area bianca nel corso del tempo (media su 100 repliche degli esperimenti). Legenda: linea grigia = apprendimento individuale (1.0); linea nera = apprendimento sociale. (b) Risultati corrispondenti del modello analitico.

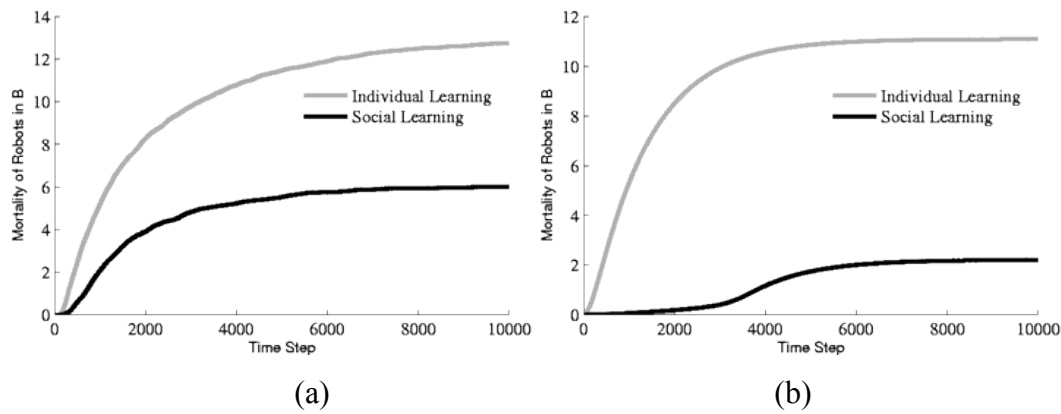


Fig. 5.8: (a) Mortalità nel corso del tempo (media su 100 repliche degli esperimenti). Legenda: linea grigia = apprendimento individuale (1.0); linea nera = apprendimento sociale. (b) Risultati corrispondenti del modello analitico.

Come si può vedere, i risultati analitici possono essere comparati abbastanza bene con i risultati degli esperimenti reali.

Le equazioni che compongono il modello possono essere utilizzate anche per ricavare gli stati di equilibrio del sistema al variare dei parametri. Nel nostro caso, abbiamo utilizzato il modello analitico per osservare l'andamento del sistema variando in modo sistematico pw e pb , ossia le probabilità di entrare nelle due aree bianca e grigia. Nell'esperimento reale queste probabilità corrispondono in modo diretto alla grandezza fisica delle due aree, che nell'esperimento non viene modificata, per ovvie ragioni di tempo¹. Queste analisi sono state effettuate per due differenti valori di δ (che corrisponde al *death rate*, ossia è inversamente proporzionale a quanto tempo un robot può rimanere nell'area nera prima di finire la sua energia ed essere sostituito).

Nella maggior parte dei casi, gli stati di equilibrio sono semplici: come nell'esperimento robotico, gli individui finiscono tutti per avere lo stesso comportamento, ossia per rimanere in una delle due aree (generalmente nell'area bianca, a meno che pb non sia estremamente alto). Tuttavia, l'analisi degli stati di equilibrio mostra che, soprattutto quando δ è basso, esistono condizioni in cui la

¹ Si potrebbe pensare a questi valori, per analogia, anche come alla *difficoltà* di imparare un dato comportamento. Se, come abbiamo detto, la velocità dell'apprendimento è proporzionale al tempo di permanenza in una data area, alte probabilità di capitare in un'area corrispondono ad una maggiore probabilità di imparare quel comportamento.

densità delle due aree raggiunge un equilibrio dinamico (non osservato nell'esperimento robotico): in altre parole, l'apprendimento sociale, in questi casi, non sembra portare all'uniformità dei comportamenti.

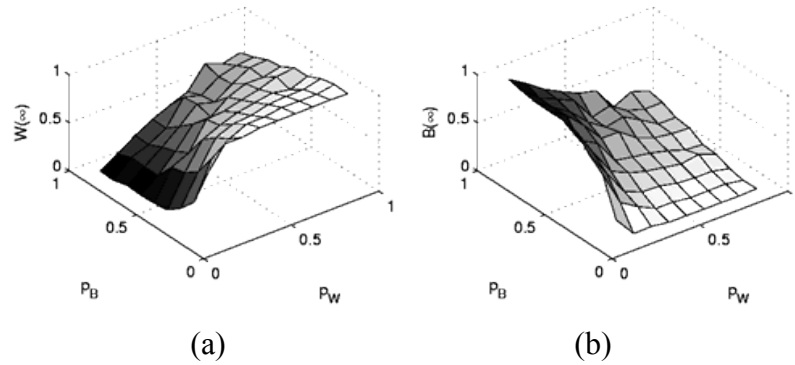


Fig. 5.9: Stati stazionari per valori bassi di δ . **(a)** area bianca. **(b)** area nera.

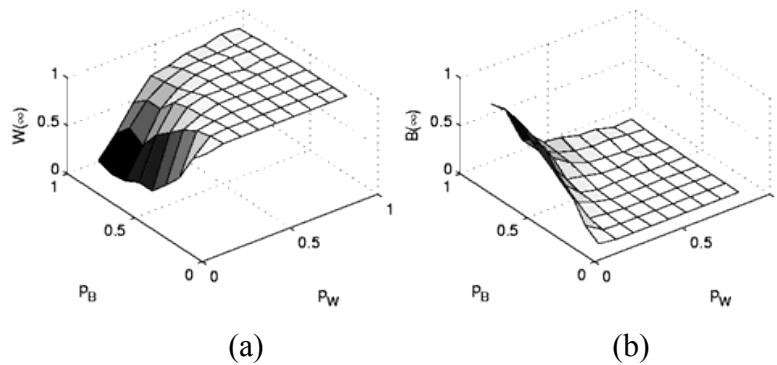


Fig. 5.10: Stati stazionari per valori alti di δ . **(a)** area bianca. **(b)** area nera.

Simulazioni e modelli matematici

Nel corso degli ultimi decenni sono stati sviluppati numerosi modelli analitici dell'apprendimento sociale. Questi modelli, che utilizzano soprattutto strumenti presi in prestito dalla genetica delle popolazioni (Cavalli – Sforza & Fedlman 1981) o dalla biologia evoluzionistica (Boyd & Richerson 1985) rappresentano sicuramente un importante contributo allo studio delle dinamiche di trasmissione ed evoluzione culturale.

Da una parte, essi possiedono molte delle caratteristiche che, nel primo

capitolo di questa tesi, sono state considerate come auspicabili per lo sviluppo di una teoria naturalizzata della cultura. Essi sono espressi in termini quantitativi, possono essere riprodotti in modo non ambiguo da altri ricercatori e permettono di ricavare predizioni empiriche precise dai loro risultati.

Dall'altra, però, sono ancora più astratti delle simulazioni ad agenti che sono state presentate in questo lavoro: il comportamento individuale non è, in genere, tenuto in considerazione, e l'interesse è quello di valutare come l'andamento di certi parametri influisca, *a livello popolazioneale*, sulle dinamiche culturali. Come si ricorderà, un argomento che è stato avanzato nel primo capitolo è che il mero realismo non è una buona guida per valutare la bontà di un modello. Perciò, il fatto che questi modelli siano molto astratti non è in sé, a mio avviso, un punto a sfavore del loro utilizzo. Tuttavia, nella maggior parte dei casi, questi modelli considerano la trasmissione culturale come un semplice processo di copia in cui, al massimo, del rumore statistico influisce sulla probabilità di acquisire o meno un tratto culturale che è, quasi sempre, una variabile e non un comportamento che richieda determinate capacità cognitive o un determinato tipo di interazioni sensomotorie con l'ambiente e via dicendo.

Come è stato discusso più volte discusso nel corso di questa tesi, ipotesi provenienti dall'antropologia cognitiva, dall'etologia e anche dagli studi di robotica e vita artificiale concordano nel sostenere che le dinamiche culturali a livello popolazioneale siano pesantemente influenzate dal tipo specifico dei comportamenti che vengono trasmessi, dai meccanismi di trasmissione che vengono utilizzati e dalle caratteristiche (cognitive) degli agenti implicati nel processo di trasmissione. Per questo motivo, ritengo che i modelli analitici, pur essendo utili, non possano rappresentare l'unico strumento per l'analisi delle dinamiche culturali. Tuttavia, come si è cercato di fare nel corso di questo capitolo, essi possono rappresentare un utile complemento di simulazioni più dettagliate. I modelli analitici possono essere utili per esplorare in modo rapido vasti spazi di parametri (cosa che non è quasi mai possibile con le simulazioni ad agenti) permettendo in questo modo: (1) di verificare che i risultati delle simulazioni non siano riconducibili ad un particolare set-up sperimentale, ma che possiedano una sufficiente generalità e (2) di *suggerire* spazi di parametri che potrebbero essere interessanti da studiare, in modo più dettagliato, con le simulazioni. Al contrario, le simulazioni, costringendo lo sperimentatore a realizzare i comportamenti che vengono trasmessi e i meccanismi specifici attraverso

i quali vengono trasmessi, possono orientare verso la costruzione di modelli matematici almeno in parte più realistici e, se possibile, più utili alla nostra comprensione.

Bibliografia

- Acerbi, A. (2007), “Le trasmissioni della cultura”, in: *Modelli della mente e processi di pensiero: il dibattito antropologico contemporaneo*, Lutri, A. (Ed.), Roma, Meltemi.
- Acerbi, A. (2006), “Antropologia cognitiva e modularità della mente: il caso della *folkbiology*”, in: *L’architettura della mente: saggi sulla modularità*, Ferretti, F. & Marraffa, M. (Eds.), Catanzaro, Abramo Editore.
- Acerbi, A. (2005). “Antropologia cognitiva: uno stato dell’arte”, *Sistemi Intelligenti*, 17 (3).
- Acerbi, A. (2003), “La mente nella cultura: cognizione ed analisi dei fatti culturali”, *XXIV Annali della Facoltà di Lettere e Filosofia dell’Università di Siena*, Firenze, Cadmo.
- Acerbi, A., Cecconi, F., Marocco, D. & Zappacosta, S. (2007), “Individual versus social learning in a population of autonomous robots”, paper presented to the workshop *Social Learning in Embodied Agents (ECAL 2007)*.
- Acerbi, A., Marocco, D. & Nolfi, S. (2007), “Social Facilitation on the Development of Foraging Behaviors in a Population of Autonomous Robots”, *Proceedings of the 9th European Conference on Artificial Life (ECAL 2007)*, Berlin, Springer.
- Acerbi, A. & Nolfi, S. (2007), “Social learning and cultural evolution in embodied and situated agents”, *Proceedings of the IEEE 2007 Symposium on Artificial Life*.
- Acerbi, A. & Nolfi, S. (2006), “Trasmissibilità culturale di tratti discreti e continui” in: *Scienze cognitive e robotica: Atti del III convegno AISC*, Greco, A. et al. (Eds.), Genova. Erga Edizioni.

- Acerbi, A. & Parisi, D. (2006), Cultural transmission between and within generations, *Journal of Artificial Societies and Social Simulation*, 9 (1).
- Adessi, E. & Visalberghi, E. (2006), “How social influences affect food neophobia in captive chimpanzees. A comparative approach”, in: Matsuzawa, T., Tomonaga, M. & Tanaka, M. (Eds.), *Cognitive development in chimpanzees*, Tokio, Springer – Verlag.
- Adams, B., Brazeal, C., Brooks, R. A. & Scassellati, B. (2000), Humanoid robots: a new kind of tool, *IEEE Intelligent Systems*, 15 (4).
- Atran, S. (1998), Folkbiology and anthropology of science: cognitive universals and cultural particulars, *Behavioral and Brain Sciences*, 21.
- Aunger, R. A. (Ed.) (2000), *Darwinizing culture*, Oxford, Oxford University Press.
- Axelrod, R. (1984), *Evolution of Cooperation*, New York, Basic Books.
- Baldassarre, G. (2001), Cultural evolution of ‘guiding criteria’ and behaviour in a population of neural-network agents, *Journal of Memetics*, 4.
- Barkow, J. H., Cosmides, L. & Tooby, J. (Eds.) (1992), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, New York, Oxford University Press.
- Blackmore, S. (2002), *La macchina dei memi*, Torino, Instar Libri.
- Boyd, R. & Richerson, P. J. (2006), *Non di soli geni: come la cultura ha trasformato l’evoluzione umana*, Torino, Codice.
- Boyd, R. & Richerson, P. J. (1995), Why does culture increase human adaptability?, *Ethology and Sociobiology*, 16.
- Boyd, R. & Richerson, P. J. (1985), *Culture and the Evolutionary Process*, Chicago, Chicago University Press.
- Boyd, R. & Richerson, P. J. (manoscritto), *The Nature of Cultures* (poi modificato e pubblicato con il titolo: *Not by genes alone*, University of Chicago Press).
- Boyer, R. (1994), *The Naturalness of Religious Ideas: Outline of a Cognitive Theory of Religion*, Los Angeles, University of California Press.
- Brooks, R. (1990), Elephants Don’t Play Chess, *Robotics and Autonomous Systems*, 6.
- Brooks, R. (1991), Intelligence without representation, *Artificial Intelligence*, 47.
- Cantwell Smith, B. (1999), “Situatedness / Embeddedness” in: *MIT Encyclopedia of the cognitive sciences*, Wilson, R. A. & Keil F. C. (Eds.), Cambridge, MIT Press.
- Cavalli – Sforza, L. L. & Feldman, M. (1981), *Cultural Transmission and Evolution:*

- a Quantitative Approach*, Princeton, Princeton University Press.
- Chiel, H. J. & Beer, R. D. (1997), The brain has a body: adaptive behaviour emerges from interactions of nervous system, body and environment, *Trends in Neurosciences*, 20 (12).
- Churchland, P. (1981), Eliminative materialism and the propositional attitudes, *Journal of Philosophy*, 78 (2).
- Cladière, N. & Sperber, D. (in stampa), The role of attraction in cultural evolution (reply to J. Henrich and R. Boyd, “On modelling cognition and culture”), *Journal of Cognition and Culture*.
- Clark, A. (1997), *Being There: Putting Brain, Body and World Together Again*, Cambridge, MIT Press.
- Dautenhahn, K. & Nehaniv, C. L. (2002), “The correspondence problem”, in: *Imitation in Animals and Artifacts*, Dautenhahn, K. & Nehaniv, C. L. (Eds.), Cambridge, MIT Press.
- Denaro, D. & Parisi, D. (1996), “Cultural evolution in a population of neural networks”, in: *Neural Networks. WIRN '96*, Marinaro, M. & Tagliaferri, R. (Eds.), New York, Springer.
- Dennett, D. C. (2005), “From Typo to Thinko: When Evolution Graduated to Semantic Norms”, in: *Evolution and Culture*, Levinson, S. C. & Jaisson, P. (Eds.), Cambridge, MIT Press.
- Dennett, D. C. (1995), *L'idea pericolosa di Darwin: l'evoluzione e i significati della vita*, Torino, Bollati – Boringhieri.
- Dennett, D. C. (1991), *Brainstorms: saggi filosofici sulla mente e la psicologia*, Milano, Adelphi, 1991.
- Dennett, D. C. & McKay, R. (2006), A continuum of mindfulness (commentary to Mesoudi *et al.*), *Behavioral and Brain Sciences*, 29 (4).
- Di Paolo, E., Noble, J. & Bullock, S. (2000), “Simulation models as opaque thought experiments”, in: Bedau, M. A., McCaskill, J. S., Packard, N. H. & Ramussen, S. (Eds.), *Artificial Life VII, Proceedings of the Seventh International Conference on Artificial Life*, Cambridge, MIT Press.
- Enquist, M., Eriksson, K. & Ghirlanda, S. (working paper), *Critical social learning: a solution to Rogers' paradox of non – adaptive culture*.
- Epstein, J. M. & Axtell, R. (1996), *Growing Artificial Societies: Social Science from the Bottom up*, Cambridge, MIT Press.

- Galef, B. G. (1996) "Social enhancement of food preferences in Norway rats. A brief review" in: Heyes, C. M. & Galef, B. G. (Eds.) *Social learning in animals: the roots of culture*, San Diego, Academic Press.
- Galef, B. G. & Laland, K. N. (2005), Social learning in animals: empirical studies and theoretical model, *BioScience*, 55 (6).
- Gandolfi, A. (1999), *Formicai, imperi e cervelli: introduzione alla scienza della complessità*, Torino, Bollati Boringhieri.
- Geertz, C. (1973), *The Interpretations of Culture*, New York, Basic Books.
- Grimm, V. & Railsback, S. F. (2005), *Individual – based Modeling and Ecology*, Princeton, Princeton University Press.
- Harris, M. (1968), *The Rise of Anthropological Theory: A History of Theories of Cultures*, New York, Crowell.
- Harris, M. (1991), "Anthropology: Ships That Crash in the Night" in: Jessor E. (Ed.), *Perspectives on Behavioral Science: The Colorado Lectures*, Boulder, Westview Press.
- Henrich, J. & Boyd, R. (2002), On modelling Cognition and Culture: Why Cultural Evolution Does Not Require Replication of Representations, *Culture and Cognition*, 2.
- Hewlett, B. S. & Cavalli – Sforza, L. L. (1986), Cultural transmission among Aka pygmies, *American Anthropologist*, 88 (4).
- Hirschfeld, L. (1996), *Race in the Making: Cognition, Culture and the Child's Construction of Human Kinds*, Cambridge, MIT Press.
- Hirschfeld, L. & Gelman, S. (Eds.) (1994), *Mapping the mind: Domain specificity in cognition and culture*, Cambridge, Cambridge University Press.
- Holland, J. H. (1995), *Hidden Order: how Adaptation Builds Complexity*, Redwood, Addison – Wesley.
- Holland, J. H. (1975), *Adaptation in Natural and Artificial Systems*, Ann Arbor, University of Michigan Press.
- Hurley, S. & Chater, N. (Eds.) (2005), *Perspective on imitation: from neuroscience to social science*, Cambridge, MIT Press.
- Hutchins, E. (1995), *Cognition in the wild*, Cambridge, MIT Press.
- Hutchins, E. & Hazlehurst, B. (1995), "How to invent a lexicon: the development of shared symbols in interaction", in: *Artificial Societies: The computer simulation of social life*", Gilbert, N. G. & Conte, R. (Eds.), London, UCL Press.

- Kohler, T. A. & Gumerman (Eds.) (2000), *Dynamics in Human and Primate Societies. Agent – based modelling of social and spatial processes*, Oxford, Oxford University Press.
- Kirkpatrick, C. D., Gelatt, J. & Vecchi, M. P. (1983), Optimization by simulated annealing, *Science*, 220 (4598).
- Knafo, A. & Schwartz, S. H. (2001), Value socialization in families of Israeli-born and Soviet-born adolescents in Israel, *Journal of Cross-Cultural Psychology*, 32 (2).
- Langton C. G. (Ed.) (1989), *Artificial Life*, Redwood, Addison – Wesley.
- Marr, D. (1982), *Vision*, New York, Freeman.
- McCulloch, W. S. & Pitts, W. H. (1943), A logical calculus of the ideas immanent in nervous activity, *Bulletin of Mathematical Biophysics*, 5.
- Meltzoff, A. & Moore, M. (1983), Newborn infants imitate adult facial gesture, *Child Development*, 54.
- Mesoudi, A., Whiten, A. & Laland, K. N. (2006), Towards a unified science of cultural evolution (with commentary), *Behavioral and Brain Sciences*, 29 (4).
- Mineka, S. & Cook, M (1988), “Social Learning and the acquisition of snake fear in monkeys”, in: Zentall, T. R. & Galef, B. G. (Eds.), *Social Learning: Psychological and Biological Perspectives*, Hillsdale, Lawrence Erlbaum Associates.
- Mitchell, M. (1996), *An introduction to genetic algorithms*, Cambridge, MIT Press.
- Nadel, L. & Stein, D. (Eds.) (1991), *1990 lectures in complex systems*, Reading, Addison – Wesley.
- Noble, J. & Todd, P. (2002), “Imitation or something simpler?”, in: *Imitation in Animals and Artifacts*, Dautenhahn, K. & Nehaniv, C. L. (Eds.), Cambridge, MIT Press.
- Nolfi, S. (2006), Behaviour as a complex adaptive system: on the role of self – organization in the development of individual and collective behaviour, *ComplexUs*, 3 (3/4).
- Nolfi, S. & Floreano, D. (2000), *Evolutionary Robotics. The Biology, Intelligence, and Technology of Self-Organizing Machines*, Cambridge, MIT Press.
- Nolfi, S. & Floreano, D. (1999), Learning and Evolution, *Autonomous Robots*, 7 (1).
- Nolfi, S., Elman, J. L. & Parisi, D. (1994), Learning and evolution in neural networks, *Adaptive Behavior*, 3 (1).

- Parisi, D. (2003), "Evolutionary Psychology and Artificial Life", in *Evolutionary Psychology: Alternative Approaches*, Scher, S. J. & Rauscher, F. (Eds.), Dordrecht, Kluwer.
- Parisi, D. (2001), *Simulazioni*, Bologna, Il Mulino.
- Parisi, D. (1997), Cultural evolution in neural networks, *IEEE Experts*, 12.
- Rogers, A. (1989), Does biology constrain culture?, *American Anthropologist*, 90.
- Rumelhart, D. E. & McClelland, J. L. (1991), *PDP. Microstruttura dei processi cognitivi*, Bologna, Il Mulino.
- Schelling, T. S. (1971), Dynamic Models of Segregation, *Journal of Mathematical Sociology*, 1.
- Sherry, D. F. & Galef, B. J. (1984), Cultural transmission without imitation: Milk bottle opening by birds, *Animal Behaviour*, 32 (3).
- Sperber, D. (1999), *Il contagio delle idee*, Milano, Feltrinelli.
- Sperber, D. (1997), "Individualisme methodologique et cognitivisme", in: *Cognition et sciences sociales*, Paris, P.U.F.
- Taylor, M. C. (2005), *Il momento della complessità*, Torino, Codice Edizioni.
- Tennie, C., Call, J. & Tomasello, M. (2006), Push or Pull: Imitation vs. Emulation in Great Apes and Human Children, *Ethology*, 112.
- Tomasello, M. (2005), *Le origini culturali della cognizione umana*, Bologna, Il Mulino.
- Tommasello, M. (1996), "Do apes ape?", in: *Social learning in animals: the roots of culture*, Heyes, C. M. & Galef, B. G. (Eds.), San Diego, Academic Press.
- Tomasello, M., Kruger, A. C. & Ratner, H. H. (1993), Cultural learning, *Behavioral and Brain Sciences*, 16.
- Tooby, J. & Cosmides, L. (1992), "The psychological foundations of culture", in: *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, cit.
- Turing, A. (1936), On computable numbers, with an application to the Entscheidungsproblem, *Proc. London Math Soc.*, 42.
- van Gelder, T. J. (1998), The dynamical hypothesis in cognitive science, *Behavioral and Brain Sciences*, 21.
- Visalberghi, E. & Fragaszy, D. (1990), "Do monkeys ape?", in: "Language" and intelligence in monkeys and apes, Parker, S. & Gibson, K. (Eds.), Cambridge, Cambridge University Press.
- Waldrop, M. (1995), *Complessità. Uomini e idee al confine tra ordine e caos*, Torino,

Instar.

Whiten, A., Horner, V. & de Wall, F. B. M. (2005), Conformity to cultural norms of tool use in chimpanzees, *Nature*, 437.

Widrow, B. & Hoff, M. E. (1960), “Adaptive switching circuits”, in: *IRE WESCON Convention Record*, IV.

Wiener, N. (1982), *La cibernetica: controllo e comunicazione nell'animale e nella macchina*, Milano, Il Saggiatore.

Wilensky, U. (1988), NetLogo Segregation Model.

<http://ccl.nortwestern.edu/netlogo/models/Segregation>

Wolf, E. (1964), *Anthropology*, Engelwood Cliff, Prentice – Hall.