

1. L'INDIVIDUALISMO METODOLOGICO E IL CONCETTO DI EQUILIBRIO NELLE SCIENZE SOCIALI.

"...the notion that a social system moved by independent actions in pursuit of different values is consistent with a final coherent state of balance, and one in which the outcomes may be quite different from those intended by the agents, is surely the most important contribution that economic thought has made to the general understanding of social processes." (Arrow - Hahn [1971], p. 3)

Probabilmente non tutti coloro che si occupano di scienze sociali sarebbero d'accordo sull'esatto significato dell'espressione "individualismo metodologico", ma concorderebbero nel collegarla in qualche modo al tentativo di spiegare i fenomeni sociali in base al comportamento degli individui:

"L'individualismo metodologico impone allo storico e al sociologo di cercar di spiegare gli sviluppi storici e i processi sociali come risultato (generalmente non intenzionale e talvolta molto al di là dei pensieri dei partecipanti) delle azioni di agenti singoli." (Watkins [1981], p 18).

Non si può negare che un approccio di questo tipo è assai diffuso nelle scienze sociali e in particolare è dominante nella scienza economica. In questo capitolo non ci si propone di difendere l'individualismo metodologico, ma soltanto di chiarirne il contenuto per poi evidenziare le forti connessioni tra l'adesione all'individualismo metodologico e l'importanza del concetto di equilibrio nelle scienze sociali.

L'adesione all'individualismo metodologico costituisce la premessa fondamentale e non argomentata di questa dissertazione.

Se si ritiene che i fenomeni sociali debbano essere spiegati come il risultato, non necessariamente intenzionale, del comportamento di individui razionali, allora è logico considerare come situazioni di equilibrio quelle nelle quali gli individui vedono confermate le loro teorie. Infatti un individuo razionale adotta il comportamento che, in base alla sua teoria, produce le conseguenze preferite. Se la teoria non viene smentita dall'esperienza, quell'individuo non ha motivi per modificare il suo comportamento.

Un'altra naturale conseguenza dell'adesione all'individualismo metodologico è l'importanza attribuita alla teoria

dei giochi, in quanto teoria o linguaggio formale delle scienze sociali. Di regola infatti, l'interazione tra più individui è tale che gli eventi, cui il singolo attribuisce importanza (gli eventi sui quali esprime delle preferenze), dipendono non solo dalle sue azioni, ma anche da quelle di qualcun altro e lo scopo principale della formalizzazione operata con la teoria dei giochi è proprio quello di cogliere e rendere esplicita questa interdipendenza.

1.1. L'individualismo metodologico inteso come "superprogramma" di ricerca nelle scienze sociali.

"Ogni persona segue il proprio fine desiderato coscientemente ed è proprio il risultato delle numerose volontà che operano in direzioni diverse, dei loro multiformi effetti sul mondo esterno, che costituisce la storia (...) le innumerevoli volontà individuali attive nella storia per la gran parte producono però risultati completamente diversi - e molto spesso opposti - da quelli intenzionali" (Engels; Marx-Engels, Selected Writings)⁽¹⁾

"E tale osservazione ci offre l'occasione per chiarire il compito principale delle scienze sociali teoriche. Esso consiste nel delineare le ripercussioni sociali, non intenzionali, che seguono alle azioni umane intenzionali". (Popper, [1972], p 63)

In questo paragrafo ci si propone di esaminare e definire il concetto di individualismo metodologico mediante alcune categorie metodologiche mutuata, direttamente o indirettamente, dal filosofo Imre Lakatos. L'uso di tali categorie è dunque eminentemente descrittivo-classificatorio e non comporta una adesione all'epistemologia di Lakatos per quanto attiene al suo contenuto normativo.

Un programma di ricerca è un sistema teorico caratterizzato da regole metodologiche che indicano al ricercatore quali vie evitare (euristica negativa) e più approssimativamente quali vie seguire (euristica positiva). L'euristica negativa opera quando il ricercatore si trova di fronte ad anomalie o controfatti che implicano logicamente la falsità dell'intero sistema costituito dalla congiunzione di proposizioni teoriche, condizioni iniziali e "di contorno", teorie osservative, regole interpretative. L'euristica negativa prescrive allora di non rigettare un insieme di ipotesi o leggi detto nucleo ("hard-core") del programma. La difesa del nucleo si attua tramite la modificazione di una "impalcatura" costituita sia

condizioni iniziali, ipotesi di contorno, regole interpretative e teorie osservative. E' tale impalcatura, detta cintura protettiva, il bersaglio del modus tollens². In altre parole rigettando qualche ipotesi della cintura protettiva ed eventualmente sostituendola con una ipotesi diversa, si ottiene una nuova teoria complessiva che è logicamente compatibile con quei "fatti", o meglio quelle proposizioni osservative, che falsificavano la vecchia. In questo modo, però, è possibile introdurre nella struttura teorica delle nuove ipotesi completamente sconnesse rispetto ad essa. Ad esempio una spiegazione delle fluttuazioni negli investimenti, secondo la quale le aspettative degli imprenditori dipendono dalle macchie solari, potrebbe anche ottenere una discreta corroborazione empirica; cionondimeno sarebbe rifiutata da qualsiasi economista "serio", perchè un'ipotesi di questo genere è assolutamente sconnessa rispetto a qualsiasi sistema teorico nel campo dell'economia. L'euristica positiva ha il compito di evitare fenomeni di questo tipo, prescrivendo almeno approssimativamente come modificare la teoria quando si producono delle anomalie (si veda Lakatos [1979], pp 208-214).

Un programma di ricerca risulta quindi completamente caratterizzato dal suo nucleo e dalla sua euristica positiva perchè questi sono gli elementi di continuità nella serie di teorie o modelli generata dal programma stesso.

Quando i nuclei e le euristiche positive di diversi programmi di ricerca hanno rispettivamente in comune almeno una ipotesi e una regola, è possibile individuare un super-programma di ricerca che li contiene³.

Si definisce allora l'individualismo metodologico come un super-programma di ricerca nelle scienze sociali, caratterizzato dal nucleo e dall'euristica seguenti:

- nucleo: il comportamento umano è costituito da azioni intenzionali, cioè volte al perseguimento di obiettivi; le azioni non sono sistematicamente scorrette rispetto agli obiettivi, date le effettive possibilità cognitive dell'agente, ovvero il comportamento è, in senso lato, razionale;
- euristica positiva: i fenomeni sociali vanno spiegati come "effetti perversi", cioè come effetti sociali non

(necessariamente) intenzionali di azioni umane intenzionali attuate in un contesto di interazione reciproca⁴.

Per giustificare in modo esaustivo questa caratterizzazione dell'individualismo metodologico sarebbe necessaria una estesa rassegna sul susseguirsi dei diversi approcci nella storia delle scienze sociali e sulle scuole attuali (soprattutto nel campo dell'economia). Ciò va ben oltre i limiti e gli scopi di questa dissertazione. Così come l'adesione all'individualismo metodologico, anche la sua definizione costituisce una premessa fondamentale che si spera essere sufficientemente condivisibile. E' comunque opportuno esporre alcune considerazioni a commento di questa definizione.

1) Questa caratterizzazione dell'individualismo metodologico fornisce una nuova base razionale alla tesi secondo cui il principio di razionalità è l'unica legge universale delle scienze sociali e, pur essendo probabilmente falso, svolge una funzione insostituibile nella spiegazione dei fenomeni sociali (Popper [1967]). Infatti una legge universale formulata nell'ambito dell'individualismo metodologico sarà, di norma, un asserto condizionale del tipo "se...allora" nel cui antecedente compaiono delle condizioni iniziali e di contorno e il cui conseguente è dedotto dal postulato di razionalità. La funzione insostituibile del postulato di razionalità è dovuta alla combinazione dell'euristica negativa con quella positiva: la prima difende il postulato dalle falsificazioni, la seconda lo chiama in causa in qualsiasi spiegazione di fenomeni sociali.

2) Il postulato di razionalità non è tautologico. Le affermazioni in senso contrario non provengono solo dai critici, ma anche dai sostenitori (Luce, Raiffa [1957], p 50). Le argomentazioni sono normalmente basate su due punti. Anzitutto le preferenze sono ritenute equivalenti alle scelte: la proposizione "a é preferita a b" è identificata con "tra a e b viene scelta a". Si procede poi dalla impossibilità di osservare le preferenze alla non falsificabilità del postulato di razionalità, visto che qualsiasi comportamento potrebbe essere razionale per qualche struttura di preferenze.

L'introduzione nel quadro delle conoscenze soggettive in base alle quali l'agente deriva le preferenze sulle azioni da quelle sulle conseguenze sembra rafforzare questa conclusione; le preferenze sulle conseguenze e le conoscenze soggettive sono infatti ancora meno osservabili delle preferenze sulle azioni. Anche se si prescinde dalla possibilità che i limiti nelle capacità di calcolo e di deduzione dell'agente producano scelte irrazionali (incompatibili con le preferenze sulle conseguenze, date le conoscenze soggettive), queste argomentazioni rimangono epistemologicamente superficiali. Il fatto che una proposizione non sia falsificabile non implica che essa sia tautologica; significa soltanto che le tecniche di controllo disponibili non sono in grado di produrre eventi osservabili che implicano la falsità della proposizione stessa, considerata isolatamente⁵. Il motivo per cui il postulato di razionalità è mantenuto anche di fronte a un'evidenza apparentemente contraria è che tale evidenza lo contraddice congiuntamente ad altre ipotesi e l'euristica negativa dell'individualismo metodologico fa sì che siano queste ultime a essere rigettate. Dunque il postulato di razionalità, pur non essendo vero a priori, risulta (metodologicamente) non falsificabile. Con ciò si accolgono nella sostanza le istanze metodologiche della scuola austriaca senza ricorrere alle discutibili tesi aprioristiche di Von Mises (si vedano Von Mises [1949], p 34, Latsis [1976], pp 3-7 e Caldwell [1982] pp 146-168)⁶.

3) Il postulato di razionalità enunciato sopra fa riferimento alle effettive possibilità cognitive dell'agente e quindi è sufficientemente generale da includere la razionalità limitata, ma in relazione ad essa sorgono almeno due difficoltà:

(a) se i limiti all'azione razionale sono dovuti a scarsità di informazioni è possibile ricorrere alla teoria dell'utilità in condizioni d'incertezza, ma, come si vedrà meglio in seguito (paragrafi 1.2 e 3.4), ciò non può risolvere il problema di come gli agenti modificano le loro teorie di fronte a previsioni errate;

(b) se invece i limiti riguardano le capacità di raccolta ed elaborazione delle informazioni, si può cercare di introdurre nei modelli di comportamento degli algoritmi decisionali semplici (algoritmi di questo tipo sono di fatto

incorporati nei programmi per il gioco degli scacchi); un procedimento di questo tipo può avere dei meriti dal punto di vista operativo-normativo, ma rischia di essere assolutamente arbitrario da un punto di vista descrittivo se tali algoritmi non sono in qualche modo equivalenti ad una massimizzazione vincolata. Inoltre l'uso di modelli di razionalità limitata è spesso giustificato in base a considerazioni sulla natura complessa dell'agente, il cui comportamento è in realtà il risultato dell'interazione di più individui (è il caso dell'impresa), ma in questo caso l'euristica positiva spinge verso un livello di analisi più profondo, in cui gli agenti corrispondono a singoli individui. In seguito si ipotizzerà che si sia già raggiunto questo livello di analisi e che le capacità deduttive e di calcolo degli individui siano illimitate nell'ambito del modello trattato. Ciò permetterà di concentrarsi sui problemi sollevati in (a). '7'

4) La coerenza delle azioni rispetto agli obbiettivi va intesa in senso pragmatico e in senso sintattico. L'agente è pragmaticamente coerente, ovvero sceglie delle azioni che producono, secondo le sue conoscenze, le conseguenze da lui preferite; è sintatticamente coerente ~~se~~, o coerente in senso proprio, se le sue conoscenze complessive non sono autocontraddittorie. Questa osservazione è fondamentale per comprendere il ruolo del concetto di equilibrio nell'ambito dell'individualismo metodologico (paragrafo 1.2) e conduce a enunciazioni formali del principio di razionalità in cui questo è suddiviso in due parti, corrispondenti alla razionalità pragmatica e a quella sintattica. X

5) La specificazione delle preferenze di base dell'agente, delle sue conoscenze soggettive (incluse eventualmente le sue valutazioni probabilistiche) e delle modalità di acquisizione delle informazioni permette di derivare dal postulato di razionalità un modello di comportamento, che è valido almeno fino al momento in cui le nuove informazioni non spingono l'agente a modificare le sue conoscenze teoriche. Ma se l'analisi si limitasse a questo nulla distinguerebbe le teorie sociali dalla teoria delle decisioni. E' l'euristica positiva che porta l'analisi a un livello realmente sociale, l'adesione ad essa differenzia le scienze sociali da quelle

del comportamento. Le scienze sociali in quanto tali sono significative e interessanti perchè esistono fenomeni che, derivando da azioni umane, non sono naturali, ma non sono neanche artificiali, perchè le azioni che li hanno prodotti non erano preordinate al verificarsi di tali fenomeni. Tali fenomeni sono il risultato di azioni umane ma non di un progetto umano (Hayek [1978a], pp 96-105). Bisogna però precisare che la non intenzionalità degli effetti sociali non esclude necessariamente che tali effetti possano essere previsti (l'opinione contraria è sostenuta da alcuni seguaci della scuola austriaca, si veda ad esempio O'Driscoll, Rizzo [1985], p 107). Basta pensare al dilemma del prigioniero per comprendere che certe situazioni di inefficienza, che ovviamente non sono desiderate né intenzionalmente perseguite, possono essere comunque previste. Va inoltre osservato che i fenomeni di cui si occupa il teorico possono essere a un livello di descrizione della realtà sociale che è irrilevante per le intenzioni e le previsioni degli agenti. Sembra essere questo il caso delle proprietà astratte, che secondo Hayek dovrebbero essere l'oggetto di studio della teoria economica (Hayek [1978a], p 210).

6) E' ormai un luogo comune affermare che le divisioni disciplinari tra le scienze sociali dipendono più dal metodo che dall'oggetto di studio. Se si accetta questa premessa, bisogna concludere che le tradizionali suddivisioni disciplinari sono arbitrarie dal punto di vista dell'individualismo metodologico, perchè gli "effetti perversi" sono un prodotto dell'interazione in generale (si veda ad esempio O'Driscoll, Rizzo [1985] p 107 e 128). Il fatto che l'individualismo metodologico sia nato nell'ambito dell'economia non deve far presupporre che esso sia caratteristico di questa disciplina. Raymond Boudon ha mostrato quanto estesamente questo approccio venga applicato in studi comunemente ritenuti sociologici o politologici (Boudon [1981] e [1983]). E' vero però che la grande maggioranza degli studi formali non completamente astratti finora prodotti hanno un contenuto di tipo economico. Questo spiega perchè nel seguito il riferimento generale sarà ai fenomeni sociali, mentre le esemplificazioni non puramente matematiche si riferiranno a fenomeni economici.

1.2. Il concetto di equilibrio nell'ambito dell'individualismo metodologico.

"Una formulazione generale del concetto di equilibrio che dipenda unicamente dagli assiomi da tutti condivisi rende meno probabile che si commettano errori quando l'analisi si fa precisa e particolare.(...) L'equilibrio è una astrazione che serve da punto di partenza, e la sua posizione centrale nella teoria economica deriva, in parte, dalla circostanza che gli assiomi fondamentali sono abbastanza potenti da permetterci di definire e studiare gli stati di equilibrio. Tuttavia questi assiomi non sono sufficienti per formulare una teoria che spieghi come l'economia raggiunga l'equilibrio (...). Per questa ragione siamo costretti a ricorrere a postulati ad hoc relativi al comportamento degli agenti economici, cioè a postulati che non derivano dagli assiomi della razionalità e dell'egoismo" (Hahn [1984], p 11).

Un sistema si dice in equilibrio quando non esistono forze endogene tendenti a modificare le sue variabili di stato. Naturalmente la descrizione dello stato del sistema può essere tale che il suo equilibrio corrisponda a una situazione di cambiamento continuo; è inoltre possibile tenere conto degli elementi stocastici senza alterare la definizione in modo essenziale.

Questa definizione non contiene nessun elemento specificamente economico o sociale, ma permette di fare alcune osservazioni utili.

1) Affermare che un certo stato del sistema è di equilibrio equivale ad asserire qualcosa sulla sua dinamica. Ciò può essere mostrato anche con un esempio matematico. Si supponga di avere un sistema, che varia in modo continuo rispetto al tempo, il cui stato al tempo t sia descritto da una variabile reale $x(t)$ e che possiede uno stato di equilibrio x^0 . Si supponga inoltre che sia definita la funzione

$$f(x) = dx/dt$$

(ciò implica che il sistema sia autonomo). La dinamica del sistema è allora rappresentata dal grafo della funzione $f(\cdot)$ e su tale grafo si può affermare che esso contiene il punto $(x^0, 0)$.

2) Quando la dinamica del sistema è completamente nota

la conoscenza dei suoi stati di equilibrio diventa un mero sottoprodotto (nell'esempio precedente gli stati di equilibrio corrispondono alle radici dell'equazione $f(x)=0$), quindi la rilevanza di una teoria dell'equilibrio dipende, un po' paradossalmente, dall'ignoranza sulla dinamica del sistema in esame (cfr. Hahn [1984], p. 35 e Von Neumann, Morgenstern [1980], p. 45).

La tesi fondamentale di questo paragrafo è che l'adesione all'individualismo metodologico non ha finora permesso e probabilmente non permetterà di elaborare teorie sufficientemente precise e rigorose sulla dinamica dei sistemi sociali in disequilibrio. Di conseguenza la fondamentale importanza del concetto di equilibrio nelle scienze sociali e particolarmente in economia è direttamente correlata all'adesione a questo tipo di approccio. Nel corso dell'argomentazione a sostegno di questa tesi saranno discussi i concetti di teoria e apprendimento di un agente, in base ai quali viene definito il significato dell'equilibrio nell'ambito dell'individualismo metodologico. Tali concetti sono mutuati, con alcune modificazioni, dall'economista matematico Frank Hahn, con particolare riferimento al saggio del 1973 *On the Notion of Equilibrium in Economics* (tradotto in italiano e pubblicato in Hahn [1984], pp. 17-53).

Secondo l'euristica positiva dell'individualismo metodologico una descrizione adeguata di un fenomeno sociale deve essere derivabile dalla specificazione delle azioni degli agenti. Si può immaginare che tali azioni siano scelte in base a piani o programmi che prescrivono quale azione adottare per ogni informazione che il singolo agente potrebbe ricevere^a. Per indicare tali programmi si userà il termine "strategia". E' dunque possibile identificare le variabili di stato di un sistema sociale con le strategie degli agenti che ne fanno parte. Ne consegue che un sistema sociale è in equilibrio quando nessun agente ha incentivi a modificare la sua strategia. Il postulato di razionalità impone che la strategia di un agente sia coerente rispetto ai suoi obiettivi (le sue preferenze) e alla sua conoscenza soggettiva alla data in cui tale strategia viene adottata. Quindi gli incentivi a modificare una strategia derivano dal fatto che questa si rivela irrazionale alla luce di nuove

informazioni (o "messaggi"). Prescindendo da variazioni nelle preferenze⁴⁷, ciò equivale a un cambiamento della teoria (ipotetica connessione tra azioni e conseguenze) in base alla quale l'agente forma le sue aspettative.

Hahn definisce la teoria di un agente nel modo seguente:

"Ad una data qualsiasi t esiste una «storia» dei messaggi ricevuti dall'agente, che dividiamo in quelli che l'agente considera indipendenti dalle sue azioni - i messaggi esogeni - e tutti gli altri (...) con l'espressione «teoria di un agente alla data t » intendo quanto segue:

- (a) l'agente ha suddiviso i messaggi nelle due categorie menzionate;
- (b) per ogni successione di messaggi esogeni a partire dalla data t l'agente ha una distribuzione di probabilità del risultato di una qualsiasi successione di azioni dal tempo t in avanti;
- (c) alla data t l'agente assegnerà al fatto di ricevere un qualsiasi messaggio esogeno, ad una qualsiasi data futura, una probabilità condizionata dai messaggi ricevuti a partire dalla data t [fino a] quella data futura" (Hahn [1984], pp 32-33, nella parentesi quadra è contenuta la correzione di una svista evidente del testo italiano).

Le definizioni di "apprendimento" e "equilibrio" scaturiscono direttamente da quanto sopra:

"Dirò che l'agente sta apprendendo se la sua teoria non è indipendente dalla data t ; e condizione perchè l'agente sia in equilibrio è che egli non stia apprendendo. (...) una economia è in equilibrio quando genera messaggi che non inducono gli agenti a modificare le teorie in cui credono o le politiche che perseguono" (Hahn [1984], pp 33 e 39).

Queste definizioni di Hahn costituiscono un punto di riferimento fondamentale della presente dissertazione. Prima di proseguire nell'argomentazione principale sono quindi necessari alcuni commenti, anche di carattere critico.

La "storia" dei messaggi ricevuti da un agente fino a una certa data t rappresenta l'informazione di quell'agente a quella data. Nel testo di Hahn è implicito che alla data t l'agente conosca tutti i messaggi ricevuti prima di t . Ciò significa che l'agente ricorda perfettamente e, attraverso il processo di memorizzazione, la sua informazione col trascorrere del tempo può solo aumentare o rimanere costante. Questa ipotesi sarà in seguito esplicitamente adottata e formalizzata nell'ambito della teoria dei giochi (paragrafo

3.1)).

Hahn definisce la teoria come un processo di elaborazione delle informazioni per formare delle aspettative condizionate da quelle informazioni. Le aspettative sono intese come distribuzioni di probabilità su uno spazio di eventi relativamente ai quali l'agente esprime delle preferenze. Questo modo di concepire la teoria di un agente ha il merito di mettere in evidenza che una variazione nelle aspettative, così come viene comunemente intesa, non implica che vi sia apprendimento. Infatti una variazione delle aspettative può riflettere semplicemente un aumento dell'informazione effettivamente in possesso dell'agente, senza che cambi il modo in cui l'agente deriva le aspettative dalle informazioni possibili. In particolare il meccanismo di revisione delle aspettative basato sul teorema di Bayes è un tipico esempio di mancanza di apprendimento nel senso di Hahn. Per chiarire ulteriormente il concetto di apprendimento è utile riportare un esempio dello stesso Hahn:

"Si supponga che alla data t l'agente abbia assegnato delle probabilità a due eventi: che pioverà e che non pioverà a Cambridge alla data $t+1$. Alla data $t+1$ l'agente saprà quale dei due eventi si è verificato. Ebbene questo incremento delle sue conoscenze non è ciò che io intendo con il termine apprendimento. Un esempio di apprendimento nel mio senso si avrebbe se alla data $t+1$ la probabilità che l'agente, avendo osservato la pioggia, attribuisce al piovere a Cambridge alla data $t+2$ è diversa dalla probabilità che egli alla data t aveva attribuito a questo evento condizionata al verificarsi della pioggia alla data $t+1$. (Hahn [1984], p 34).

Tuttavia il fatto che una teoria generi un processo di elaborazione delle informazioni e formazione delle aspettative non implica che quella teoria debba essere identificata con tale processo. Lo stesso Hahn sembra implicitamente ammettere che i) la teoria di un agente è costituita da una congiunzione di proposizioni (alcune delle quali possono essere probabilistiche) e ii) quando una teoria implica che un certo evento non si verificherà (cioè l'agente non riceverà i relativi messaggi) e tale evento si verifica la teoria viene falsificata e l'agente apprende (cfr Hahn [1984], pp 39-40). Da ciò consegue che una teoria non deve necessariamente specificare le probabilità subordinate ad

eventi impossibili secondo la teoria stessa. Per chiarire ciò proseguiamo con l'esempio di Hahn. Supponiamo che l'agente debba decidere in data t se prendersi un giorno di ferie in data $t+2$ per fare una gita a Cambridge e preferisca fortemente una gita col bel tempo. Se in base alla sua teoria prevede con probabilità unitaria (con certezza) che a Cambridge ci sarà bel tempo in data $t+1$, egli non avrà nessun interesse a valutare la probabilità del bel tempo in data $t+2$ subordinata alla pioggia in data $t+1$, perchè ciò è irrilevante nel calcolo delle probabilità sul tempo a Cambridge in data $t+2$ subordinate alle informazioni note alla data t , quella in cui l'agente deve decidere. E' vero che l'agente potrebbe prendere alla data $t+1$ delle decisioni rilevanti, ma egli non ha bisogno di specificare nel suo piano cosa farà se pioverà alla data $t+1$ finchè ritiene quell'evento impossibile. Se poi in $t+1$ l'agente scopre che effettivamente piove la sua teoria viene falsificata ed egli ne adotta un'altra per la quale la pioggia in data $t+1$ è possibile e che specifica le probabilità di eventi futuri subordinate a questo evento. Dunque il requisito (c) della definizione di Hahn va leggermente indebolito:

(c') alla data t l'agente suddivide i messaggi esogeni in due categorie, quelli possibili e quelli impossibili, e assegna al fatto di ricevere un qualsiasi messaggio esogeno ad una qualsiasi data futura una probabilità condizionata dai messaggi possibili ricevuti a partire dalla data t fino a quella data futura.

Un altro aspetto discutibile della definizione di Hahn è che la teoria di un agente al tempo t sembra essere identificata con le previsioni condizionate che l'agente fa relativamente ad eventi che possono verificarsi da t in poi. Non è quindi chiaro in che senso si possa parlare di una teoria dipendente o indipendente da t . Se l'espressione è intesa in senso stretto, l'eterogeneità temporale delle teorie implicata dalla definizione rende una teoria dipendente dal tempo per costruzione. L'interpretazione più ovvia sembra essere quella secondo cui una teoria è indipendente dal tempo se per ogni t le previsioni condizionate ad eventi con data $T \geq t + \tau$,

derivate dalla teoria al tempo t coincidono con le previsioni derivate dalla teoria al tempo $t+r$, condizionate agli stessi eventi. Ma da ciò consegue che possono verificarsi al tempo $t+r$ eventi impossibili secondo la teoria al tempo t , senza che vi sia apprendimento: è sufficiente che l'agente continui a fare le stesse previsioni condizionate relativamente agli eventi che possono verificarsi da $t+r$ in poi. Dal punto di vista qui adottato questa conclusione non è accettabile. Se è vero che in questo caso l'agente continua a seguire la strategia precedentemente adottata, è anche vero che, probabilmente, non risceglierebbe certe azioni prescritte da quella strategia per le date precedenti a $t+r$ e sembra intuitivo considerare ciò come una forma di apprendimento.

Questa obiezione perde gran parte del suo peso se si suppone che l'agente si trovi in un contesto ripetitivo, almeno per alcuni aspetti importanti. In questo caso la situazione sopra delineata, cioè quella di un agente che in $t+r$ vede falsificata la teoria del tempo t e ciononostante mantiene in $t+r$ una teoria congruente (nelle previsioni condizionate) con la teoria al tempo t , sembra essere contraria a un intuitivo principio di razionalità cognitiva o semantica, che impone di non ripetere gli stessi errori:

"On dira encore: dans ces conditions, que devient la distinction entre rationalité et irrationalité? (...) La différence principale, me semble-t-il, tient à ce que les idées d'une personne en bonne santé mentale ne sont pas impossible à corriger; une personne en bonne santé montre toujours une certaine disposition à réviser ses opinions" (Popper [1967], pp 148-149).

Il principio di razionalità semantica afferma in sostanza che gli agenti, essendo razionali, "non chiudono gli occhi davanti ai fatti". Questo principio è fondamentale ad un livello euristico-interpretativo e a tale livello deve considerarsi parte del più generale principio di razionalità. Tuttavia, al livello logico formale cui sarà svolta l'indagine successiva (cap. 3 e 4), da un lato non sarà possibile riconoscere una ripetitività nei contesti, dall'altro gli agenti saranno "costretti", per costruzione, a tenere conto dei fatti, resi loro noti sotto forma di informazioni. Ciò, nell'ambito dei modelli studiati, renderà la razionalità semantica indistinguibile da quella sintattica (cfr paragrafo

1.1, osservazione 4). La scelta di questo tipo di formalizzazione permetterà di rappresentare il "tempo reale", in cui l'aumento di informazione degli individui rende ogni evento diverso da quelli che lo precedono e perciò unico (cfr O'Driscoll, Rizzo [1985], p 61). Ma è chiaro che gli eventi possiedono anche delle caratteristiche "tipiche", temporalmente omogenee (ad esempio l'attributo "pivosità" del tempo meteorologico a Cambridge), e la rappresentazione di tali caratteristiche è importante nel definire l'apprendimento (cfr O'Driscoll, Rizzo [1985], pp 76-79). La scelta suddetta non è quindi priva di costi e ci si troverà costretti a introdurre a livello interpretativo una nozione di ripetitività relativa alle strategie.

In conclusione, il particolare approccio qui adottato impone una leggera modificazione delle definizioni date da Hahn:

1) la teoria di un agente al tempo t è costituita da una congiunzione di proposizioni, possibilmente probabilistiche, che l'agente ritiene valide al tempo t ; tale teoria genera un processo di elaborazione delle informazioni e formazione delle aspettative che soddisfa i requisiti (a) e (b) sopra enunciati; inoltre tutti gli eventi logicamente possibili¹⁰, compresi quelli anteriori a t , sono suddivisi in eventi possibili e impossibili secondo la teoria stessa e da essa è derivabile la probabilità di ogni evento condizionata agli eventi teoricamente possibili (ciò implica che è soddisfatto anche il requisito (c')).

2) un agente apprende se per qualche t la teoria al tempo t è diversa dalla teoria al tempo $t-1$; condizione sufficiente, ma non necessaria, affinché l'agente apprenda è che gli sia noto il verificarsi di un evento teoricamente impossibile.

A questo punto è possibile riprendere l'argomentazione principale. Perché la nozione di equilibrio è importante?

Da quanto detto sopra consegue che per conoscere la dinamica di un sistema sociale in disequilibrio è necessario specificare il processo di apprendimento degli agenti. Un'informazione preziosa sull'apprendimento viene fornita dal

principio di razionalità, da cui si può derivare almeno la "ipotesi imprecisa che un agente abbandona la sua teoria quando questa viene sufficientemente e sistematicamente falsificata" (Hahn [1984], p 39). Ma almeno per ora non sembra possibile andare oltre, perchè una teoria rigorosa "di livello superiore" sul processo di apprendimento non è attualmente disponibile nell'ambito dell'individualismo metodologico. Ciò non contrasta col fatto che esiste una grande quantità di studi matematici sulla dinamica del disequilibrio e in particolare sulla stabilità; questi studi infatti hanno un carattere ad hoc (cfr Hahn [1984], p 11 e [1982a], p 747) nel senso preciso che si basano su ipotesi in contrasto con l'euristica dell'individualismo metodologico (si tratta di ipotesi ad hoc 3 secondo la terminologia di Lakatos). Inoltre non si vuole qui affermare che non esistono teorie informali interessanti sull'apprendimento in un contesto sociale, prodotte nell'ambito di questo programma di ricerca. A questo proposito la teoria austriaca dell'imprenditorialità e della concorrenza come processo di scoperta sembra perlomeno degna di considerazione (cfr O'Driscoll, Rizzo [1985], pp 38-39). Ma quanto più una teoria è informale, tanto più le sue inferenze sono incerte e i suoi concetti poco nitidi e ciò riduce le possibilità critica e di deduzione di conclusioni interessanti. Dunque l'inesistenza di una teoria formale e precisa sull'apprendimento resta un problema fondamentale. L'individualismo metodologico permette di specificare le condizioni in cui gli individui apprendono, ma non di precisare (neanche in senso stocastico) come apprendono, quale nuova teoria sostituisce la vecchia. Ma se l'equilibrio consiste nell'assenza di apprendimento ed è possibile specificare le condizioni in cui gli individui apprendono, allora è anche possibile precisare quali configurazioni di teorie e strategie sono di equilibrio. Di norma la teoria sociale sarà in grado di spiegare la persistenza o ripetitività di certi fenomeni, mostrando che dipendono da comportamenti "di routine" (ovvero da strategie di equilibrio), ma non sarà in grado di prevederli. Tuttalpiù potrà prevedere la transitorietà di certe configurazioni sociali e l'impossibilità di alcune dinamiche.

Questa formulazione generale della nozione di equilibrio mostra dunque le possibilità e i limiti di una analisi rigorosa e formale nell'ambito dell'individualismo metodologico:

"...una volta tradotte la <<regolarità di comportamento>> nella definizione qui adottata di <<comportamento di equilibrio>>, abbiamo raggiunto il limite a cui, allo stato attuale delle conoscenze, si può spingere un economista. Questa è la ragione dell'interesse e dell'importanza della nozione di comportamento di equilibrio" (Hahn [1984], p 42).

NOTE

(1) La citazione italiana è riportata in Howard-King [1980], p 63.

(2) Il modus tollens è uno schema logico, una regola d'inferenza che permette di asserire la falsità di una proposizione "p", basandosi sulla falsità di una proposizione "q" implicata da "p" (Copi [1964] p 319). Adottando la formalizzazione del calcolo proposizionale, il modus tollens può essere fatto corrispondere a una proposizione tautologica o derivazione categorica (si veda Bencivenga [1984], p 60) della seguente forma:

$$((p \Rightarrow q) \wedge (\neg q)) \Rightarrow (\neg p).$$

I simboli non letterali diversi dalle parentesi sono connettivi logici:

" \Rightarrow "	implicazione	"se...allora"
" \wedge "	coniunzione	"e"
" \neg "	negazione	"non"

gli altri connettivi sono:

" \Leftrightarrow "	coimplicazione	"se e solo se"
" \vee "	disgiunzione	"o...o...oppure entrambe" (corrispondente al "vel" latino)

Le parentesi indicano in che ordine risolvere il calcolo logico, inferendo il valore di verità delle proposizioni composte da quello delle preposizioni elementari (indicate dalle lettere).

Il problema in questione è che le teorie "p" implicanti un evento "q" e falsificate dall'evento " $\neg q$ " mediante modus tollens, sono costituite dalla congiunzione di più proposizioni r,s,t... che, prese isolatamente, possono essere compatibili con " $\neg q$ ". Formalmente:

$$(\neg(r \wedge s \wedge t \dots)) \equiv ((\neg r) \vee (\neg s) \vee (\neg t) \dots)$$

quali proposizioni eliminare tra r,s,t...?

(3) E' possibile definire il concetto di superprogramma di ricerca mediante l'uso della logica formale.

Siano $PDR_1, PDR_2, \dots, PDR_K$ dei programmi di ricerca con nuclei n_i e euristiche positive e_i , $i=1,2,\dots,K$; i nuclei n_i sono proposizioni verofunzionali (cioè assumono valore di verità o falsità in funzione del valore di verità o falsità delle proposizioni relativamente più elementari che li compongono), le euristiche e_i sono proposizioni normative che possono essere operative (messe in atto) o non operative (disattese) in funzione della operatività delle proposizioni normative più elementari che le compongono. I valori di verità e operatività vengono inferiti secondo le

stesse regole astratte della logica formale. In particolare se "p" e "q" sono proposizioni verofunzionali (normative), valgono le seguenti tabelle di verità (operatività):

p	q	(p ∧ q)	(p ∨ q)
V	V	V	V
V	F	F	V
F	V	F	V
F	F	F	F

V = vero/operativo
F = falso/non operativo

Se esistono una proposizione verofunzionale "p" e una proposizione normativa "r" per le quali

$$n_i \equiv p \wedge p_{11} \wedge p_{12} \dots \wedge p_{1k(i)}$$

i=1,2...K, (k(i)+1 e l(i)+1 sono i numeri di proposizioni
relativamente elementari in n_i e e_i),

$$e_i \equiv r \wedge r_{11} \wedge r_{12} \dots \wedge r_{1l(i)}$$

allora si definisce super-programma minimale relativo ai programmi PDR₁, PDR₂,...PDR_K il programma di ricerca seguente:

$$SPDR(PDR_1, PDR_2 \dots PDR_K) = (sn, se),$$

$$sn \equiv n_1 \vee n_2 \dots \vee n_K, se \equiv e_1 \vee e_2 \dots \vee e_K$$

Si osservi che il super-nucleo "sn" è falso quando è falsa la proposizione comune "p" e la super-euristica è non operativa quando non lo è la regola comune "r", ma in generale non è detto che vi sia equivalenza tra "sn" e "p" o tra "se" e "r", vi è solo una relazione di implicazione. Il super-programma è quindi detto minimale, perchè possono esistere programmi più generali che implicano "p" e "q".

(4) L'espressione "effetti perversi" è usata dal sociologo Raymond Boudon e non indica necessariamente effetti negativi, per Boudon anche l'operare della "mano invisibile" è un effetto perverso (Boudon [1981], pp. 11 - 20, 177, 189)

(5) Nell'esposizione si prescinde dai problemi epistemologici legati al controllo di ipotesi probabilistiche e alla teoricità degli asserti osservativi. La considerazione di tali problemi mostra che la falsificazione, intesa come rifiuto di una teoria incompatibile con certi asserti osservativi, ha sempre un carattere convenzionale (si veda Lakatos [1979]). Dunque chi identifica tautologia e non falsificabilità si trova di fronte all'assurda conseguenza che ogni proposizione

empirica non logicamente falsa è tautologica, perchè non falsificabile in senso stretto.

(6) Latsis attribuisce a Von Mises la tesi di tipo kantiano secondo cui il principio di razionalità è un giudizio sintetico a priori del teorico sociale (Latsis [1976], p 4).

Una opinione, analoga a quella esposta nel testo, sulla non falsificabilità del postulato di razionalità è sostenuta da Caldwell (Caldwell [1982], p 55).

(7) Si può anche sostenere che un comportamento ottimizzante non implica un processo di ottimizzazione cosciente da parte dell'agente. L'ottimizzazione può essere intesa come un vero processo analitico di determinazione dell'azione migliore e questa potrebbe essere anche il risultato di un processo di successivi miglioramenti per esperienza. Probabilmente nessun campione di biliardo conosce completamente le leggi fisiche necessarie per calcolare come colpire la palla per ottenere gli effetti voluti; eppure è assai plausibile che un calcolo del genere fornisca soluzioni molto simili a quelle messe in atto dal giocatore. Allo stesso modo, nella meccanica, il moto può essere rappresentato analiticamente come minimizzazione dell'energia potenziale, senza presumere alcuna capacità di calcolo dei corpi materiali.

(8) Anche se gli agenti non adottano coscientemente tali programmi, esiste sempre un programma coerente con una qualsiasi successione di azioni di un certo agente e questo è sufficiente per ciò che segue nel testo.

(9) E' plausibile pensare che le preferenze relative ad eventi "osservabili" di uno spazio X, ad esempio quelle riferite a panieri di beni, siano parzialmente endogene. Tuttavia ciò può essere spiegato in base al fatto che l'agente ha delle preferenze "fondamentali" (ed esogene) relative ad eventi di un certo spazio Z e che le preferenze su X siano dedotte da quelle su Z in base a una teoria che, per ogni informazione h, assegna a ogni evento di X un evento di Z (o più in generale una distribuzione di probabilità su Z). Si può allora ipotizzare che le preferenze sono costanti, identificando correttamente lo spazio delle conseguenze rilevanti Z e inserendo la teoria suddetta nella più ampia connessione tra azioni e conseguenze ipotizzata dall'agente (cfr O'Driscoll, Rizzo [1985] p 45).

(10) In mancanza di ulteriori precisazioni per "eventi logicamente possibili" si intendono quegli eventi che corrispondono a proposizioni fattuali non logicamente false (ad esempio "p o \neg p" è logicamente falsa e quindi corrisponde a un evento logicamente impossibile). Nella teoria formale dei capitoli 3 e 4 gli eventi logicamente possibili sono invece specificati dal modello.