

# Giochi, teoria dei

di Pierpaolo Battigalli

SOMMARIO: **1.** Introduzione: a) *campo predicativo della teoria*; b) *cenni storici*; c) *nota sui riferimenti bibliografici*. □ **2.** Gli elementi di base della teoria: a) *forma estesa e forma strategica*; b) *conoscenza comune e informazione completa*; c) *giochi noncooperativi e cooperativi*; d) *teoria delle decisioni e ragionamento strategico*; e) *equilibrio, concetti di soluzione*. □ **3.** Il paradigma dominante: a) *la “vulgata cantabrigense”*; b) *raffinamenti*; c) *riflessioni critiche*. □ **4.** Gli sviluppi più recenti: a) *approccio epistemico*; b) *approccio adattivo e razionalità limitata*; c) *studi empirici ed esperimenti*. □ **5.** Conclusioni. □ Bibliografia.

## 1. Introduzione

### a) *Campo predicativo della teoria dei giochi*

La *teoria dei giochi* (TG) è una disciplina che analizza in modo formale le *situazioni decisionali interattive*, cioè situazioni nelle quali i guadagni (economici o di altro tipo) di ogni individuo dipendono dalle azioni altrui, oltre che dalle proprie. I giochi propriamente detti, come gli Scacchi, il Poker, o il Bridge costituiscono esempi paradigmatici di situazioni decisionali interattive. Ma la precedente descrizione generale si adatta anche a una miriade di situazioni economiche, politiche e sociali: concorrenza e collusione tra imprese, organizzazione d'impresa, contrattazione, aste, gare d'appalto, interazione tra scelte di politica economica di un governo e scelte di consumatori, lavoratori e imprese, determinazione delle politiche economiche (monetarie, fiscali, commerciali) da parte di più stati sovrani, scelte di piattaforme elettorali, votazioni in comitati e assemblee, scelte di politica internazionale, conflitti militari, attività criminose organizzate, insorgere di convenzioni sociali. Di seguito indicherò col termine “gioco” qualunque situazione decisionale interattiva e con “giocatori” gli individui che vi partecipano.

Si chiama *strategia* il modo in cui un giocatore reagirebbe alle diverse circostanze che si possono verificare durante l'interazione. Un giocatore che si sforza di fare scelte razionali formula congetture sulle strategie altrui per poter prevedere le conseguenze delle proprie scelte e quindi determinare il piano d'azione migliore. In generale, queste congetture non sono arbitrarie; si può cercare di dedurle da ipotesi sulla razionalità e la conoscenza altrui, oppure di ricavarle mediante generalizzazione induttiva da altre esperienze in situazioni simili. La derivazione di congetture sulle strategie altrui per determinare il miglior piano di gioco è detta *ragionamento strategico*.

L'analisi del ragionamento strategico permette, almeno nei casi più favorevoli, di fare previsioni sull'esito delle interazioni sociali. E' quindi possibile utilizzare la TG per studiare problemi di progettazione istituzionale, cioè problemi in cui un cosiddetto “pianificatore sociale” deve scegliere le “regole del gioco” a cui saranno soggetti gli individui di una certa comunità di riferimento in modo da perseguire i propri obiettivi (il pianificatore può

essere un rappresentante della stessa o di altra comunità). Esempi di progettazione istituzionale (“mechanism design”) sono la scelta di regole costituzionali, dello statuto di una associazione, delle regole di un’asta, o delle regole di un mercato borsistico.

Come traspare da queste prime anticipazioni, la TG offre una metodologia di analisi che attraversa i tradizionali confini disciplinari nell’ambito delle scienze umane. Non sorprende quindi che la TG sia utilizzata da studiosi di filosofia, sociologia, storia e, soprattutto, di scienze politiche (v. Morrow, 1994). Il principale campo di applicazione resta tuttavia la scienza economica, come prefigurato nel primo trattato sistematico sull’argomento: *Theory of Games and Economic Behavior* di John von Neumann e Oscar Morgenstern (di seguito abbreviato in *Theory of Games*).

#### b) *Cenni storici*

La TG è una disciplina giovane, che solo di recente ha acquisito una configurazione relativamente stabile e che comunque continua a svilupparsi e a cambiare. Lo sviluppo non è solo cumulativo, cambia anche la prospettiva sulla teoria stessa e sul significato dei suoi concetti formali. In questo paragrafo mi limiterò a menzionare alcune tappe che ritengo fondamentali dal punto di vista della teoria moderna e degli sviluppi correnti. Nel far ciò, renderò poca giustizia ad alcuni contributi molto importanti. Per un resoconto più completo degli sviluppi della TG fino all’inizio degli anni ’80 si veda l’ottima rassegna di Aumann (1987); Walker (1995) fornisce una cronologia più aggiornata.

Esempi di analisi di specifici giochi sono rintracciabili negli scritti di matematici ed economisti del XVIII e XIX secolo, ma la vera e propria nascita della TG si può far risalire al 1928, anno di pubblicazione del saggio *Zur Theorie der Gesellschaftsspiele* del matematico ungherese von Neumann, che presenta la prima trattazione generale e sistematica sull’argomento. L’articolo mostra come rappresentare matematicamente le regole di qualunque gioco, definisce formalmente il concetto di strategia e dimostra l’esistenza di una soluzione essenzialmente unica per tutti i giochi finiti *antagonistici*, cioè con due giocatori e somma delle vincite costante. Il contributo di von Neumann comincia ad avere un impatto con la pubblicazione di *Theory of Games*, scritto in collaborazione con l’economista austriaco Morgenstern, in cui la TG è presentata come strumento analitico fondamentale per la formulazione di modelli nelle scienze sociali e in particolare in economia (v. von Neumann e Morgenstern, 1947<sup>2</sup>). Il trattato distingue tra i giochi antagonistici e tutti gli altri giochi. Per questi ultimi si propone un’analisi di tipo *cooperativo*, in cui l’oggetto dell’indagine è la formazione di coalizioni che redistribuiscono tra i propri membri la massima vincita complessiva che questi sono in grado di ottenere agendo in modo coordinato. L’appendice del trattato contiene un’elegante teoria assiomatica della scelta in condizioni di rischio, che di per sé costituisce una pietra miliare della teoria economica.

*Theory of Games* riceve una discreta accoglienza tra i teorici economici, ma l’impostazione di von Neumann e Morgenstern rende difficile il “corteggiamento” tra TG e scienza economica. Il punto forte del trattato è infatti l’analisi dei giochi antagonistici, che sono tuttavia di limitato interesse per gli economisti. D’altra parte, l’analisi cooperativa proposta per i giochi non antagonistici trascura i problemi di incentivazione e comunicazione all’interno delle coalizioni; inoltre, il concetto di soluzione proposto dagli autori si rivela poco trattabile e non ha nulla in comune con quelli utilizzati dagli economisti matematici nell’analisi di specifici giochi economici.

La situazione cambia radicalmente con la pubblicazione di alcuni saggi del matematico americano John Nash all'inizio degli anni '50 (v. Nash, 1996). Nash propone un concetto di equilibrio non-cooperativo (che ancora oggi porta il suo nome) applicabile a *tutte* le situazioni di gioco: una combinazione di strategie  $s$  costituisce un *equilibrio* se la strategia di ogni giocatore massimizza la sua vincita (attesa) *date* le strategie degli altri giocatori in  $s$ . Il fatto di considerare *date* le scelte altrui è giustificato dalla stessa definizione di strategia come piano d'azione contingente, scelto prima che cominci la "partita". L'equilibrio strategico di Nash è un concetto unificante, che comprende come casi particolari la soluzione di maximin proposta da von Neuman per i giochi antagonistici e alcuni concetti di equilibrio oligopolistico già analizzati nella teoria economica, in particolare da Augustin Cournot (1838). Nash inoltre propone un argomento secondo il quale la formazione di coalizioni e stipulazione di accordi possono essere analizzati con l'approccio non-cooperativo e quindi con il concetto di equilibrio strategico (v. cap. 2, § c). Un primo esempio lo fornisce egli stesso, formulando una elegante teoria assiomatica per risolvere problemi di contrattazione bilaterale (il suo contributo più brillante, anche se forse non il più importante) e poi derivando la stessa soluzione come equilibrio strategico. Il *programma di Nash*, secondo cui tutta la TG è riconducibile ad un approccio non-cooperativo basato sull'equilibrio strategico, riceverà un nuovo impulso trent'anni più tardi (v. Osborne e Rubinstein, 1994, capp. 7 e 15) ed è oggi dominante, soprattutto nelle applicazioni economiche.

Nonostante il contributo di Nash, è l'approccio cooperativo quello che domina la TG negli anni '50 e '60, facendola crescere in modo impetuoso. Vengono introdotti, analizzati e applicati concetti di soluzione diversi da quello proposto da von Neumann e Morgenstern (v. Owen, 1995<sup>3</sup>, capp.10, 13 e 15). Vengono inoltre proposti degli indici di potere derivati da pochi assiomi apparentemente ragionevoli, in modo analogo alla teoria della contrattazione bilaterale di Nash (v. Owen, 1995<sup>3</sup>, cap. 12). I risultati forse più notevoli riguardano le relazioni tra questi concetti e la teoria dell'equilibrio economico concorrenziale, ottenibile come soluzione limite quando ogni singolo agente economico è, in un senso matematicamente ben definito, "trascurabile" rispetto alla dimensione del mercato (v. Osborne e Rubinstein, cap. 13). La teoria dei giochi cooperativi è oggi un po' trascurata (come testimonia lo spazio ristretto o nullo dedicato nei più recenti libri di testo, per esempio, Fudenberg e Tirole, 1991, Gibbons 1992), ma rimane comunque una componente importante della TG, anche per le sue numerose applicazioni all'economia e alle scienze politiche (v. Owen, 1995<sup>3</sup>).

Affinché l'approccio non-cooperativo possa definitivamente affermarsi è necessario superare almeno due ostacoli: (1) i ragionamenti intuitivi spesso utilizzati per motivare il concetto di equilibrio sembrano presupporre che tutti i giocatori siano consapevoli delle regole del gioco e delle altrui preferenze sui possibili esiti (v. cap. 2, § b), un'ipotesi questa che appare troppo forte per molte potenziali applicazioni, (2) il concetto di equilibrio strategico non tiene adeguatamente conto degli aspetti dinamici del gioco ed è quindi compatibile con "minacce e promesse non credibili". Questi problemi vengono (parzialmente) risolti da due fondamentali contributi di John Harsanyi e Reinard Selten. Il primo mostra come rappresentare formalmente i possibili stati di conoscenza soggettiva dei giocatori (includere le loro conoscenze e credenze sulle conoscenze e credenze altrui) e introduce un concetto di equilibrio, detto *bayesiano*, che modifica l'equilibrio strategico per tener conto della incompleta conoscenza delle regole del gioco e delle preferenze altrui (v. Harsanyi, 1967-68). Il secondo affronta il problema della credibilità proponendo un *raffinamento* del concetto di equilibrio, detto *equilibrio perfetto*, secondo il quale le strategie devono prescrivere un

comportamento “razionale” per tutte le circostanze di gioco, comprese quelle che secondo l’equilibrio non si dovrebbero verificare (v. Selten, 1975). Negli anni ’70 aumentano in modo esponenziale le applicazioni della nuova TG alla teoria della concorrenza imperfetta e dell’organizzazione industriale (v. Tirole, 1988). Inoltre il concetto di equilibrio strategico (bayesiano e perfetto) permette di organizzare in uno schema di riferimento teorico coerente e, allo stesso tempo, di sviluppare ed estendere i pionieristici lavori sulle aste e i mercati con informazione asimmetrica apparsi negli anni ’60 e ’70. Oramai una buona parte dei recenti libri di testo di economia teorica (soprattutto quelli avanzati) è dedicata alla TG e alle sue applicazioni ai modelli di mercato con concorrenza imperfetta e informazione asimmetrica (v. Mas Colell *et al.*, 1994). Si può dire che questo moderno corpus teorico nasca dalla sintesi dei contributi di Nash, Harsanyi e Selten, la cui importanza è stata recentemente riconosciuta con l’assegnazione a questi tre studiosi, nel 1994, del premio della Banca di Svezia per le scienze economiche in memoria di Alfred Nobel.

Negli anni ’80 si assiste a uno sviluppo un po’ caotico della teoria. Le applicazioni dell’equilibrio strategico bayesiano e perfetto portano al centro dell’attenzione le credenze dei giocatori sulle informazioni e le strategie dei loro avversari. Emerge la consapevolezza che l’analisi di equilibrio incorpora, in modo implicito e spesso poco trasparente, ipotesi su tali credenze e sul modo in cui possono essere modificate durante lo svolgimento del gioco. Da un lato, si ritiene desiderabile aggiungere ulteriori ipotesi ritenute “ragionevoli” o “intuitive”, ottenendo così svariati “raffinamenti” che permettono di operare una selezione tra i molteplici esiti di equilibrio dei modelli applicati (v. cap. 3, § b); dall’altro, le ipotesi sulle credenze vengono messe in questione e ciò porta però a proporre dei concetti di soluzione più deboli dell’equilibrio strategico (bayesiano, perfetto). La risultante pleora di soluzioni provoca un comprensibile sconcerto tra i fruitori della TG e finisce per scalfirne il prestigio da poco acquisito.

Gli sviluppi teorici degli ultimi quindici anni sono, almeno in parte, una reazione a questo stato di cose. Il linguaggio formale della TG viene arricchito e reso più espressivo per poter analizzare rigorosamente le conoscenze e le aspettative reciproche dei giocatori. Ciò permette di valutare i diversi concetti di soluzione esaminando le ipotesi epistemiche da cui possono essere derivati (v. cap. 4, § a). Questo tipo di analisi è particolarmente adatto allo studio di una interazione ideale tra individui altamente sofisticati che però non hanno modo di imparare da precedenti esperienze in situazioni simili. Tuttavia la nuova TG è anche molto interessata all’interazione tra individui limitatamente razionali che procedono per tentativi ed errori. Si considerano allora dinamiche di tipo adattivo in situazioni d’interazione strategica “tipiche” e ripetitive tra individui non (necessariamente) sofisticati e informati (v. cap. 4, § b). Gli equilibri corrispondono a stati stazionari di tali dinamiche e vengono valutati in base alle loro proprietà di stabilità. Questo secondo approccio è in parte collegato a sorprendenti applicazioni della TG alla biologia evolutivistica (v. Maynard Smith, 1982). Le predizioni di molti modelli di TG vengono inoltre sempre più confrontate con osservazioni ottenute in esperimenti di laboratorio (v. cap. 4, § c). E’ presumibile che i risultati di questi esperimenti stimoleranno nuovi interessanti sviluppi della teoria.

#### c) *Nota sui riferimenti bibliografici*

La letteratura sulla TG è ormai vastissima. Oltre a migliaia di articoli, sono stati pubblicati molti libri di testo e monografie, soprattutto negli ultimi quindici anni, dopo che

nella disciplina si è formato qualcosa di simile a un paradigma dominante. Piuttosto che menzionare i contributi originali in cui sono apparsi nuovi concetti e risultati, è più utile in questa sede fornire riferimenti ad alcuni testi e rassegne. Fanno eccezione solo le opere di alcuni “fondatori” menzionati nel precedente paragrafo. Tra i libri di testo menzionati in bibliografia, quello di Dixit e Skeath (1999) è accessibile anche a chi non ha dimestichezza con la matematica, quelli di Binmore (1992) e Gibbons (1992) sono di difficoltà intermedia, quelli di Myerson (1991), Fudenberg e Tirole (1991), Osborne e Rubinstein (1994) e Owen (1995<sup>3</sup>) sono avanzati. I testi di Fudenberg, Tirole e Gibbons sono più focalizzati sulle applicazioni economiche, quello di Owen dedica molto spazio ai giochi cooperativi.

## 2. Elementi di base della teoria

### a) *Forma estesa e forma strategica*

Il punto di partenza della TG consiste nel descrivere i giochi utilizzando concetti matematici di natura insiemistica (insiemi, funzioni, relazioni d'ordine) e, secondariamente, di natura numerico-quantitativa (vincite, probabilità, utilità). Ecco un semplice esempio, a cui farò ripetutamente riferimento in seguito. Sia  $I$  (Primo) il proprietario di un certo oggetto e  $II$  (Seconda) una potenziale acquirente. Semplificando al massimo, si consideri la seguente situazione di contrattazione: Primo può chiedere un Euro, oppure per due Euro in cambio dell'oggetto, Seconda può solo accettare ( $a$ ) o rifiutare ( $r$ ). Indicando con  $V_i$  il valore monetario dell'oggetto per il giocatore  $i$  ( $i = I, II$ ), si può rappresentare il gioco con il seguente diagramma ad albero:

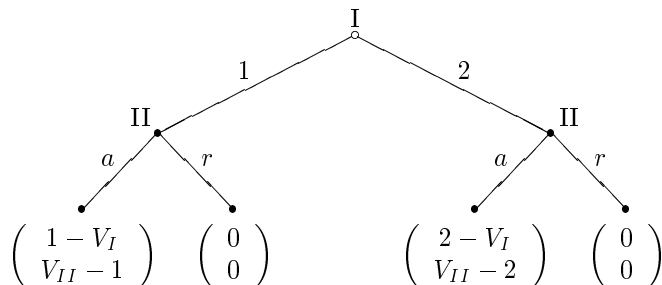


Figura 1. – Diagramma ad albero di un mini-gioco di contrattazione.

In generale la rappresentazione in *forma estesa* di un gioco permette di catturare tutti gli aspetti essenziali delle sue regole specificando i seguenti elementi: (i) un insieme di giocatori (ii) un insieme di sequenze di azioni, dette *storie*, permesse dalle regole del gioco, ripartibile in storie *parziali* (per esempio un'offerta) e storie complete, o *terminali* (per esempio  $(2, a)$ , cioè un'offerta seguita da una risposta), (iii) una regola che, per ogni storia parziale, determina il giocatore che farà la mossa successiva e l'insieme delle azioni per lui possibili (che può dipendere dalla storia precedente), (iv) una funzione che per ogni storia terminale determina i guadagni (*payoff*) dei giocatori. L'insieme delle storie compatibili con le regole

del gioco ha una struttura ad albero, come l'esempio in Fig. 1 (formalmente, l'immediato predecessore di una sequenza  $(a^1, \dots, a^K, a^{K+1})$  è  $(a^1, \dots, a^K)$ ; la radice dell'albero è la "successione vuota" e corrisponde all'inizio del gioco). Questa descrizione matematica può essere arricchita per permettere di rappresentare *mosse simultanee* (come nel gioco "Sasso, Forbice, Carta"), *mosse casuali* (lanci di dadi, distribuzioni di carte), *imperfetta informazione* su mosse precedenti (le carte scartate da un avversario) ed esiti di gioco non numerici che ogni giocatore valuta secondo una *funzione di utilità* personale, che ne esprime le preferenze. Nel caso sia importante valutare esiti incerti, la funzione di utilità può essere specificata in modo tale da rappresentare anche le preferenze su questi ultimi, purché tali preferenze soddisfino certi assiomi apparentemente ragionevoli (v. von Neumann e Morgenstern, 1947<sup>2</sup>, Appendix). In definitiva, per ogni storia terminale  $z$  e ogni giocatore  $i$ , risulta assegnata un'utilità  $u_i(z)$ . Un gioco si dice *antagonistico* se ci sono due giocatori e la somma delle loro utilità è costante, cioè  $u_I(z) + u_{II}(z) = K$  per ogni  $z$  (il gioco in Fig. 1 è antagonistico se e solo se l'oggetto ha lo stesso valore per entrambi i giocatori).

Secondo von Neumann, è sufficiente analizzare una rappresentazione più parsimoniosa del gioco, detta forma strategica, o forma normale. Si è già accennato al concetto di strategia (v. cap. 1, § a), che può ora essere definito più formalmente: una *strategia* per il giocatore  $i$  è una funzione (interpretabile come un piano di gioco contingente) che specifica una delle azioni permesse a  $i$  per ogni storia parziale immediatamente seguita da una mossa di  $i$  (se il giocatore  $i$  non è perfettamente informato delle mosse precedenti, la scelta prescritta può dipendere solo da ciò che questi osserva). Invece di scegliere azioni quando è il suo turno, un giocatore potrebbe scegliere una strategia prima che cominci il gioco e farla eseguire meccanicamente da un agente fidato, o da un programma. Un *profilo di strategie* (una per ogni giocatore)  $s = (s_I, s_{II}, \dots)$  determina una particolare storia terminale, che si indicherà con  $\zeta(s)$ , e quindi un profilo di utilità  $(u_I(\zeta(s)), u_{II}(\zeta(s)), \dots)$ . La *forma strategica* del gioco specifica, per ogni giocatore  $i$ , l'insieme delle strategie  $S_i$  e l'applicazione  $U_i(s) = u_i(\zeta(s))$  che assegna a ogni profilo di strategie la corrispondente utilità di  $i$ . Se i giocatori sono due si può rappresentare la forma strategica con una matrice di coppie di numeri. Per esempio, nel mini-gioco di contrattazione della Fig. 1, il profilo strategico in cui Primo offre l'oggetto per due Euro e Seconda accetta un'offerta se e solo se il prezzo è basso induce la storia terminale  $(2, r)$  e quindi il profilo di utilità  $(0, 0)$ . La forma strategica del gioco è data dalla matrice nella Tabella 1. Le strategie di Primo coincidono con le sue azioni perché Primo sceglie solo all'inizio del gioco. Seconda invece dispone di due azioni possibili (accettare o rifiutare) per ognuna delle due situazioni di scelta (prezzo alto o prezzo basso) e quindi le strategie, cioè i modi di assegnare una azione per ogni situazione di scelta, sono  $2 \times 2 = 4$ . Nella matrice,  $x.y$  rappresenta la strategia di  $II$  che risponde con  $x$  al prezzo 1 e con  $y$  al prezzo 2 (per esempio,  $a.r$  prescrive di accettare se e solo se il prezzo è basso):

	$a.a$	$a.r$	$r.a$	$r.r$
1	$1 - V_I, V_{II} - 1$	$1 - V_I, V_{II} - 1$	0, 0	0, 0
2	$2 - V_I, V_{II} - 2$	0, 0	$2 - V_I, V_{II} - 2$	0, 0

Tabella 1 – Forma strategica del gioco in Fig. 1.

### b) *Conoscenza comune e informazione completa*

La descrizione formale di un gioco non è sufficiente per formulare una teoria sensata su come il gioco viene (o dovrebbe essere) giocato. E' infatti necessario fare delle ipotesi su ciò che i giocatori sanno delle regole del gioco e delle preferenze altrui. Si consideri il gioco in Fig. 1. Assumendo che Seconda conosca  $V_{II}$  (ipotesi non necessariamente scontata), la sua strategia razionale è quella di rifiutare il prezzo offerto se è maggiore di  $V_{II}$  e accettare se è minore. La strategia *r.a.* (accettare solo se il prezzo è alto) è chiaramente irrazionale, ma quale delle altre tre sia razionale dipende dal valore di  $V_{II}$ . Se Primo conosce  $V_{II}$  (almeno approssimativamente) e *sa che anche Seconda conosce  $V_{II}$* , allora egli è in grado di dedurre dall'ipotesi di razionalità le reazioni di Seconda, altrimenti no. Dalla conoscenza di  $V_I$  Primo potrà infine dedurre l'offerta migliore da fare.

Da questo esempio risulta evidente che è necessario fare ipotesi sulle conoscenze *interattive* dei giocatori, ovvero ciò che essi sanno con riguardo al gioco e anche ciò che essi sanno delle conoscenze altrui. Si dice che un evento  $E$  (per esempio, " $V_{II} = 3$ ") è di *conoscenza comune* tra gli individui di una data comunità di riferimento se tutti sanno che  $E$ , tutti sanno che tutti sanno che  $E$ , tutti sanno che ... tutti sanno che  $E$ , ecc. (v. Lewis, 1969). Si dice che in un gioco vi è *informazione completa* se le regole del gioco e le preferenze dei giocatori sono di comune conoscenza, altrimenti l'informazione è incompleta. Per esempio, vi è informazione incompleta se Primo non conosce  $V_{II}$ , oppure se  $V_I$  e  $V_{II}$  sono noti a entrambi i giocatori, ma Primo non sa se Seconda conosce  $V_{II}$ . Anche le ipotesi sulle conoscenze interattive possono essere espresse in modo formale (v. Binmore, 1992, cap. 10). L'ipotesi di informazione completa (da non confondersi con la perfetta informazione sulle mosse passate) è una idealizzazione adeguata per l'analisi dei giochi propriamente detti e di molte interessanti situazioni d'interazione economica, politica e sociale.

### c) *Giochi non-cooperativi e cooperativi*

La distinzione tra giochi non-cooperativi e cooperativi, così come è intesa oggi, si riferisce alla possibilità di stipulare, prima della partita, accordi *vincolanti* sulle strategie da adottare. Se qualunque accordo può essere reso vincolante, allora si dice che il gioco è *cooperativo*. Se non è possibile stipulare accordi vincolanti, si dice che il gioco è *non-cooperativo*. La TG moderna è focalizzata sui giochi non-cooperativi per diversi motivi. In primo luogo, non è detto che sia possibile comunicare prima del gioco. In secondo luogo, anche se la comunicazione è possibile, non è detto che sia possibile stipulare accordi vincolanti, fatti rispettare da un agente esterno (ad esempio, non sono vincolanti gli accordi collusivi tra imprese, gli accordi tra "boss mafiosi" per spartirsi un territorio, o gli accordi tra stati sovrani). Infine, se gli accordi vincolanti sono possibili, allora – seguendo il programma di Nash – si può analizzare un meta-gioco non-cooperativo le cui regole sono date dal protocollo di contrattazione e i cui esiti sono gli accordi stessi. Si può quindi sostenere che l'analisi non-cooperativa, su cui si sofferma questo articolo, è più fondamentale di quella cooperativa. Va però sottolineato che per molte applicazioni risulta arbitrario, o troppo complicato, specificare un preciso protocollo di comunicazione e contrattazione; in tal caso l'analisi di tipo cooperativo può essere illuminante (v. Nash, 1996, Intr.)

d) *Teoria delle decisioni e ragionamento strategico*

La teoria delle decisioni (TD) può essere vista come la teoria dei giochi con una persona (giochi di pura abilità, giochi contro il caso), oppure come una teoria a sé stante, i cui risultati sono utilizzabili nella TG. Il secondo approccio è oggi prevalente. La TD determina quali strategie siano razionali, cioè coerenti con le preferenze del giocatore rispetto a esiti certi e incerti, e con le sue credenze sulle variabili che egli non controlla, inclusive del comportamento altrui. Le credenze sul comportamento altrui, pur avendo natura soggettiva, sono in tutto o in parte derivabili dal ragionamento strategico.

Il problema decisionale di un giocatore  $i$  ammette la seguente rappresentazione generale: scegliere un piano d'azione (strategia) tenendo conto che le conseguenze della scelta dipendono anche da una combinazione di variabili non note e non controllabili o influenzabili da  $i$ , incluse le strategie degli altri giocatori. In base ai postulati "standard" della TD, è possibile riformulare la questione come segue: il giocatore  $i$  è *razionale* se sceglie una strategia che *massimizza l'utilità attesa* di  $i$ , calcolata assegnando probabilità *soggettive* ai possibili valori delle variabili non controllate da  $i$ , date le informazioni di cui  $i$  dispone (v. Myerson, 1991, cap. 1).

La differenza fondamentale tra TD e TG è che la TG cerca di spiegare, almeno in parte, le credenze soggettive usando il ragionamento strategico. Si consideri per esempio il gioco della Fig. 1, assumendo informazione completa e identificando il guadagno di un giocatore con la sua utilità. Il problema decisionale di Primo è rappresentato dalla Tabella 1. Supponiamo che Primo creda che Seconda sia razionale. Se  $V_{II} > 2$ , allora Primo è sicuro che Seconda accetterebbe entrambi i prezzi (cioè assegna il 100% delle probabilità alla strategia *a.a*). Se  $1 < V_{II} < 2$ , allora Primo si aspetta che Seconda accetti solo il prezzo più basso. Riassumendo: nella TG le credenze soggettive sono *endogene*.

Una conseguenza importante di questa osservazione è che alcuni risultati comparativi della TD non sono applicabili quando il problema decisionale è interattivo. Per esempio, un risultato apparentemente ovvio della TD è che la situazione del decisore peggiora (o comunque non migliora) se egli viene privato di una o più possibilità di scelta. Invece *in un contesto interattivo ridurre le proprie opzioni può avere un valore strategico*. Si consideri il gioco in Fig. 1, con informazione completa,  $V_I < 1$  e  $V_{II} > 2$ : si può verificare che Seconda risulterebbe avvantaggiata se potesse privarsi della possibilità di accettare il prezzo più alto (purché Primo venga a conoscenza di questa limitazione), perché nel gioco modificato per Primo sarebbe ottimale richiedere un prezzo basso, mentre nel gioco originale Primo ritiene ottimale richiedere un prezzo alto.

e) *Equilibrio, concetti di soluzione.*

La TG "standard" spiega il comportamento dei giocatori in due passi: (1) formulazione matematica delle regole del gioco e delle preferenze dei giocatori su esiti certi e incerti (descrizione del gioco), (2) applicazione al gioco in esame di un concetto di equilibrio, o soluzione. Formalmente, un *concetto di soluzione* è una corrispondenza  $\mathcal{S}$  che associa a ogni gioco  $G$  un suo sottoinsieme  $\mathcal{S}(G)$  di profili strategici che soddisfano certe proprietà "desiderabili". I concetti di soluzione più utilizzati nelle applicazioni sono delle varianti dell'equilibrio di Nash. Si ricordi che  $U_i(s)$  indica l'utilità del giocatore  $i$  in funzione del profilo strategico adottato  $s = (s_I, s_{II}, \dots)$  (v. § a).



**Definizione.** Un profilo  $s^* = (s_I^*, s_{II}^*, \dots)$  è un *equilibrio strategico*, o di Nash, se

$$U_i(s^*) \geq U_i(s_i, s_{-i}^*) \text{ per ogni } i \text{ e ogni } s_i \text{ in } S_i,$$

dove  $(s_i, s_{-i}^*)$  indica il profilo ottenuto da  $s^*$  sostituendo  $s_i^*$  con  $s_i$ .

Si dimostra che nei giochi antagonistici, in cui  $U_{II}(s) = K - U_I(s)$ , l'equilibrio strategico equivale alla soluzione di maxmin di von Neumann e Morgenstern. Ne discende che nei giochi antagonistici la molteplicità degli equilibri è irrilevante: se vi sono più equilibri, essi danno la stessa utilità ai giocatori (quella di maxmin) e ogni coppia di strategie di equilibrio è essa stessa un equilibrio (v. Osborne e Rubinstein, 1994, cap. 2, § 5).

Queste proprietà non valgono per i giochi non antagonistici, come mostra il gioco della Tabella 1: se  $V_I < 1$  e  $V_{II} > 2$  ci sono due equilibri  $(2, a.a)$  e  $(1, a.r)$ , il primo equilibrio è più favorevole a  $I$ , il secondo a  $II$ ; inoltre né  $(1, a.a)$  né  $(2, a.r)$  sono equilibri (un profilo di strategie  $(s_I, s_{II})$  è un equilibrio se muovendosi in verticale dalla corrispondente casella è impossibile aumentare l'utilità di  $I$  e muovendosi in orizzontale è impossibile aumentare l'utilità di  $II$ ). Si vede dunque che un gioco può avere una molteplicità di equilibri non equivalenti.

Esistono giochi che non hanno alcun equilibrio strategico. Tuttavia, se si espandono le possibilità di scelta in un gioco finito permettendo che ogni giocatore selezioni una qualunque distribuzione di probabilità sulle proprie strategie (associando a profilo di distribuzioni le corrispondenti utilità attese), *il gioco così ottenuto possiede sempre almeno un equilibrio* (v. Nash, 1996, cap. 2). Sul significato di queste scelte probabilistiche, dette *strategie miste*, tornerò in seguito (v. cap. 4, § b).

I concetti di soluzione sono tipicamente motivati con argomenti informali riguardanti il ragionamento strategico. Per l'equilibrio di Nash si forniscono due argomenti indipendenti: (1) se esiste un "modo ovvio di giocare" (per esempio, se i giocatori, essendo completamente informati, razionali e intelligenti, ragionano tutti allo stesso modo e giungono ad una conclusione *univoca* su come si deve giocare), allora deve trattarsi di un equilibrio, altrimenti il "modo ovvio di giocare" violerebbe l'ipotesi di razionalità di almeno un giocatore; (2) se i giocatori cercano di accordarsi prima del gioco, ma non possono stipulare accordi vincolanti (premessa dell'analisi non-cooperativa), allora una condizione necessaria affinché l'accordo di giocare un profilo strategico  $s^*$  venga rispettato è che  $s^*$  sia un equilibrio. Altri concetti di soluzione verranno discussi nel seguito (v. capp. 3 e 4).

### 3. Il Paradigma dominante

#### a) La "vulgata cantabrigense"

Il paradigma oggi dominante nella TG è stato formulato alla fine degli anni '80 nel tentativo di sintetizzare i contributi di Nash, Harsanyi e Selten e organizzare in un corpus teorico coerente e facilmente accessibile le molte varianti utilizzate in specifiche applicazioni. In questa impresa si sono particolarmente distinti alcuni economisti che in quel periodo insegnavano al Massachusetts Institute of Technology, producendo quella che si potrebbe definire una "vulgata cantabrigense" (v. Gibbons, 1992, Fudenberg e Tirole, 1991).

La "vulgata" si sofferma principalmente sui giochi con *azioni osservabili* (detti anche giochi con *informazione quasi-perfetta*), cioè i giochi con uno o più stadi, in cui dal secondo

stadio in poi si osservano le mosse di tutti gli stadi precedenti (si noti che la durata effettiva del gioco può dipendere dal comportamento dei giocatori). Tra questi giochi si distinguono quattro categorie, ad ognuna delle quali è associata una variante di base del concetto di equilibrio strategico di Nash, secondo la seguente tabella riassuntiva:

Giochi	con informazione completa	con informazione incompleta
statici (uno stadio)	equilibrio di Nash	equilibrio bayesiano
dinamici (più stadi)	equilibrio perfetto	equilibrio bayesiano perfetto

Tabella 2. – Classificazione dei giochi con azioni osservabili ed equilibri.

Nei *giochi statici* i giocatori scelgono una sola volta e simultaneamente, quindi le strategie coincidono con le azioni possibili. I giochi statici con *informazione completa*, due giocatori e un numero finito di azioni possibili sono rappresentabili con “matrici dei guadagni” come le seguenti:

DP	$c$	$d$	SD	$s$	$d$	FC	$c$	$f$
$c$	3,3	0,4	$s$	1,1	0,0	$c$	2,2	1,4
$d$	4,0	1,1	$d$	0,0	$x, x$	$f$	4,1	0,0

Tabella 3. – Esempi di giochi statici (nella matrice SD si assume  $x > 0$ )

I giochi rappresentati nella Tabella 3 sono semplici esempi di “dilemmi sociali” che illustrano alcune applicazioni e interpretazioni del concetto di equilibrio di Nash. Per convenzione assumo che Primo scelga la riga e Seconda la colonna, come nella Tabella 1.

DP è il ben noto “Dilemma del Prigioniero”, in cui il comportamento razionale non-cooperativo è “defezionare” ( $d$ ) e quindi induce un esito inefficiente ( $c$  significa “cooperare”, cioè non confessare,  $d$  è la strategia dominante). Molte situazioni economiche hanno una struttura analoga. Per esempio, le imprese in un mercato oligopolistico preferirebbero contenere la produzione per mantenere un prezzo alto, ma ognuna di esse ha un forte incentivo a deviare unilateralmente da tale accordo collusivo aumentando la produzione (ciò mostra, tra l’altro, che l’esito più efficiente dal punto di vista dei giocatori può essere indesiderabile dal punto di vista di una comunità più ampia, in questo caso quella che comprende i compratori).

SD (“Sinistra o Destra”) è un gioco di puro coordinamento: i giocatori hanno un comune interesse a scegliere la stessa azione (per esempio, tenere la destra), nel qual caso si determina un equilibrio, interpretabile come una convenzione. Alcune convenzioni possono essere superiori ad altre (caso  $x \neq 1$ ), ma non necessariamente prevalgono. Fattori, anche casuali, non catturati dalla rappresentazione formale del gioco possono determinare un coordinamento più o meno efficiente.

FC (“Falco o Colomba”) è un gioco con equilibri asimmetrici: se i due giocatori si comportano “da falchi”, cioè in modo aggressivo, ne scaturisce un conflitto che danneggia entrambi, ma il comportamento aggressivo è vantaggioso contro un comportamento “da colomba”, cioè timido. Gli equilibri sono dunque  $(f, c)$  e  $(c, f)$ . Nei giochi SD e FC esiste anche un terzo equilibrio, in *strategie miste*, che verrà discusso nel seguito.

I giochi dinamici con informazione completa (e azioni osservabili), hanno la seguente caratteristica: la situazione immediatamente successiva ad ogni storia parziale  $h$  corrisponde a un gioco, detto *sottogioco* con radice  $h$ . La Fig. 1 rappresenta un gioco dinamico molto semplice con due sottogiochi propri (cioè non coincidenti col gioco originario), quello successivo alla richiesta di un Euro (sottogioco con radice  $h = (1)$ ) e quello successivo alla richiesta di due Euro (sottogioco con radice  $h = (2)$ ). Si tratta di giochi banali con un solo giocatore attivo (Seconda) che deve scegliere tra due alternative con conseguenze note. Si è visto che il gioco con  $V_I < 1$  e  $V_{II} > 2$  ammette due equilibri:  $(2, a.a)$  è l'equilibrio "ragionevole" in cui Seconda è disposta a comprare ad entrambi i prezzi, mentre  $(1, a.r)$  è basato sulla "minaccia non credibile" di rifiutare la richiesta più alta. Questo secondo profilo strategico richiederebbe un comportamento incompatibile con l'equilibrio nel sottogioco preceduto da tale richiesta, perché il valore dell'oggetto per Seconda è superiore anche al prezzo più alto.

Questo esempio sembra suggerire che le "minacce e promesse non credibili" possono essere eliminate eliminando gli equilibri che non soddisfano la seguente condizione di "perfezione":

**Definizione.** Un equilibrio  $s^*$  si dice *perfetto (nei sottogiochi)* se, per ogni storia parziale  $h$ ,  $s^*$  induce un equilibrio nel sottogioco con radice  $h$ .

In molti giochi dinamici con un numero finito di stadi esiste un unico equilibrio perfetto, che si può calcolare con un algoritmo di riduzione iterativa del gioco in esame.

*Procedura di induzione a ritroso:* Dato il gioco  $G$ , si considerano tutte le sue storie *quasi-terminali*, cioè le storie parziali a partire dalle quali il gioco si conclude con una sola ulteriore mossa; per ogni storia quasi-terminale  $h$ , si determina l'azione  $a(h)$  che massimizza l'utilità del giocatore attivo ad  $h$ ; l'esito che ne consegue viene direttamente associato alla storia  $h$ ; in questo modo si determina un gioco ridotto  $G'$ , le cui storie terminali coincidono con le storie quasi-terminali di  $G$ . Poi l'operazione si ripete con il gioco  $G'$ , ottenendo un gioco ulteriormente ridotto  $G''$  e così via, procedendo a ritroso fino alla radice dell'albero. La procedura assegna un'azione  $a(h)$  ad ogni storia parziale  $h$  di  $G$ , perciò risulta determinato un profilo strategico, cioè l'equilibrio perfetto di  $G$ . L'induzione a ritroso può essere facilmente applicata al mini-gioco di contrattazione della Fig. 1, ottenendo l'unico equilibrio intuitivamente ragionevole. Più in generale, l'induzione a ritroso consente di determinare un unico equilibrio, necessariamente perfetto, in "quasi tutti" i giochi con azioni osservabili senza mosse simultanee (detti anche giochi con *informazione perfetta*). La procedura può anche essere generalizzata per tenere conto di eventuali mosse simultanee, purché per ogni storia quasi-terminale (del gioco originale o di quelli ridotti) il corrispondente sottogioco abbia un unico equilibrio.

Naturalmente l'induzione a ritroso non è applicabile a giochi con durata (almeno potenzialmente) infinita, per i quali esistono tecniche di analisi differenti. Un esempio notevole è costituito dai *giochi infinitamente ripetuti*, cioè i giochi ottenuti dalla ripetizione infinita di un gioco statico, detto gioco *costituente*. Nei giochi infinitamente ripetuti l'utilità associata ad una storia completa (una sequenza infinita di profili di azioni) è una sorta di media delle utilità risultanti in ogni stadio. Eleganti teoremi matematici mostrano come la molteplicità degli equilibri perfetti sia una caratteristica intrinseca di questi giochi (v. Osborne e Rubinstein, 1994, cap. 8). Una implicazione interessante di questi risultati è che nelle relazioni di lunga durata è possibile sostenere un alto livello di cooperazione anche se non si possono

stipulare accordi vincolanti. Il “Dilemma del Prigioniero” infinitamente ripetuto fornisce un esempio paradigmatico. Se se i giocatori sono sufficientemente “pazienti” (cioè se, in termini formali, il loro tasso di sconto è piccolo), vi sono molti altri equilibri perfetti oltre a quello che prescrive di giocare sempre e comunque l’equilibrio statico  $(d, d)$ . In particolare, è un equilibrio perfetto il profilo strategico  $(s_I^*, s_{II}^*)$  così determinato: nel primo periodo i giocatori “cooperano”, cioè scelgono entrambi  $c$ , nei periodi successivi al primo ogni giocatore continua a cooperare se e solo se in passato non ci sono state deviazioni  $(d)$  dal profilo di azioni  $(c, c)$ , per cui una singola deviazione è sufficiente a distruggere per sempre il tacito accordo a cooperare. E’ proprio questo meccanismo che rende la cooperazione compatibile con gli incentivi dei giocatori. Se Primo è convinto che Seconda segua la strategia  $s_{II}^*$  e non è ancora avvenuta una deviazione da  $(c, c)$ , egli (essendo paziente) trova più vantaggioso scegliere  $c$  e quindi continuare a godere in futuro dei benefici della cooperazione, piuttosto che dissolverla per sempre con una deviazione che fornisce solo un beneficio immediato (un analogo ragionamento si applica a Seconda). Lo stesso tipo di intuizione è applicabile a una miriade di situazioni economiche, politiche e sociali in cui accordi taciti o espliciti, ma comunque non vincolanti, vengono rispettati nonostante sia possibile ottenere benefici immediati non rispettandoli. Per esempio, è possibile sostenere la collusione tra imprese in un mercato oligopolistico anche se gli accordi collusivi sono illegali (v. Tirole, 1988, cap. 6).

Si noti che, per la molteplicità degli equilibri nei giochi ripetuti, l’ipotesi che la durata dell’interazione non abbia un termine comunemente noto è cruciale. Per esempio, se è comunemente noto che il gioco DP sarà ripetuto  $T$  volte, si può applicare l’induzione a ritroso e verificare che l’unico equilibrio perfetto del gioco ripetuto è quello in cui non si coopera mai.

Nei *giochi con informazione incompleta* alcuni elementi del gioco non sono di comune conoscenza. Si presenta quindi il problema di rappresentare le conoscenze e credenze interattive dei giocatori in modo tale da ottenere qualcosa di analogo ad un gioco in forma estesa e poi definire un appropriato concetto di equilibrio. Il problema è concettualmente complesso e la TG lo affronta con un approccio molto sofisticato, basato sul fondamentale contributo di Harsanyi (1967-68). Qui mi limiterò a presentare una versione semplificata tipica della “vulgata”.

Supponiamo che ci siano due giocatori, Primo ( $I$ ) e Seconda ( $II$ ) e che le utilità corrispondenti alle storie terminali  $(z)$  dipendano da un profilo di parametri  $\theta = (\theta_I, \theta_{II})$ ; formalmente  $u_i = u_i(\theta_I, \theta_{II}, z)$ . L’informazione è incompleta perché almeno uno dei due giocatori *non* conosce il vero valore di  $\theta$ . Più precisamente, per ogni giocatore  $i$ , il vero valore di  $\theta_i$  appartiene ad un insieme di valori possibili  $\Theta_i = \{\alpha_i, \beta_i, \dots\}$ , i due insiemi  $\Theta_I$  e  $\Theta_{II}$  sono comunemente noti, e ogni  $i$  conosce  $\theta_i$ , ma sa soltanto che  $\theta_j$  (con  $j$  diverso da  $i$ ) appartiene all’insieme  $\Theta_j$ . Quindi *i non conosce*  $\theta_j$ , a meno che  $\Theta_j$  non contenga un unico elemento ( $\Theta_j$  contiene un unico elemento se e solo se tutto ciò che  $j$  sa sul gioco è noto anche a  $i$ ). In un certo senso, l’insieme  $\Theta_I$  rappresenta l’ignoranza di  $II$  su ciò che è noto a  $I$  e, simmetricamente,  $\Theta_{II}$  rappresenta l’ignoranza di  $I$  su ciò che è noto a  $II$ . Il parametro  $\theta_i$  è detto *tipo*, o *informazione privata*, del giocatore  $i$ . Nel gergo della TG spesso ci si riferisce ai diversi tipi come se fossero degli individui, usando espressioni come “il tipo  $\beta_i$  sceglie l’azione  $a_i$ ”. Adotterò, talvolta, questo genere di espressioni perché permette di semplificare il linguaggio. Ma va sempre tenuto presente che i tipi rappresentano soltanto stati di conoscenza dei giocatori.

Secondo il cosiddetto approccio “bayesiano” alla teoria delle decisioni, un individuo assegna una distribuzione *sogettiva* di probabilità a qualunque variabile non nota che sia (direttamente o indirettamente) rilevante per le sue scelte. (v. cap. 2, § d). Seguendo tale approccio, si assume che il giocatore  $i$ , all’inizio del gioco, assegni una distribuzione di probabilità  $\pi_i = (\pi_i(\alpha_j), \pi_i(\beta_j), \dots)$ , detta *credenza iniziale*, sui possibili tipi dell’altro giocatore,  $j$ . Se si ipotizza, per semplicità, che le distribuzioni soggettive  $\pi_I$  e  $\pi_{II}$  siano parametri comunemente noti, si ottiene un modello matematico, detto *gioco bayesiano*, che è sufficientemente completo per studiare il ragionamento strategico mediante l’analisi di equilibrio.

Per fissare le idee, è utile riferirsi a un esempio numerico rappresentato in forma estesa nella Fig. 2. L’unico giocatore informato è  $I$ , che può essere di tre tipi,  $\alpha_I$ ,  $\beta_I$ , o  $\gamma_I$ . Formalmente  $\Theta_I = \{\alpha_I, \beta_I, \gamma_I\}$ ,  $\Theta_{II} = \{\alpha_{II}\}$ . Poiché il tipo di  $II$ ,  $\alpha_{II}$ , è comunemente noto, non è necessario alcun riferimento esplicito ad esso.  $I$  fa la prima mossa e può terminare il gioco immediatamente ( $f$ =“fuori”) o far continuare  $II$  ( $d$ =“dentro”). Se  $I$  sceglie  $d$ ,  $II$  ha a sua disposizione tre azioni,  $a$ ,  $b$  e  $c$ , e guadagna un Euro se “indovina” il tipo di  $I$  (ad esempio, se sceglie  $a$  e nel caso in cui  $\theta_I = \alpha_I$ ). La figura mostra in dettaglio come i guadagni dei due giocatori dipendono da  $\theta_I$  e dalle azioni; per esempio  $u_i(\theta_I, f) = 1$  per ogni tipo  $\theta_I$  e ogni giocatore  $i$ ,  $u_I(\gamma_I, d, b) = u_I(\gamma_I, d, c) = 2$ ,  $u_{II}(\beta_I, d, b) = 1$ ,  $u_{II}(\beta_I, d, c) = 0$ . I numeri in parentesi accanto ai tipi rappresentano la credenza iniziale di  $II$  su  $I$  (per esempio,  $\pi_{II}(\alpha_I) = \frac{1}{2}$ ), che si assume comunemente nota. Per ogni tipo del giocatore  $I$  c’è un corrispondente albero di gioco. I nodi contrassegnati da  $II$  sono uniti da una linea tratteggiata per enfatizzare che  $II$  osserva l’azione  $d$  ma non il tipo  $\theta_I$ .

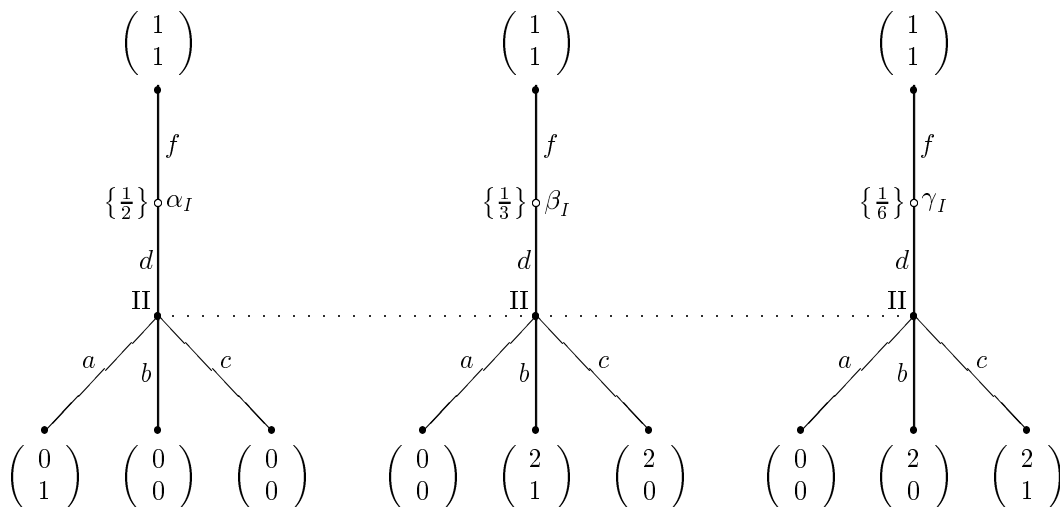


Figura 2. – Un gioco bayesiano.

Prima di passare all’analisi degli equilibri è opportuna una digressione. L’ipotesi di comune conoscenza delle credenze iniziali  $\pi_I$  e  $\pi_{II}$  è molto forte. Trattandosi di creden-

ze soggettive, è infatti più plausibile assumere che il giocatore  $i$  non conosca  $\pi_j$ . Seguendo l'approccio bayesiano, sarebbe allora necessario specificare una distribuzione soggettiva sulle credenze altrui, cioè delle credenze "secondo ordine", poi una distribuzione di "terzo ordine" sulle altrui credenze di "secondo ordine" e così via, modellando un regresso infinito di credenze sulle credenze altrui. Harsanyi (1967-68) ha però mostrato che è possibile eliminare l'ipotesi di comune conoscenza delle credenze iniziali senza dover modellare esplicitamente questo regresso infinito. La rappresentazione matematica introdotta da Harsanyi ha inoltre il pregio di essere strutturalmente simile a quella più semplice e meno generale che viene qui analizzata.

Nel contesto di un gioco bayesiano, è naturale aspettarsi che le scelte di un giocatore dipendano dal suo tipo. Per esempio, nel gioco della Fig. 2, l'azione  $d$  è dominata per il tipo  $\alpha_j$  ( $f$  dà al tipo  $\alpha_I$  un'utilità superiore a  $d$ , qualunque sia la risposta attesa di  $II$ ), ma non lo è per gli altri due tipi. Quindi, assumendo che  $I$  sia razionale, si può escludere che il tipo  $\alpha_I$  scelga  $d$ , ma non si può escludere che un diverso tipo scelga  $d$ . In generale, il ragionamento strategico consiste, per ogni giocatore  $i$ , nel cercare di mettersi nei panni dell'altro giocatore  $j$  e immaginare cosa farebbe  $j$  se fosse del tipo  $\alpha_j$ , cosa farebbe se fosse del tipo  $\beta_j$  e così via. Un equilibrio del gioco bayesiano è costituito da un profilo di congetture, o funzioni di scelta,  $(\sigma_I, \sigma_{II})$  che specificano, per ogni giocatore  $j$  e ogni tipo  $\theta_j$ , la strategia  $s_j = \sigma_j(\theta_j)$  che in base alle conclusioni del ragionamento strategico il tipo  $\theta_j$  dovrebbe adottare.

Per dare una definizione formale è possibile (anche se non strettamente necessario) ricorrere ad un "trucco" che permette di ricondurre l'analisi di equilibrio dei giochi bayesiani a quella dei giochi con informazione completa. Si consideri il seguente gioco fittizio: esiste uno stadio preliminare, detto *stadio ex ante*, in cui i giocatori ignorano il loro tipo, poi una mossa casuale sceglie il tipo  $\theta_I$  con probabilità  $\pi_{II}(\theta_I)$  e un'altra mossa casuale sceglie il tipo  $\theta_{II}$  con probabilità  $\pi_I(\theta_{II})$  (l'ordine di queste due mosse casuali è irrilevante) infine i giocatori  $I$  e  $II$  osservano (soltanto) il proprio tipo e interagiscono secondo le regole fissate dal gioco bayesiano. Si osservi che le funzioni di scelta  $\sigma_I$  e  $\sigma_{II}$  corrispondono alle strategie di questo gioco fittizio. Per esempio, nel gioco della Fig. 2 una funzione di scelta  $\sigma_I$  corrisponde a una strategia dell'associato gioco fittizio, in base alla quale  $I$  pianifica, allo stadio ex ante, di scegliere l'azione  $\sigma_I(\theta_I)$  se osserverà  $\theta_I$ . Si tenga però presente che la corrispondenza è puramente formale: lo stadio ex ante, in cui il "vero" tipo  $\theta_I$  è ancora indeterminato, è soltanto un utile artificio teorico. Ciò premesso, un equilibrio del gioco bayesiano può essere definito come segue:

**Definizione.** Sia dato un gioco bayesiano  $GB$ . Un profilo di funzioni di scelta  $(\sigma_I, \sigma_{II})$  è un *equilibrio bayesiano* di  $GB$  se corrisponde ad un equilibrio strategico (di Nash) del gioco fittizio associato a  $GB$ .

Il gioco della Fig. 2 ha due equilibri bayesiani: (1)  $\sigma_I^*(\alpha_I) = f$ ,  $\sigma_I^*(\beta_I) = \sigma_I^*(\gamma_I) = d$ ,  $\sigma_{II}^* = b$  e (2)  $\sigma_I^\circ(\alpha_I) = \sigma_I^\circ(\beta_I) = \sigma_I^\circ(\gamma_I) = f$ ,  $\sigma_{II}^\circ = a$ . Per verificarlo, basta calcolare le utilità attese corrispondenti a ogni profilo  $(\sigma_I, \sigma_{II})$  ottenendo così la forma strategica del gioco fittizio mostrata nella Tabella 4 (la matrice ha 8 righe perché ci sono  $8 = 2^3$  possibili modi di associare una delle 2 azioni  $d$  o  $f$  ad ognuno dei 3 tipi  $\alpha_I, \beta_I, \gamma_I$ ; le utilità attese sono calcolate utilizzando la distribuzione  $\pi_{II}(\theta_I)$ , per esempio:  $U_I(\sigma_I^*, b) = \frac{1}{2} \times 1 + \frac{1}{3} \times 2 + \frac{1}{6} \times 2 = \frac{3}{2}$ ).

	$a$	$b$	$c$
$d.d.d$	$0, \frac{1}{2}$	$1, \frac{1}{3}$	$1, \frac{1}{6}$
$d.f.d$	$\frac{1}{3}, \frac{5}{6}$	$\frac{2}{3}, \frac{1}{3}$	$\frac{2}{3}, \frac{1}{2}$
$d.d.f$	$\frac{1}{6}, \frac{2}{3}$	$\frac{5}{6}, \frac{1}{2}$	$\frac{5}{6}, \frac{1}{6}$
$d.f.f$	$\frac{1}{2}, 1$	$\frac{1}{2}, \frac{1}{2}$	$\frac{1}{2}, \frac{1}{2}$
$\sigma_I^* = f.d.d$	$\frac{1}{2}, \frac{1}{2}$	$\frac{3}{2}, \frac{5}{6}$	$\frac{3}{2}, \frac{2}{3}$
$f.d.f$	$\frac{2}{3}, \frac{2}{3}$	$\frac{4}{3}, 1$	$\frac{4}{3}, \frac{2}{3}$
$f.f.d$	$\frac{5}{6}, \frac{5}{6}$	$\frac{7}{6}, \frac{5}{6}$	$\frac{7}{6}, 1$
$\sigma_I^\circ = f.f.f$	$1, 1$	$1, 1$	$1, 1$

Tabella 4. – Matrice associata al gioco bayesiano della Fig. 2.

Per comprendere meglio la definizione di equilibrio bayesiano, si immagini che un ragionamento strategico trasparente, basato (anche) sull'ipotesi di razionalità, porti ad una conclusione univoca sulla scelta di ogni giocatore *per ogni* suo possibile *tipo*. Risultano allora individuate una congettura  $\sigma_I$  di  $II$  su  $I$  e una congettura  $\sigma_{II}$  di  $I$  su  $II$ . Poiché il ragionamento è trasparente, il giocatore  $j$  conclude altresì che  $\sigma_j$  è la congettura di  $i$  nei suoi confronti. Dunque  $j$  si aspetta che  $i$ , qualunque sia il suo tipo, adotti una strategia ottima rispetto alla congettura  $\sigma_j$ . In conclusione, deve valere la seguente condizione: *per ogni giocatore  $i$  e ogni possibile tipo  $\theta_i$ , la strategia (congetturata)  $\sigma_i(\theta_i)$  massimizza l'utilità attesa di  $i$ , dato  $\theta_i$  e data la congettura  $\sigma_j$  sul giocatore  $j$* , altrimenti vorrebbe dire che le conclusioni del ragionamento strategico violano l'ipotesi di razionalità di almeno un giocatore, per almeno un suo tipo possibile. Analizzando l'esempio, si può verificare che un profilo di funzioni di scelta ha questa proprietà se e solo se è un equilibrio strategico del gioco fittizio associato. Harsanyi ha mostrato che questo risultato è valido per ogni gioco bayesiano.

La nozione di equilibrio bayesiano è adeguata per l'analisi dei giochi statici e anche di alcuni interessanti giochi dinamici. Per esempio, è stata applicata con grande successo allo studio delle aste (v. Klemperer, 1999) e più in generale dei meccanismi di scambio (v. Myerson, cap. 6, § 4 e 5). Ma ci sono molti giochi dinamici in cui la nozione di equilibrio bayesiano, analogamente all'equilibrio strategico di Nash, non consente di eliminare "minacce o promesse non credibili". Per eliminare questi equilibri poco plausibili, la "vulgata" adotta una variante del concetto di equilibrio perfetto. Intuitivamente, si dovrebbe associare una specie di "sottogioco bayesiano"  $GB(h)$  ad ogni storia parziale  $h$  e richiedere che il profilo di equilibrio  $\sigma = (\sigma_I, \sigma_{II})$  determini un equilibrio bayesiano di  $GB(h)$ , per ogni  $h$ . Per avere un ben definito gioco bayesiano a partire dalla storia  $h$  è però necessario specificare le *credenze* di ogni giocatore  $i$  sul tipo altrui *condizionate* all'osservazione di  $h$ ,

indicate col simbolo  $\mu_i(\theta_j|h)$ . Le credenze condizionate dipendono tuttavia dalle congetture di equilibrio e quindi, contrariamente alle credenze iniziali sul tipo altrui, sono *endogene*.

Questo punto si può chiarire considerando l'equilibrio  $\sigma^*$  nel gioco della Fig. 2. Se la congettura di *II* è  $\sigma_I^*$  e *II* osserva l'azione  $d$ , allora *II* esclude il tipo  $\alpha_I$ , assegnandogli probabilità condizionata nulla e riscalando equi-proporzionalmente le probabilità degli altri due tipi in modo che sommino a uno. Poiché  $\beta_I$  è inizialmente considerato due volte più probabile di  $\gamma_I$ , il risultato è  $\mu_{II}^*(\alpha_I|h) = 0$ ,  $\mu_{II}^*(\beta_I|h) = \frac{2}{3}$ ,  $\mu_{II}^*(\gamma_I|h) = \frac{1}{3}$  (questo calcolo può essere visto come un'applicazione della cosiddetta "regola di Bayes"). Se invece si considera l'equilibrio  $\sigma^\circ$ , la regola non è applicabile: infatti non si può dedurre la credenza condizionata all'azione  $d$  dalla congettura  $\sigma_I^\circ$  perché  $\sigma_I^\circ$  esclude che  $d$  possa verificarsi. Si può però ipotizzare che la credenza condizionata  $\mu_{II}(\theta_I|d)$  sia determinata in base a considerazioni differenti, per esempio un qualche principio di ragione insufficiente.

Si può dunque supporre che il ragionamento strategico porti a determinare in modo univoco, oltre ad un equilibrio bayesiano  $\sigma = (\sigma_I, \sigma_{II})$ , anche un *sistema di credenze*  $\mu$  che assegna, per ogni storia parziale  $h$ , e ogni giocatore  $i$ , una distribuzione di probabilità condizionate  $\mu_i(\theta_j|h)$ . Se la congettura  $\sigma_j$  non esclude che una certa storia  $h$  possa verificarsi, allora  $\mu_i(\theta_j|h)$  è derivata da  $\sigma_j$  e dalla credenza iniziale  $\pi_i$ , come nel precedente esempio. Un profilo  $(\sigma, \mu)$  che soddisfa questa condizione è detto *coerente*.

**Definizione.** Un profilo coerente  $(\sigma, \mu)$  è un *equilibrio bayesiano perfetto* se  $\sigma$  è un equilibrio bayesiano e inoltre, per ogni storia parziale  $h$ ,  $\sigma$  induce un equilibrio del corrispondente sottogioco bayesiano determinato dalle credenze condizionate  $\mu_I(\theta_{II}|h)$  e  $\mu_{II}(\theta_I|h)$ .

Entrambi gli equilibri bayesiani del gioco della Fig. 2 possono essere resi perfetti "completandoli" con credenze condizionate appropriate. Per quanto riguarda l'equilibrio  $\sigma^*$  il completamento è dato dalla credenza  $\mu_{II}^*(\theta_I|d)$  derivata da  $\sigma_I^*$  e  $\pi_{II}$ . Per l'altro equilibrio  $\sigma^\circ$  esistono invece infiniti completamenti possibili, perché la credenza condizionata a  $d$  non è deducibile da  $\sigma_I^\circ$  e  $\pi_{II}$ . L'unica condizione che  $\mu_{II}^\circ(\theta_I|d)$  deve soddisfare affinché  $(\sigma^\circ, \mu^\circ)$  sia un equilibrio bayesiano perfetto è che l'azione  $a$  massimizzi l'utilità attesa di *II* calcolata in base alla distribuzione  $\mu_{II}^\circ(\theta_I|d)$  stessa [ciò equivale a richiedere che  $\mu_{II}^\circ(\alpha_I|d)$  sia maggiore o al più uguale a  $\mu_{II}^\circ(\beta_I|d)$  e  $\mu_{II}^\circ(\gamma_I|d)$ ].

Questa estensione del concetto di equilibrio è stata fondamentale per almeno due ragioni. Per quanto riguarda le applicazioni, essa ha permesso di studiare rigorosamente quelle situazioni in cui è possibile inferire, in tutto o in parte, l'informazione privata degli altri giocatori dall'osservazione delle loro azioni. Ciò ha contribuito alla comprensione di importanti fenomeni economici (Mas Colell, cap. 13, Tirole, cap. 9). Dal punto di vista teorico, il concetto di equilibrio bayesiano perfetto è un primo passo verso un'analisi esplicita e formale delle credenze endogene.

#### b) *Raffinamenti*

L'articolo di Selten (1975) è il primo di una nutrita letteratura in cui si propongono dei "raffinamenti" del concetto di equilibrio strategico che hanno lo scopo di eliminare equilibri considerati poco plausibili (v. Van Damme, 1991<sup>2</sup>). Gli studiosi si dividono tra coloro che ritengono essenziale analizzare la forma estesa del gioco e coloro che, seguendo von



Neumann, ritengono sufficiente analizzare la forma strategica. Anche se Selten adotta il primo di questi due approcci, il suo contributo è stato fondamentale per entrambi.

L'idea centrale di Selten è che per eliminare gli equilibri poco plausibili sia necessario "perturbare" il gioco assumendo che tutte le scelte abbiano una "probabilità minima" positiva, anche se molto piccola. Per esempio, nel mini-gioco di contrattazione della Fig. 1 (con  $V_I < 1$  e  $V_{II} > 2$ ) il profilo  $(1, a.r)$  è di equilibrio solo perché Seconda, essendo assolutamente certa che Primo non offra l'oggetto per il prezzo di 2 Euro, è indifferente tra la strategia intuitivamente non credibile  $a.r$  e la strategia "ragionevole"  $a.a$  (v. Tabella 1). Ma se esiste una probabilità positiva, anche se molto piccola, che l'offerta da 2 Euro sia fatta "per sbaglio", allora la strategia  $a.a$  dà un'utilità attesa superiore. Se ad esempio Seconda valuta l'oggetto 3 Euro ( $V_{II} = 3$ ) e ritiene che con 1% di probabilità Primo offrirà l'oggetto per 2 Euro, Seconda attribuisce alla strategia  $a.a$  un'utilità attesa pari a  $(3-1) \times 0,99 + (3-2) \times 0,01 = 1,99$  e alla strategia  $a.r$  un'utilità attesa pari a  $(3-1) \times 0,99 + 0 \times 0,01 = 1,98$ .

Esiste un modo equivalente di vedere la questione, cioè quello di verificare se esistono relazioni di dominanza tra strategie. Si dice che una strategia  $s_i^*$  *domina debolmente* un'altra strategia  $s_i$  se

$$U_i(s_i^*, s_{-i}) \geq U_i(s_i, s_{-i})$$

per ogni profilo di strategie altrui  $s_{-i}$ , e almeno una di queste diseguaglianze è stretta (cioè vale con segno  $>$ ). Si può verificare che se  $s_i^*$  domina debolmente  $s_i$  e se tutti i profili strategici altrui hanno probabilità positiva, allora l'utilità attesa di  $s_i^*$  è maggiore dell'utilità attesa di  $s_i$ . E' inoltre immediato osservare che nel suddetto esempio  $a.a$  domina debolmente  $a.r$ . Questa relazione di dominanza riflette una caratteristica della forma estesa del gioco, cioè che la scelta  $a$  è superiore alla scelta  $r$  anche se l'oggetto viene offerto per 2 Euro. L'esempio quindi suggerisce che alcune caratteristiche essenziali della forma estesa sono catturate da relazioni di dominanza debole nella forma strategica, e che è possibile eliminare gli equilibri basati su minacce non credibili postulando piccole probabilità di errore nella scelta delle strategie. L'approccio dei raffinamenti basati sulla forma strategica si fonda su questa intuizione.

Il raffinamento proposto da Selten (1975), detto "trembling hand perfect equilibrium" ("equilibrio perfetto della mano tremante"), è però definito in relazione alla forma estesa. Usando un'immagine, Selten suppone che i giocatori facciano le loro scelte schiacciando dei bottoni di un terminale e che, a causa di un tremore alle mani, possano scegliere per errore qualunque azione (permessa dalle regole) in qualunque circostanza di gioco. Le strategie di equilibrio specificano, per ogni circostanza di gioco, le azioni che i giocatori *intendono* scegliere, ma le scelte effettive possono risultare differenti. La probabilità che una qualunque azione venga scelta per errore è però piccolissima; inoltre gli errori in differenti stadi del gioco sono tra loro indipendenti. Ciò implica che quando i giocatori osservano deviazioni dalle strategie di equilibrio da parte dei loro avversari, essi le attribuiscono ad errori che quasi certamente non si ripeteranno in futuro. Dunque le strategie di equilibrio continuano ad essere utilizzate per prevedere il comportamento altrui negli stadi successivi del gioco. Si dimostra che, nei giochi a più stadi con azioni osservabili e informazione completa, il raffinamento di Selten è "quasi" equivalente all'equilibrio perfetto nei sottogiochi. Un risultato analogo vale anche sotto l'ipotesi di informazione incompleta se ci sono due giocatori. Nei giochi con tre o più giocatori la questione è più complessa e non è qui possibile entrare

nei dettagli. Basti sapere che non esiste un completo accordo tra i vari testi sulla precisa definizione di equilibrio bayesiano perfetto e che l'equilibrio "trembling hand perfect" essenzialmente corrisponde alla definizione più restrittiva tra quelle apparse in letteratura. Alla base delle differenze tra le varie versioni del concetto di equilibrio bayesiano perfetto stanno differenti ipotesi, spesso implicite e poco trasparenti, su come i giocatori interpretano le deviazioni dall'equilibrio dei loro avversari.

Queste ipotesi hanno un ruolo determinante anche nello sviluppo di ulteriori raffinamenti del concetto di equilibrio, nati soprattutto dalle applicazioni della teoria dei giochi con informazioni incompleta. Molti importanti modelli applicati hanno la struttura di *giochi di segnalazione*, cioè giochi a due stadi in cui il giocatore informato ( $I$ ) muove per primo e il giocatore dis informato ( $II$ ) risponde alla mossa del primo giocatore. Il nome di questi giochi deriva dal fatto che l'azione di  $I$  può essere interpretata come un segnale sulla sua informazione privata. Il gioco della Fig. 2 costituisce un esempio: nell'equilibrio  $\sigma^*$  l'azione  $d$  segnala che il tipo di  $I$  non è  $\alpha_I$ . Gli esempi concreti abbondano: un potenziale lavoratore decide quanto istruirsi e se gli individui più abili fanno meno fatica a istruirsi, il grado di istruzione può segnalare al datore di lavoro l'abilità di un candidato all'assunzione; un'impresa monopolista ( $I$ ) che subisce la minaccia di una potenziale concorrente ( $II$ ) può vendere a un prezzo relativamente basso per segnalare che ha bassi costi di produzione e quindi che l'entrata nel mercato da parte di  $II$  risulterebbe poco profittevole; un venditore può segnalare la qualità del suo prodotto offrendo garanzie che risulterebbero troppo onerose se la qualità fosse scadente.

Spesso questi modelli presentano una grande molteplicità di equilibri bayesiani perfetti, ma gli studiosi non li considerano tutti egualmente plausibili e sono stati suggeriti dei ragionamenti che portano a scartarne la maggior parte. Il tema comune di questi ragionamenti è quello della cosiddetta *induzione in avanti*: il giocatore  $II$  interpreta l'azione osservata  $a_I$  come una scelta intenzionale di  $I$  e dunque considera possibili solo i tipi  $\theta_I$  per i quali  $a_I$  non è una scelta patentemente irrazionale (l'induzione a ritroso si basa invece sull'ipotesi che le azioni *future* degli altri giocatori saranno "razionali"). Per esempio, nel gioco della Fig. 2 l'azione  $d$  è chiaramente irrazionale per il tipo  $\alpha_I$ , mentre per gli altri due tipi può essere giustificata dall'aspettativa che  $II$  scelga  $b$  o  $c$ . Perciò, osservando  $d$ , il giocatore  $II$  dovrebbe inferire che il tipo di  $I$  può essere solo  $\beta_I$  o  $\gamma_I$ . Se  $I$  crede che  $II$  ragioni in questo modo, si aspetta che  $II$  risponda a  $d$  con  $b$  o  $c$ ; quindi i tipi  $\beta_I$  e  $\gamma_I$  ritengono ottimale scegliere  $d$ . Si ottiene così l'equilibrio  $(\sigma^*, \mu^*)$  esaminato nel precedente paragrafo, scartando gli altri equilibri bayesiani perfetti.

Il principio dell'induzione in avanti è chiaro in alcuni casi semplici, come quello appena esaminato, ma risulta in generale più elusivo. Sono stati perciò proposti diversi raffinamenti dell'equilibrio (bayesiano) perfetto che cercano di formalizzare l'induzione in avanti. Alcuni di questi sono basati sulla forma strategica e richiedono, almeno come condizione necessaria, che un profilo di equilibrio sopravviva all'eliminazione *iterata* delle strategie debolmente dominate [si può verificare che nella Tabella 4 tale procedura elimina tutti i profili diversi da  $(\sigma_I^*.b)$ ].

### c) *Riflessioni critiche*

Il paradosso del paradigma dominante, ereditato dalla TG "classica" degli anni '50, consiste nella pretesa di analizzare il ragionamento strategico senza disporre di un linguaggio

matematico sufficientemente espressivo per formulare ipotesi su come ragionano i giocatori. Il linguaggio con cui viene formulata la teoria permette di rappresentare accuratamente le regole del gioco, le preferenze, le strategie, le credenze sugli aspetti non noti del gioco, e può spingersi fino a rappresentare le congetture sul comportamento degli altri giocatori. Ma manca un elemento essenziale per una teoria che aspira a modellare giocatori razionali ed estremamente sofisticati: la rappresentazione delle credenze sul modo di ragionare altrui.

Ciò costringe a formulare teorie sul ragionamento e comportamento strategico con un approccio indiretto: si presuppone informalmente che il ragionamento strategico porti a conclusioni univoche sul comportamento e le congetture dei giocatori e si argomenta, sempre informalmente, che il comportamento e le congetture così ottenuti devono soddisfare delle proprietà matematiche, evocativamente denominate “coerenza”, “stabilità”, “credibilità”. Tale formulazione matematica definisce un concetto di equilibrio e le ipotesi sul ragionamento strategico sono *implicite* nella definizione.

Per esempio, si dice che un equilibrio bayesiano perfetto  $(\sigma^*, \mu^*)$  in un gioco di segnalazione soddisfa il *raffinamento dei messaggi dominati* se vale la seguente condizione di induzione in avanti (v. § b): per ogni azione  $a_I$  e ogni tipo  $\theta_I$  tali che  $a_I$  è dominata per  $\theta_I$  da un'altra azione  $a'_I$  (ma  $a_I$  è non dominata per qualche altro tipo  $\theta'_I$ ),  $\mu_{II}(\theta_I|a_I) = 0$ . Si può arguire che questa definizione intende catturare le seguenti ipotesi: (1) nel secondo stadio del gioco *II* crederà che *I* è razionale, a meno che l'azione osservata non sia in contraddizione con tale credenza, (2) all'inizio del gioco *I* crede nell'ipotesi (1). Tuttavia il linguaggio matematico della TG tradizionale non è sufficientemente ricco per esprimere queste ipotesi in modo formale, quindi ci si accontenta di catturarle in modo indiretto definendo il raffinamento dei messaggi dominati.

Utilizzando questo genere di approccio indiretto, la TG aspira a definire il concetto di equilibrio “giusto”, quello buono per tutti gli usi. I quattro concetti di equilibrio fondamentali della “vulgata cantabrigense” vengono in effetti visti come specificazioni, per particolari classi di giochi, di un concetto generale applicabile a qualunque gioco: l'equilibrio perfetto di Selten. Ma l'estendersi delle applicazioni della TG a campi sempre più vasti delle scienze economiche, politiche e sociali propone con sempre maggior frequenza il problema della molteplicità degli equilibri. Di qui la gara a proporre nozioni di equilibrio sempre più “raffinate”. Ciò non risolve in modo generale il problema della molteplicità, ma semmai crea un “meta-problema”, cioè quello della molteplicità dei concetti di equilibrio [nel testo di Van Damme (1991<sup>2</sup>) sono analizzati più di trenta concetti di equilibrio]. Il modo in cui è avvenuta questa proliferazione di concetti di equilibrio illustra bene i limiti di questo approccio. Il tipico articolo che propone un nuovo raffinamento comincia presentando alcuni esempi in cui gli  $n$  concetti di equilibrio precedenti falliscono nell'eliminare equilibri che l'autore sostiene essere intuitivamente irragionevoli. L'articolo quindi prosegue illustrando, sempre intuitivamente, gli argomenti strategici che intende catturare con la sua nuova definizione. Infine viene proposta la definizione matematica del raffinamento  $n+1$ , e si dimostrano alcuni teoremi, uno di esistenza, e uno o più che mettono in relazione il raffinamento  $n+1$  con i raffinamenti da 1 a  $n$ .

La molteplicità degli equilibri, quale che sia il concetto d'equilibrio adottato, pone due questioni: 1) Come è possibile usare la TG per fare previsioni precise e per capire come cambierà l'equilibrio quando cambiano alcuni parametri del gioco? 2) Se ci sono molteplici equilibri, come può il ragionamento strategico, che si presume univoco, individuarne uno solo?

La risposta alla prima domanda è che forse la precisione delle previsioni è un obiettivo troppo ambizioso. Tanto per cominciare, anche nei casi in cui l'equilibrio è unico, esso dipende da una serie di variabili esogene (i parametri che definiscono il gioco), che tipicamente non sono note ad un osservatore esterno. Ne consegue che la previsione non può essere precisa. Solo in esperimenti controllati è possibile specificare in modo preciso le regole del gioco e i guadagni monetari, ma restano comunque fuori del controllo dello sperimentatore le preferenze dei giocatori rispetto ai possibili esiti del gioco. Non è infatti ovvio che gli individui si limitino a massimizzare i propri guadagni monetari, né è ovvio quali siano le loro attitudini verso il rischio (v. cap. 4, § c). La molteplicità degli equilibri non fa che accentuare un problema generale che esisterebbe comunque, cioè l'esistenza di molteplici esiti compatibili con la teoria, data l'incertezza degli osservatori esterni su certe variabili esogene che caratterizzano le situazioni interattive.

La risposta alla seconda domanda, dopo quanto detto in precedenza, dovrebbe essere abbastanza ovvia. Contrariamente a quanto sembra indicare la retorica utilizzata da molti studiosi nel commentare verbalmente le applicazioni della TG, il ragionamento strategico *non* è costituito dal calcolo degli equilibri. Questi ultimi si limitano a rappresentare i possibili risultati di un ipotetico ragionamento che *non* viene esplicitamente modellato e che si assume porti a conclusioni univoche. Forse è proprio quest'ultima assunzione che bisognerebbe rimuovere.

#### 4. Gli sviluppi più recenti

##### a) *Approccio epistemico*

L'approccio epistemico alla TG è un tentativo di analizzare in modo formale il ragionamento strategico esplicitando le ipotesi sulle conoscenze dei giocatori con riguardo al gioco e alle conoscenze altrui e le ipotesi concernenti le credenze dei giocatori sugli aspetti non noti del gioco, sulle strategie altrui e sulle credenze altrui (v. Osborne e Rubinstein, 1994, cap. 5; Battigalli e Bonanno, 1999). Le ipotesi sulle conoscenze e le credenze dei giocatori sono dette *epistemiche*. Non è qui possibile entrare nei dettagli della formalizzazione matematica delle ipotesi epistemiche. Perciò mi limiterò ad presentare alcuni concetti e risultati in modo sommario e informale.

Un esempio già visto di ipotesi epistemica è la conoscenza comune del gioco, o informazione completa (v. cap. 2, § b). Un altro esempio importante è la *comune credenza* nella razionalità: si indichi con  $R$  l'evento "tutti i giocatori massimizzano l'utilità attesa, date le loro credenze", allora vi è comune credenza nella razionalità se tutti credono  $R$  (cioè assegnano probabilità uno a  $R$ ), tutti credono che tutti credono  $R$ , tutti credono che tutti credono che ... tutti credono  $R$ , e così via. Queste ipotesi epistemiche vengono rappresentate in modo formale ed *esplicito* nella teoria, che è anche in grado di analizzare le conseguenze di ipotesi più generali, o comunque diverse. Va osservato che l'impianto teorico comunemente utilizzato per rappresentare esplicitamente queste ipotesi si sorregge *implicitamente* sulla ulteriore assunzione di tipo cognitivo che gli individui razionali sono dotati di capacità deduttive e di calcolo illimitate. Nell'ambito delle TG, l'analisi dell'interazione tra individui con limitate capacità cognitive utilizza in prevalenza un approccio di tipo adattivo (v. § b).

Il risultato più elementare ottenibile con l'approccio epistemico mette in relazione la comune conoscenza del gioco e la comune credenza nella razionalità con l'eliminazione iterata delle azioni "non giustificabili" nei giochi statici.

Una qualunque distribuzione di probabilità sui profili di azioni degli avversari di un certo giocatore  $i$  è chiamata *congettura* di  $i$  e indicata col simbolo  $\mu^i$ . Si dice che un'azione  $a_i^*$  è giustificata da una congettura  $\mu^i$  se  $a_i^*$  è una delle azioni che massimizzano l'utilità attesa di  $i$  data  $\mu^i$ . Un'azione è *giustificabile* se esiste almeno una congettura che la giustifica. Sia  $G$  un gioco statico e si consideri la seguente procedura:

*Eliminazione iterata delle azioni non giustificabili. Passo 1* Si eliminano tutte le azioni non giustificabili di ogni giocatore  $i$ ; si indichi con  $A_i^1(G)$  il rimanente insieme di azioni giustificabili, si definisce allora il gioco ristretto  $G^1$  in cui l'insieme delle azioni di ogni giocatore  $i = I, II, \dots$  è  $A_i^1(G)$  (a parte cioè, le funzioni di utilità sono come in  $G$ ). *Passo 2*, per ogni  $i$  si eliminano tutte le azioni non giustificabili di  $i$  nel gioco ristretto  $G^1$ ; si ottiene così un gioco ulteriormente ristretto  $G^2$  con insiemi di azioni  $A_i^2(G)$  ( $i = I, II, \dots$ ). Si procede in modo analogo finché non è più possibile alcuna eliminazione, ottenendo degli insiemi di azioni  $A_i^*(G)$  ( $i = I, II, \dots$ ) dette *iterativamente giustificabili*.

Per esempio, nel gioco  $G$  rappresentato nella Tabella 5, solo l'azione  $c_I$  non è giustificabile (infatti dà un'utilità inferiore a  $b_I$  qualunque sia la scelta di  $II$ ). L'unica azione non giustificabile nel gioco ristretto  $G^1$  è  $c_{II}$ . Ogni altra azione è iterativamente giustificabile, cioè  $A_I^*(G) = \{a_I, b_I\}$  e  $A_{II}^*(G) = \{a_{II}, b_{II}\}$ .

$G$	$a_{II}$	$b_{II}$	$c_{II}$
$a_I$	4, 0	0, 4	1, 0
$b_I$	1, 3	3, 1	1, 0
$c_I$	0, 0	0, 0	0, 4

Tabella 5. – Un gioco statico.

Assumiamo che valga l'ipotesi di conoscenza comune di  $G$ . E' vero per definizione che  $A_i^1(G)$  è l'insieme delle azioni di  $i$  compatibili con l'ipotesi di razionalità. Credere nella razionalità dunque equivale a considerare impossibili le azioni non contenute nel gioco ristretto  $G^1$ . Perciò le azioni compatibili con la razionalità e la credenza nella razionalità (data l'informazione completa) sono quelle giustificabili in  $G^1$ , cioè le azioni in  $A_i^2(G)$ ,  $i = I, II, \dots$ . Proseguendo con questo ragionamento si ottiene il seguente risultato:

*L'insieme delle azioni di un giocatore  $i$  del gioco statico  $G$  che sono compatibili con la conoscenza comune di  $G$ , la razionalità e la comune credenza nella razionalità è  $A_i^*(G)$ , ovvero ciò che sopravvive all'eliminazione iterata delle azioni non giustificabili.*

Il gioco della Tabella 5 mostra che i giocatori possono disporre di più di una azione iterativamente giustificabile. Dunque il ragionamento strategico basato "soltanto" sull'informazione completa, la comune credenza nella razionalità e illimitate capacità deduttive, non necessariamente porta a conclusioni univoche sul comportamento dei giocatori. Si è in precedenza argomentato che il ragionamento strategico porta all'equilibrio quando le sue conclusioni sono univoche. Invece le conclusioni di questo tipo di ragionamento strategico, in questo gioco, non sono univoche e infatti le azioni non escluse non sono di equilibrio.

D'altra parte, si può dimostrare che se un'azione è parte di un profilo di equilibrio, allora è iterativamente giustificabile. Ne consegue che se in un dato gioco ogni giocatore ha un'unica azione iterativamente giustificabile, il risultante profilo di azioni è necessariamente l'unico equilibrio del gioco. Ciò conferma l'argomento informale che il ragionamento strategico,

quando è univoco, porta all'equilibrio. Va osservato che questo è il caso in molti giochi interessanti dal punto di vista applicativo.

In generale, adottando l'approccio epistemico è possibile dimostrare risultati della seguente forma: se il gioco  $G$  soddisfa la proprietà  $\mathcal{P}$  (per es.,  $G$  è un gioco a mosse simultanee), allora l'insieme dei profili strategici compatibili con l'ipotesi di razionalità e con le ipotesi epistemiche  $\mathcal{E}$  (per es., conoscenza comune di  $G$  e credenza comune nella razionalità) è  $\mathcal{S}(G)$  (per es., l'insieme dei profili di azioni iterativamente giustificabili). Ciò mostra che, nell'ambito dei giochi che soddisfano  $\mathcal{P}$ , il concetto di soluzione  $\mathcal{S}$  cattura (oltre all'ipotesi di razionalità) le ipotesi epistemiche  $\mathcal{E}$ . Naturalmente, il concetto di soluzione  $\mathcal{S}$  può essere già noto nella letteratura di TG. In tal caso il risultato ha fornito una "fondazione epistemica" di  $\mathcal{S}$ . Per esempio, sono state formulate ipotesi epistemiche che giustificano l'induzione a ritroso nei giochi finiti ad informazione perfetta e un certo tipo di induzione in avanti nei giochi di segnalazione. Va osservato che tali ipotesi sono più forti della comune credenza nella razionalità; si può infatti mostrare che nella maggior parte dei giochi ad informazione perfetta con più di due stadi esistono esiti compatibili con la comune conoscenza del gioco e la comune credenza nella razionalità e tuttavia diversi dall'esito (di equilibrio perfetto) calcolato mediante induzione a ritroso (v. Battigalli e Bonanno, 1999, § 4).

In ogni caso, la "solidità" delle cosiddette fondamenta epistemiche dei concetti di soluzione dipende dal grado di accettabilità delle ipotesi, o per il loro interesse come idealizzazioni teoriche, o per la loro plausibilità empirica in certi contesti interattivi. Il merito dell'approccio epistemico è soprattutto quello di rendere esplicite, dunque valutabili criticamente, le ipotesi sulle conoscenze e credenze interattive dei giocatori.

Inoltre l'approccio epistemico non si limita a "fondare" concetti di soluzione già proposti dalla TG più tradizionale. Esso rende possibile, almeno in linea di principio, un'analisi del comportamento strategico non basata su concetti di soluzione preconfezionati, ma sull'adozione delle ipotesi (sulle regole del gioco, comportamentali, epistemiche) che si ritengono più adeguate al caso in esame, deducendone le conseguenze in termini di esiti possibili.

#### b) *Approccio adattivo e razionalità limitata*

Come si è visto, l'approccio epistemico mostra che l'equilibrio di Nash non è in generale giustificabile in base all'ipotesi che i giocatori completamente informati e razionali compiano sofisticati ragionamenti strategici. Inoltre la completa razionalità e la sofisticazione strategica sono solo un'idealizzazione teorica. I giocatori reali sono infatti soggetti a limitazioni cognitive: l'informazione è spesso molto incompleta, le capacità deduttive sono limitate, la stessa attività di deduzione utilizza risorse mentali preziose. Ma tutto ciò non significa che l'analisi di equilibrio sia priva di fondamento. E' infatti possibile che l'apprendimento da esperienze passate, proprie e/o altrui, permetta agli individui di adattare reciprocamente le proprie strategie in modo tale che le scelte di ognuno tendano ad essere ottimali, date le scelte altrui. Naturalmente questo presuppone che gli individui affrontino ripetutamente la stessa situazione d'interazione.

L'approccio adattivo analizza una varietà di modelli dinamici d'interazione strategica ripetuta in cui le strategie adottate in un certo periodo dipendono, tramite un processo di retroazione, dalle strategie adottate nei periodi precedenti (v. Fudenberg e Levine, 1998, e Van Damme, 1991<sup>2</sup>, cap. 9). Questi modelli dinamici possono essere sommariamente divisi in due classi: i modelli evolutivi e i modelli di apprendimento.

I *modelli evolutivi* assumono l'esistenza di grandi popolazioni di individui, i quali si incontrano a caso, interagiscono secondo le regole di un gioco tipico (in cui ogni individuo assume il ruolo di un particolare giocatore) e poi si separano. Questi modelli cercano di spiegare solo le dinamiche aggregate, non i comportamenti individuali. L'ipotesi di base è che *le strategie che hanno più successo vengono adottate con più frequenza*; ciò può essere dovuto a dinamiche di apprendimento sociale, oppure a vere e proprie dinamiche di evoluzione biologica in cui il comportamento è geneticamente determinato e gli individui portatori di geni di successo producono una prole più numerosa, che replica il comportamento dei genitori. I primi modelli di giochi evolutivi hanno infatti avuto origine nella biologia e si proponevano di analizzare la distribuzione di certe caratteristiche di specie biologiche animali o vegetali non facilmente spiegabili come effetti dell'adattamento della specie all'ambiente esterno (v. Maynard Smith, 1982).

Un esempio può essere utile a illustrare questi modelli e anche a chiarire il significato di un concetto finora solo brevemente accennato: l'equilibrio in strategie miste (v. cap. 2, § e). Si consideri il gioco FC ("Falco o Colomba") della Tabella 3. In base all'interpretazione biologica, quando i membri di una certa specie animale si incontrano e si contendono una risorsa, possono avere un atteggiamento aggressivo ( $f$ ) o timido ( $c$ ). I numeri nella matrice del gioco FC rappresentano allora gli incrementi di "fitness" (numero atteso di figli) delle due tipologie di individui in funzione del rispettivo comportamento. La "fitness" complessiva corrispondente a una certa strategia dipende dalla composizione della popolazione. Se, per esempio, il 20% (un quinto) degli individui si comporta in modo timido e il resto in modo aggressivo, la "fitness" *media* di  $c$  è  $2 \times \frac{1}{5} + 1 \times \frac{4}{5} = \frac{6}{5}$ , e quella di  $f$  è  $4 \times \frac{1}{5} + 0 \times \frac{4}{5} = \frac{4}{5}$ . Dunque in tale situazione i timidi tenderanno a procreare di più degli aggressivi e la proporzione dei timidi aumenterà. Calcoli analoghi mostrano che in questo esempio si ha un processo di retroazione negativa: la proporzione dei timidi tende ad aumentare quando questi costituiscono una frazione strettamente positiva della popolazione, ma minore di  $\frac{1}{3}$ , e tende a diminuire nel caso opposto (frazione inferiore a 1, ma maggiore di  $\frac{1}{3}$ ). Se invece la popolazione contiene  $\frac{1}{3}$  di timidi e  $\frac{2}{3}$  di aggressivi, le due strategie hanno la stessa "fitness" media e quindi le proporzioni rimangono costanti; in altri termini, questa composizione della popolazione è *stazionaria*. Inoltre, se questa composizione stazionaria viene alterata (per esempio a causa di mutazioni) il processo di retroazione negativa tende a ristabilirla; quindi la composizione stazionaria è anche *stabile*. Se la popolazione fosse composta solo di "falchi" (individui aggressivi), potrebbero nascere solo "falchi" e quindi la composizione della popolazione rimarrebbe costante anche in questo caso; ma se una mutazione introducesse nella popolazione una piccola frazione di colombe, queste comincerebbero a riprodursi velocemente. Quindi una popolazione composta di soli "falchi" sarebbe stazionaria, ma non stabile. Un discorso analogo vale per una popolazione di sole "colombe".

Il gioco FC è simmetrico: i ruoli dei due giocatori sono interscambiabili. E' quindi naturale considerare dinamiche evolutive in cui gli individui che si incontrano e giocano provengono dalla stessa popolazione. Nel caso di giochi asimmetrici è più plausibile considerare dinamiche co-evolutive in cui esiste una popolazione distinta per ogni ruolo (per esempio, nel mini-gioco di contrattazione gli individui che giocano nel ruolo di  $I$  sono estratti da una popolazione di venditori e quelli che giocano nel ruolo di  $II$  sono estratti da una popolazione di compratori).

Una *strategia mista* per il giocatore  $i$  è una distribuzione di probabilità  $p_i$  sull'insieme  $S_i$  delle sue strategie, dette anche *pure*. Un equilibrio in strategie miste, o più semplicemente,

*equilibrio misto*, di un gioco a due persone è un profilo  $(p_I, p_{II})$  tale che  $p_i$  ( $i = I, II$ ) massimizza l'utilità attesa di  $i$  data la distribuzione  $p_j$  (interpretabile anche come una congettura di  $i$  su  $j$ ). Questa definizione generalizza il concetto di equilibrio. Infatti, ogni strategia pura  $s_i$  è interpretabile come una strategia mista  $p_i$  che assegna probabilità 1 a  $s_i$ ; quindi gli equilibri in strategie pure, o più in breve gli *equilibri puri* sono un caso particolare di equilibri in strategie miste. Per esempio, il gioco FC ha due equilibri puri e asimmetrici,  $(c, f)$  e  $(f, c)$ , e un equilibrio misto in senso stretto, cioè il profilo simmetrico che per ogni giocatore  $i$  assegna probabilità  $\frac{1}{3}$  a  $c$  e  $\frac{2}{3}$  a  $f$ . Perciò l'equilibrio misto simmetrico del gioco corrisponde alla composizione stazionaria e stabile della dinamica evolutiva (nei modelli simmetrici con giocatori estratti da un'unica popolazione ha senso considerare solo gli equilibri simmetrici, mentre nei modelli, simmetrici o asimmetrici, in cui esiste una popolazione distinta per ogni ruolo  $i = I, II$  ha senso considerare come possibili stati stabili anche gli equilibri asimmetrici).

Sfruttando le proprietà dell'utilità attesa, si può dimostrare che  $(p_I, p_{II})$  è un equilibrio in strategie miste se e solo se, data la strategia mista dell'avversario  $p_j$ , tutte le strategie di  $i$  a cui  $p_i$  assegna probabilità positiva sono ottime, e quindi danno la stessa utilità attesa di  $p_i$ . Per esempio, nell'equilibrio misto simmetrico di FC le strategie  $f$  e  $c$  danno utilità attesa  $\frac{4}{3}$ , dato che l'avversario sceglie  $c$  con probabilità  $\frac{1}{3}$ ; ciò implica che qualunque strategia mista del giocatore  $i$  dà la stessa utilità attesa (cioè  $\frac{4}{3}$ ), data la strategia mista di equilibrio dell'avversario  $j$ . Questa condizione di indifferenza tra le strategie lascia perplessi. E' già abbastanza strano pensare a individui che decidono di scegliere a caso (cosa significa infatti "scegliere  $c$  con probabilità  $\frac{1}{3}$ "?). Ma è ancora più bizzarro ipotizzare che lo facciano esattamente con le probabilità specificate dalla distribuzione di equilibrio, visto che sono indifferenti (se  $f$  e  $c$  danno la stessa utilità attesa non sarebbe più naturale assegnare ad entrambe la stessa probabilità  $\frac{1}{2}$ ?).

L'approccio adattivo fornisce un'interpretazione soddisfacente degli equilibri misti: la distribuzione di probabilità di equilibrio (in questo caso, simmetrico)  $(p(f), p(c)) = (\frac{2}{3}, \frac{1}{3})$  non va intesa come la rappresentazione matematica di una scelta casuale da parte di un decisore, ma piuttosto come una distribuzione stazionaria di azioni *pure* in una popolazione di individui, in cui due terzi scelgono  $f$  e un terzo sceglie  $c$ . D'altra parte si può dimostrare che le distribuzioni stazionarie *stabili* di dinamiche evolutive come quelle sommariamente descritte sopra corrispondono sempre a equilibri (puri o) misti.

I modelli qui classificati come "evolutivi" descrivono dinamiche aggregate senza spiegare in dettaglio come tali dinamiche siano generate da comportamenti individuali. I *modelli di apprendimento* sono invece più focalizzati sui singoli individui. Alcuni modelli considerano un gruppo fisso di individui che giocano ripetutamente un certo gioco  $G$ . Altri modelli, più simili a quelli evolutivi, postulano l'esistenza di grandi popolazioni di individui (una per giocatore/ruolo di  $G$ ) da cui i giocatori sono estratti a caso per interagire una volta sola e poi separarsi. Per "apprendimento" si intende un meccanismo di retroazione in base al quale le scelte correnti di un individuo dipendono dalle sue precedenti esperienze, che a loro volta dipendono anche dalle scelte degli altri individui. I meccanismi di retroazione possono essere distinti in base al grado di sofisticazione attribuito agli individui. Alcuni modelli di apprendimento assumono che gli individui si limitino a scegliere con più frequenza le azioni che in passato sono risultate associate ai guadagni più alti. Altri modelli si collocano invece all'interno del paradigma della scelta massimizzante. Mi soffermerò brevemente su questi



ultimi. L'ipotesi di base è che ogni giocatore formula una congettura (probabilistica) sulle strategie degli avversari nel periodo corrente e sceglie una risposta ottima; la congettura è poi modificata prima del periodo successivo in base a ciò che il giocatore ha potuto osservare. Di norma si suppone che i giocatori ignorino gli effetti delle loro scelte correnti sulle scelte future degli avversari. Ciò può essere dovuto a semplice miopia, oppure viene giustificato in base all'ipotesi che gli avversari cambino da una partita all'altra e che la probabilità di incontrare nuovamente gli stessi avversari sia trascurabile. L'ipotesi che ogni giocatore  $i$  scelga una risposta ottima alla sua congettura presuppone implicitamente che la funzione di utilità  $u_i$  gli sia nota; le funzioni di utilità altrui potrebbero invece non essergli note.

Nei modelli di apprendimento la specificazione delle possibilità di osservazione è cruciale. Il caso più semplice e più studiato è quello in cui un gioco a mosse simultanee  $G$  viene giocato ripetutamente e dopo ogni partita vengono *perfettamente osservate le azioni altrui*. Se il processo di revisione delle congetture viene specificato in modo plausibile, esso deve soddisfare una proprietà di *empirismo asintotico*: nel lungo periodo le probabilità soggettive che ogni individuo assegna alle azioni altrui tendono a coincidere con le frequenze osservate. Sotto queste ipotesi si possono trarre due conclusioni:

(1) *Un profilo di azioni  $a^*$  può essere stazionario (cioè riprodursi indefinitamente) solo se è un equilibrio di Nash*; infatti la congettura di ogni giocatore  $i$  tenderà ad assegnare probabilità 1 a  $a_{-i}^*$  e se  $a_i^*$  non fosse una risposta ottima a  $a_{-i}^*$  il giocatore  $i$  finirebbe per cambiare azione, contraddicendo l'ipotesi che  $a^*$  sia stazionario (un risultato analogo vale per i profili di strategie miste). Ciò implica che se una dinamica di apprendimento converge (a partire da una condizione iniziale data), il limite deve essere un equilibrio di Nash.

(2) *Nel lungo periodo saranno scelte solo azioni iterativamente giustificabili* (v. § a); infatti le azioni non giustificabili di  $G$  non sono mai scelte perché non sono risposte ottime ad alcuna congettura, quindi (per l'ipotesi di empirismo asintotico) dopo un numero sufficientemente elevato di ripetizioni le congetture dei giocatori assegneranno probabilità trascurabili a tali azioni, quindi i giocatori si comporteranno come se giocassero il gioco ristretto  $G^1$  e non sceglieranno azioni non giustificabili in  $G^1$ . Dopo un certo tempo anche a tali azioni si finirà per assegnare una probabilità trascurabile e quindi i giocatori si comporteranno come se giocassero il gioco ristretto  $G^2$ . Il processo di riduzione continuerà finché non saranno giocate solo azioni iterativamente giustificabili. Si noti che il risultato è valido anche se la dinamica non converge ad un comportamento stabile (un risultato simile vale per le dinamiche evolutive, per le quali si dimostra che la proporzione di individui che utilizzano azioni iterativamente giustificabili tende a zero).

Sotto le suddette ipotesi si può quindi affermare che, dopo un processo di adattamento, i giocatori si comporteranno *come se* vi fosse informazione completa e comune certezza della razionalità, e che può essere stabile solo una situazione che soddisfa le condizioni di equilibrio specificate dal paradigma dominante (vd. cap. 3) per il gioco  $G$  (trattandosi di un gioco a mosse simultanee, il concetto di equilibrio rilevante è quello di Nash). Tuttavia questi risultati non valgono in generale. In un gioco a mosse simultanee azioni e strategie coincidono, perciò osservare le azioni equivale ad osservare le strategie. Se invece il gioco  $G$  è dinamico, risulta impossibile osservare, alla fine della partita, le strategie degli avversari. Il motivo è semplice: non è possibile osservare come un avversario si sarebbe comportato in circostanze che durante la partita non si sono verificate, al massimo si possono osservare le azioni effettivamente scelte durante la partita. Per esempio, nel mini-gioco di contrattazione della Fig. 1, se  $I$  ha paura che  $II$  possa rispondere negativamente alla richiesta di un prezzo

alto (magari perché non conosce il valore che  $II$  assegna all'oggetto), allora  $I$  offrirà l'oggetto al prezzo basso (che verrà accettato) e non avrà modo di osservare come  $II$  si sarebbe comportata nell'altro caso. Non vi sarà quindi alcuna ragione di modificare la congettura iniziale e il sistema dinamico rimarrà intrappolato nell'esito  $(1, a)$  che non è compatibile con l'equilibrio perfetto e nemmeno con le ipotesi di informazione completa e comune certezza della razionalità. L'esito  $(1, a)$  nel precedente esempio è comunque compatibile con l'equilibrio di Nash [è infatti generato dall'equilibrio imperfetto  $(1, a, r)$ ]. Tuttavia esempi più complessi mostrano che perfino esiti incompatibili con l'equilibrio di Nash possono risultare stazionari e stabili: si tratta dei cosiddetti *equilibri auto-confermanti*, cioè di situazioni in cui ogni individuo sceglie una risposta ottima alla sua congettura e tale congettura risulta compatibile con ciò che l'individuo osserva "a fine partita" (v. Fudenberg e Levine, 1998, cap. 6).

In conclusione, si possono fare alcune considerazioni sulla rilevanza dei concetti di equilibrio proposti dal paradigma dominante nei contesti in cui è applicabile l'approccio adattivo. Tale approccio tende a giustificare l'utilizzazione del concetto di equilibrio di Nash per spiegare configurazioni stabili in situazioni d'interazione ripetuta, anche se i giocatori non sono completamente razionali, informati e sofisticati. Nel caso delle dinamiche di apprendimento è però necessario ipotizzare un alto grado di osservabilità delle azioni altrui. L'utilizzazione del concetto di equilibrio a fini previsivi è invece più problematica. Infatti in certe situazioni interattive ripetute il comportamento dei giocatori potrebbe non convergere ad uno stato stazionario (si può però affermare che le azioni iterativamente non giustificabili saranno progressivamente eliminate). L'approccio adattivo dà invece poco supporto al concetto di equilibrio perfetto per giochi dinamici.

### c) *Studi empirici ed esperimenti*

Come si è visto in precedenza, le predizioni dei modelli basati sulla TG dipendono da quattro tipi di parametri: (*REG*) le *regole del gioco* (ordine delle mosse, vincoli, informazioni su mosse precedenti, relazione tra azioni dei giocatori e conseguenze), (*PREF*) le *preferenze* dei giocatori rispetto alle conseguenze (certe e incerte) dell'interazione, (*C-REG*) le *conoscenze* interattive dei giocatori *sulle regole del gioco*, (*C-PREF*) le *conoscenze* interattive dei giocatori *sulle preferenze*. Dati tali parametri, l'ipotesi di razionalità e le ipotesi sul ragionamento strategico (normalmente "condensate" in un concetto di equilibrio) determinano delle previsioni sul comportamento dei giocatori. Il problema principale di un osservatore esterno, o "investigatore", che vuole confrontare tali modelli con la realtà è che i suddetti parametri sono solo parzialmente osservabili. Risulta quindi difficile stabilire se i comportamenti dei giocatori (o certe loro conseguenze osservabili) sono coerenti con le predizioni della TG. Tutto ciò che si può fare è valutare congiuntamente un insieme di ipotesi che comprende assunzioni ausiliarie sui parametri non osservabili unitamente alle ipotesi sul comportamento strategico proprie della TG.

Questo problema è particolarmente marcato negli studi che, per mancanza di una migliore terminologia, definirò *empirici*, cioè quelli in cui l'investigatore si limita a raccogliere dati generati dal "mondo esterno" indipendentemente dalla propria indagine (un tipico esempio è lo studio dell'andamento dei prezzi in un certo mercato oligopolistico). In questi studi

infatti è frequente che nessuno dei parametri che caratterizzano il gioco sia adeguatamente osservabile. Per ovviare a questo problema alcuni studiosi si limitano ad analizzare le predizioni *robuste* di certe classi di modelli di TG (ad esempio, di certi modelli di oligopolio), cioè le predizioni che valgono indipendentemente dalla specificazione dei parametri non osservabili (v. Sutton, 2000, cap. 3). Altri studiosi si concentrano sui casi in cui le regole del gioco sono esplicitamente codificate, come ad esempio nella vendita all'asta dei diritti di estrazione di petrolio o gas naturale da determinati tratti di mare (v. Sutton, cap. 2). Tuttavia anche nei casi più favorevoli l'approccio empirico è soggetto ad un altro importante limite: non potendo influenzare il processo di generazione dei dati, è infatti molto difficile per l'investigatore ottenere nuovi dati che permettano di scegliere tra due ipotesi alternative entrambe compatibili con le osservazioni.

Per ovviare ai limiti dell'approccio empirico, negli ultimi vent'anni si è definitivamente affermato, accanto a questo, anche un approccio *sperimentale*: le osservazioni sui comportamenti dei giocatori vengono generate "in laboratorio" sotto condizioni controllate. In particolare, negli esperimenti di TG l'investigatore ha pieno controllo dei parametri REG e C-REG. Questo consente di limitare le ipotesi ausiliarie e focalizzarsi maggiormente sulle ipotesi riguardanti il comportamento strategico. Gli esperimenti sono anche usati per valutare il possibile impatto di nuove regole e istituzioni (ad esempio, nuovi formati d'asta per vendere licenze di trasmissione). L'importanza del metodo sperimentale in economia è stata riconosciuta con l'assegnazione a Daniel Kahneman e Vernon Smith del premio Nobel per questa disciplina nell'anno 2002.

Per avere un'idea delle problematiche e dei risultati dell'approccio sperimentale è utile considerare il caso paradigmatico della serie di esperimenti sui giochi di contrattazione (v. Roth, 1995). Il gioco di contrattazione più studiato negli esperimenti, detto *Gioco dell'Ultimatum* (GU), ha una struttura molto simile al gioco della Fig. 1:  $I$  e  $II$  devono suddividere una somma di denaro  $X$ ,  $I$  propone a  $II$  una suddivisione di tale somma e  $II$  può accettare o rifiutare; se  $II$  accetta, ognuno percepisce l'importo proposto, se rifiuta entrambi i giocatori ottengono 0. Nel tipico esperimento sul GU,  $X$  è compresa tra 10 e 30 Euro, la quantità offerta da  $I$  a  $II$  è un multiplo di un certo importo minimo, ad esempio 5 centesimi di Euro (per semplificare la discussione supporrò che  $I$  non possa offrire a  $II$  meno di 5 centesimi), all'esperimento partecipa un numero elevato di individui che vengono suddivisi tra i due ruoli ( $I$  e  $II$ ) e accoppiati per giocare il GU. Il gioco viene ripetuto una decina di volte, *accoppiando gli individui ogni volta con un avversario diverso*. Ciò permette ai soggetti sperimentali di acquisire esperienza, ma non introduce incentivi a cercare di modificare il comportamento futuro dell'avversario, visto che non lo si incontrerà di nuovo. La comunicazione avviene in modo *anonimo* (per esempio tramite terminale). Gli individui percepiscono alla fine dell'esperimento la somma degli importi delle dieci partite giocate. Le regole del gioco GU e gli altri aspetti del protocollo sperimentale appena descritti sono spiegati pubblicamente. Dunque l'investigatore determina le regole e la comune conoscenza di tali regole. Ma per ottenere una previsione teorica è necessario fare ipotesi ausiliarie su PREF e C-PREF. Le ipotesi ausiliarie "di riferimento", cioè quelle più naturali come punto di partenza dell'indagine, sono:

(PREF- $M_I$ )  $I$  è esclusivamente interessato al (valore atteso del) proprio guadagno monetario,

(PREF- $M_{II}$ )  $II$  è esclusivamente interessato al proprio guadagno monetario,

( $C_I$ -PREF $_{II}$ )  $I$  è certo di PREF- $M_{II}$ .

Date tali ipotesi ausiliarie, la previsione teorica è quella di *equilibrio perfetto*:  $I$  offre a  $II$  il minimo (5 centesimi) e  $II$  accetta. Questo risultato è ottenibile per induzione a ritroso e si basa sulle ipotesi di razionalità dei giocatori e di certezza, da parte di  $I$ , della razionalità di  $II$ .

L'esperimento è stato ripetuto molte volte con soggetti diversi e in paesi diversi. I dati osservati presentano due notevoli regolarità: (1) nella quasi totalità dei casi la somma offerta a  $II$  è tra il 20% e il 50% di  $X$  con un picco di frequenza vicino al 40% di  $X$ , (2) circa il 20% delle offerte inferiori a  $\frac{X}{2}$  è *rifiutato*. Quindi la previsione teorica è nettamente falsificata.

Come interpretare tale falsificazione? Alcuni studiosi hanno sostenuto che le osservazioni sperimentali falsificano la teoria dell'equilibrio perfetto e indicano che la contrattazione è primariamente influenzata da considerazioni non strategiche, quali il desiderio di raggiungere un accordo ritenuto "equo" (si vedano i riferimenti in Roth, 1995, pp. 258-260). Ma senza le ipotesi ausiliarie l'equilibrio perfetto non è né quantitativamente determinabile, né teoricamente giustificabile. In realtà, le uniche ipotesi falsificate dall'evidenza sperimentale sono le ipotesi ausiliarie  $PREF-M_{II}$  e  $C_I-PREF-M_{II}$ . I risultati osservati sono invece pienamente compatibili con la seguente teoria "strategica": l'utilità di  $II$  dipende positivamente dal suo guadagno monetario e negativamente dalla differenza tra il guadagno di  $I$  e il proprio, ciò determina una "soglia di accettazione" che varia da individuo a individuo (un'offerta è accettata se e solo se è superiore alla soglia);  $I$  è incerto sulla soglia di accettazione e fa un'offerta che massimizza il suo guadagno atteso.

Per meglio valutare l'interazione tra considerazioni strategiche e di equità, sono stati condotti esperimenti con regole di contrattazione differenti, che però, mantenendo le ipotesi ausiliarie "di riferimento" sulle preferenze puramente monetarie, danno luogo a equilibri perfetti con distribuzioni dei guadagni tanto estreme quanto l'equilibrio del GU. Per esempio, è stata studiata una variante del GU in cui dieci giocatori,  $I_1, I_2, \dots, I_{10}$  fanno (simultaneamente e indipendentemente) un'offerta a  $II$  e  $II$  decide quale offerta accettare; in ogni equilibrio perfetto con preferenze puramente monetarie  $II$  adotta la semplice strategia di accettare sempre l'offerta più alta e la concorrenza spinge ad offrire a  $II$  (quasi) tutta la somma  $X$ .

Se il comportamento in questi giochi fosse guidato in modo preponderante da considerazioni di equità, si dovrebbero osservare in prevalenza accordi su distribuzioni intermedie e rifiuti di distribuzioni estreme (non è forse iniquo che un giocatore sfrutti le asimmetrie insite nelle regole di contrattazione per ottenere quasi tutto il guadagno complessivamente generato da un accordo?). L'evidenza sperimentale invece mostra che, dopo un aggiustamento nelle iterazioni iniziali, si raggiunge nella quasi totalità dei casi l'accordo (estremo) di equilibrio perfetto. Questa serie di esperimenti sui giochi di contrattazione indica quindi che considerazioni di tipo distributivo possono influenzare il comportamento (come accade nel GU), ma interagiscono in modo complesso con considerazioni strategiche.

## 5. Conclusioni

La TG non fornisce una teoria unitaria delle situazioni decisionali interattive, quanto piuttosto un quadro di riferimento concettuale e un linguaggio matematico che permettono di modellare tali situazioni in modo rigoroso. La TG consente di formulare ipotesi precise sull'interazione strategica e derivarne le implicazioni in termini di comportamento dei giocatori.

Il compito dei teorici “puri” è di raffinare i concetti, rendere il linguaggio sempre più espressivo e derivare relazioni generali tra ipotesi e conseguenze. I fruitori della TG hanno invece il compito cruciale di costruire modelli che aiutino a comprendere determinati aspetti dell’interazione economica, politica e sociale, valutando le ipotesi in base al contesto applicativo e testandole col metodo empirico e sperimentale. In questo senso è auspicabile che in futuro si faccia un uso meno acritico del “ricettario” proposto dal paradigma oggi dominante (v. cap. 3) e che si sfrutti maggiormente la ricchezza concettuale ed espressiva degli sviluppi recenti della TG.

### **Bibliografia**

AUMANN, R., *Game Theory*, in *The New Palgrave Dictionary of Economics* (a cura di J. Eatwell, M. Milgate e P. Newman), Vol. II, London-Basingstoke 1987, 460-482.

BATTIGALLI, P., BONANNO, G., *Recent results on belief, knowledge and the epistemic foundations of game theory*, in “Research in Economics”, 1999, DIII, 149-226.

BINMORE, K., *Fun and Games: a Text on Game Theory*, Lexington, MA, 1992.

COURNOT, A., *Recherches sur les principes mathématiques de la théorie des richesses*, Paris, 1838.

DIXIT, A., SKEATH, S., *Games of strategy*, New York, 1999.

FUDENBERG, D., LEVINE, D., *The theory of learning in games*, Cambridge MA, 1998.

FUDENBERG, D., TIROLE, J., *Game theory*, Cambridge MA, 1991.

GIBBONS, R., *Game theory for applied economists*, Princeton, N.J., 1992.

GIBBONS, R., *An Introduction to Applicable Game Theory*, in “Journal of Economic Literature”, 1997, XI, pp127-149.

HARSANYI, J., *Games with incomplete information played by Bayesian players*, in “Management Science”, 1967-68, XIV, pp. 159-182, 320-334, 486-502.

KLEMPERER, P., *Auction theory: a guide to the literature*, in “Journal of Economic Surveys”, 1999, XIII, pp. 227-286.

LEWIS, D., *Convention*, Cambridge MA, 1969.

MAS COLELL, A., WHINSTONM., GREEN, J., *Microeconomic theory*, Oxford, 1995.

MAYNARD SMITH, J., *Evolution and the theory of games*, Cambridge, MA, 1982.

- MORROW, J., *Game Theory for Political Scientists*, Princeton, N.J., 1994.
- MYERSON, R., *Game theory. Analysis of conflict*, Cambridge, MA, 1991.
- NASH, J., *Essays on game theory* (Intr. di K. Binmore), Cheltenham, U.K., 1996.
- NEUMANN, J. VON, *Zur Theorie der Gesellschaftsspiele*, in “Mathematische Annalen”, 1928, C, pp. 295-320.
- NEUMANN, J. VON, MORGENSTERN, O., *Theory of games and economic behavior*, Princeton, N.J., 1947<sup>2</sup>.
- OSBORNE, M., RUBINSTEIN, A., *A course in game theory*, Cambridge, MA, 1994.
- OWEN, G., *Game Theory*, San Diego, CA, 1995<sup>3</sup>.
- ROTH, A., *Bargaining Experiments*, in *Handbook of Experimental Economics* (a cura di J. Kagel e A. Roth), Princeton, N.J., 1995.
- TIROLE, J., *Theory of Industrial Organization*, Cambridge, MA, 1988
- SELTEN, R., *Re-examination of the Perfectness Concept for Equilibrium Points in Extensive Games*, in “International Journal of Game Theory”, 1975, IV, pp. 25-55.
- SUTTON, J., *Marshall's tendencies. What can economists know?*, Cambridge, MA, 2000.
- VAN DAMME, E., *Stability and perfection of Nash equilibria*, Berlin, 1991<sup>2</sup>.
- WALKER, P., *An outline of the history of game theory*,  
<http://william-king.www.drexel.edu/top/class/histf.html>.