

# Informazione, ordine individuale, ordine sociale

di Giorgio Rampa

## 1. Introduzione

In buona parte dell'analisi economica recente il problema dell'informazione a disposizione di un decisore viene impostato nel modo seguente (o in modi equivalenti)<sup>1</sup>. Dato un insieme di stati del mondo possibili, un canale invia al decisore uno o più segnali che dipendono da quale fra i possibili mondi è davvero quello in cui il decisore vive senza saperlo con certezza. Esempi di canale possono essere quelli che danno luogo a partizioni (il decisore viene edotto circa quale tra i diversi eventi che la sua partizione gli consente di discriminare contenga il vero stato del mondo); oppure quelli che, al variare dello stato del mondo vigente, danno luogo a distribuzioni differenti su un preassegnato insieme di segnali (segnali con rumore). Se si suppone che il decisore possenga una distribuzione di probabilità iniziale sugli stati del mondo, un segnale «affina» la sua percezione del mondo (nel senso che concentra la sua distribuzione posteriore sugli stati)<sup>2</sup>; in questo senso si dice che un segnale è informazione. Notiamo che il contesto è dinamico; la ricezione del segnale è logicamente e cronologicamente successiva alla definizione degli stati possibili e del tipo di canale.

Introducendo d'altra parte un opportuno insieme di scelta (per

*Ringrazio Paolo Bertoletti e Nicolò De Vecchi per gli utili commenti; sono grato anche ad un referee anonimo, le cui osservazioni hanno messo in evidenza gli aspetti ingenui o oscuri di una precedente versione del lavoro; costoro non sono ovviamente responsabili del risultato finale. La ricerca ha fruito di un finanziamento MURST 40%.*

<sup>1</sup> Si veda per esempio Laffont [1989, cap. 4]; Rasmusen [1989, cap. 2]; Hirschleifer e Riley [1992, cap. 5]. Si veda anche Luini [1994].

<sup>2</sup> Nel caso di partizione finita, in seguito all'arrivo di un segnale non rumoroso alcuni eventi della partizione hanno probabilità posteriore nulla, e la massa si concentra sui rimanenti. Nel caso invece di segnale con rumore, per una famiglia molto ampia di distribuzioni iniziali sugli stati del mondo, ogni nuovo segnale provoca generalmente la riduzione di qualche misura di variabilità della distribuzione posteriore: si rinvia a De Groot [1970].

esempio lotterie, cioè distribuzioni, su conseguenze rilevanti per il payoff; oppure atti, cioè funzioni da stati del mondo a conseguenze) e su di esso un ordinamento di preferenze che risponda a certi assiomi (rispettivamente, quelli di von Neumann e Morgenstern, oppure quelli di Savage)<sup>3</sup>, si perviene alla caratterizzazione del comportamento razionale del decisore come massimizzatore dell'utilità attesa: l'opzione prescelta (lotteria oppure atto) è quella che massimizza la media delle utilità delle conseguenze, con pesi dati dalle loro probabilità (queste ultime sono definite direttamente dalle lotterie, oppure sono indotte dagli atti a partire dalla distribuzione sugli stati del mondo).

Va osservato che la rappresentazione in termini di lotterie è una sorta di forma ridotta della rappresentazione in termini di atti e stati del mondo. Quest'ultima prevede (nel senso che caratterizza il comportamento come equivalente a) quanto segue: il decisore sa descrivere l'insieme dei mondi possibili; ha una distribuzione su di essi (una distribuzione iniziale, il prior, oppure una distribuzione posteriore in seguito all'arrivo di segnali tramite il canale); e sa definire le diverse conseguenze che corrispondono ai diversi stati, dato ogni atto da lui prescelto. Dunque in ultima istanza, poiché la scelta di un atto induce una distribuzione sulle conseguenze, dal punto di vista della decisione un atto equivale a una lotteria. Ma la rappresentazione diretta in termini di lotterie contrae tutti i passaggi precedenti; considerarli in modo esplicito sembra invece importante quando si discuta di conoscenza e informazione. In effetti, la maggior parte delle applicazioni a contesti specifici consiste esattamente nel prendere come punti di partenza una descrizione probabilistica del mondo esterno, una tecnologia che connette le azioni individuali e le circostanze esterne alle conseguenze per il decisore, e una funzione di utilità definita su queste ultime. Questo pare il modo più naturale di descrivere un comportamento di scelta: la descrizione del problema in forma «strutturale» è usualmente preferita a quella in forma «ridotta» (tranne nei casi in cui l'oggetto di scelta siano proprio biglietti di lotterie diverse).

Queste pagine offrono una serie di osservazioni sulla prima com-

<sup>3</sup> Von Neumann e Morgenstern [1944]; Savage [1954]. Esistono caratterizzazioni della scelta ottimale sotto incertezza in termini di utilità attesa che precedono temporalmente le due citate, come quelle di Ramsey, Kolmogorov e de Finetti: per una rassegna si rinvia a Fishburn [1982] e Muliere e Parmigiani [1993].

<sup>4</sup> Recentemente si è creato un certo interesse per alcune generalizzazioni nel senso dell'utilità «non attesa» (l'indicatore di utilità di un prospetto incerto non è cioè lineare nelle probabilità delle conseguenze); non è tuttavia questo l'aspetto su cui vorremmo concentrare l'attenzione in questo lavoro. Su tali aspetti si rinvia a Machina [1982; 1987]; si veda inoltre il contributo di Bernasconi al presente volume.

ponente del discorso, quella relativa alla descrizione del mondo esterno (e al modo in cui un canale ne può «svelare» degli aspetti); non ci occuperemo invece degli aspetti che riguardano più da vicino la tecnologia e l'utilità. In particolare, nel paragrafo 2 si argomenta che la rappresentazione del problema dell'informazione usualmente adottata nei modelli economici possa risultare insoddisfacente in molte circostanze importanti, specie quando si tratti di interazione sociale e non solo di decisione individuale. Si sostiene che le ipotesi accettate da quella modellistica presuppongono un eccessivo grado di conoscenza da parte dei decisori: dunque, queste note si inseriscono nel filone della «razionalità limitata» (ma meglio sarebbe dire «non illimitata»). Nel paragrafo 3, prendendo spunto dalla «teoria della decisione basata sui casi» recentemente proposta da Gilboa e Schmeidler, si suggerisce che la rappresentazione del comportamento dei decisori in termini di reti neurali può costituire una linea di analisi promettente. Passando all'interazione sociale, il paragrafo 4 deriva da questa impostazione una nozione di equilibrio, o *ordine sociale*, fondata principalmente sugli elementi congetturali dell'interazione; di conseguenza, il concetto di informazione assume una connotazione diversa da quella usualmente adottata da molti economisti: non «disvelamento del vero», ma *linguaggio* che produce coerenza fra i mondi individuali. Come prodotto congiunto, si argomenta che la nozione di ordine congetturale qui prospettata implica la possibilità (se non la necessità) logica della rottura dell'ordine stesso.

## 2. Informazione: tra ingegneria e teoria economica

Nella teoria tradizionale dell'informazione, la definizione dei mondi possibili è un dato non problematico per il decisore, che è supposto in grado di enumerarli e descriverli con precisione, anche se non sa con certezza quale sia lo stato vero. Il problema è invece costituito dalla natura tecnologica del canale che invia segnali, i quali sono tipicamente imprecisi (non esiste una funzione biunivoca fra stati e segnali).

Lo studio dei problemi di informazione dal punto di vista ingegneristico<sup>3</sup> ha infatti incentrato l'attenzione su quelle proprietà dei canali che li rendono più o meno efficienti. Per fare un esempio, ci si

<sup>3</sup> A partire almeno da Shannon e Weaver [1949]; si veda anche Gallager [1968]. Per una interessante rassegna si rinvia a Aspray [1985].

aspetta intuitivamente che codificare la descrizione degli stati mondo<sup>6</sup> e trasmettere tale codifica sia meno impegnativo, in termini di requisiti tecnici del canale e quindi di risorse fisiche, nel caso in cui alcuni stati ricorrano più frequentemente di altri, rispetto al caso in cui essi siano equi-probabili: questo perché in presenza di molti stati relativamente poco probabili gli errori o le ambiguità di trasmissione sono più tollerabili (la probabilità di non conoscere lo stato vero è più bassa). Dualmente, un canale con preassegnate caratteristiche tecniche può convogliare informazioni attendibili solo quando il numero degli stati possibili (o maggiormente probabili) non ecceda un certo limite. Osservazioni di questo tipo hanno per esempio dato luogo all'adozione dell'entropia quale misura dell'ammontare di informazione che un canale è in grado di fornire<sup>7</sup>: più precisamente, l'entropia misura il requisito minimo (in termini di cardinalità, cioè numerosità dei simboli diversi, del dizionario codificatore) necessario per trasmettere segnali non ambigui circa gli stati del mondo. Dunque, si tratta di una misura indiretta della difficoltà di rappresentare il mondo: talora è stato usato anche il termine «incertezza» a tale proposito<sup>8</sup>.

NA) Come ha osservato ripetutamente Marschak<sup>9</sup>, non è detto che l'ammontare di informazione, così come misurato dall'entropia, sia una corretta valutazione del valore informativo<sup>10</sup> che un canale ha per il decisore: in effetti, molto dipende da come è configurata la funzione di utilità del decisore stesso. Piuttosto, misure quali l'entropia (o la capacità, o ancora il tasso di trasmissione per unità di tempo) possono essere utilizzate per valutare il *costo* di un canale, che ragione-

<sup>6</sup> Naturalmente questo tipo di letteratura non parla di stati del mondo da identificare ma di messaggi da inviare. Non è tuttavia difficile rappresentare il problema in termini di stati.

<sup>7</sup> Nel caso di distribuzioni  $p$  su un supporto finito  $X$ , l'entropia è definita come  $H(p) = -\sum_{x \in X} p(x) \log(p(x))$ , una funzione che raggiunge un massimo per  $p$  uniforme:

in questo caso si ha  $H(p) = \log(|X|)$ , dove  $|X|$  è la cardinalità (il numero di elementi) di  $X$ .

<sup>8</sup> Contro questa identificazione *tout court* di entropia, incertezza e informazione si sono schierati più volte Georgescu e Roegen: si veda in particolare Georgescu e Roegen [1975].

<sup>9</sup> Si vedano i contributi raccolti in Marschak [1974, vol. II].

<sup>10</sup> Per valore informativo di un segnale si deve intendere, secondo Marschak, l'incremento di utilità attesa che il decisore può ottenere utilizzando il canale anziché non utilizzandolo: si tratta del valore lordo, dal quale occorre poi detrarre il costo del canale.

volmente dipende dalla precisione con cui si vogliono trasmettere i segnali.

La letteratura ingegneristica si è occupata anche di un problema più vicino agli interessi degli economisti, vale a dire quello della scelta di un canale ottimale fra quelli disponibili per trasmettere segnali relativi ad un insieme preassegnato di stati possibili. Quando tale problema venga impostato in modo da dare un giudizio indipendente da qualsiasi specifica funzione di utilità del decisore (e quindi in termini nuovamente tecnico-ingegneristici), la soluzione è quella individuata da Blackwell<sup>11</sup>. Nel caso di canali rumorosi (con numero finito di stati del mondo e di segnali), ogni canale è caratterizzato da una matrice di verosimiglianze (cioè di probabilità dei diversi segnali condizionatamente ai diversi stati); si dice allora che il canale A è almeno tanto informativo quanto il canale B se la matrice delle verosimiglianze di B può essere vista come una trasformazione lineare di quella di A<sup>12</sup>. Nel caso invece di canali che danno luogo a partizioni, un canale A è almeno tanto informativo quanto un canale B se la partizione indotta da A è non meno fine di quella indotta da B<sup>13</sup>. Ordinamenti parziali di questo genere consentono naturalmente di eliminare i canali dominati<sup>14</sup>. Queste nozioni costituiscono la base di molta teoria economica dell'informazione, che di suo aggiunge ovviamente l'ingrediente della funzione di utilità del decisore: quest'ultima informa la scelta del canale individualmente ottimale tra quelli non dominati nel senso di Blackwell.

Il tratto essenziale di molte applicazioni economiche è che il decisore, oltre ad avere una percezione chiara e non problematica degli stati del mondo possibili, conosce perfettamente le caratteristiche dei diversi canali disponibili (partizioni o verosimiglianze). Ciò è vero per le due branche in cui si può suddividere l'economia dell'informazione: quella che si occupa del problema se comprare un canale, e quale comprare; e quella che studia gli effetti di strutture alternative dell'in-

<sup>11</sup> Blackwell [1953].

<sup>12</sup> In altri termini, i segnali trasmessi da B non possono dire niente di più di quelli provenienti da A, essendone in un certo senso una combinazione lineare; possono anche dire di meno, se la matrice che combina le linee della matrice della verosimiglianze di A per ottenere quella di B è di rango non pieno: in tal caso si dice anche che il canale B è una «distorsione» (garbling) di A.

<sup>13</sup> La partizione P' è non meno fine di P'' se ogni elemento di P' è sottoinsieme di un solo elemento di P''.

<sup>14</sup> Non è difficile dare anche caratterizzazioni delle classi di equivalenza di questi ordinamenti, cioè individuare proprietà degli insiemi dei canali ugualmente informativi: si veda Arrow [1992].

formazione (cioè della distribuzione dei canali fra i decisori) sulle scelte delle azioni individuali<sup>15</sup>. Per esempio fa parte di questa seconda branca lo studio delle proprietà di efficienza paretiana degli equilibri che si generano in corrispondenza di differenti strutture informative. In entrambe le branche<sup>16</sup> sorge un problema di informazione solo in quanto il canale non è costituito da una funzione biunivoca dall'insieme degli stati a quello dei segnali; viceversa, la descrizione di tali insiemi e le proprietà dei canali sono conosciute con certezza. Per quanto riguarda in particolare la seconda branca di ricerca, si inseriscono in questo paradigma le varie definizioni di informazione in qualche modo difettosa: informazione incompleta (non si conosce con certezza il corso intrapreso dall'agente esogeno «Natura», tra quelli che si sa essere possibili)<sup>17</sup>; informazione imperfetta (la versione, nel linguaggio dei giochi dinamici, del fatto che la partizione a disposizione di un agente è a qualche nodo meno fine della collezione degli stati elementari, pur essendo nota la struttura dell'albero); informazione asimmetrica (agenti diversi, concordando sulla descrizione delle alternative possibili, hanno accesso a canali diversi tra loro non equivalenti). In particolare, quest'ultima nozione è stata ampiamente utilizzata di recente sia in microeconomia (teoria dei contratti) sia in macroeconomia (nuovi keynesiani); essa non fa eccezione rispetto al paradigma di cui sopra: la definizione di contratti in presenza di informazione asimmetrica è modellata supponendo che le parti conoscano in cosa consiste esattamente l'asimmetria (i canali informativi sono, come si dice, conoscenza comune).

Ora, non si può negare che questo paradigma, oltre ad essere pienamente coerente con la visione ingegneristica (dove la rappresentazione del mondo ed i suoi significati non costituiscono oggetto di in-

<sup>15</sup> Si noti che il problema del comprare (oppure vendere) o meno «informazioni» (per esempio brevetti tecnologici, oppure notizie sull'andamento del raccolto dall'altra parte del pianeta) ricade nella seconda branca. Infatti questo problema si pone se decisori diversi hanno già a disposizione canali privati, che sanno essere diversi tra loro (in caso contrario non vi sarebbe l'asimmetria che giustifica, appunto, la compravendita di informazioni).

<sup>16</sup> La stessa bipartizione si ripropone nel caso in cui il decisore sia una squadra: in tal caso si pone sia il problema della progettazione ottima della rete di comunicazioni interna all'organizzazione, sia quello della scelta delle azioni ottimali in risposta ad ogni diverso segnale che passa per una rete data. Si veda per esempio Marschak [1954, in Marschak 1974, cap. 20]; Marschak e Radner [1972], e, più recentemente, Green e Laffont [1986] e Radner [1993].

<sup>17</sup> Questa caratterizzazione dell'incompleta informazione non è la stessa che si può trovare in altri contesti: essa è mutuata dalla teoria bayesiana dei giochi.

dagine)<sup>18</sup>, possa ben adattarsi a situazioni di scelta non problematiche (ben codificate, appunto). Ma il punto è un altro: sembra infatti legittimo chiedersi in che modo quelle situazioni di scelta siano diventate non problematiche; occorre cioè giustificare l'assunto che la rappresentazione dei mondi possibili non è una difficoltà per il decisore.

Tale questione, come è facile immaginare, diventa particolarmente pressante nel caso di molti decisori che interagiscono, dove sorgono problemi di intersoggettività. Infatti, se ci si limita alla teoria della decisione individuale e si assume una prospettiva soggettivista (coerente con la teoria bayesiana della decisione individuale), si può ben sostenere che la definizione dei mondi possibili e della loro relazione con i segnali potenzialmente osservabili siano un atto di libera scelta teorica da parte del decisore: è la formulazione di un modello interpretativo. Eventualmente, si può richiedere (come fa appunto l'impostazione bayesiana) che tale modello rispetti certi criteri di coerenza interna, ma nulla più. Non sorge alcun problema di comparabilità interpersonale della descrizione degli stati del mondo possibili.

A dire il vero, una parte della letteratura recente in tema di decisione individuale ha messo in discussione la nozione stessa di coerenza interna del modello interpretativo; ciò a partire dall'osservazione che in molti casi può essere difficile definire in modo preciso l'insieme dei mondi possibili, oppure specificare in modo convenzionale le implicazioni conoscitive dell'arrivo di un segnale. Questi fenomeni vengono incasellati nella categoria della razionalità limitata (rispetto ai canoni standard della teoria bayesiana).

La razionalità limitata può trarre origine da circostanze diverse. Per esempio, essa può derivare da problemi di non computabilità, dovuti al fatto che molte questioni aritmetiche (che servono per esempio a definire le proprietà di certi stati del mondo) non possono essere risolte in modo completo tramite un processo finito di calcolo. Si pensi, banalmente, all'espansione decimale di un numero irrazionale: si noti che talora il problema non può essere risolto postulando

<sup>18</sup> «Il termine *informazione*, in questa teoria, è utilizzato in un senso speciale che non deve essere confuso con il suo utilizzo nel linguaggio ordinario. In particolare, "informazione" non va confuso con "significato". Infatti, due messaggi, dei quali uno sia profondamente denso di significato e l'altro sia puro nonsenso, possono essere esattamente equivalenti, dal nostro punto di vista, in termini di informazione» [Shannon e Weaver 1949, 9].

<sup>19</sup> In questa prospettiva cade l'interpretazione ingegneristica dell'informazione come fenomeno tipicamente intersoggettivo: uno stesso segnale può significare cose diverse per decisori diversi. Su questo punto torneremo in seguito.

che una soluzione approssimata sia equivalente a quella precisa, perché può accadere che il problema venga semplicemente spostato al livello della ricerca dell'algoritmo che effettua la migliore approssimazione per la classe di questioni aritmetiche in questione (tale ricerca un altro processo di calcolo). Questo punto è strettamente connesso con il problema gödeliano della indecidibilità<sup>20</sup>. La conseguenza è che la descrizione degli stati possibili può rimanere incompleta, e dunque alcuni segnali possono non avere «significato» per il decisore<sup>21</sup>.

Un'altra via per motivare la razionalità limitata consiste nell'ammettere che il decisore non sia in grado di trarre tutte le implicazioni logiche (teoremi) conseguenti alle conoscenze e informazioni in suo possesso: ciò perché è in grado di tenere sotto controllo solo un limitato numero di passaggi logici concatenati<sup>22</sup>. Questo punto è solo parzialmente connesso con il precedente, poiché non deriva da impossibilità di tipo aritmetico, ma da limiti logici, desumibili (senza ragionevole possibilità di contestazione, crediamo) dall'introspezione psicologica o anche dall'osservazione di ciò che accade nelle comunità scientifiche: neppure il più brillante economista ha scritto sin dall'inizio della propria carriera un unico libro contenente tutti i teoremi ricavati in seguito da lui stesso o potenzialmente ricavabili dalla sua scuola.

Una implicazione di questo stato di cose è che il decisore potrebbe trarre dall'informazione disponibile deduzioni tra loro logicamente contraddittorie<sup>23</sup>, oppure deduzioni incompatibili con l'usuale rappresentazione per partizioni (per esempio, se io non so *E* non ne segue che so di non saperlo)<sup>24</sup>. Mentre queste ultime considerazioni sembrano venire incontro al disagio di chi dubita della generale fattibilità di una rappresentazione completa e coerente degli stati del mondo, c'è in alcune trattazioni un aspetto che certamente disturba: infatti, nel tentativo di costruire una teoria più generale che renda conto di questa possibilità (e di cui l'usuale approccio bayesiano sia un caso particolare), ci si costringe a considerare situazioni decisionali in cui l'insieme degli stati possibili è addirittura ampliato rispetto all'analisi

<sup>20</sup> Per una trattazione esauriente si veda Cutland [1980].

<sup>21</sup> Questo conduce per esempio alla formulazioni di contratti incompleti dovuti non ad asimmetrie informative ma alla verificabilità computazionale di certi eventi: per esempio Anderlini e Felli [1994], oppure Lipman [1995a].

<sup>22</sup> Su questa via si è mosso per esempio Bacharach [1992].

<sup>23</sup> Lipman [1995b] considera questa possibilità.

<sup>24</sup> Alcune implicazioni di queste strutture informative «generalizzate» nel caso di interazione strategica sono considerate in Brandenburger, Dekel e Geanakoplos [1992].

usuale!<sup>25</sup> Non si vede come si possa parlare di razionalità limitata su queste basi.

Molto più importante agli occhi di chi scrive è invece una diversa causa di difficoltà per la definizione coerente dei mondi possibili.

Questa causa fa la sua apparizione non appena si studino problemi di interazione anziché di decisione individuale. Qui (posto di non assumere informazione completa, cioè assenza di problemi informativi), ogni stato del mondo possibile agli occhi di un giocatore deve includere la descrizione delle caratteristiche degli altri giocatori. La parabola dei «tipi» di Harsanyi<sup>26</sup> affronta esattamente questo problema, con l'obiettivo di ridurre la teoria della scelta razionale nei giochi a caso particolare della teoria bayesiana della decisione individuale. L'incompletezza dell'informazione viene rappresentata da Harsanyi assumendo che ogni giocatore possa essere, per capriccio della Natura, uno fra molteplici tipi. Poniamo che l'insieme dei tipi possibili sia ben specificato e noto a ogni giocatore; allora, se un giocatore possiede una distribuzione di probabilità sull'insieme dei tipi, che induce una distribuzione sulle azioni degli altri giocatori, la sua scelta individualmente razionale in senso bayesiano è ben definita. Inoltre, se esiste una distribuzione «oggettiva» dei tipi che sia conoscenza comune e se quelle soggettive dei giocatori sono opportunamente vincolate a quella oggettiva, allora le scelte individualmente razionali costituiscono un equilibrio di Bayes e Nash<sup>27</sup>.

Tralasciamo le altre (ipotesi) (che certamente non sono di poco conto), e riflettiamo su quella secondo cui l'insieme dei tipi possibili è ben definito e noto a tutti. La descrizione di un tipo da parte di un decisore deve includere tutte le caratteristiche di quel tipo che siano rilevanti per la valutazione del guadagno atteso del decisore; la descrizione deve dunque includere anche il modo in cui quel tipo rappresenta la propria incertezza sugli altri tipi. In altri termini, l'insieme dei tipi possibili, diciamo  $\Theta$ , deve includere l'insieme delle distribuzioni sull'insieme dei possibili tipi,  $\Delta(\Theta)$ . Si intuisce facilmente, ed è stato dimostrato formalmente<sup>28</sup>, che se la definizione dello spazio di

<sup>25</sup> I medesimi lavori citati nelle precedenti due note sono un esempio di questo modo di procedere.

<sup>26</sup> Harsanyi [1967-1968].

<sup>27</sup> Aumann [1987] ha generalizzato questa idea; nel suo lavoro la conoscenza comune dell'insieme dei tipi possibili è sostituita dall'ipotesi che le partizioni individuali sugli stati del mondo sono conoscenza comune.

<sup>28</sup> Si veda per esempio Mertens e Zamir [1985] e Tan e Werlang [1989]; in un contesto diverso, analoga dimostrazione è data da Lipman [1991].

tutti i tipi possibili deve avere una forma chiusa (deve cioè essere condivisa da tutti i tipi), questo non può avvenire in uno spazio con un numero finito di dimensioni. Infatti, sia data una qualsiasi descrizione probabilistica  $\Delta_a$  dello spazio dei tipi che il giocatore  $a$  possa formulare circa i suoi partner; questa descrizione è una fra le varie possibili appartenenti ad una famiglia  $D_a$ . Una descrizione probabilistica  $\Delta_b(D_a)$  delle possibili configurazioni di  $\Delta_a$  (cioè dei tipi di  $a$ ) formulata dal giocatore  $b \neq a$  non può che essere più «potente»:  $\Delta_b(D_{-b}) = \Delta_b(\prod_{a \neq b} D_a)$  è una teoria più elaborata della teoria  $\Delta_a$ . Lo

spazio degli stati del mondo possibili sulla cui descrizione tutti i giocatori possano concordare deve essere così esteso da includere le distribuzioni di probabilità su se stesso<sup>29</sup> (non senza qualche assonanza con paradossi russelliani).

Poiché chi scrive ha già tentato di discutere altrove di questi problemi<sup>30</sup>, qui non vi insisteremo oltre. Notiamo solo che la definizione dello spazio «universale» dei tipi è tale che, prima ancora di poter passare a formulare una distribuzione di probabilità su di esso, per nominare un solo tipo occorre elencare un vettore con infinite coordinate. Se si prendono sul serio le precedenti osservazioni sulle limitazioni delle capacità di calcolo e di deduzione logica dei decisori, questo è semplicemente incredibile.

### 3. Decidere per casi, conoscere per casi

Recentemente, Gilboa e Schmeidler [1995a] hanno proposto una teoria della decisione «cognitivamente più plausibile» (secondo le parole degli stessi Gilboa e Schmeidler) rispetto alla teoria bayesiana della decisione individuale<sup>31</sup>: la «teoria della decisione basata sui casi». Con ciò si deve intendere, per questi autori, una teoria che non richieda ai

<sup>29</sup> In effetti, sia  $\Delta \in D$  il vettore dei  $\Delta_a \in D_a$ , cioè una configurazione di tipi, uno per ogni giocatore; sia poi  $\Delta_2(D) \in D_2$  il vettore degli operatori  $\Delta_b(D_{-b})$  definiti nel testo, cioè una possibile configurazione di opinioni probabilistiche sui tipi. Ognuna di tali configurazioni deve anche essere una configurazione di tipi, data la definizione di tipo. Se dunque siamo alla ricerca di uno spazio dei tipi chiuso rispetto all'operazione di costruzione delle distribuzioni di probabilità su se stesso, stiamo cercando di fatto un punto fisso di  $\Delta(\cdot)$ , un operatore che genera spazi via via più ampi.

<sup>30</sup> Si rinvia a Rampa [1991] e Lughini e Rampa [1996].

<sup>31</sup> Per una introduzione e una panoramica degli sviluppi si veda anche Gilboa e Schmeidler [1995b].

decisori di compiere ragionamenti ipotetici che coinvolgono qualsiasi stato possibile o immaginabile. Più semplicemente, si postula che il decisore valuti i diversi corsi di azione fattibili in un certo caso problematico sulla base della sola sua esperienza passata di casi simili.

La memoria del passato ( $M$ ) è costituita da un vettore di lunghezza finita i cui elementi sono terne del tipo  $m = (p, a, r)$ , che indicano rispettivamente il problema affrontato, l'azione prescelta e il risultato ottenuto in ogni singolo caso passato. Nella versione più semplice si suppone inoltre che il decisore abbia una funzione di utilità definita sui risultati effettivamente ottenuti.

Supponiamo che, di fronte ad un nuovo problema  $p'$  e in presenza di una memoria  $M$ , il decisore sappia esprimere un ordinamento  $\succ_{p', M}$  di preferenza sugli atti disponibili. Se tale ordinamento soddisfa alcuni requisiti minimali, si può dimostrare che esso è equivalente ad ordinare gli atti secondo un funzionale come il seguente:

$$U_{p', M}(a) = \sum_{(p, a, r) \in M} s([\ast])p', p \cdot u(r)$$

dove la funzione  $s(\cdot)$ , che è sostanzialmente unica, è interpretabile come una misura scalare di similarità fra problemi. Si noti che nella precedente sommatoria  $a$  è costante. In altri termini, la valutazione di ogni atto  $a$ , in relazione al nuovo problema  $p'$  che si pone, dipende dai risultati che l'uso di quell'atto ha consentito di ottenere in passato in presenza di problemi più o meno simili a  $p'$ .

Come osservano Gilboa e Schmeidler [1993], questa rappresentazione della scelta incorpora una teoria della conoscenza molto diversa da quella presupposta dalla teoria bayesiana della decisione individuale: in quest'ultima (vista nel caso dinamico) si deve ipotizzare che il decisore sappia scegliere «alla nascita» un singolo piano onnicomprensivo che tenga conto di tutte le possibili storie future. Viceversa, nella decisione basata sui casi le «coordinate» in cui inquadrare un nuovo problema vengono definite solo dall'esperienza passata, e la metrica tramite cui la similarità fra problemi viene valutata è indotta dal successo che l'uso di determinati atti ha consentito di ottenere in ciascuno dei problemi effettivamente affrontati. Ciò implica che, nel funzionale scritto più sopra, la funzione di similarità in realtà non va pensata costante nel tempo<sup>32</sup>, almeno nel caso in cui si supponga che ogni nuova esperienza sia in grado di «insegnare» qualcosa di nuovo: il caso di funzione di similarità invariante corrisponde ad una situazione

<sup>32</sup> Ciò dovrebbe risultare chiaro dal fatto che l'ordinamento  $\succ_{p', M}$  è parametrizzato allo specifico problema  $p'$  che si sta affrontando di volta in volta ed alla memoria  $M$  accumulata in precedenza.

in cui il decisore sa (o ritiene di sapere) già inquadrare ogni nuovo possibile problema in categorie note.

Come risulta evidente, il contributo della teoria della decisione basata sui casi contiene molti suggerimenti che si situano direttamente sul piano della teoria della conoscenza sottostante. In particolare, <sup>(non)</sup> si suppone che inizialmente il decisore disponga di una batteria di teorie alternative da mettere alla prova per definire il corso di azione per lui migliore; una «teoria» emerge di fatto come solo risultato dei problemi affrontati e delle azioni effettivamente intraprese in ognuno di essi. In effetti, il soggetto non conosce la funzione che lega il problema  $p$  e l'azione prescelta  $a$  al risultato  $r$  ottenuto, posto che tale funzione esista. Semplicemente, tramite l'esperienza egli classifica quegli elementi tramite una funzione di similarità che è *path-dependent*.

Ora, secondo chi scrive questi suggerimenti possono essere esemplificati e approfonditi facendo riferimento a recenti sviluppi che si registrano nelle scienze cognitive, in particolare nel campo delle reti neurali<sup>33</sup>. In questa sede non facciamo riferimento all'utilizzo delle reti neurali come strumenti di controllo, tipico della letteratura ingegneristica, o come strumenti di previsione, uso quest'ultimo che si sta diffondendo in econometria<sup>34</sup>. Piuttosto, le reti neurali possono essere viste come modelli per la rappresentazione della conoscenza, dell'apprendimento e della decisione da parte di soggetti (che la teoria bayesiana della decisione definirebbe di certo «limitatamente razionali»): in altri termini, si suggerisce un'interpretazione secondo cui i decisori non usano, ma sono reti neurali<sup>35</sup>.

Una rete neurale è un vettore di unità elementari («neuroni») tra loro connesse, in generale in modo non lineare. In particolare, ogni singola unità assume in un dato periodo un valore che dipende da una combinazione lineare dei valori delle altre unità nel periodo precedente. Il valore assunto da ogni unità è usualmente binario, e risulta «acceso» o «spento» a seconda che quella combinazione lineare superi o non superi una determinata soglia: qui, cioè in questa funzione «a gradino», sta la non linearità. Spesso la dipendenza di ogni

<sup>33</sup> Per una introduzione a questi argomenti, si rinvia a Serra e Zanarini [1990], oppure a Hertz, Krogh e Palmer [1991].

<sup>34</sup> Si veda Gallant e White [1993].

<sup>35</sup> Come è noto, il primo esempio, per così dire *ante litteram*, di un tale tipo di teoria cognitiva è il lavoro di Hayek [1952]. In effetti, gli stessi autori oggi accreditati di essere gli iniziatori dello studio delle reti neurali avevano presente questo contributo; per una interessante introduzione a questo tipo di lettura di Hayek si veda Birner [1996]; si veda anche Rampa [1997a].

unità dalle altre è formulata per mezzo di una funzione non lineare differenziabile anziché discontinua. Se  $x_t \in C^n \equiv [-1, 1]^n$  rappresenta il vettore composto dalle  $n$  unità al periodo  $t$ , e  $x_{i,t}$  è una delle unità, si può per esempio scrivere:

$$(1) \quad x_{i,t} = \frac{1 - e^{g_i(x_{t-1})}}{1 + e^{g_i(x_{t-1})}}, \quad \text{con } g_i(x_{t-1}) = -\gamma_i \cdot \left[ \sum \alpha_{ij} x_{j,t-1} - \beta_i \right], \quad \text{sigmoidale}$$

dove  $\gamma_i, \beta_i, \alpha_{ij}, i, j = 1, \dots, n$  sono parametri, dei quali  $\beta_i$  rappresenta la soglia  $\alpha_{ij}$  sono i pesi delle connessioni tra unità, e  $\gamma_i$  misura la reattività dell'unità  $i$  alla circostanza che lo stimolo ricevuto dalle altre unità,  $\sum \alpha_{ij} x_{j,t-1}$ , superi o meno la soglia  $\beta_i$ . Si può infatti verificare agevolmente che il grafico della funzione (1) è una sigmoide compresa fra  $-1$  e  $1$  (dove  $-1$  corrisponde allo stato «spento» di  $x_{i,t}$ , e  $1$  a quello «acceso»); quanto più  $\gamma_i$  è grande, tanto più la sigmoide approssima la funzione discontinua «a gradino». A tutti gli effetti, una rete neurale come qui rappresentata è un sistema dinamico: infatti tutte le unità della rete assumono in un certo periodo valori che dipendono dal valore assunto da tutte le unità nel periodo precedente. Ne segue che tale sistema dinamico possiede un certo numero di attrattori, a cui lo stato della rete (il vettore delle variabili  $x$ ) può convergere asintoticamente: a quale degli attrattori il vettore  $x$  converga dipende dalla condizione iniziale; la condizione iniziale, nel caso delle reti neurali, è costituita da qualche stimolo esterno, nel senso seguente.

Possono esistere alcune unità particolari che vengono denominate di *input*: in presenza di uno stimolo proveniente dal mondo esterno, esse codificano tale stimolo assumendo particolari valori iniziali. Esistono altre unità considerate di *output*: dopo che le connessioni fra tutte le altre unità hanno cessato di operare le trasformazioni descritte sopra e la dinamica di rete ha raggiunto un attrattore, il valore finale delle unità di *output* viene trasferito all'esterno per determinare, via qualche funzione preassegnata, un'azione. Il valore degli *input* in seguito allo stimolo esterno costituisce dunque la condizione iniziale, mentre il valore finale degli *output* (con la conseguente azione) è il corrispondente attrattore del sistema dinamico.<sup>36</sup> Utilizzando il linguaggio della teoria di Gilboa a Schmeidler, possiamo definire *problema* lo stimolo esterno (o la sua rappresentazione in termini di condizione iniziale), mentre possiamo chiamare *atto* l'azione intrapresa in

<sup>36</sup> In molte circostanze si fa uso delle cosiddette reti *feed-forward*. In queste ultime l'insieme delle unità è partizionato in  $K$  «strati», escluso quello di input e incluso

seguito all'asestarsi della rete nell'attrattore corrispondente a quello stimolo: evidentemente, l'insieme dei problemi che danno luogo ad un medesimo atto costituisce il bacino d'attrazione di quell'attrattore. Dobbiamo ora affrontare la questione della numerosità degli attrattori della rete.

Il tipo specifico di non linearità incorporato nella funzione (1) di «accensione» delle unità fa sì che lo spazio dei possibili input dell'unità  $x_i$ , sia partizionato in due semispazi; in corrispondenza di tutti gli input localizzati nel medesimo semispazio il valore dell'unità  $x_i$ , è il medesimo (e dunque il valore trasmesso a valle alle altre unità è il medesimo): tale valore differisce invece in seguito a input che si situano in semispazi diversi. In particolare, la rapidità di transizione del valore della cella  $x_i$ , da «acceso» a «spento» dipende dal valore del parametro  $\gamma$ .<sup>37</sup> Di conseguenza, elevati valori di tale parametro accrescono la capacità dell'unità in questione di discriminare fra input tra loro differenti; peraltro, il genere di differenza tra due input diversi della unità è molto semplice: appartenenza o meno ad uno dei due semispazi. Consideriamo tuttavia molte unità congiuntamente: se ciascuna di esse opera una buona discriminazione, il vettore di valori passati a valle alle altre unità nel periodo successivo rimane ben differenziato in corrispondenza di stimoli tra loro differenziati. Si può infatti dimostrare che il numero degli attrattori della rete cresce al crescere del numero delle sue unità ad elevata capacità di discriminazione: in un senso molto preciso, l'insieme dei possibili stimoli esterni

quello di output, che sono ordinati gerarchicamente: le unità di ciascuno strato ricevono stimoli solo dallo strato precedente e ne inviano solo allo strato successivo (in generale si considera un solo strato «intermedio», e dunque  $K=2$ ). In tal caso lo jacobiano della trasformazione dei valori delle unità da un periodo all'altro è una matrice triangolare a blocchi per la parte relativa agli strati di input e intermedi, mentre coincide con la matrice identica per le unità di output; gli «attrattori» di questa dinamica sono i valori dello strato di output letti  $K$  o più periodi dopo che un input è arrivato. Il caso da noi trattato nel testo è quello delle reti cosiddette ricorrenti: in tal caso lo jacobiano è una matrice in generale non decomponibile, ed il concetto di attrattore è quello usuale dei sistemi dinamici: i ruoli di input e output non sono più riferiti a particolari unità, ma alla condizione iniziale e all'attrattore, rispettivamente. Chiaramente, da questo punto di vista le reti *feed-forward* sono un caso particolare di quelle ricorrenti. Per una più approfondita analisi di questi aspetti si rinvia a Rampa [1997b].

<sup>37</sup> Infatti la funzione di accensione (1) individua due semispazi separati da un iperpiano: a seconda che il valore racchiuso tra parentesi quadre sia positivo o negativo, per valori di  $\gamma$  «grandi» l'unità assume valori radicalmente diversi, e invia dunque stimoli radicalmente diversi alle altre unità. Se  $\gamma$  è «piccolo», esiste una zona intermedia di indecisione (in analogia con la logica *fuzzy*); zona che è tanto più ristretta quanto più forte è l'approssimazione alla funzione «a gradino» (valore di  $\gamma$  elevato).

viene partizionato in modo più fine, cioè in un numero più elevato di celle (bacini di attrazione) elementari<sup>38</sup>.

Il funzionamento della rete, come ora descritto, può per un verso somigliare a quanto previsto dalla teoria tradizionale della scelta. Infatti la rete induce una partizione dell'insieme dei problemi possibili. Tuttavia ciò non vuol dire che ciascuno stimolo esterno elementare abbia «significato» quando considerato singolarmente, come accade invece nella teoria tradizionale dell'informazione (dove ogni possibile stato del mondo è individuabile per il decisore). Sono solo le classi di problemi (le celle) ad avere significato, in quanto «riconoscibili» e cioè associate ad un medesimo attrattore (e conseguente atto): da questo punto di vista la somiglianza con la teoria di Gilboa e Schmeidler è maggiore.

Se i parametri delle funzioni di transizione (1) sono costanti nel tempo, il comportamento della rete è ben definito: ogni problema viene sempre classificato in una ben determinata cella, e ne segue un ben determinato atto. In particolare, se il peso relativo alla non-linearità di una certa unità ( $\gamma$ ) è prossimo a zero, la capacità discriminante di quella unità è trascurabile (è come se non fosse operante): dunque la partizione operata dalla rete neurale è meno di fine di quanto potrebbe altrimenti essere (ma non è detto che questo sia dannoso in termini di guadagno, se si considerano anche i costi di calcolo).

Diverso, e più interessante, è il caso in cui i parametri delle funzioni di transizione delle unità siano variabili nel tempo. Supponiamo che la rete venga sottoposta ad una sequenza di stimoli esterni; ad ogni tornata avviene una trasformazione dallo stimolo esterno (problema) ad un attrattore della rete (cioè un atto). Ipotizziamo inoltre

<sup>38</sup> Nel caso di reti *feed-forward*, il numero  $n$  di unità presenti nello strato intermedio definisce ovviamente la numerosità degli iperpiani separatori, posto che essi siano linearmente indipendenti: l'operare congiunto di varie unità (cioè iperpiani separatori) dà luogo ad una partizione in celle dello spazio degli stimoli esterni. Dunque, il numero di unità dello strato intermedio determina il numero di celle diverse in cui lo spazio dei problemi è partizionato: naturalmente il numero di queste cresce con  $n$  (ci si avvede facilmente che il numero di celle è  $2^n$ ). Ne segue che al crescere del numero di unità dello strato intermedio la partizione diviene più fine, e il valore delle unità di output può essere modulato più finemente al variare del valore degli input. Questa è intuitivamente l'essenza del teorema di «approssimazione universale»: qualsiasi funzione può essere approssimata a piacere da una rete neurale con un numero opportunamente elevato di unità intermedie. Discorso analogo, sebbene lievemente più complesso, vale per le reti ricorrenti, dove la distinzione fra input, intermedio e output viene sostituita da quella fra condizione iniziale, fase transiente e attrattore: in questo caso è il numero degli attrattori a variare al variare delle unità ad elevata capacità di discriminazione. Per una trattazione più precisa si veda Rampa [1997b].

che ad ogni coppia problema-atto sia associato un risultato, e dunque un *payoff*, analogamente a quanto accade nella rappresentazione di Gilboa e Schmeidler. Questo aspetto è standard nella teoria delle reti neurali, e costituisce la cosiddetta fase di addestramento. La funzione di *payoff* può essere nota o meno alla rete (decisore); essa è da interpretarsi come un meccanismo di «punizione» o «rinforzo» delle azioni attuate nelle diverse circostanze. Si può trattare delle reazioni di un «insegnante», oppure di un processo di «selezione naturale»<sup>39</sup>. Ciò che conta, tuttavia, è che il decisore non conosce completamente la funzione che associa la coppia problema-atto al risultato (e dunque al *payoff*) che gliene deriva. Il decisore viene informato soltanto circa il valore assunto da quella funzione in pochi punti: quelli corrispondenti ai problemi già sperimentati in passato. Se il decisore congetture una qualche approssimazione locale della funzione composta da problemi-atti a *payoff*, è possibile calcolarne il gradiente rispetto ai pesi e modificare questi ultimi nelle direzioni desiderate, cioè quelle che inducono un aumento del *payoff* che si potrebbe ottenere se si ribresentassero i problemi già sperimentati. Si tratta dunque di fare quanto segue: ad ogni tornata di addestramento, i pesi sono modificati in modo tale che le risposte (attrattori, atti) che verrebbero associati (dopo la modifica) ai problemi sperimentati in passato diano un *payoff* congetturato maggiore<sup>40</sup>.

Tutto ciò induce un secondo sistema dinamico, che opera su una scala temporale più ampia di quello che governa la trasformazione da problemi a atti da parte della rete, (dati i pesi). Tale nuovo sistema dinamico è definito sul vettore dei pesi ed è forzato dalla sequenza esogena degli *input* e dei risultati: si tratta di una *dinamica di apprendimento*.

Non è detto che la dinamica di apprendimento sia convergente, né si può dire *a priori* quali siano i suoi attrattori. Infatti una rete neurale non può essere pensata come un catalogo di ipotesi formulate *a priori*, le quali vengano man mano confutate sino ad isolare l'ipotesi

<sup>39</sup> Nel caso di molte attività basilari degli organismi biologici superiori, pare accertato che tale ruolo venga svolto da un insieme di ghiandole della zona ipotalamica connesse ai sensi del piacere: si tratta di un meccanismo interno, ma usualmente al di fuori della consapevolezza soggettiva. In effetti, tali ghiandole si situano in una zona cerebrale «antica», evolutivamente antecedente la corteccia cerebrale. Su questo si veda Edelman [1993, cap. 10].

<sup>40</sup> Nella pratica econometrica, il meccanismo è ancora più semplice: la rete viene *ex post* informata esattamente circa la risposta corretta (in termini previsivi) ai singoli problemi (valore della variabili predeterminate), e la funzione di *payoff* è l'opposto della somma degli errori al quadrato.

«vera» o comunque meno in contraddizione con i dati sperimentati<sup>41</sup>. Si può dimostrare che l'apprendimento neurale non avviene per mezzo di aggiustamenti su uno spazio parametrico «liscio», bensì consiste principalmente in salti discontinui: ciò deriva essenzialmente dalla elevata non linearità delle funzioni di accensione delle unità, che provoca l'improvvisa comparsa di nuove celle nella partizione operata<sup>42</sup>. Di conseguenza, sarà la tipologia della storia effettivamente sperimentata, assieme alla complessità della rete (cioè la numerosità delle sue unità), a determinare il grado di finezza della classificazione operata dalla rete stessa sugli stimoli esterni. Se la storia passata (problemi e performances richieste alla rete) è sufficientemente stazionaria e la rete è sufficientemente complessa, la dinamica di apprendimento tenderà a convergere a qualche configurazione stazionaria dei pesi. Se invece la complessità della storia sperimentata è superiore a quella della rete, si assisterà a frequenti ristrutturazioni globali della configurazione dei pesi alternate a periodi di relativa stabilità. Può cioè accadere che una configurazione di pesi rimanga stabile per periodi di tempo lunghi ma finiti, per poi subire improvvise alterazioni dovute all'accumularsi di esempi che non erano stati sperimentati durante la precedente fase<sup>43</sup>. Le configurazioni che rimangono stabili per un certo periodo sono dunque fortemente *path-dependent*.

Analogamente, la configurazione dei pesi cui la rete può convergere dipende anche dalla configurazione iniziale: reti che inizialmente hanno capacità di discriminazione diversa e/o struttura della partizione in celle diversa cominceranno a interpretare gli stimoli in modo diverso. Data la forte dipendenza dalla storia, le configurazioni cui

<sup>41</sup> Per esempio, un modello econometrico usuale (lineare o meno) è un catalogo di ipotesi *a priori*, che sono simili tra loro quanto a forma funzionale e differiscono quanto a valori dei parametri. Viceversa una rete neurale, come si è visto più sopra, non implementa funzioni alternative definite su uno spazio euclideo, bensì partizioni alternative di quello spazio a fini di classificazione degli stimoli.

<sup>42</sup> Nella rappresentazione delle reti da noi adottata in termini di sistemi dinamici, il mutamento dei parametri della rete dà luogo tipicamente a *biforcazioni*, vale a dire cambiamenti nelle caratteristiche degli attrattori: il loro numero, la stabilità e la dimensionalità dei bacini di attrazione.

<sup>43</sup> Che tali alterazioni siano improvvise deriva da quanto segue. Come argomentato nella precedente nota 38, il numero di neuroni che compongono una rete governa la finezza della sua capacità di partizionare lo spazio dei problemi; di fronte ad una certa storia sperimentata, la rete può dunque giungere a saturare le proprie capacità, e dunque non essere in grado di adattarsi a prosecuzioni di quella storia che presentino novità. Come ben sa chi ha applicato le reti neurali, l'esito di questa situazione è un'instabilità che provoca una revisione non marginale di tutti i pesi della rete: si veda anche Serra e Zanarini [1990].

reti diverse convergono in seguito ad una medesima esperienza possono essere assai diverse tra loro.

Supponiamo che la configurazione dei pesi di una rete converga (anche solo temporaneamente) a qualche stato. Conveniamo di chiamare *teoria* un attrattore della dinamica di apprendimento: si tratta di un modo di classificare e interpretare le esperienze dell'ambiente esterno. Come si è detto, in generale esiste una molteplicità di teorie ammissibili per un certo ambiente esterno: questo a causa sia della potenziale molteplicità delle sequenze di esempi effettivamente sperimentabili, sia della diversità delle teorie possedute inizialmente da una rete.

Un ulteriore elemento che contribuisce ad ampliare il numero delle teorie che possono interpretare con successo una data storia è il fatto che le reti neurali di tipo naturale (i sistemi nervosi di organismi viventi, specie quelli dotati di una corteccia cerebrale) sono costituite da un numero molto elevato di unità elementari. Ciò implica che una rete siffatta può costituire partizioni con numeri di celle che sono di molti ordini superiori rispetto alla complessità delle storie effettivamente sperimentate in relazione ai vari problemi e ambienti con cui la rete interagisce<sup>44</sup>.

A questo punto, pare emergere una notevole affinità fra questa impostazione e la teoria della decisione basata sui casi: tale affinità non è dovuta alla sola presenza di analogie fra i due approcci, ma anche al fatto che dal punto di vista logico essi hanno molti elementi in comune. I principali sono: (i) in entrambe le impostazioni gli atti sono valutati sulla base dei soli casi effettivamente sperimentati, anziché di tutti quelli concepibili; (ii) la valutazione degli atti avviene tramite i *payoffs* ottenuti e tramite qualche nozione di similarità fra problemi affrontati; (iii) più in particolare, nel caso delle reti neurali la similarità fra problemi è, per definizione, indotta dalle celle della partizione operata dai pesi (problemi appartenenti alla stessa cella hanno

<sup>44</sup> In un certo senso, una rete possiede moltissimi «gradi di libertà» rispetto alla sequenza di stimoli esterni ricevuti. Si noti che stiamo usando questo concetto in modo opposto al suo utilizzo nel linguaggio econometrico, dove la presenza di gradi di libertà ha a che fare con l'uso «parsimonioso» dei parametri della teoria rispetto alla lunghezza del campione osservabile. Rispetto alla prescrizione, molto diffusa, di puntare a teorie parsimoniose, possiamo replicare che una rete neurale naturale non controlla mai la porzione di se stessa che viene dedicata ad ogni problema particolare, e inoltre che nessuna rete neurale naturale sperimenta mai storie di lunghezza infinita, cosicché il riferimento a «teoremi del limite centrale» o di convergenza della distribuzione posteriore non è pertinente. A ciò si deve aggiungere quanto diremo nel paragrafo 4 circa la non stazionarietà degli ambienti di interazione fra reti neurali.

distanza nulla, e dunque similarità massima); (iv) la definizione della partizione, cioè dei pesi, di una rete deriva direttamente dall'obiettivo di associare ai vari problemi effettivamente sperimentati gli atti che rendono massimo il payoff congetturato, e a problemi simili atti simili; (v) la similarità fra i diversi problemi non è definita a priori, ma dipende dalla storia sperimentata. Naturalmente, tutte queste osservazioni devono essere suffragate da un'analisi più precisa, e sono per il momento esse stesse congetture: tuttavia costituiscono ipotesi molto ragionevoli per una futura ricerca.

Si può concludere dicendo che una rete neurale conosce per casi: instaura un ordine individuale («sensoriale», secondo la terminologia di Hayek<sup>43</sup>), cioè una classificazione che consente di dare un significato, per vicinanza a quanto già sperimentato, alle nuove esperienze. Si tratta di un ordine personale e dipendente dalla storia, in quanto: (a) parità di sequenza dei problemi sperimentati, congetture iniziali (strutture dei pesi) diverse conducono a teorie diverse; (b) parità di congettura iniziale sequenze di problemi sperimentati diverse conducono a teorie diverse. Utilizzando il linguaggio delle teorie dinamiche o evolutive, ogni teoria corrisponde ad un «minimo locale» (uno tra moltissimi) di un paesaggio lungo il quale l'apprendimento neurale si muove.

Ma quanto siamo venuti dicendo ha anche implicazioni per la teoria dell'informazione. Mentre sussiste un'analogia formale fra le partizioni definite dalla teoria bayesiana, da una parte, e quelle operate da una rete neurale o dalla funzione di similarità di Gilboa e Schmeidler, dall'altra, esistono profonde differenze tra i due approcci. Innanzitutto la conoscenza-decisione basata sui casi non prevede che il decisore sappia definire le conseguenze dei propri atti in corrispondenza di tutti gli stati elementari; anzi quest'ultimo concetto non è neppure definibile da parte del decisore, che percepisce classi di problemi, e comunque solo quelle formatesi in seguito alla sua esperienza<sup>44</sup>. In secondo luogo, le «partizioni» (o classi di similarità) sono endogene, poiché dipendono dalla storia sperimentata; dunque non fanno parte della descrizione iniziale delle caratteristiche del decisore, e di conseguenza non possono essere conoscenza comune per even-

<sup>43</sup> Hayek [1952]; per un tentativo di connessione fra le idee psicologiche di Hayek e quanto siamo qui venuti dicendo, si rinvia a Rampa [1997a].

<sup>44</sup> Si potrebbe affermare che si tratta qui di un caso particolare della teoria tradizionale, un caso in cui le conseguenze di un atto sono costanti per tutti gli stati appartenenti alla stessa classe (elemento della partizione). Ma non si capisce allora la rilevanza del definire gli stati elementari.

tuali altri decisori. In terzo luogo, prendiamo diversi decisori, e poniamo che ciascuno abbia appreso una sua «teoria» (nel senso definito sopra): non v'è alcuna presunzione che la presenza di «asimmetrie informative» (dovute al fatto che le partizioni individuali sono diverse tra loro) possano «svelare», via le opportune intersezioni e induzioni all'indietro<sup>47</sup>, lo stato «vero». Infatti quelle asimmetrie non sono conoscenza comune tra i decisori, e le implicazioni conoscitive dei «segnali» esterni per i diversi decisori non sono in generale comunicabili in modo completo<sup>48</sup>.

Si potrebbe tuttavia obiettare che, trascorso un tempo sufficientemente lungo, la storia (anziché il ragionamento) sveli lo stato «vero» del mondo. Ciascun decisore individuale può così pervenire all'attrattore globale della dinamica di apprendimento (la Teoria), avendo confutato tutti quelli erronei. Discutiamo di questa obiezione nel paragrafo successivo.

#### 4. Ordine sociale

Torniamo ora al comportamento di un decisore non illimitatamente razionale. Un tale decisore non è in grado in generale di concepire tutte le possibili evenienze, e quindi di formulare un piano onnicomprensivo (specie relativamente alle storie future). Ciò non toglie che il suo comportamento non possa essere interpretato come un comportamento intenzionale. Se qualche criterio guida il suo comportamento, come abbiamo supposto, egli arriverà a deliberare un corso di azioni che lo convince in relazione alla sua visione del mondo. Inoltre se il decisore è appena previdente, il suo piano includerà almeno qualche variante da mettere in atto se si presenterà la tale o tal altra circostanza.

Per quanto in forma debole (rispetto alla teoria bayesiana della scelta), abbiamo di fatto descritto una strategia una funzione che prescrive un'azione per ogni possibile storia tra quelle che il decisore riesce a immaginare. Questo potrebbe apparire sufficiente per una teoria che badi alla sola razionalità individuale: semplicemente, il mondo va avanti, con ogni decisore che mette in atto il proprio piano, nel senso che a ogni stimolo risponde obbedendo alla strategia

<sup>47</sup> Nel senso di Aumann [1976].

<sup>48</sup> Su questa impossibilità di conoscere «nel dettaglio» il funzionamento di una rete neurale diversa dalla propria (o anche solo la propria) si veda Hayek [1952, 263 ss.], e gli argomenti svolti in Rampa [1997a].

prescelta. Se il procedere della storia non gli crea sorprese cognitive, la sua teoria gli è sufficiente per orientarsi nel proprio ambiente, e possiamo dire che quel decisore si trova in una situazione di equilibrio soggettivo.

Ora, la differenza fondamentale rispetto alla teoria dei giochi bayesiana è che in un contesto di conoscenza-decisione basata sui casi i soggetti (non) possono convincersi di aver inizialmente prescelto ciascuno la coordinata di un vettore di strategie di equilibrio (complessivo, non solo individuale) di Bayes e Nash. Questo perché, per calcolare un equilibrio di Bayes e Nash, bisogna quantomeno fermarsi davanti al proprio computer per un tempo infinito solo per digitare la descrizione di un «tipo» harsanyiano. Se invece i decisori sono usciti di casa e hanno cominciato ad attuare delle decisioni, significa che hanno seguito qualche altra via<sup>49</sup>: la loro posizione cognitiva deve essere lontana da quella dei giocatori bayesiani. Sicuramente essi agiscono senza la presunzione di aver progettato ex ante un equilibrio sociale.

Se si accetta questo suggerimento, e se si continua a supporre che ogni decisore sia alla ricerca di un ordine individuale (una teoria che gli consenta di dar conto di ciò che vede attorno a sé), allora ne segue che qualunque nozione di equilibrio sociale non può che essere impostata (e accettata da ogni decisore) dal punto di vista ex post. Tale nozione dovrà essere del tipo seguente: una sequenza di decisioni congiunte di vari agenti forma un equilibrio se le azioni intraprese sulla base delle diverse teorie individuali non mettono in discussione tali teorie, cioè se ciascuno vede ciò che si aspettava di vedere. Questa è evidentemente la nozione di equilibrio, o meglio di ordine sociale, suggerita da Hayek [1945, 50-53]: essa è però del tutto equivalente alla nozione di equilibrio di Keynes<sup>50</sup>, o di equilibrio congetturale di Hahn<sup>51</sup>: uno stato di affari dove le azioni di tutti non disconfermano le congetture (teorie, aspettative) di alcuno. Se la condizione di equilibrio non è soddisfatta, almeno un decisore sperimenta un disordine cognitivo individuale e reagisce modificando la propria teoria e di

<sup>49</sup> Si potrebbe pensare che abbiano deliberato un equilibrio correlato puramente «soggettivo» [Aumann 1987]: ma qui si tratta di pura questione di nominalismo; infatti questo tipo di «equilibrio» non piace ai teorici bayesiani, proprio in quanto in esso non sussiste alcuna condivisione dell'interpretazione della realtà.

<sup>50</sup> Questa opinione è motivata in Rampa [1990; 1991]. Per un esempio specifico si rinvia anche a Rampa [1989].

<sup>51</sup> Si veda Hahn [1973]; la contrapposizione fra questa nozione e quella bayesiana evoca quella di Binmore [1987-1988] tra fondazione «evolutiva» ed «eduttiva» dell'equilibrio.

conseguenza la propria strategia. Tale modificazione di un comportamento individuale può dare luogo alla rimessa in discussione delle proprie teorie da parte degli altri individui. Si mette dunque in moto un processo di apprendimento congiunto da parte di tutti gli individui: un equilibrio congetturale è da interpretarsi quindi come punto stazionario di questo processo. Ora, si dovrebbe adottare il punto di vista dinamico e studiare se tale processo sociale converge a qualche attrattore (per definizione, se converge non può che convergere ad una posizione di ordine sociale, ovvero di equilibrio congetturale). La dinamica può essere indotta dall'applicazione ricorsiva della regola di Bayes, dei minimi quadrati ordinari, oppure del metodo del gradiente applicato ai pesi di una rete neurale, a seconda del contesto. Si intuisce che uno studio di questo tipo non possa che essere estremamente complesso, tranne casi molto speciali e artificiali.

Supponiamo che questo processo converga a qualche attrattore. Si potrebbe sostenere che se ciò accade, questo non possa che essere uno stato in cui tutta l'informazione è «svelata»: quindi, se si ottiene questo risultato, si potrebbe ritenere di aver fondato per via di apprendimento l'equilibrio di Bayes e Nash. Alcuni autori<sup>32</sup> mostrano di prediligere questa impostazione. Essi assumono che l'oggetto di distribuzioni individuali sia l'insieme delle storie future di lunghezza infinita e che l'apprendimento sia «razionale» (cioè, consista nell'applicazione sistematica della regola di Bayes ad ogni data, così come previsto dalla strategia ottima decisa inizialmente). Questi autori sono in grado di dimostrare che le previsioni individuali, da un certo periodo in poi, coincidono praticamente con le effettive realizzazioni; dunque, poiché le strategie scelte da ciascuno sono comunque razionali nel senso bayesiano, si converge ad un equilibrio di Nash. Per ottenere il risultato occorre tuttavia un'ipotesi «tecnica»: che cioè ogni evento che abbia probabilità positiva nella distribuzione «oggettiva» la abbia anche in ogni distribuzione soggettiva. Ciò significa che lo spazio di incertezza soggettiva è comune a tutti gli agenti e coincide con lo spazio di incertezza «oggettivo»: dato il punto di vista che stiamo proponendo in questa sede, tale ipotesi appare cruciale e non credibile.

Si consideri invece il seguente argomento. In generale esistono molte posizioni di ordine sociale o equilibrio congetturale; anzi si può dimostrare che questa molteplicità è un risultato generico. Ciò accade sia quando l'apprendimento è modellato in termini di rete neurale, sia nel caso di modelli individuali di tipo statistico più classico, a pat-

<sup>32</sup> Per esempio, Kalai e Lehrer [1993].

di considerare in questo secondo caso agenti sufficientemente «sofisticati». In entrambi i casi la ragione della molteplicità consiste sostanzialmente in questo: si osservi che i segnali o stimoli esterni che arrivano ai decisori sono principalmente il risultato delle azioni degli altri decisori; nel contempo le decisioni di ognuno possono essere in generale interpretate come trasformazioni dagli stimoli esterni alle azioni intraprese. Ora, se un decisore vuole «comprendere» il comportamento degli altri per prevederlo, deve tentare di costruire una rappresentazione di quelle trasformazioni, e non solo degli stimoli esterni. Evidentemente, lo spazio delle trasformazioni (funzioni) da stimoli ad azioni è più complesso (dimensionalmente più ampio) dello spazio dei soli stimoli. Dunque la rappresentazione di come un altro decisore reagisce agli stimoli è più complessa di quella dei suoi stimoli. Ne segue che le teorie individuali dei decisori (quando costoro si rendono conto dell'intenzionalità altrui e ne vogliono comprendere le ragioni: parliamo di agenti «sofisticati») sono più elaborate dei segnali emessi dall'economia, specie per periodi di tempo circoscritti. Questa ricchezza dà luogo tipicamente ad una molteplicità di teorie in grado di interpretare una medesima storia<sup>33</sup>

Ora, se il gioco dei singoli agenti è cercare di avere ciascuno un modello che riesca a «contenere» quelli altrui, il risultato è un'espansione indefinita (sino al punto di divenire impraticabile): il grado di molteplicità degli equilibri aumenta senza limite. Viceversa, se viene allungata l'estensione temporale su cui mettere alla prova un insieme di teorie individuali limitate, l'insieme degli equilibri si assottiglia. Infatti il prolungamento di una storia finita tende genericamente a produrre una maggiore complessità degli stimoli che arrivano ai decisori: qualche teoria individuale potrebbe essere non più in grado di interpretare coerentemente l'ambiente<sup>34</sup>.

Esistono dunque molteplici equilibri congetturali, cioè configurazioni delle teorie individuali in grado di soddisfare i giocatori per un numero finito (che può essere molto grande) di periodi; per «soddisfazione» dei giocatori si intende la conferma delle aspettative sogget-

<sup>33</sup> Nel caso dell'apprendimento di tipo neurale, questo argomento è già presente in Hayek [1952], ed un tentativo di analisi più rigorosa è svolto in Rampa [1997a]. Nel caso di teorie di stampo statistico più tradizionale, si rinvia all'analisi svolta in Rampa [1990].

<sup>34</sup> Né giova appellarsi al fatto che prima o poi la storia manifesterà un tipo suo proprio di «stazionarietà»: per sottoporre a test tale ipotesi nel caso di sistemi sociali non artificiali occorrerebbe un'ampiezza campionaria ben maggiore di quella sperimentabile da uno qualsiasi dei decisori.

tive. Se il sistema incappa in una di queste posizioni di equilibrio, evidentemente non c'è ragione di uscirne: semplicemente, ciò che i decisori vedono attorno a sé non li induce a cambiare le proprie teorie ed il proprio comportamento. (Ma ciò non significa che essi abbiano «scoperto» la verità ultima: è sufficiente che le aspettative siano corrette lungo il sentiero che l'economia sta percorrendo (non si manifestano aspetti che, se scoperti, indurrebbero una modificazione delle teorie e dei piani di azione)».

Cosa accade lungo un equilibrio di questo genere? Accade che l'osservazione di una storia finita non fornisce sufficienti gradi di libertà per discriminare fra differenti modelli sufficientemente elaborati. Detto in altri termini, gli agenti avrebbero sì i numeri per interpretare oggetti anche più complessi rispetto alla storia sinora realizzatasi, ma non è dato loro di osservare alcunché di più elaborato di quanto l'economia sta mettendo a loro disposizione. Possiamo interpretare le azioni reciprocamente osservate lungo un sentiero di ordine sociale come segnali, messaggi, o parole che circolano fra le persone: si noti che non circolano né si osservano le teorie individuali, bensì solo le loro implicazioni in termini di azioni, le quali appartengono in generale ad uno spazio «meno ricco» di quello delle teorie individuali. Allora questo linguaggio, che si produce in una situazione di ordine sociale, è di fatto più semplice della capacità di elaborazione delle persone stesse, e quindi può essere interpretato in vari modi alternativi. Esso è peraltro stabile, nel senso che l'effetto sul resto del sistema dell'emissione di quelle parole risponde alle aspettative di ognuno.

Aggiungendo a ciò il tipico fenomeno della potenziale molteplicità degli ordini sociali (tra i quali quello in cui stiamo vivendo è stato selezionato dalla storia), si deve concludere che il concetto di «informazione» assume una valenza ben diversa da ciò che si è visto nel paragrafo 2. Innanzitutto, come si è già detto, il «significato» dei segnali per un individuo è endogeno, nel senso che dipende dall'esperienza passata. In secondo luogo, pur in assenza di rumore, il significato di un medesimo segnale può benissimo essere differente in termini di rappresentazione interna di individui differenti. In terzo luogo, anche l'insieme dei segnali ammissibili, cioè il potenziale informativo di un linguaggio, è endogeno: non esiste un canale dato esogenamente.

Supponiamo ora che l'economia sia su un sentiero di equilibrio congetturale. Supponiamo anche che un osservatore esterno (un eco-

<sup>33</sup> Hayek [1945, 53].

nometrico) si metta a raccogliere dati osservabili, cioè registri la storia in corso. Se tale storia non è troppo instabile, l'econometrico riuscirà a trovare un modello che si adatti bene ai dati: si noti che verosimilmente sarà un modello «parsimonioso», non come quello dei giocatori sofisticati che usano molti parametri non identificabili. In effetti, il modello econometrico fa riferimento alle azioni osservabili, e non alle teorie individuali non osservabili. Poniamo infine che, per queste ragioni, quel modello abbia un discreto successo tra i colleghi econometrici ed economisti, per cui molti di loro cominciano a crederci. Dato il successo ottenuto, l'econometrico avvia ora una prosecuzione della ricerca: tramite interviste dirette, raccoglie dati sulle aspettative degli agenti, per vedere se siano statisticamente significative quando inserite in un modello più generale del precedente. Per definizione (del concetto di ordine sociale), trova che gli agenti formulano previsioni che sono in linea con le previsioni del suo modello. Infine pubblica i suoi risultati e la vicenda finisce con un grande successo accademico: anche molti agenti (inclusi i responsabili della politica economica) cominciano a convincersi della sua versione, rafforzando ulteriormente la tenuta del vigente ordine sociale.

Si confronti tutto ciò con la seguente frase: «L'ipotesi di aspettative razionali [...] afferma che le aspettative delle imprese (o, più in generale, le distribuzioni di probabilità soggettive degli esiti) tendono ad essere distribuite, per lo stesso insieme di informazione, attorno alle predizioni della teoria (o la distribuzione di probabilità "oggettiva" degli esiti)» [Muth 1961, 316].

La morale che si può trarre dal ragionamento precedente è che, mentre è certamente vero che l'informazione contribuisce a produrre ordine (individuale), diversi ordini (sociali) possono produrre strutture molto diverse dell'informazione. Le teorie economiche (degli agenti in generale) e la Teoria Economica (di alcuni agenti particolari) non sono esterne a questo circuito.

## 5. Conseguenze: evoluzione per rotture degli ordini

Gli argomenti sollevati in queste pagine sono, secondo chi scrive, in sintonia con alcune opinioni di Hayek. Da una parte i singoli individui posseggono conoscenze parziali rispetto al complesso delle teorie esistenti nell'economia: è impossibile che una sola mente possa concepire tutta la conoscenza e le teorie incorporate in varie altre menti. Insistiamo sul fatto che questa impossibilità è un problema logico, non quantitativo: non si tratta del fatto che la buca nella sabbia

è troppo piccola per ricevere l'oceano, ma del fatto che in un certo senso i tentativi di ampliare la buca fanno ampliare anche l'oceano. Si osservi che la confusione tra l'aspetto quantitativo e quello logico del problema costituisce un'interessante questione di interpretazione del lavoro dello stesso Hayek: che la concorrenza, via i prezzi, conduca «infine» a svelare tutta l'informazione necessaria è una affermazione hayekiana sicuramente ambigua. Mentre gran parte delle interpretazioni successive propende per il versante quantitativo (convergenza asintotica ad uno stato di informazione pienamente diffusa), pare a chi scrive che la posizione di Hayek sia molto più articolata di ciò. Ricordiamo solo la sua opinione (contenuta in *The Meaning of Competition*, in Hayek [1945]) secondo cui «la concorrenza è il modo in cui le opinioni si formano», non il mezzo per far convergere le opinioni alla verità<sup>56</sup>.

La sola versione attendibile di fronte all'impossibilità logica di un modello vero è che ogni singolo tenti di mettere ordine, in maniera del tutto personale e path-dependent, sui dati sensoriali. Ma ciò non produce anarchia e disordine: il processo di interazione fra le azioni economiche e le attività cognitive individuali può raccordare le varie tensioni esistenti e incanalarle in uno fra i molti possibili ordini sociali: «Dunque è possibilissimo che prevalga una posizione di equilibrio per il solo fatto che alcuni non hanno la possibilità di apprendere cose che, se fossero loro note, li indurrebbero a modificare i loro piani» [Hayek 1945, 53].

Tutto ciò posto, occorre riconoscere che in Hayek è quasi assente un secondo fondamentale aspetto dell'interazione sociale (il cui studio era invece caro a Schumpeter): il verificarsi di repentine e radicali interruzioni di un preesistente ordine, che si osservano a diverse scale del sistema economico e sociale. Non si tratta di concepire tali interruzioni come derivanti banalmente da disturbi esogeni. In effetti, chiunque è capace di concepire un qualsiasi shock sufficientemente importante (chiamandolo magari innovazione tecnologica), inserirlo in un modello di equilibrio e generare così un cambiamento improvviso; con una opportuna sequenza di tali shock si può anche ottenere una sovrapposizione di onde brevi, medie e lunghe. Non sembra questa la prospettiva di Schumpeter, il quale era interessato al cambiamento

<sup>56</sup> Su questo punto si vedano le osservazioni di Zappia [1994], che smontano la lettura di Hayek in termini di teoria delle aspettative razionali e dei mercati pienamente efficienti (pur concedendo forse troppo al versante delle «asimmetrie informative» volume. Si veda anche il contributo di Zappia al presente

come tratto tipico di un'economia complessa. Si consideri invece il seguente argomento.

Come si è detto, allungando la durata temporale su cui vogliamo si estenda un equilibrio congetturale a parità di complessità delle teorie individuali, l'insieme degli equilibri si assottiglia (si riduce la sua dimensionalità). Dunque esiste una durata tale per cui nessun equilibrio di questo tipo esiste: tale allungamento di durata fa sì che la storia sperimentata possa diventare sufficientemente complessa da non essere più comprensibile per gli agenti. Supponiamo che l'economia sia stata su un sentiero di equilibrio sino alla data precedente: come reagiranno gli agenti alla data successiva? Tutto dipende da quanto la prosecuzione diverge dalle tendenze precedenti, ovvero da quanto queste erano «stazionarie»: più lo erano, più è grande la sorpresa. Supponiamo che lo fossero abbastanza per rendere necessaria una radicale revisione delle teorie (ovviamente non esiste una teoria bayesiana della revisione meccanica delle distribuzioni che sia applicabile a questa circostanza).

Esistono sicuramente alcuni agenti la cui teoria era in precedenza più sofisticata (usava più parametri) di altre: costoro saranno gli ultimi a scomporsi. Si assisterà inizialmente a fluttuazioni scoordinate, perché vari agenti tenteranno aggiustamenti indipendenti gli uni dagli altri. Se però qualcuno degli agenti «sofisticati» riesce a rivedere la propria teoria in modo non troppo drammatico (effettuando solo un processo di generalizzazione, così da mantenere una certa fiducia nel proprio modo di pensare), tale agente potrebbe continuare a comportarsi in modo ordinato, lucido, razionalizzabile, benché naturalmente diverso da prima. In altri termini, sarà per lui più agevole tentare di «spiegare» agli altri agenti cosa sta succedendo. Se inoltre questi agenti sofisticati riescono a non farsi trascinare dal rumore che li circonda, essi cominceranno a emettere segnali che possono suonare come «meccanismi» di coordinamento per gli altri.

Chiaramente non si può prevedere cosa accadrà con precisione («nel dettaglio», come dice Hayek). Ma, posto che i precedenti «se» siano verificati, è probabile che la fluttuazione che ne segue, anziché rimanere disordinata, assuma nel tempo una forma più facilmente identificabile: niente è infatti più adatto - a indurre gli uomini a rassicurarsi imitando gli altri - di un periodo di alta incertezza. Se il sistema, a partire da quella fluttuazione, produce nel tempo un nuovo ordine sociale, cognitivo, culturale, allora si riconosceranno gli individui che abbiamo chiamato sofisticati come gli «innovatori», precursori dei tempi nuovi. Se invece l'economia non riesce a coordinarsi sul loro comportamento iniziale, non si riuscirà a dare alcun significato

alle azioni da loro intraprese durante la fase di transizione («pazzi») Si osservi che il punto di partenza per gli innovatori non è uno shock esogeno (o degli *animal spirits*, nell'interpretazione deteriore della locuzione), bensì una freddezza sufficiente per inventare qualche futuro possibile a partire da un presente considerato dai più incomprensibile. Più un pizzico di fortuna, necessario per intercettare le altrui teorie e farsi «imitare».

Queste possono essere, secondo chi scrive, basi per una costruzione teorica che ci metta in grado di concepire le crisi come prodotto congiunto di una teoria dell'ordine.

### Riferimenti bibliografici

- Anderlini, L. e Felli, L. (1994), *Incomplete Written Contracts: Undescribable States of Nature*, in «The Quarterly Journal of Economics», CIX, pp. 1085-1124.
- Arrow, K.J. (1992), *Informational Equivalence of Signals*, in P. Dasgupta, D. Dale, O. Hart e E. Maskin (a cura di), *Economic Analysis of Market and Games*, Cambridge (Mass.), The MIT Press.
- Aspray, W. (1985), *The Scientific Conceptualization of Information: A Survey*, in «Annals of the History of Computing», VII, pp. 117-140.
- Aumann, R.J. (1976), *Agreeing to Disagree*, in «Annals of Statistics», IV, pp. 1236-1239.
- (1987), *Correlated Equilibrium as an Expression of Bayesian Rationality*, in «Econometrica», LV, pp. 1-18.
- Bacharach, M. (1992), *The Acquisition of Common Knowledge*, in C. Bicchieri e M.L. Valla (a cura di), *Knowledge, Belief and Strategic Interaction*, Cambridge, Cambridge University Press.
- Bacharach, M., Shin, H.S. e Williams, M. (1992), *Sophisticated Bounded Agents Play the Repeated Dilemma*, Oxford Project on Bounded Rationality in Economic Behaviour, Oxford University, mimeo.
- Binmore, K. (1987-8), *Modeling Rational Players, Parts I & II*, in «Economics and Philosophy», III, pp. 179-214; IV, pp. 9-55.
- Birner, J. (1996), *Mind, Market and Society. Network Structures in the Work of F.A. Hayek*, University of Maastricht, mimeo.
- Blackwell, D. (1953), *Equivalent Comparisons of Experiments*, in «Annals of Mathematical Statistics», XXIV, pp. 265-272.
- Brandenburger, A., Dekel, E. e Geanakoplos, J. (1992), *Correlated Equilibrium with Generalized Information Structures*, in «Games and Economic Behavior», IV, pp. 182-201.
- Cutland, N. (1980), *Computability*, Cambridge, Cambridge University Press.
- DeGroot, M.H. (1970), *Optimal Statistical Decision*, New York, McGraw-Hill.
- Edelman, G.M. (1993), *Sulla materia della mente*, Milano, Adelphi.

- Fishburn, P.C. (1982), *Subjective Expected Utility: A Review of Normative Theories*, in «Theory and Decision», XIII, pp. 139-199.
- Gallager, R. (1968), *Information Theory and Reliable Communication*, New York, Wiley.
- Georgescu-Roegen, N. (1975), *The measure of information: a critique*, in J. Rose e C. Bilciu (a cura di), *Modern Trends in Cybernetics and Systems*, Berlin, Springer Verlag, vol. III.
- Gilboa, I. e Schmeidler, D. (1993), *Case-Based Knowledge Representation: An Expository Note*, Università di Northwestern e Tel Aviv, mimeo.
- (1995a), *Case-Based Decision Theory*, in «Quarterly Journal of Economics», C, pp. 605-640.
- (1995b), *An Overview of Case-Based Decision Theory*, in *Proceedings of the International School of Economic Research on «Decision under Uncertainty»*, Università di Siena, mimeo.
- Green, J.R. e Laffont, J.J. (1986), *Incentive theory with data compression*, in W.P. Heller, R.M. Starr e D.A. Starret (a cura di), *Uncertainty, Information and Communication. Essays in honor of K.J. Arrow*, Cambridge, Cambridge University Press, vol. III.
- Hahn, F.H. (1973), *On the Notion of Equilibrium in Economics. An Inaugural Lecture*, Cambridge, Cambridge University Press.
- Harsanyi, J.C. (1967-8), *Games with Incomplete Information Played by Bayesian Players*, parti I-III, in «Management Science», XIV, pp. 159-182, 320-334, 486-502.
- Hayek, A.F. von (1945), *Individualism and Economic Order*, Chicago, The University of Chicago Press.
- (1952), *The Sensory Order. An Inquiry into the Foundations of Theoretical Psychology*, Chicago, The University of Chicago Press.
- Hertz, J., Krogh, A. e Palmer, R.G. (1991), *Introduction to the Theory of Neural Computation*, Redwood City (Calif.), Addison Wesley.
- Hirschleifer, J. e Riley, J.G. (1992), *The Analytics of Uncertainty and Information*, Cambridge, Cambridge University Press.
- Kalai, E. e Lehrer, E. (1993), *Rational Learning Leads to Nash Equilibrium*, in «Econometrica», LXI, pp. 1019-1045.
- Laffont, J.J. (1989), *The Economics of Uncertainty and Information*, Cambridge (Mass.), The MIT Press.
- Lipman, B.L. (1991), *How to Decide How to Decide How to... Modeling Limited Rationality*, in «Econometrica», LIX, pp. 1105-1125.
- (1995a), *Limited Rationality and Endogenous Incomplete Contracts*, Department of Economics, Queen's University and University of Pennsylvania, mimeo.
- (1995b), *Decision Theory without Logical Omniscience: Toward an Axiomatic Framework for Bounded Rationality*, Department of Economics, Queen's University and University of Pennsylvania, mimeo.
- Luini, L. (1994), *Economia dell'informazione. Mercato, tecnologia, organizzazione*, Firenze, La Nuova Italia Scientifica.

- Lunghini, G. e Rampa, G. (1996), *Conoscenza, equilibrio e incertezza endogena*, di prossima pubblicazione in «Economia Politica», VIII.
- Machina, M. (1982), «Expected Utility» Analysis without the Independence Axiom, in «Econometrica», L, pp. 277-323.
- (1987), *Choice Under Uncertainty: Problems Solved and Unsolved*, in «Economic Perspectives», I, pp. 121-154.
- Marshall, T. (1971), *Economic Information, Decision, and Prediction. Selected Essays*, Dordrecht, Reidel.
- Marschak, T. e Radner, R. (1972), *Economic Theory of Teams*, New Haven, Yale University Press.
- Mertens, J.F. e Zamir, S. (1985), *Formulation of Bayesian Analysis for Games with Incomplete Information*, in «International Journal of Game Theory», XIV, pp. 1-29.
- Muliere, P. e Parmigiani, G. (1993), *Utility and Means in the 1930s*, in «Statistical Sciences», VIII, pp. 421-432.
- Muth, J.F. (1961), *Rational Expectations and the Theory of Price Movements*, in «Econometrica», XXIX, pp. 315-335.
- Neumann, J. von e Morgenstern, O. (1944), *Theory of Games and Economic Behavior*, Princeton, Princeton University Press.
- Radner, R. (1993), *The Organization of Decentralized Information Processing*, in «Econometrica», LXI, pp. 1109-1146.
- Rampa, G. (1989), *Conjectures, Learning and Equilibria in Monopolistic Competition*, in «Journal of Economics», XLIX, pp. 139-163.
- (1990), *Apprendimento interattivo «non razionale» ed equilibri conseguenti*, in «Giornale degli Economisti e Annali di Economia», XLVIII, pp. 1-28.
- (1991), *Giochi, concorsi di bellezza e slot machines. Riflessioni sui processi di apprendimento*, in «Economia Politica», III, pp. 429-457.
- (1997a), *Trovare in ordine, mettere in ordine: il difficile rapporto tra Hayek e gli economisti*, in «Materiali per una storia della cultura giuridica», XXVII (1), pp. 105-126.
- (1997b), *Social Systems as Nets of Networks and the Problem of Information Transmission*, in «Quaderni dell'Istituto di Scienze Economiche», Genova.
- Rasmusen, E. (1989), *Games and Information. An Introduction to Game Theory*, Oxford, Basil Blackwell.
- Savage, L.J. (1954), *The Foundations of Statistics*, New York, Wiley.
- Serra, R. e Zanarini, G. (1990), *Complex Systems and Cognitive Processes*, Berlin, Springer Verlag.
- Shannon, C. e Weaver, W. (1949), *The Mathematical Theory of Communication*, Urbana, University of Illinois Press.
- Tan, T.C. e Werlang, S.R.C. (1988), *The Bayesian Foundations of Solution Concepts of Games*, in «Journal of Economic Theory», XLV, pp. 370-391.
- Zappia, C. (1994), *The Notion of Private Information in a Modern Perspective: A Reappraisal of Hayek's Contribution*, Università di Siena, mimeo.