



**MIMESIS
ETEROTOPIE**

N. 209

Collana diretta da Salvo Vaccaro e Pierre Dalla Vigna

COMITATO SCIENTIFICO

Pierandrea Amato (Università degli Studi di Messina)

Antonio Caronia (NABA)

Pierre Dalla Vigna (Università degli Studi "Insubria" Varese)

Giuseppe Di Giacomo (Università di Roma La Sapienza)

Maurizio Guerri (Università degli Studi di Milano)

Salvo Vaccaro (Università degli Studi di Palermo)

José Luis Villacañas Berlanga (Universidad Complutense de Madrid)



ROBERTO FESTA E GUSTAVO CEVOLANI

GIOCHI DI SOCIETÀ

Teoria dei giochi
e metodo delle scienze sociali



MIMESIS
Eterotopie

© 2013 – MIMESIS EDIZIONI (Milano – Udine)
Collana: Eterotopie n. 209
Isbn 9788857518770
www.mimesisedizioni.it
Via Risorgimento, 33 – 20099 Sesto San Giovanni (MI)
Telefono +39 02 24861657 / 24416383
Fax: +39 02 89403935
E-mail: mimesis@mimesisedizioni.it

INDICE

1. INTRODUZIONE	7
2. COS'È LA TEORIA DEI GIOCHI	13
2.1 Breve storia della teoria dei giochi	15
2.2 Oltre la teoria classica dei giochi	17
3. ELEMENTI DI TEORIA CLASSICA DEI GIOCHI	27
3.1 Giochi contro natura	27
3.2 Giocatori, mosse e strategie	32
3.3 Coordinazione e conflitto	34
3.4 Strategie dominanti, equilibri e soluzioni	39
3.5 Due problemi per la teoria classica dei giochi	44
4. BENI PUBBLICI E ALTRI DILEMMI SOCIALI	51
4.1 Il dilemma dei beni pubblici	52
4.2 Dilemma dei beni pubblici e dilemma del prigioniero	55
4.3 Beni pubblici nelle teorie non classiche dei giochi	60
4.4 Beni pubblici, teoria dei giochi e filosofia della politica	67
4.5 Oltre il dilemma del prigioniero	74
4.6 Dilemmi della sfiducia	80
5. COMUNICAZIONE, DETERRENZA E STRATEGIA DEL CONFLITTO	85
5.1 Teoria strategica e deterrenza	86
5.2 Strategie dissuasive e teoria del conflitto	91
5.3 Logica della deterrenza	93
5.4 Metodologia della deterrenza	108
5.5 Paradossi della deterrenza	112
6. NATURA, GENESI ED EVOLUZIONE DELLE NORME SOCIALI	121
6.1 L'approccio giochistico alle norme sociali	122
6.2 Norme, altruismo e origini evolutive della cooperazione	130

6.3 L'evoluzione della cooperazione e delle norme di giustizia	137
6.4 Società artificiali, teorie ABM e complessità della cooperazione	145
7. TEORIA DEI GIOCHI, SCUOLA AUSTRIACA E METODO DELLE SCIENZE SOCIALI	149
7.1 Le origini "austriache" della teoria dei giochi	151
7.2 Teoria dei giochi e concezione hayekiana dell'evoluzione culturale	154
7.3 Spiegazione e complessità nelle scienze sociali	163
7.4 Teoria dei giochi, individualismo e filosofia della politica	171
ELENCO DELLE FIGURE	177
BIBLIOGRAFIA	179

1. INTRODUZIONE

La teoria dei giochi studia le decisioni di individui coinvolti in interazioni “strategiche” – situazioni, cioè, in cui il comportamento ottimale di ogni individuo dipende anche dalle scelte degli altri individui coinvolti. Una partita di scacchi, una battaglia o una compravendita sono tutti esempi di “gioco” in questo senso: cioè, di situazioni che coinvolgono due o più individui, ognuno dei quali sceglie la propria linea di condotta sulla base di ciò che si attende facciano gli altri.

Giochi di società

Le origini della teoria dei giochi risalgono al tentativo di analizzare matematicamente le strategie vincenti di giochi da tavolo come gli scacchi o il poker.¹ Nella sua forma compiuta, fu poi presentata – nell’epocale volume *Theory of Games and Economic Behavior* di von Neumann e Morgenstern (1944) – come contributo alla teoria del comportamento economico. Tuttavia, alla luce della generalità dei suoi metodi e del numero dei suoi impieghi, considerare la teoria dei giochi come una branca della matematica applicata o dell’economia matematica sarebbe riduttivo. In ogni caso, non è questa la prospettiva del presente volume. Se infatti si astrae dalle diverse formulazioni matematiche e dagli specifici dettagli delle situazioni che la teoria analizza, essa si presenta come una teoria generale delle decisioni interdipendenti o del comportamento strategico, cioè della cooperazione e del conflitto fra agenti di qualsiasi tipo (Schelling 1960). Dato che l’interazione fra individui è alla base di ogni fenomeno sociale, l’insieme delle scienze sociali potrebbe addirittura essere considerato una branca della te-

1 Per una felice coincidenza, quest’anno ricorrono sia il terzo centenario della lettera (1713) in cui il “gentiluomo inglese” Waldegrave propose la prima discussione moderna di una strategia per un gioco di carte (cfr. Bellhouse 2007) sia il centenario del primo teorema di teoria dei giochi, enunciato da Zermelo nel 1913 (cfr. Schwalbe e Walker 2001).

oria dei giochi (cfr. Binmore 1990, p. 5). Senza arrivare a questo, occorre comunque riconoscere che l'approccio "giochistico" (*game theoretic*) fornisce uno strumento ideale per analizzare i fenomeni della società, tanto da venir sempre più ampiamente considerato come l'approccio privilegiato ai problemi delle scienze sociali, in grado di offrire una "teoria unificata" nell'ambito di queste discipline (cfr. Gintis 2009, Buchanan 2002).

Nel corso della sua storia, la teoria dei giochi ha conosciuto un enorme sviluppo ed è stata applicata all'analisi di problemi sempre più complessi e numerosi in una grande varietà di campi, dalla strategia militare (Schelling 1960) alla biologia evolutivista (Maynard Smith 1982), dalla teoria politica (Morrow 1994) all'etica (Verbeek e Morris 2010), dalla logica alla filosofia (de Bruin 2005). Lo scopo principale di questo libro è offrire una rassegna dettagliata degli sviluppi più recenti della teoria dei giochi. Negli ultimi decenni, infatti, la versione "classica" di questa teoria, sviluppata da von Neumann, Morgenstern e altri, è stata affiancata, e in parte soppiantata, da un certo numero di approcci alternativi che mirano a evitare alcuni suoi fondamentali difetti teorici e a migliorarne la capacità esplicativa e l'applicabilità a problemi concreti. Per questo motivo, è ormai più corretto parlare di "teorie", piuttosto che di "teoria", dei giochi, includendo fra queste la teoria evolutivista dei giochi, le teorie ad agenti basate su simulazioni al computer (o teorie ABM, da *Agent Based Modelling*), le teorie cognitive e comportamentali dei giochi, e altri approcci ormai consolidati.

Perché questo libro, e come usarlo

Sebbene esistano numerosi e autorevoli contributi, sia di livello introduttivo, sia di livello specialistico, a molte di queste teorie,² manca

2 A parte il volume di von Neumann e Morgenstern (1944), la più nota introduzione tecnica alla teoria classica dei giochi è probabilmente il vecchio, ma sempre utile, manuale di Luce e Raiffa (1957). Il libro di Schelling (1960), recentemente tradotto in italiano, rimane una lettura imprescindibile, a metà strada fra teoria classica e teorie non classiche (cognitive) dei giochi; un altro libro importante è Sugden (1986). I libri sulla teoria classica sono ormai innumerevoli; citiamo solo Hargreaves Heap *et al.* (1992) e Hargreaves Heap e Varoufakis (2004), per una discussione critica dei fondamenti concettuali e filosofici, rispettivamente, della teoria della scelta razionale e dei giochi; le introduzioni elementari di Davis (1983) e Binmore (2007); e il manuale tecnico di Osborne e Rubinstein (1994), disponibile gratuitamente in rete (i libri di Hargreaves Heap *et al.* e di Binmore sono disponibili anche in italiano). Per una rassegna della bibliografia essenziale, continuamente aggiornata, si veda Ross (2012). Con l'eccezione dell'ormai con-

al momento una panoramica aggiornata che ne illustri le caratteristiche essenziali, i comuni intenti teorici e metodologici e le principali differenze. Il presente volume si propone appunto di colmare questa lacuna nella letteratura non solo di lingua italiana, offrendo al lettore gli elementi formali e concettuali alla base dello studio dei fenomeni sociali fondato sulle teorie dei giochi.

Il libro si rivolge sia a studenti e studiosi di filosofia e di scienze sociali sia al lettore interessato a capire come sempre più economisti, sociologi e politologi affrontano i problemi di queste discipline. Lo stile è discorsivo e mira, evitando qualsiasi tecnicismo, a presentare gli elementi della teoria della scelta razionale e della teoria dei giochi, utilizzando soprattutto esempi concreti della loro applicazione ad alcuni problemi centrali nelle scienze sociali. L'ampia bibliografia permette al lettore non solo di orientarsi nell'ormai vastissima letteratura, ma soprattutto di programmare uno studio più approfondito sia degli strumenti formali utilizzati sia dei temi discussi nel libro, nonché di altri che rimangono ai margini della trattazione.

Sebbene presenti i risultati del lavoro di ricerca compiuto dagli autori nel corso degli ultimi anni, il libro non presuppone che il lettore abbia compiuto altre letture sull'argomento, né richiede alcuna specifica preparazione matematica o filosofica. Per questo, si presta per essere utilizzato – con le necessarie integrazioni segnalate in nota e in bibliografia – come manuale per corsi di filosofia della scienza, epistemologia e metodologia delle scienze sociali, in corsi di laurea sia triennale sia specialistica di qualsiasi facoltà. In effetti, negli ultimi anni i contenuti esposti sistematicamente in questo volume sono stati usati con successo per presentare la teoria della scelta razionale e la teoria dei giochi sia a studenti di filosofia, economia e scienze umane, sia a medici, veterinari, infermieri e altre figure professionali in occasione di corsi di formazione e di aggiornamento.

solidata teoria evolutivista dei giochi (Maynard Smith 1982; Weibull 1997; Gintis 2000), non esistono manuali introduttivi alle teorie non classiche dei giochi, per cui occorre ancora rivolgersi alla letteratura specialistica. Si vedano però almeno Camerer (2003), Rasmussen (2006) e de Bruin (2010). In italiano, oltre a quelle già citate, si trovano le traduzioni dei manuali di Kreps (1990) e Gibbon (1992) e di alcuni contributi di Morgenstern (1969) e Aumann e Hart (2008), oltre a diverse introduzioni alla teoria classica dei giochi. Per una rassegna sui problemi metodologici e filosofici relativi alle teorie dei giochi, si veda Grüne-Yanoff e Lehtinen (2012); sulla metodologia delle scienze sociali e sulla filosofia dell'economia, ci limitiamo a menzionare Elster (1989b), Di Nuoscio (2006) e Guala (2006).

Struttura e contenuti del libro

Oltre alla presente introduzione, il volume si compone di altri sei capitoli, che possono venire idealmente divisi in tre parti. Nella prima (capitoli 2 e 3) si introducono la terminologia e i concetti essenziali delle diverse versioni della teoria dei giochi; nella seconda (capitoli 4, 5 e 6) si presenta l'applicazione della teoria ad alcuni problemi fondamentali delle scienze sociali; infine, nella terza (capitolo 7) si discutono alcuni aspetti metodologici dell'approccio giochistico allo studio della società.

Più in particolare, nel capitolo 2 introduciamo la teoria dei giochi come una branca della teoria della scelta razionale. Come la teoria delle decisioni studia le scelte di un individuo "in isolamento", la teoria dei giochi analizza le scelte di un individuo che interagisce con altri individui in modo strategico. La teoria dei giochi può essere quindi caratterizzata come la teoria delle decisioni "interattive" o "interdipendenti". Dopo aver ricostruito a grandi linee lo sviluppo storico della teoria "classica" dei giochi – quella di von Neumann e Morgenstern (1944) e dei primi teorici dei giochi, fino agli anni Sessanta –, ne definiamo gli scopi e i limiti, distinguendola da diverse, e successive, teorie "non classiche" dei giochi. Vale la pena avvertire fin d'ora il lettore che, dato che tali teorie sono tuttora in pieno sviluppo, la terminologia in questo campo è tutt'altro che uniforme e stabilita. Così, alcune scelte terminologiche – come teorie "razionalistiche", "cognitive", "comportamentali", "adattive" dei giochi – hanno solo parziale riscontro in letteratura, e vanno considerate come proposte preliminari per una tassonomia delle teorie dei giochi.

Nel capitolo 3 presentiamo tutti i concetti essenziali della teoria classica dei giochi che verranno utilizzati nel seguito. La discussione è quasi interamente informale, e l'uso della matematica si limita a quello delle quattro operazioni. Dopo aver presentato le basi della teoria delle decisioni, introduciamo il concetto tecnico di gioco, i diversi tipi di gioco e la nozione di "equilibrio di Nash". Il capitolo si chiude con la discussione di due problemi che mettono in crisi l'idea, centrale nella teoria classica, di poter individuare la soluzione di qualsiasi tipo di gioco: cioè, l'esistenza di (molti) giochi con più di una soluzione plausibile e l'esistenza di giochi – come il famoso dilemma del prigioniero – la cui unica soluzione è intuitivamente "sbagliata".

I capitoli 4, 5 e 6 mostrano come le diverse versioni della teoria dei giochi sono state applicate all'analisi di alcuni importanti fenomeni sociali: i cosiddetti dilemmi sociali (capitolo 4), la comunicazione e la deterrenza (capitolo 5) e le norme sociali (capitolo 6).

Il capitolo 4 considera il problema della produzione di un bene pubblico come tipico esempio di dilemma sociale (e in particolare di dilemma del prigioniero). Secondo la teoria economica e politica “ortodossa”, il dilemma dei beni pubblici è insolubile dal punto di vista della teoria classica dei giochi e dimostra l’impossibilità della coordinazione spontanea fra “egoisti razionali” e quindi la necessità dello stato per rendere possibile la cooperazione sociale. Le teorie non classiche dei giochi – in particolare le teorie cognitive, comportamentali, evolucionistiche e adattive – hanno invece esplorato diverse vie d’uscita da questo dilemma, mostrando come la cooperazione possa emergere e diffondersi spontaneamente, anche in assenza di un’autorità centrale che regoli il comportamento individuale.

Le teorie cognitive dei giochi – che risalgono almeno al pionieristico lavoro di Schelling (1960) – mettono in luce l’importanza del “contesto cognitivo” di un gioco, cioè delle capacità e credenze dei diversi individui coinvolti. Il capitolo 5 prende spunto dal famoso *The Strategy of Conflict* (Schelling 1960) per mostrare, utilizzando l’esempio della deterrenza atomica, la funzione essenziale che la comunicazione e la (s)fiducia reciproca – e quindi le promesse, le minacce, gli avvertimenti e così via – svolgono in una grande varietà di situazioni sociali.

La genesi e l’evoluzione delle norme, delle convenzioni e delle altre istituzioni sociali è un problema centrale per una tradizione di studi che risale almeno a David Hume (1739; cfr. Hardin 2007). Il capitolo 6 ricostruisce i principali risultati dell’approccio giochistico alle norme sociali, che, sulla linea di Hume, tenta di spiegare come potrebbe essersi evoluto, e come può continuare ad evolversi, l’implicito “contratto sociale” che regola la nostra vita in società. Le teorie evolucionistiche dei giochi (biologiche e adattive) e le teorie ABM offrono un raffinato strumento per comprendere come si siano prodotte le nostre effettive predisposizioni alla giustizia, alla cooperazione e all’altruismo.

I risvolti di questo approccio per la metodologia delle scienze sociali costituiscono il tema dell’ultima parte del libro. Il capitolo 7 parte prende le mosse dalle somiglianze – a prima vista sorprendenti – fra l’analisi giochistica dei fenomeni sociali e le tesi della cosiddetta Scuola Austriaca di economia. Tale confronto – e in particolare la discussione delle posizioni teoriche e metodologiche di Friedrich von Hayek – permette di mettere in luce alcuni presupposti metodologici dell’approccio giochistico, la natura e il ruolo della spiegazione dei fenomeni complessi nelle scienze sociali, e il nesso fra l’analisi sociale e la filosofia della politica.

Ringraziamenti

I materiali su cui si basa il libro sono stati presentati a diversi seminari, convegni e conferenze e discussi in queste sedi con molte persone. In particolare, gli autori desiderano ringraziare Carlo Buttasi, Vincenzo Crupi, Raimondo Cubeddu, Michele Di Francesco, Enzo Di Nuoscio, Pierdaniele Giaretta, Carlo Lottieri, Pierpaolo Marrone, Ruggero Rangoni e Luca Tambolo per le utili conversazioni sui temi del libro. Come mostrano i riferimenti in bibliografia, il volume nasce da una prolungata consuetudine di lavoro congiunto fra i due autori, che ha fatto sì che i contributi di ognuno siano pressoché inseparabili e indistinguibili. Tuttavia, a fini squisitamente accademici, i capitoli 2, 4 e 6 sono da attribuire a Gustavo Cevolani e i capitoli 3, 5 e 7 a Roberto Festa.³

3 Il lavoro di ricerca di Roberto Festa è stato parzialmente finanziato dal progetto MIUR-PRIN 2008 *Probabilità, conferma e verosimilitudine. Le strutture cognitive dell'opinione e della decisione nelle scienze empiriche, nelle pratiche esperte e nelle interazioni sociali*.

2. COS'È LA TEORIA DEI GIOCHI

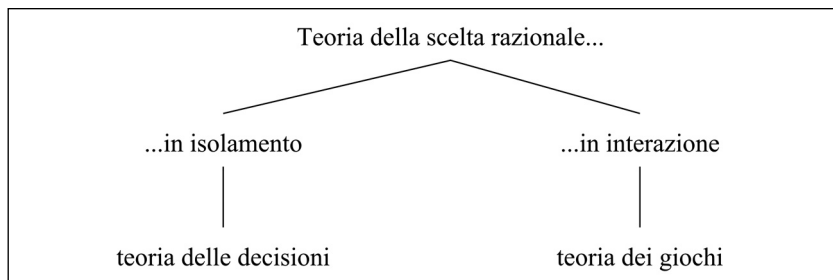
Per gli esseri umani, scegliere è quasi come nutrirsi: bene o male, non possono farne a meno. L'esigenza di prendere decisioni, cioè scegliere fra diverse alternative, ci è imposta dalla limitatezza delle risorse a nostra disposizione, a cominciare dal tempo e dal denaro. Dalla più piccola delle scelte quotidiane a quella di sposarsi, dall'acquisto di un determinato modello di automobile alla decisione di accettare o meno un posto di lavoro, il problema ricorrente è la necessità di individuare uno specifico corso d'azione da perseguire a dispetto di altri possibili.

L'esperienza comune mostra quanto le decisioni possano essere fra loro diverse sotto molti punti di vista. Possono avere conseguenze più o meno serie. Possono richiedere la considerazione di pochi dati semplici e chiari, o di una gran quantità di informazioni complesse e di difficile interpretazione. Possono riguardare solo, o principalmente, noi stessi, o dipendere dalle decisioni altrui, e avere ricadute sugli altri. Eppure, nonostante questa innegabile varietà, è spesso possibile analizzare le decisioni in base a un quadro concettuale unitario e ormai consolidato, noto come *teoria della scelta razionale*.

La teoria della scelta razionale si propone di analizzare in che modo le scelte di un individuo o "agente razionale" siano determinate dalle sue preferenze e dalle sue credenze relative ai vari contesti in cui egli si trova ad operare. Tali contesti sono di due tipi fondamentali:

- 1) i "giochi contro natura", cioè le situazioni in cui un individuo deve scegliere tra diverse azioni, senza conoscere con certezza il reale "stato di natura" e, quindi, neppure i risultati prodotti dalle sue possibili azioni;
- 2) i "contesti interattivi", cioè le situazioni in cui un agente è impegnato nell'interazione con altri individui.

Figura 2.1. Teoria della scelta razionale, teoria delle decisioni e teoria dei giochi



Le scelte razionali nei giochi contro natura vengono analizzate dalla *teoria delle decisioni*; quelle effettuate nei contesti interattivi costituiscono l'oggetto della *teoria dei giochi* nelle sue diverse varianti (si veda la figura 2.1).

La teoria delle decisioni è più “semplice” rispetto alla teoria dei giochi nel senso che, in un gioco contro natura, l'esito delle azioni di un agente razionale dipende solo dall'azione scelta e dal reale stato di cose, che l'agente può avere più o meno correttamente anticipato. Per esempio, la scelta di prendere l'ombrello uscendo di casa si rivelerà più o meno saggia a seconda del fatto che piova o meno. Nei contesti interattivi, invece, il risultato dell'azione di un agente dipenderà sia dalla sua scelta sia da quelle degli agenti con i quali interagisce. Per esempio, nell'interazione tra due agenti, la “bontà” di qualunque decisione presa dal primo normalmente dipenderà dalla decisione del secondo: se anche decido di impegnarmi seriamente in una relazione sentimentale, il buon esito della stessa dipenderà dalle intenzioni e dalle scelte del mio *partner*. Di conseguenza, la scelta del primo agente dovrà basarsi, in qualche modo, sulle sue *attese* circa la decisione che verrà presa dal secondo, che dipenderà, a sua volta, dalle attese di quest'ultimo circa la decisione del primo. Si è soliti dire che interazioni di questo tipo, che costituiscono l'oggetto privilegiato della teoria dei giochi, hanno carattere *strategico*.

Il termine “teoria dei giochi” è legato al fatto che gli scacchi, i giochi con le carte, la morra cinese e molti altri giochi da tavolo rientrano nell'ambito della teoria, che comprende però anche molti altri tipi di interazioni strategiche. Sono giochi, infatti, anche una compravendita, un accordo commerciale, una trattativa politica o sindacale, una battaglia o una conferenza internazionale: sono, insomma, giochi tutte le situazioni in cui due “giocatori” – ove questo termine può applicarsi non solo a persone, ma anche a organizzazioni, imprese, governi e così via – sono coinvolti in

qualche interazione strategica. Non tutti i giochi, tuttavia, sono “giochi” nel senso della teoria dei giochi. Come nota Schelling (1960, p. 3, nota 1), la natura strategica delle situazioni di cui si occupa la teoria le distingue sia dai giochi di abilità (come una corsa podistica) sia dai giochi d'azzardo o “di fortuna” (come la *roulette* o il lotto, il cui esito è determinato dal caso). Nell'analisi dei giochi di strategia, cioè dei giochi *tout court* di cui ci occuperemo nel resto del libro, l'attenzione si concentra sull'interdipendenza delle decisioni di tutti i giocatori coinvolti e sulle loro attese circa il comportamento altrui.

Negli ultimi decenni, la teoria dei giochi ha conosciuto un impetuoso sviluppo e si è ramificata in diverse direzioni, tanto che oggi sembra più corretto parlare, al plurale, di *teorie* dei giochi più che di un'unica teoria dei giochi. Tale sviluppo è tuttora in corso ed è spesso difficile distinguere e classificare i diversi approcci che vengono comunemente individuati nell'ambito della teoria. Dopo un rapido sguardo storico alle origini della teoria (paragrafo 2.1), proponiamo una classificazione che crediamo utile per orientarsi nell'ampio campo d'indagine delle teorie dei giochi (paragrafo 2.2).⁴

2.1 Breve storia della teoria dei giochi

I primi elementi della teoria della scelta razionale si possono rintracciare nell'opera del filosofo e matematico Blaise Pascal (1623-1662), che è anche uno dei padri riconosciuti della teoria della probabilità. (Come si vedrà, esistono importanti legami fra le due teorie.) Le indagini di Pascal su questo tema furono ispirate da preoccupazioni che all'epoca impegnavano i giocatori d'azzardo, particolarmente sensibili – per ovvie ragioni – all'elaborazione di metodi affidabili per determinare quanto fossero vantaggiose le scommesse (per un'avvincente ricostruzione storica, si veda Devlin 2008).

Pascal, che aveva intuito le enormi potenzialità della teoria, non esitò a presentarne un'applicazione divenuta celebre: analizzò in termini di scommessa una scelta che ogni uomo è chiamato a fare, mettendo letteralmente in gioco la propria vita. Nel paragrafo 233 dei suoi *Pensieri*, egli sostenne infatti che adottare una condotta di vita cristiana è la scelta più razionale anche per chi si senta attratto dai piaceri del libertinaggio. Il suo argomento

4 Sulla storia della teoria dei giochi, si vedano in particolare i contributi raccolti in Dimand e Dimand (1996) e Dimand (2014) e la cronologia curata da Walker (2012).

ha una forma matematica: tutto considerato, sostiene Pascal, “scommettere su Dio” presenta un bilancio di benefici e costi attesi enormemente più vantaggioso della scelta opposta, perché la possibilità di ottenere l’eterna beatitudine – per quanto incerta – ha un valore incommensurabile.⁵ Come avremo modo di vedere più nei dettagli, la teoria della scelta razionale ha continuato a servirsi comunemente di questo parallelismo fra scommesse e problemi decisionali della vita reale.

Successivi contributi di figure centrali della storia della filosofia e della matematica, come Jeremy Bentham (1748-1832) e Jakob Bernoulli (1654-1705), testimoniano il persistente intreccio fra le due discipline che caratterizza l’elaborazione dei principi razionali della scelta. Ancora negli anni Venti del Novecento, sono due giovani geni della filosofia e della matematica a mettere a punto in una forma compiuta gli elementi fondamentali della teoria. Nel 1926, il ventitreenne Frank Plumpton Ramsey (1903-1930), originale seguace di Russell e Wittgenstein, elabora la prima trattazione assiomatica della scelta in condizioni di incertezza in un importante articolo rimasto inedito fino alla sua prematura scomparsa (Ramsey 1931). Due anni più tardi, il suo coetaneo ungherese John von Neumann (1903-1957) pubblica una teoria matematica dei giochi in cui il comportamento strategico ottimale emerge da un calcolo basato sulle conoscenze e gli obiettivi di agenti perfettamente razionali (von Neumann 1928). Questo scritto di von Neumann – assieme ad alcuni altri del logico tedesco Ernst Zermelo (1871-1953) e del matematico francese Émile Borel (1871-1956) dedicati, negli anni Dieci e Venti, agli scacchi, al poker e a altri giochi da tavolo – rappresentano i primi contributi alla nascente teoria dei giochi.

Per veder maturare i frutti di questi lavori inaugurali sarà però necessario attendere la metà del secolo, con la pubblicazione dei due epocali volumi *Theory of Games and Economic Behavior* (von Neumann e Morgenstern, 1944), ancora di von Neumann con la collaborazione dell’economista austriaco Oskar Morgenstern (1902-1977), e *The Foundations of Statistics* dello statistico Leonard Savage (1917-1971) (Savage, 1954). È da notare, peraltro, che alcuni dei principi della teoria avevano nel frattempo trovato una originale trattazione nell’opera del grande matematico italiano Bruno de Finetti (1906-1985) (De Finetti 1937), del quale Savage non esitò a considerarsi seguace.

5 Aldilà del suo perdurante interesse filosofico (si veda Jordan, 1994), la “scommessa di Pascal” è rimasta un vivido caso illustrativo della teoria delle decisioni, ripreso perciò dalla manualistica, come in Baron (1988, p. 230). Del resto, il filosofo e storico della scienza Ian Hacking ha identificato proprio nell’argomento di Pascal “il primo contributo convincente alla teoria delle decisioni” (Hacking, 1975, p. viii).

Nella scienza contemporanea, la teoria della scelta razionale è divenuta il cuore della cosiddetta “economia neoclassica” (si veda in proposito Motterlini e Piattelli Palmarini 2005). Essa si è altresì rivelata un importante punto di riferimento per una varietà di discipline interessate al comportamento, dalla psicologia cognitiva alle scienze politiche, dalle neuroscienze fino all'etologia, arrivando infine a toccare lo studio delle decisioni in virtualmente qualsiasi campo dell'azione umana e non. Anche gli animali, infatti, come gli uomini, operano continuamente delle scelte che, pur non basandosi su un calcolo razionale, sono state attentamente studiate dalle versioni più recenti della teoria dei giochi. Così, sia la teoria delle decisioni sia la teoria dei giochi vengono oggi utilizzate per analizzare le scelte di individui nelle più svariate situazioni, dai laboratori scientifici alle contese animali per una preda o un territorio.

Mentre, nel corso degli ultimi decenni, il campo di applicazione della teoria della scelta razionale si estendeva progressivamente, i fondamenti della teoria venivano riconsiderati, criticati e in parte modificati, sia alla luce dei problemi posti dalla loro applicazione sia a causa degli sviluppi delle altre discipline con cui la teoria interagisce. Per esempio, l'analisi giochistica del comportamento animale, inaugurata nella sua forma moderna dal biologo inglese John Maynard Smith (1920-2004) e dai suoi collaboratori, ha richiesto, come vedremo, una profonda revisione dei fondamenti concettuali della teoria dei giochi. Una revisione simile è seguita agli sviluppi dell'economia e alla sua crescente interazione con la psicologia e con altre scienze del comportamento umano. Sia l'economia “cognitiva” o “comportamentale” (Kanheman 2011; Motterlini e Piattelli Palmarini 2005) sia l'economia “sperimentale” (Smith 2008) si distaccano in più punti dal paradigma dell'economia neoclassica e hanno di conseguenza influenzato anche la teoria della scelta razionale.⁶ Nel prossimo paragrafo, passiamo in rassegna alcune di queste versioni recenti della scelta razionale, e in particolare della teoria dei giochi, che verranno più ampiamente trattate nel resto del libro.

2.2 Oltre la teoria classica dei giochi

Come abbiamo detto, la prima esposizione organica della teoria dei giochi si deve a von Neumann e Morgenstern, che nel 1944 pubblicarono

6 Si veda Motterlini e Guala (2011) per un'introduzione all'economia cognitiva e sperimentale.

l'influente libro *The Theory of Games and Economic Behavior*. Lo scopo della loro teoria era quello di individuare la soluzione di un gioco, cioè la combinazione delle strategie ottimali che dovrebbero essere adottate da ciascuno dei giocatori. Nelle intenzioni di von Neumann e Morgenstern, la teoria dei giochi aveva carattere sostanzialmente "normativo" (cfr. Binmore 1990, pp. 9 ss.; Skyrms 1994; Selten 1998). In altre parole, lo scopo della teoria non era descrivere come di fatto si comportano individui impegnati in interazioni strategiche, ma determinare quali *dovrebbero* essere le scelte di un agente perfettamente razionale in ogni possibile tipo di gioco adeguatamente specificato. Grazie alla natura generale e astratta della nozione di gioco, tuttavia, è stato possibile applicare con successo la teoria dei giochi nei campi più disparati, dalle scienze sociali all'economia, dagli studi di strategia militare all'etica e alla filosofia della politica. La teoria ha così trovato numerose applicazioni descrittive, nella spiegazione e nella previsione di svariati comportamenti umani; tali applicazioni si basano sul presupposto che, in molti casi, gli agenti impegnati nelle interazioni sociali sono almeno approssimativamente razionali.

Nel seguito, chiameremo teoria *classica* dei giochi la teoria elaborata nel solco dell'opera di von Neumann e Morgenstern. Gli elementi essenziali di questa teoria verranno presentati nel prossimo capitolo, ma è utile fin d'ora descriverne alcune caratteristiche generali, per sottolinearne i tratti peculiari che la distinguono dalle successive teorie dei giochi non classiche. La teoria classica dei giochi si fonda su cinque presupposti concettuali, che vengono quasi universalmente, anche se spesso tacitamente, accettati fino all'inizio degli anni Sessanta. Tali presupposti o postulati, che chiameremo TGC1-TGC5, sono elencati di seguito.⁷

TGC1. *Preferenze autointeressate*. Le preferenze di un giocatore tra i possibili risultati di un gioco sono *autointeressate*, cioè sono determinate unicamente dai vantaggi "materiali" da lui ottenuti con ciascun risultato. Tali vantaggi possono essere costituiti da denaro, posizioni di potere, onori e da ogni altro genere di beni o servizi di cui il giocatore può usufruire. Ciò implica che la preferenza di un giocatore per un determinato risultato non è influenzata dai vantaggi materiali ottenuti, con quel risultato, da altri giocatori, né dal "processo" che ha condotto a quel risultato, cioè dalla specifica combinazione di strategie adottate dai vari giocatori.

7 Occorre avvertire il lettore che né la specifica formulazione scelta per i presupposti TGC1-TGC5, né tanto meno i loro nomi, sono standard in letteratura.

TGC2. *Perfezione razionale*. I giocatori sono *perfettamente razionali*, cioè sono dotati di illimitate risorse di memoria e di una perfetta capacità di ragionamento.

TGC3. *Perfezione pratica*. I giocatori hanno *perfette capacità pratiche*, cioè hanno una volontà d'acciaio e la capacità di eseguire alla perfezione tutte le possibili strategie del gioco; quindi, essi attueranno sempre, e senza errori, la strategia che avranno scelto di attuare.

TGC4. *Conoscenza comune*. Ogni giocatore è a conoscenza della perfezione razionale (vedi TGC2) e pratica (vedi TGC3) del concorrente. Inoltre, tutti i giocatori sanno che ciascuno di loro è a conoscenza dell'altrui perfezione, dell'altrui conoscenza della propria perfezione, e così via, all'infinito.

TGC5. *Universalismo aprioristico*. Date le sue preferenze (vedi TGC1), le sue informazioni circa le preferenze altrui, la sua perfetta razionalità (vedi TGC2) e la sua conoscenza dell'altrui perfezione razionale e pratica (vedi TGC4), ogni giocatore può determinare la propria strategia ottimale affidandosi a principi di razionalità determinati sulla base di considerazioni *a priori* e, di conseguenza, *universalmente applicabili*.

I presupposti appena descritti, e in particolare i primi quattro, identificano il protagonista della teoria classica dei giochi (e dell'economia neoclassica) nel famoso (o famigerato) "egoista illuminato" o *homo æconomicus*: un astratto agente perfettamente razionale e perfettamente egoista, nel senso tecnico del postulato TGC1, che calcola infallibilmente le proprie scelte al fine di ricavare il massimo beneficio atteso da ogni possibile contesto (interattivo o contro natura). Il sogno dei teorici "classici" dei giochi, espresso dal postulato TGC5, era (e in parte è tuttora) di poter identificare l'esito, cioè la "soluzione", di qualsiasi gioco che coinvolge due o più esemplari di *homo æconomicus*.

Secondo la maggior parte degli studiosi, il sogno della teoria classica dei giochi si è infranto sia contro alcune formidabili difficoltà matematiche implicite nell'impostazione analitica della teoria classica sia a causa della caricatura, troppo astratta e irrealistica, che l'*homo æconomicus* fornisce dei veri agenti sociali. Come ha notato il poliedrico economista e premio Nobel (1978) americano Herbert Simon (1916-2001), le teorie classiche della razionalità – come la teoria delle decisioni e la teoria dei giochi – si basano su un "modello olimpico" di razionalità; sono, cioè, teorie più adatte agli dei che agli uomini, poiché consigliano l'uso di strategie che, per la loro attuazione, presuppongono capacità letteralmente sovrumane di analisi e calcolo (Simon 1983; 1982-1997). Al contrario, il modello di

razionalità limitata proposto da Simon – e da molti altri dopo di lui – mira a individuare strategie del genere effettivamente utilizzabili da parte di agenti con tutte le limitazioni proprie degli esseri umani. Come osserva Simon (1983, p. 56)

[tale modello] rinuncia a molte delle magnifiche proprietà formali del modello olimpico; ma in cambio fornisce un approccio alla razionalità che chiarisce come le creature dotate delle nostre capacità mentali – o anche dotate delle nostre capacità mentali aumentate di quelle di tutti i computer di Silicon Valley – possono avere successo in un mondo che è fin troppo complicato da capire...

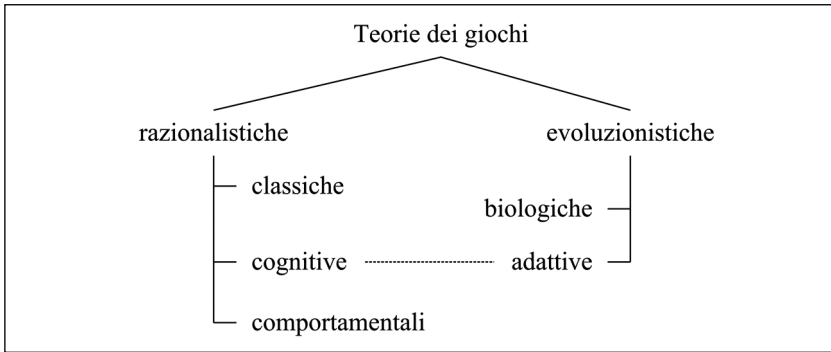
Poiché l’obiettivo di queste “teorie limitate” è quello di formulare norme e strategie di comportamento realmente attuabili, tali teorie devono basarsi su un’adeguata *descrizione* delle nostre effettive capacità e delle diverse classi di situazioni in cui ci troviamo a operare.

A partire dagli anni Sessanta del secolo scorso, la teoria dei giochi si è ramificata in diverse direzioni, in parte coincidenti con quella auspicata da Simon. Accanto a importanti sviluppi della teoria classica, sono stati elaborati nuovi approcci, tra loro molto diversi ma accomunati dal fatto di abbandonare almeno uno dei presupposti TGC1-TGC5 della teoria classica. La figura 2.2 illustra schematicamente le relazioni tra le diverse versioni della teoria dei giochi.⁸ Rimanderemo la presentazione più dettagliata degli elementi essenziali della teoria classica al capitolo 3 di questo libro; per quanto riguarda le teorie non classiche, saranno discusse, relativamente ad alcune loro applicazioni, nei capitoli 4-6. Per il momento, sarà sufficiente commentare brevemente la figura 2.2 per spiegare le differenze fondamentali fra i diversi tipi di teoria dei giochi.

In primo luogo, conviene distinguere le teorie che possiamo chiamare “razionalistiche” dalle teorie “evoluzionistiche” dei giochi (cfr. Sugden 2001). Parliamo di *teorie razionalistiche dei giochi* con riferimento alle versioni della teoria dei giochi sviluppate sulla scia di von Neumann e Morgenstern, che interpretano il comportamento dei giocatori come il risultato di scelte razionali. Verso la fine degli anni Settanta le teorie razionalistiche sono state affiancate dalle *teorie evoluzionistiche dei giochi*, di carattere squisitamente descrittivo, sviluppate a partire dalle ricerche di

8 Conviene avvisare nuovamente il lettore che la classificazione e la terminologia qui suggerite non sono standard e che in letteratura gli stessi nomi e altri ancora (come “teoria epistemica dei giochi”, “teoria sperimentale dei giochi”, ecc.) vengono spesso usati, in modo tutt’altro che rigoroso, per designare diversi approcci alla teoria dei giochi.

Figura 2.2. Teorie dei giochi



Maynard Smith (1982). Tali teorie non concepiscono le strategie adottate dai giocatori come il prodotto di scelte razionali, bensì come “schemi di comportamento” acquisiti mediante i meccanismi della trasmissione genetica o culturale. Dopo il 1982, le teorie evoluzionistiche dei giochi sono state ampiamente impiegate nelle scienze sociali, per esempio nell’analisi dell’evoluzione delle norme e delle istituzioni sociali, che è l’oggetto del capitolo 6 di questo libro. Sia le teorie razionalistiche sia quelle evoluzionistiche hanno prodotto diversi approcci all’analisi del comportamento strategico, che descriviamo brevemente di seguito.

Le teorie razionalistiche dei giochi condividono con la teoria classica il fine di individuare, per ogni possibile gioco, la sua soluzione, cioè la combinazione delle strategie ottimali che dovrebbero essere adottate da ciascuno dei giocatori. Chiameremo *teorie classiche dei giochi* tutti gli approcci, che integrano o modificano la teoria originale di von Neumann e Morgenstern, ma che condividono i presupposti TGC1-TGC5 e in particolare l’ultimo (universalismo aprioristico), secondo il quale la razionalità individuale, assieme alla conoscenza comune delle preferenze e della razionalità altrui, è sufficiente a individuare la soluzione di ogni gioco. Ci sono però buoni motivi per dubitare dell’ottimismo razionalistico delle teorie classiche. Uno di questi è legato alla circostanza che, nelle nostre interazioni sociali, non sempre abbiamo a che fare con agenti *perfettamente* razionali. In realtà, siamo quasi sempre alle prese con persone che condividono le nostre stesse limitazioni cognitive e che non di rado agiscono sulla base di abitudini o schemi fissi.

Considerazioni di questo tipo hanno ispirato lo sviluppo di alcune versioni della teoria dei giochi in grado di tenere conto delle effettive “condi-

zioni cognitive” o “epistemiche” in cui si svolgono le interazioni sociali. In particolare, diversi studiosi – a partire dal lavoro pionieristico dell’economista americano Thomas Schelling (nato nel 1921, Nobel nel 2005) negli anni Sessanta del secolo scorso – hanno abbandonato il postulato TGC5, ritenendo che, in generale, la soluzione di un gioco possa venire identificata solo impiegando, assieme ai principi di razionalità della teoria classica, anche appropriati principi contestuali, formulati con riferimento agli aspetti essenziali del “contesto cognitivo” entro il quale si svolge il gioco. Ciò significa, per esempio, che si dovranno considerare le caratteristiche percettive e mentali dei giocatori, il loro orizzonte culturale, le loro opinioni circa il “tipo” di concorrente che devono affrontare, e così via. Queste teorie, che chiamiamo *teorie cognitive dei giochi*, hanno già trovato interessanti applicazioni a diversi problemi delle scienze sociali (si vedano in particolare i capitoli 4 e 5 di questo libro).⁹

Alla famiglia delle teorie razionalistiche dei giochi appartengono anche quelle che, in figura 2.2, abbiamo chiamato *teorie comportamentali dei giochi*. Tali teorie sono caratterizzate dal rifiuto del presupposto TGC1 della teoria classica dei giochi, secondo il quale i giocatori hanno sempre preferenze autointeressate. Soprattutto nel corso degli anni Novanta del secolo scorso, economisti e psicologi hanno cominciato ad accumulare una quantità di dati ormai sterminata, basata su studi sperimentali e ricerche sul campo, sull’effettivo comportamento di individui impegnati in decisioni strategiche.¹⁰ I dati raccolti mostrano come le scelte degli individui in contesti reali si distacchino sistematicamente da quelle che la teoria dei giochi prevedrebbe per agenti razionali astratti. In particolare, i fautori della teoria comportamentale sostengono che gli esseri umani hanno tipicamente preferenze *sociali* o, per dirlo con un termine equivalente, *eterointeressate*. Ciò significa che, di solito, le preferenze di un giocatore non tengono conto solo dei vantaggi materiali da lui acquisiti con un determinato risultato del gioco, ma anche dei vantaggi materiali degli altri giocatori, oppure del processo che ha condotto a quel risultato. Nel primo caso si parla di preferenze altruistiche, nel secondo di preferenze basate sulla reciprocità. Come vedremo nel capitolo 4, entrambi i tipi di preferenza giocano un ruolo essenziale nello spiegare a livello teorico il comportamento decisionale effettivamente osservato negli esperimenti o sul campo.

9 Sulle teorie cognitive dei giochi si vedano soprattutto Schelling (1960, 1978 1984), Sugden (1986) e Bicchieri (1993, 2006).

10 Si vedano Camerer (2003), Bowles (2004), Guala (2005) e Smith (2008) per la discussione dei dati sperimentali e per una rassegna dei principali contributi.

L'altra grande famiglia di teorie dei giochi è quella delle teorie evolucionistiche dei giochi, che può essere caratterizzata dall'abbandono del postulato TGC2 di perfezione razionale della teoria classica e delle sue conseguenze. La prima formulazione di una teoria di questo tipo, indicata in figura 2.2 come *teoria evolucionistica-biologica dei giochi*, si deve a Maynard Smith e ai suoi collaboratori i quali, nell'intento di far luce sull'evoluzione del comportamento animale, combinarono la teoria classica dei giochi con la biologia evolucionistica.¹¹ Tale approccio si basa sull'idea che le interazioni fra animali presentano notevoli somiglianze e altrettanto significative differenze con le interazioni strategiche umane, analizzate nell'ambito delle teorie razionalistiche.

La somiglianza fondamentale consiste nel fatto che in un'interazione animale, proprio come in un'interazione strategica umana, il risultato dell'interazione dipende dalle strategie adottate dai partecipanti. Due importanti differenze riguardano invece la natura delle strategie impiegate e quella delle preferenze relative ai diversi risultati del gioco. In primo luogo, infatti, mentre i partecipanti a un'interazione strategica umana scelgono razionalmente le proprie strategie e valutano i possibili risultati dell'interazione secondo le loro preferenze soggettive, nelle interazioni animali le strategie adottate dai partecipanti sono schemi di comportamento istintivi acquisiti per via ereditaria. Inoltre, gli animali non hanno vere preferenze per i diversi risultati del gioco, ma tutte le ricompense (cibo, territorio, *partner* sessuali e così via) che essi possono trarne si traducono nel complessivo *vantaggio riproduttivo* – cioè nell'accresciuta numerosità della prole – determinato da quelle ricompense.

Queste differenze non hanno impedito di applicare i concetti fondamentali della teoria evolucionistica dei giochi all'analisi del comportamento umano. Al contrario, diversi problemi centrali delle scienze sociali, come l'origine e lo sviluppo della cooperazione sociale e dell'altruismo che discuteremo nel capitolo 6, sono stati affrontati con successo proprio a partire dallo studio dell'evoluzione biologica delle preferenze e dei comportamenti umano. Inoltre, riflettendo sulle identità strutturali tra interazioni animali e umane, alcuni studiosi hanno mostrato come l'apparato formale e i metodi di indagine della teoria evolucionistica dei giochi possono essere applicati anche all'analisi dell'evoluzione *culturale* del comportamento strategico umano. Le loro ricerche hanno permesso di sviluppare versioni

11 Si veda in particolare il pionieristico articolo di Maynard Smith e Price (1973). La prima trattazione sistematica della teoria evolucionistica-biologica dei giochi si trova in Maynard Smith (1982).

delle teorie evoluzionistiche che potremmo chiamare *teorie (evoluzionistico-)adattive dei giochi*.¹²

Le teorie adattive dei giochi si fondano sull'osservazione che, in molte interazioni strategiche umane, il grado di razionalità esibito dai partecipanti si colloca a un livello intermedio tra quelli ipotizzati, rispettivamente, dalle teorie evoluzionistiche e razionalistiche. Ciò significa che, diversamente da quanto accade nelle interazioni animali, descritte dalle teorie evoluzionistiche, nelle interazioni umane i partecipanti sono quasi sempre consapevoli delle loro scelte strategiche. D'altra parte, diversamente da quanto ipotizzato dalle teorie razionalistiche, accade piuttosto spesso che essi non abbiano la possibilità di basare le loro scelte strategiche su una ponderata valutazione razionale della struttura del gioco e del tipo di concorrenti con cui hanno a che fare. Vi sono, infatti, molte situazioni in cui gli esseri umani devono limitarsi a una rapida valutazione intuitiva delle prospettive di successo delle strategie disponibili, a cominciare dalle più semplici e, su questa base, scegliere la strategia da adottare, almeno per un certo lasso di tempo, in qualunque interazione. Vale la pena notare che, come mostra la figura 2.2, le teorie adattive e le teorie cognitive sono strettamente connesse, poiché entrambe mirano a fornire una descrizione più realistica, rispetto a quella fornita dalla teoria classica, dei meccanismi cognitivi che governano il comportamento strategico umano.

Anche se, come abbiamo visto nella nostra breve rassegna, la teoria classica dei giochi è ormai solo uno degli approcci esplorati dagli studiosi interessati al comportamento strategico, e probabilmente non il principale, essa costituisce tuttora un punto di riferimento per tutte le altre teorie dei giochi. Inoltre, tali teorie, lungi dal rovesciare la teoria classica, ne rappresentano un indispensabile completamento. Infatti, i modelli normativi elaborati dalla teoria classica andrebbero intesi come *idealizzazioni*, formulate con l'intento di dirci come dovrebbero comportarsi agenti ideali, con preferenze autointeressate, dotati di una totale perfezione razionale e pratica, e con una conoscenza comune dell'altrui perfezione. Poiché gli esseri umani in carne ed ossa sono molto diversi dagli agenti ideali postulati dalla teoria classica, non sorprende che la diretta applicazione delle idealizzazioni classiche all'analisi del comportamento strategico umano abbia prodotto risultati deludenti. Tuttavia, questo non significa che gli esseri umani sono irrazionali, ma solo che la complessità, emotiva e cognitiva, delle loro interazioni viene in larga parte ignorata dalla teoria classica. Per

12 Si vedano in particolare Axelrod (1984, 1997), Sugden (1986/2004), Bowles (2004) e Gintis (2000, 2009).

comprendere i meccanismi del comportamento strategico umano occorre, quindi, completare la teoria classica con un'adeguata rappresentazione della struttura emotiva e del contesto cognitivo delle interazioni umane. Questo compito è stato affrontato con discreto successo dalle teorie non classiche (specialmente comportamentali e cognitive) dei giochi, le quali hanno consentito di comprendere che il comportamento strategico umano esibisce un grado di ragionevolezza di gran lunga maggiore di quanto non appaia dalla prospettiva della teoria classica. Tali teorie non vanno, quindi, intese come alternative alla teoria classica, bensì come *concretizzazioni*, diverse e complementari, di tale teoria.¹³ Nel prossimo capitolo presentiamo quindi gli elementi essenziali della teoria classica, che costituiscono gli strumenti di base nella cassetta degli attrezzi di qualsiasi teorico dei giochi.

13 Sui concetti di idealizzazione e concretizzazione, qui usati per descrivere le relazioni tra la teoria classica, da un lato, e le teorie comportamentali e cognitive, dall'altro, si veda Niiniluoto (1987, p. 113 e par. 7.4) e Kuipers (2000, par. 10.4).



3. ELEMENTI DI TEORIA CLASSICA DEI GIOCHI

Come abbiamo visto nel capitolo precedente, la teoria dei giochi può venire considerata come una branca della teoria della scelta razionale. Quest'ultima è una teoria di carattere normativo che si propone di determinare le scelte che un individuo razionale dovrebbe compiere, nelle più diverse situazioni, dati i suoi fini e le sue opinioni. La teoria dei giochi analizza tali scelte quando l'individuo è impegnato in un'interazione strategica con altri agenti. Tuttavia, gli strumenti di base di tale analisi sono gli stessi utilizzati per studiare le decisioni di un individuo isolato, in quelli che abbiamo chiamato giochi contro natura. Nel prossimo paragrafo, introduciamo gli elementi essenziali della teoria delle decisioni (seguendo Festa e Crupi 2009), per poi estendere la discussione alla teoria dei giochi nei paragrafi successivi.

3.1 *Giochi contro natura*

Abbiamo già visto (nel capitolo 2) che la teoria della scelta razionale è considerata uno strumento di analisi essenziale nello studio dei fenomeni economici. Non è un caso, quindi, che il contesto delle transazioni commerciali rappresenti un ambito di applicazione privilegiato della teoria e fornisca efficaci illustrazioni dei suoi principi. Supponiamo, per esempio, che siate interessati all'acquisto di una casa e che abbiate a disposizione due *opzioni* alternative, vale a dire l'acquisto di uno fra gli appartamenti *A* e *B*. Supponiamo anche che le uniche differenze rilevanti fra *A* e *B* riguardino la metratura e il prezzo di vendita (si tratta, evidentemente, di una semplificazione alquanto drastica, introdotta a scopi puramente illustrativi). L'appartamento *A* è, poniamo, di 100 mq e costa 300 mila euro; *B* è di 75 mq e costa 200 mila euro. Le opzioni *A* e *B* si distinguono quindi per i loro *esiti*: perseguendo l'una o l'altra, vi troverete a vivere in un ambiente più o meno spazioso affrontando una spesa più o meno impegnativa.

È naturale presumere che la vostra scelta dipenderà da quanto le due opzioni vi appaiono attraenti in base alla desiderabilità dei loro esiti. La teoria della scelta razionale rappresenta questo tipo di valutazione attribuendo al decisore una *funzione di utilità* (solitamente indicata con U) definita sugli esiti delle opzioni fra cui scegliere. L'utilità complessiva di un'opzione sarà poi data dalla somma delle utilità dei suoi esiti, cosicché:

$$(1) \quad \begin{aligned} U(A) &= U(100 \text{ mq}) + U(\text{spesa di 300 mila euro}) \\ U(B) &= U(75 \text{ mq}) + U(\text{spesa di 200 mila euro}) \end{aligned}$$

È bene notare che i valori di utilità sono interpretati in termini del tutto soggettivi, tali quindi da riflettere i desideri e gli obiettivi di un particolare decisore, che possono differire notevolmente da quelli di un altro. La teoria non impone alcun vincolo a questo livello. Ciò che essa prevede, piuttosto, è che un decisore, se agisce razionalmente, sceglierà invariabilmente l'opzione che *massimizza* la sua utilità. Sceglierà di acquistare l'appartamento A , per esempio, se la somma dell'utilità dei suoi esiti è maggiore di quella degli esiti di B . In condizioni ordinarie, naturalmente, l'utilità soggettiva crescerà a fronte di una metratura maggiore e di un prezzo più conveniente (cioè più basso). Nel problema degli appartamenti A e B , tuttavia, non è possibile ottenere il meglio su entrambi i fronti. È invece necessario stabilire quale dei due aspetti far prevalere. In termini tecnici, si dice che il problema decisionale presenta un *trade-off* fra i due attributi rilevanti (l'ampiezza degli appartamenti e la loro convenienza).¹

L'esempio appena illustrato illustra un caso di decisione in *condizioni di certezza*: il problema consiste nel ponderare costi e benefici di due opzioni alternative, ma, una volta valutato il *trade-off* fra le due, l'esito della scelta è sicuro. In altre parole, se l'individuo decide che l'opzione A massimizza la sua utilità, si suppone che la sua scelta determinerà esattamente l'esito atteso, cioè che l'individuo si troverà a vivere in un ambiente più spazioso anche se più costoso. In molti casi, tuttavia, le decisioni che ci troviamo a

1 È tanto semplice spiegare che cosa sia un *trade-off* quanto è difficile tradurre il termine in italiano, al punto che non sembra esservi alcuna traduzione generalmente accettata. In certi contesti, può essere utile riferirsi alla nozione di "conflitto" decisionale; il conflitto, cioè, fra due possibili ragioni per la scelta che spingono in direzioni opposte. Un altro termine talvolta impiegato è quello di "compromesso", come nell'espressione "compromesso qualità-prezzo". Un'ulteriore possibilità è quella di far riferimento al termine "equilibrio", considerando che il decisore deve valutare quale fra le opzioni di scelta realizzi l'"equilibrio" più soddisfacente fra gli attributi rilevanti.

dover prendere sono *in condizioni di incertezza*. Ciò significa che il decisore può ancora valutare l'utilità di ogni esito e quindi di ogni opzione, ma non è sicuro di quale esito si verificherà data la sua scelta. Uno degli aspetti più interessanti della teoria della scelta razionale è che può essere elegantemente estesa a situazioni più complesse di questo tipo, che comprendono elementi di incertezza o di *rischio*.²

Per illustrare questo tipo di casi, è comune rappresentare i problemi di scelta in relazione a diverse possibili scommesse. Facciamo ancora un esempio di scelta fra due opzioni.

(A) Vinci 100 euro se ottieni un sei con un dado bilanciato; altrimenti non vinci nulla.

(B) Vinci 20 euro se ottieni testa con una moneta bilanciata; altrimenti non vinci nulla.

Quale delle due scommesse un decisore razionale dovrebbe scegliere di giocare? In questo scenario, gli esiti possibili delle due opzioni sono in tutto tre: non vincere nulla, vincere 20 euro oppure vincerne 100. A ciascuno di essi un particolare decisore potrà attribuire un certo valore di utilità. Le utilità in questione, tuttavia, non sono di per sé sufficienti per una valutazione complessiva delle due opzioni, perché gli esiti di queste ultime non sono *certi*, bensì dipendenti dall'avverarsi di stati di cose (il risultato del lancio del dado, oppure della moneta) che non sono sotto il controllo del decisore. In questi casi, la quantità che orienta la scelta razionale è detta utilità *attesa* (solitamente indicata con *EU*, per *expected utility*). L'utilità attesa di un'opzione è definita come la somma delle utilità dei suoi possibili esiti, ciascuna moltiplicata per la corrispondente probabilità. Secondo questa definizione:

$$(2) \quad \begin{aligned} EU(A) &= [1/6 \times U(100 \text{ euro})] + [5/6 \times U(0 \text{ euro})] \\ EU(B) &= [1/2 \times U(20 \text{ euro})] + [1/2 \times U(0 \text{ euro})] \end{aligned}$$

2 A partire da Frank Knight (1885-1972), gli economisti distinguono fra i concetti di "incertezza" e "rischio". Il rischio caratterizza situazioni di incertezza, come il lancio di un dado o una puntata alla *roulette*, in cui è possibile specificare le probabilità dei diversi esiti. In una situazione di "reale" incertezza, o di incertezza "radicale", nemmeno le probabilità degli esiti sono note. Di conseguenza, solo il rischio, ma non l'incertezza, è suscettibile di trattamento matematico. La teoria della scelta razionale tende ad ignorare questa distinzione, trattando qualsiasi decisione in condizioni di incertezza come una situazione "rischiosa" nel senso di Knight.

Se i valori di utilità fossero identici alle corrispondenti vincite monetarie, la scommessa *A* risulterebbe maggiormente vantaggiosa, cioè associata a una più elevata utilità attesa (per rendersene conto, è sufficiente notare che $1/6 \times 100$ è maggiore di $1/2 \times 20$). Occorre precisare, tuttavia, che dato il carattere soggettivo della funzione di utilità non c'è motivo per cui tale rapporto di identità debba necessariamente valere (e anzi è in contrasto con le effettive valutazioni di molte persone). Se per esempio vi interessasse soltanto guadagnare abbastanza per pagarvi una pizza stasera, 20 euro sarebbero più che sufficienti. Di conseguenza, la differenza fra la vincita di 20 e quella di 100 euro sarebbe molto ridotta (o nulla) in termini di utilità soggettiva. La scommessa *B* risulterebbe così la più attraente semplicemente per il fatto che offre una più elevata probabilità di vincita. Ancora una volta, quindi, la teoria non definisce la razionalità dei decisori vincolando i loro valori di utilità, ma piuttosto assumendo che essi scelgano in maniera conseguente, vale a dire massimizzando la loro utilità attesa, quale che sia.

Un altro esempio di gioco contro natura in condizioni di incertezza è la scelta che affronta un individuo che deve decidere se prendere o meno l'ombrello uscendo di casa. In questo caso, il decisore può scegliere fra due opzioni:

- (O) Prendi l'ombrello
- (N) Non prendere l'ombrello

Naturalmente, la bontà della scelta fra *O* e *N* dipende dal fatto che, nel corso della giornata, piova (*P*) o sia sereno (*S*). Così, i possibili esiti di questo gioco contro natura sono quattro: prendi l'ombrello e piove (*O,P*), prendi l'ombrello ma non piove (*O,S*), non prendi l'ombrello e piove (*N,S*), non prendi l'ombrello e non piove (*N,S*). Ogni esito del gioco ha, in generale, una diversa utilità per il decisore, che può venir utilmente inserita in una tabella o *matrice* come quella in figura 3.1.

La scelta del decisore dipenderà dalle sue preferenze relative ai diversi esiti del gioco. Un possibile (e forse per molti di noi plausibile) ordinamento delle quattro utilità nel gioco dell'ombrello è il seguente:

$$(3) \quad U(N,S) > U(O,S) > U(O,P) > U(N,P)$$

In altre parole, un decisore con l'ordinamento di utilità qui sopra preferisce sempre il sereno alla pioggia ma, nel caso in cui sia sereno preferisce non avere l'ombrello, mentre nel caso in cui piova preferisce averlo.

	Piove (P)	Non piove (S)
Ombrello (O)	$U(O,P)$	$U(O,S)$
Niente ombrello (N)	$U(N,P)$	$U(N,S)$

Figura 3.1.
Un gioco contro natura

Secondo la teoria delle decisioni, la scelta dovrebbe cadere sull'opzione che massimizza l'utilità attesa del decisore. A sua volta, l'utilità attesa di un'opzione è la somma delle utilità dei suoi possibili esiti, ciascuna moltiplicata per la corrispondente probabilità, cioè:

$$(4) \quad \begin{aligned} EU(O) &= [p(P) \times U(O,P)] + [p(S) \times U(O,S)] \\ EU(N) &= [p(P) \times U(N,P)] + [p(S) \times U(N,S)] \end{aligned}$$

Abbiamo indicato con $p(P)$ e $p(S)$ rispettivamente la probabilità che piova e che sia sereno. Se, per semplificare, supponiamo che “sereno” significhi “non piove”, allora la somma $p(P) + p(S)$ deve essere uguale a 1, cioè al massimo grado di probabilità. Per calcolare l'effettiva utilità attesa delle sue opzioni, il decisore deve fornire dei valori numerici sia alle sue utilità sia alle due probabilità in gioco. Per quanto riguarda queste ultime, per esempio, il decisore potrebbe aver sentito al telegiornale che la probabilità di pioggia è solo del 30% – cioè che $p(P) = 0,3$ e quindi $p(S) = 1 - p(P) = 0,7$. Per quanto riguarda le seconde, il decisore può assegnare alle proprie utilità qualsiasi valore numerico che rispetti l'ordinamento (3) considerato sopra, per esempio:

$$(5) \quad \begin{aligned} U(N,S) &= 50 \\ U(O,S) &= 20 \\ U(O,P) &= 10 \\ U(N,P) &= 1 \end{aligned}$$

Inserendo questi valori nel calcolo dell'utilità attesa (4) otteniamo così (approssimando i risultati):

$$(6) \quad EU(O) = [0,3 \times 10] + [0,7 \times 20] = 17$$

$$EU(N) = [0,3 \times 1] + [0,7 \times 50] = 35$$

Dato che l'utilità attesa di N è maggiore di quella di O , un decisore razionale, sulla base delle sue preferenze soggettive e delle previsioni meteorologiche a sua disposizione, dovrebbe uscire di casa senza ombrello.

3.2 Giocatori, mosse e strategie

La teoria dei giochi utilizza gli strumenti della teoria delle decisioni – e in particolare il principio di massimizzazione dell'utilità attesa – per studiare le scelte di individui impegnati in interazioni strategiche. Tali scelte sono sempre prese in condizioni di incertezza, per il semplice motivo che un individuo non può controllare, né prevedere con certezza, le scelte degli altri individui, e quindi essere sicuro dell'esito di ogni sua decisione (cfr. Luce e Raiffa 1957, p. 14).

Un classico esempio di interazione strategica di questo tipo è il familiare gioco della “morra cinese”, o “sasso-carta-forbici”. In questo gioco, molto diffuso tra i bambini, si fronteggiano due giocatori che devono scegliere simultaneamente, annunciandola ad alta voce, una delle tre mosse “sasso”, “carta” o “forbici”. Se entrambi scelgono la stessa mossa, il turno di gioco finisce in pareggio (0 punti per entrambi). Se invece i due giocatori scelgono mosse diverse, il vincitore guadagna 1 punto e l'avversario ne perde 1, secondo la regola: “sasso” batte “forbici” (le spunta), “carta” batte “sasso” (lo avvolge) e “forbici” batte “carta” (la taglia). Il gioco della morra cinese può quindi essere rappresentato dalla matrice in figura 3.2.

		COLONNA		
		Sasso	Carta	Forbici
RIGA	Sasso	0 0	1 -1	-1 1
	Carta	-1 1	0 0	1 -1
	Forbici	1 -1	-1 1	0 0

Figura 3.2.
La morra cinese

Nella figura 3.2, i due giocatori sono chiamati, per comodità, “Riga” e “Colonna”. Così, ogni riga della tabella rappresenta una possibile mossa del primo giocatore, Riga, e ogni colonna una mossa del secondo, Colonna. Le nove celle della matrice rappresentano i possibili esiti o risultati del gioco, cioè le nove possibili combinazioni delle mosse di Riga e di Colonna. All’interno di ogni cella compaiono due numeri che rappresentano le utilità o i *payoff* ottenuti da ognuno dei due giocatori nel caso in cui quella cella corrisponda al risultato del gioco; il numero in basso a sinistra è il payoff di Riga, mentre quello in alto a destra è il payoff di Colonna.³

La morra cinese è un esempio perfetto di semplice interazione strategica, in cui le scelte di ogni giocatore dipendono dalle attese circa la scelta altrui. Infatti, è chiaro che, a ogni turno di gioco, Riga ha un’unica scelta ottimale, che è però determinata dalla scelta di Colonna, e dalle attese di Riga su di essa; e viceversa. Se per esempio Riga sapesse che Colonna giocherà “forbici”, la sua unica scelta razionale sarebbe chiaramente scegliere “sasso”. Per questo motivo, chiunque abbia mai giocato alla morra cinese sa bene che non conviene mai “fissarsi” su un’unica mossa, perché, appena se ne accorgesse, l’avversario vincerebbe facilmente a ogni turno di gioco, scegliendo la contromossa migliore. Né è sufficiente cambiare a ogni turno la propria mossa secondo una regola fissa. Se per esempio Colonna si accorge che Riga gioca sempre, nell’ordine, “sasso-carta-forbici”, può facilmente vincere scegliendo di giocare sempre, nell’ordine, “carta-forbici-sasso”.

L’unica cosa ragionevole da fare è allora giocare, a ogni turno, una mossa “a caso”. Per esempio, Riga potrebbe lanciare un dado e scegliere “sasso” se esce 1 o 2, “carta” se esce 3 o 4 e “forbici” se esce 5 o 6. In questo caso, in cui Riga gioca con una certa probabilità ognuna delle mosse a sua disposizione, si dice che Riga adotta una *strategia mista*. Se invece, a ogni turno di gioco, Riga sceglie direttamente una delle sue mosse, diciamo che adotta una *strategia pura*. Si noti che, mentre il numero delle strategie pure è determinato dalla struttura del gioco, un giocatore ha sempre a disposizione un’infinità di strategie miste. Nel caso della morra cinese, ogni giocatore ha solo 3 strategie pure (cioè “sasso”, “carta” e “forbici”) ma infinite strategie miste, una delle quali è giocare ogni mossa con probabilità $1/3$. Qualsiasi strategia che assegna alle probabilità $p(\text{sasso})$, $p(\text{carta})$ e

3 Il termine *payoff* viene talvolta tradotto con “premio”, “incentivo” o “vincita”. Tuttavia, dato che i payoff possono essere negativi – come nel caso della morra cinese – preferiamo usare il termine inglese originale, che può riferirsi sia alle “vincite” sia alle “perdite” dei giocatori.

p (forbici) valori che assommino a 1 – per esempio non giocare mai carta e tirare una moneta per scegliere fra sasso e forbici – è una strategia mista. Tuttavia, come vedremo fra poco, non tutte le strategie (pure o miste) sono ugualmente vantaggiose dal punto di vista del giocatore che le adotta. Nella morra cinese, per esempio, si può dimostrare che l'unica strategia razionale di gioco è, per ogni giocatore, scegliere ogni mossa con $1/3$ di probabilità (von Neumann 1928; cfr. Osborne 2003, p. 65-66). Si tratta di un caso in cui le previsioni teoriche si accordano con, e confermano, le nostre intuizioni su quale sia la miglior strategia di gioco da adottare.

3.3 Coordinazione e conflitto

La morra cinese è un gioco con due giocatori, ognuno dei quali ha le stesse tre mosse a disposizione. La teoria dei giochi studia però qualsiasi tipo di gioco, con un qualsiasi numero di giocatori, ognuno dei quali può adottare un qualsiasi numero di mosse, eventualmente diverso dal numero di mosse di ogni altro giocatore. A ogni turno di una partita di scacchi, per esempio, i due giocatori possono applicare le varie mosse previste dalle regole del gioco ai diversi pezzi rimasti sulla scacchiera. In un mercato, un numero indefinito di venditori e compratori adotta diverse strategie commerciali; allo stesso modo, diversi automobilisti hanno varie mosse a disposizione per destreggiarsi nel traffico cittadino; e così via. Come vedremo, ci sono diversi modi di ordinare e classificare la grande varietà di situazioni studiate dalla teoria dei giochi.

Un primo criterio riguarda il numero dei giocatori coinvolti; in particolare, possiamo distinguere fra giochi a due e giochi a tre o più giocatori. Lo studio dei giochi a n giocatori – con $n > 2$ – è complicato dalla necessità di prendere in considerazione non solo le strategie di ogni giocatore e tutti i possibili esiti del gioco determinati dalle loro combinazioni, ma anche le possibili coalizioni fra i giocatori. Per questo motivo, una trattazione elementare della teoria dei giochi si concentra solitamente sui giochi a due giocatori. Per illustrare tutti i concetti essenziali della teoria, è anzi sufficiente considerare il più semplice tipo di gioco in assoluto: cioè, giochi a due giocatori con sole due mosse a disposizione (spesso uguali fra loro). Nel seguito del capitolo, faremo quindi riferimento quasi esclusivamente a giochi di questo tipo.

Un ulteriore, importante criterio di classificazione dei giochi è quello relativo alla cosiddetta *struttura dei payoff*. La struttura dei payoff di un gioco è determinata dall'*ordinamento delle preferenze* dei giocatori rela-

tivo ai possibili esiti del gioco. In alcuni casi potrà accadere che entrambi i giocatori abbiano lo stesso ordinamento di preferenza rispetto a tutti gli esiti del gioco; in altri casi potranno avere ordinamenti parzialmente coincidenti o anche totalmente opposti. Per esempio, nel gioco della morra cinese (figura 3.2) le preferenze dei due giocatori sono opposte rispetto a tutti gli esiti del gioco, eccetto nei casi di pareggio. In molti giochi da tavolo, come gli scacchi o il poker, i giocatori sono *avversari*, nel senso che, qualsiasi sia l'esito del gioco (escludendo il caso del pareggio), un giocatore vince e l'altro perde. D'altra parte, in moltissime altre situazioni, come una contrattazione commerciale o la guida nel traffico, un giocatore ha spesso alcune preferenze opposte, ma altre coincidenti, con quelle degli altri giocatori.

Giochi con totale coincidenza di interessi

Chiameremo giochi con *totale coincidenza di interessi* i giochi caratterizzati da una struttura di payoff simile a quella del gioco in figura 3.3. I due giocatori, Riga e Colonna, hanno ognuno due mosse a disposizione: R1 ed R2 per Riga, C1 e C2 per Colonna. (Non è importante specificare quali siano le mosse dei giocatori; vedremo un esempio fra un momento.)

In questo gioco i giocatori hanno lo stesso ordinamento di preferenza rispetto a tutti i quattro possibili esiti del gioco: per entrambi il risultato preferito è (R2,C2), seguito da (R2,C1), da (R1,C1) e, infine, da (R1,C2), che è il risultato peggiore per entrambi. In altre parole, i giocatori sono sempre concordi nel valutare i diversi esiti del gioco, nel senso che per nessun esito del gioco c'è fra di loro un conflitto di interessi. Giochi di questo tipo vengono anche chiamati "giochi di pura collaborazione" o "giochi di coordinazione pura" (Schelling 1960, pp. 98-104).

Un semplice gioco di pura coordinazione è quello che potremmo chiamare il "gioco della guida". Supponiamo che due automobilisti, Riga e Colonna, si incontrino, provenendo da direzioni opposte, lungo una strada stretta. Per evitare lo scontro, occorre che entrambi i giocatori accostino sul lato della carreggiata. Per entrambi è indifferente se accostare a destra (D) o a sinistra (S), ma è essenziale scegliere lo stesso lato dell'altro giocatore. La matrice del gioco della guida è illustrata in figura 3.4.

In questo gioco, entrambi i giocatori concordano su quali siano gli esiti peggiori – (D,D) e (S,S) – e quelli migliori – (D,S) e (S,D) e sono indifferenti rispetto allo specifico esito di ogni coppia. In altre parole, c'è perfetta coincidenza di interessi fra i giocatori, il cui solo problema è coordinarsi fra loro in modo da ottenere una delle due soluzioni migliori per entrambi.

		COLONNA	
		C1	C2
RIGA	R1	20 1	10 0
	R2	30 2	40 3

		COLONNA	
		Destra (D)	Sinistra (S)
RIGA	Destra (D)	0 0	10 10
	Sinistra (S)	10 10	0 0

Figura 3.3.
Gioco con totale
coincidenza di interessi

Figura 3.4.
Il gioco della guida

Giochi con totale conflitto di interessi

Il caso opposto a quello dei giochi di pura coordinazione è quello dei giochi con *totale conflitto di interessi*, cioè con una struttura di payoff simile a quella del gioco in figura 3.5.

In questo gioco, l'ordinamento delle preferenze di un giocatore è esattamente opposto a quello dell'altro: il risultato preferito da Riga corrisponde al risultato peggiore per Colonna, e viceversa. Nei giochi con totale conflitto di interesse qualsiasi forma di collaborazione fra i giocatori risulta quindi impossibile. Per questo, tali giochi vengono talvolta anche chiamati "giochi di puro conflitto" o "giochi (puramente) competitivi" (Schelling, 1960).

Un interessante caso storico di gioco di puro conflitto è quello della Battaglia del Mare di Bismarck fra Alleati e Giapponesi (1943) nella seconda guerra mondiale, che viene rappresentato in forma di gioco nella figura 3.6.⁴

4 Questo gioco è stato originariamente descritto dal Colonnello Olivier G. Haywood e viene spesso utilizzato come esempio dai teorici dei giochi (cfr. per esem-

		COLONNA	
		C1	C2
RIGA	R1	30 1	10 3
	R2	40 0	20 2

Figura 3.5.
Gioco con totale
conflitto di interessi

		COLONNA	
		Rotta Nord (N)	Rotta Sud (S)
RIGA	Rotta Nord (N)	-2 2	-2 2
	Rotta Sud (S)	-1 1	-3 3

Figura 3.6.
La battaglia del Mare
di Bismarck: Alleati
(Riga) contro Giapponesi
(Colonna)

Per rinforzare il proprio contingente, i Giapponesi (Colonna) devono scegliere se inviare la propria flotta lungo una rotta a Nord (N) dell'isola di Nuova Britannia, oppure a Sud (S). Contemporaneamente, il comando alleato (Riga) deve scegliere dove inviare (N o S) il grosso delle proprie forze aeree per cercare di intercettarli. Entrambe le rotte richiedono tre giorni di navigazione, ma con un'importante differenza: a Sud il tempo è sereno, mentre a Nord è coperto e piovoso. Le diverse condizioni meteorologiche influiscono sulla capacità dei bombardieri alleati di colpire la flotta. Per esempio, se la flotta giapponese venisse intercettata dal grosso dello stormo alleato – cioè se l'esito del gioco fosse (N,N) o (S,S) – dovrebbe subire 3 interi giorni di bombardamenti sulla rotta Sud, ma, grazie alla scarsa visibilità, solo 2 sulla

pio Luce e Raiffa 1957, pp. 64-65; Davis 1983, pp. 13-14; Binmore 1990, p. 2). Haywood (1954, pp. 366-371) ricostruisce in termini giochistici la vera battaglia avvenuta presso l'isola di Nuova Britannia, concludendo che i due comandati, il Generale alleato Kenney e il contrammiraglio giapponese Masatomi Kimura, si comportarono esattamente come previsto dalla teoria di von Neumann e Morgenstern (si veda più avanti, nel paragrafo 3.4, la soluzione del gioco secondo la teoria classica).

rotta Nord. Nella matrice in figura 3.6, i payoff dei due giocatori sono quantificati come giorni di bombardamento che Colonna deve subire nei diversi casi. La matrice mostra chiaramente come Riga e Colonna abbiano interessi diametralmente opposti: Riga vorrebbe fare la stessa scelta di Colonna, che invece vorrebbe evitarlo a tutti i costi. Qualsiasi sia l'esito delle loro scelte finali, il payoff di Riga è esattamente l'opposto di quello di Colonna.⁵

Giochi con parziale coincidenza di interessi

I giochi con totale conflitto di interessi e quelli con totale coincidenza di interessi rappresentano i due casi estremi di un *continuum* di diverse strutture dei payoff. Fra questi due estremi, si trovano tutti i giochi *con parziale coincidenza di interessi*, nei quali le preferenze dei giocatori coincidono rispetto ad alcuni esiti, ma non rispetto ad altri. La struttura dei payoff è simile, in questo caso, a quella del gioco in figura 3.7, in cui, per esempio, entrambi i giocatori preferiscono (R2,C2) a (R2,C1), ma Riga preferisce (R2,C1) a (R1,C1) mentre Colonna preferisce il secondo esito al primo.

In questo gioco, gli interessi dei giocatori sono in parte coincidenti e in parte opposti. Si realizza quindi una sorta di situazione intermedia fra i due estremi della totale coincidenza e del totale conflitto di interessi. Per questo, Schelling (1960, p. 104) chiama questi casi “giochi a interessi misti”, o “giochi misti” *tout court*. È importante notare che la parziale coincidenza (o, equivalentemente, parziale conflitto) di interessi caratterizza la struttura dei payoff della maggior parte dei giochi più comuni e, ciò che più qui ci interessa, dei giochi usati nell'analisi delle interazioni sociali.

Il gioco noto come “la battaglia dei sessi” (Luce e Raiffa, pp. 90-91) è un esempio di gioco misto. In questo gioco (figura 3.8), Riga e Colonna sono due fidanzati che devono decidere dove passare assieme la serata.

Riga preferirebbe andare al museo (M), per l'inaugurazione di una mostra di arte concettuale, mentre Colonna preferirebbe andare al cinema (C),

5 Per questo motivo, in letteratura, i giochi di puro conflitto vengono spesso chiamati “giochi a somma costante” o “giochi a somma zero”, con riferimento alla *somma* dei payoff dei giocatori in ogni esito del gioco. Per esempio, in giochi come il poker, un giocatore vince esattamente la somma delle perdite degli altri giocatori. In effetti, si può dimostrare che, in qualsiasi gioco di puro conflitto, è possibile assegnare ai payoff dei giocatori valori tali che (i) rispettano l'ordinamento di utilità del gioco, cioè la struttura dei payoff e (ii) assommano a 0 in ogni cella del gioco (cfr. Luce e Raiffa 1957, pp. 63-64). Questa terminologia può però risultare fuorviante, poiché, come mostra un confronto fra le matrici nelle figure 3.5 e 3.6, a essere rilevante non è la somma numerica dei payoff quanto appunto il loro ordinamento.

		COLONNA	
		C1	C2
RIGA	R1	30 1	20 0
	R2	10 2	40 3

Figura 3.7.
Gioco con parziale
coincidenza di interessi

		COLONNA	
		Museo (M)	Cinema (C)
RIGA	Museo (M)	2 3	1 1
	Cinema (C)	1 1	3 2

Figura 3.8.
La battaglia dei sessi

a vedere l'ultimo film di Bruce Willis appena uscito. Entrambi tuttavia, preferiscono passare assieme la serata piuttosto che uscire separati. Come mostra la matrice, gli interessi di Riga e Colonna sono in parte in accordo e in parte in conflitto. I giocatori concordano infatti sui risultati peggiori del gioco – cioè (M,C) e (C,M), in cui i due escono da soli – ma sono in disaccordo sul risultato migliore – (M,M) per Riga e (C,C) per Colonna. Il problema è quindi coordinarsi in modo da scegliere uno fra questi due risultati.

3.4 Strategie dominanti, equilibri e soluzioni

Lo scopo della teoria classica dei giochi è quello di identificare, sulla base di appropriati criteri generali di razionalità, la *strategia ottimale* di ciascun giocatore in ogni possibile gioco. La combinazione delle strategie ottimali di tutti i giocatori viene chiamata *soluzione* del gioco. Possiamo quindi dire che l'ambizioso (e non ancora raggiunto) obiettivo della teoria

dei giochi è quello di identificare la soluzione di qualunque gioco possibile, compresi ovviamente quelli discussi nel precedente paragrafo.

Strategie dominanti

Che cos'è una strategia ottimale? Un esempio molto chiaro è la cosiddetta *strategia dominante*. Supponiamo che, in un gioco a due giocatori, X e Y siano due fra le strategie a disposizione di Riga. Diremo allora che X domina Y se, *qualunque sia* la strategia adottata da Colonna, il risultato ottenuto da Riga adottando X non è mai peggiore, e in alcuni casi è persino migliore, di quello che otterrebbe adottando Y . Ciò significa che, adottando X , Riga ottiene payoff almeno pari e, in qualche caso, superiori ai payoff che avrebbe ottenuto adottando Y .

Consideriamo per esempio il gioco della battaglia fra Alleati e Giapponesi, che per comodità riportiamo in figura 3.9.

Osservando la matrice del gioco, possiamo vedere che Colonna (cioè il comando giapponese) ha una strategia dominante (segnalata in grassetto), cioè scegliere la rotta a Nord. Infatti, se Riga (cioè gli Alleati) scegliessero la rotta a Nord, Colonna potrebbe scegliere indifferentemente la rotta a Nord o quella a Sud (subirebbe in entrambi i casi 2 giorni di bombardamento); mentre se gli Alleati scegliessero la rotta Sud, a Colonna converrebbe senz'altro scegliere la rotta a Nord (1 giorno di bombardamento invece di 3). In altre parole, la strategia N di Colonna domina la sua strategia S , perché, rispetto a questa, garantisce a Colonna lo stesso payoff se Riga sceglie N e un payoff migliore se Riga sceglie S . Poiché in questo caso N domina qualsiasi altra strategia a disposizione di Colonna, N è la sua strategia dominante: qualsiasi cosa faccia l'avversario, ai giapponesi conviene prendere la rotta a Nord. Si noti, invece, che Riga non ha alcuna strategia

		COLONNA	
		Rotta Nord (N)	Rotta Sud (S)
RIGA	Rotta Nord (N)	2 -2	2 -2
	Rotta Sud (S)	1 -1	3 -3

Figura 3.9.
La battaglia del Mare di Bismarck: Colonna ha una strategia dominante (N)

dominante: infatti, se Colonna sceglie N , Riga otterrà un payoff migliore scegliendo N , mentre se Colonna gioca S a Riga converrà scegliere S .

Un principio di scelta largamente accettato è il *principio di dominanza*, secondo il quale un giocatore razionale non dovrebbe mai scegliere una strategia dominata da qualche altra sua strategia. Scegliere strategie dominate significa, infatti, garantirsi payoff sicuramente non migliori, e talvolta peggiori, di quelli che si potrebbe ottenere scegliendo un'altra strategia. Dal principio di dominanza segue che, se un giocatore possiede una strategia dominante, allora dovrebbe adottarla, indipendentemente dalle sue opinioni su quello che farà l'altro giocatore. Possiamo quindi dire che, se un giocatore ha una strategia dominante, questa rappresenta la sua strategia ottimale.

Si noti che, quando almeno uno dei due giocatori possiede una strategia dominante, il problema di determinare la soluzione del gioco è facilmente risolvibile. Supponiamo infatti che, come nel gioco della battaglia fra Alleati e Giapponesi, Colonna abbia una strategia dominante. Se ammettiamo che entrambi i giocatori siano perfettamente razionali (si ricordi il postulato TGC2 della teoria classica), per il principio di dominanza Colonna adotterà la sua strategia dominante. Inoltre, dato che sia le preferenze sia la razionalità dei giocatori è conoscenza comune fra di essi (postulato TGC4), Riga può prevedere che Colonna sceglierà tale strategia e quindi adotterà la strategia che costituisce la risposta ottimale alla strategia del primo. La combinazione di queste due strategie determina quindi l'esito del gioco. L'unica soluzione razionale del gioco fra Alleati e Giapponesi in figura 3.9 è così quella in cui entrambi scelgono la rotta Nord (abbiamo evidenziato in grassetto i payoff della cella che rappresenta la soluzione del gioco).

Questo esempio mostra l'importanza del postulato di conoscenza comune TGC4. Infatti, dato che la "bontà" della strategia adottata da Riga dipende dalle strategie adottate da Colonna (e viceversa), nella scelta razionale della propria strategia, Riga dovrà in qualche modo basarsi sulle sue attese circa la scelta di Colonna, che a sua volta dipenderà dalle attese di Colonna circa la scelta di Riga. È allora naturale chiedersi se la dipendenza delle scelte dei giocatori dalla loro attese circa le attese degli avversari non dia inizio a una sequenza infinita di *attese di attese*, dalla quale non si capisce come un giocatore possa giungere all'identificazione delle sua strategia ottimale. Il postulato TGC4 permette appunto ai giocatori, in molti tipi di gioco, di anticipare con certezza la strategia adottata dagli avversari e, di conseguenza, di scegliere la loro strategia senza lasciarsi impigliare nella spirale delle attese di attese.

L'equilibrio di Nash

In moltissime situazioni, nessun giocatore dispone di una strategia dominante. Non è quindi altrettanto facile determinare la soluzione del gioco, poiché la risposta ottimale di ogni giocatore alla strategia dell'avversario dipende dalla scelta di quest'ultimo, che è imprevedibile. Per esempio, è facile controllare che sia nella morra cinese (figura 3.2), sia nel gioco della guida (figura 3.4), sia in quello della battaglia dei sessi (figura 3.8), nessuna delle strategie dei due giocatori domina l'altra. Per analizzare situazioni di questo tipo si ricorre al concetto di *equilibrio di Nash*, che è la nozione centrale della teoria classica dei giochi, poiché definisce cosa si intende per "soluzione" di un gioco in generale. Questo concetto è stato introdotto negli anni Cinquanta dal matematico americano, e Nobel per l'economia nel 1994, John Nash (nato nel 1928), uno dei pionieri della teoria classica dei giochi.

In un gioco a due giocatori, la strategia *X* di Riga e la strategia *Y* di Colonna formano un equilibrio di Nash – o, semplicemente, sono in equilibrio – quando l'una rappresenta la risposta ottimale all'altra. Se *X* e *Y* sono in equilibrio, nessuno dei giocatori, dopo essere venuto a conoscenza della strategia adottata dall'altro, avrà motivo di pentirsi della propria scelta. In altre parole, se anche i giocatori avessero la possibilità di cambiare *unilateralmente* la propria scelta dopo aver visto quella dell'altro giocatore, nessuno dei due avrebbe interesse a farlo.

Consideriamo, per esempio, il gioco in figura 3.10 (Hargreaves Heap e Varoufakis 2004, p. 52).

		COLONNA		
		C1	C2	C3
RIGA	R1	9 10	0 0	10 9
	R2	0 0	1 1	0 0
	R3	10 9	0 0	9 10

Figura 3.10.
Gioco con un unico
equilibrio di Nash (R2,C2)

È facile controllare che nessuno dei due giocatori ha una strategia dominante. Tuttavia, possiamo notare che la cella (R2,C2) – i cui payoff sono evidenziati in grassetto – rappresenta l'unico equilibrio di Nash del gioco. Supponiamo infatti che, dopo aver scelto le proprie mosse, i giocatori si trovino nella cella (R2,C2) e che a ognuno dei due venga offerta la possibilità di cambiare, indipendentemente dall'altro, la propria mossa. Data la mossa C2 di Colonna, a Riga non conviene modificare la propria scelta (R2), perché con entrambe le altre sue mosse otterrebbe un payoff inferiore (4 e 1, rispettivamente per R2 e R3, invece di 7). In altre parole, per Riga R2 è la miglior risposta a C2. Allo stesso modo, si può vedere che per Colonna C2 è la miglior risposta a R2. Di conseguenza, nessun giocatore ha interesse a modificare unilateralmente la propria scelta, una volta che il risultato (R2,C2) sia stato raggiunto: questo risultato rappresenta quindi un "punto d'equilibrio" del gioco. (Con un pò di pazienza, si può controllare che è anche l'unico equilibrio nel gioco in figura 3.10.)

In un equilibrio di Nash, la strategia adottata da ciascuno dei giocatori è la risposta ottimale alla strategia dell'altro. Per questo motivo, appare naturale richiedere che la soluzione di un gioco debba essere un equilibrio di Nash. Questo è il significato di quello che potremmo chiamare *principio di Nash*: se un gioco ha un unico equilibrio, tale equilibrio è la soluzione del gioco. Uno dei principali risultati ottenuti da Nash (nel 1951) è la dimostrazione del fatto che qualsiasi gioco *finito* – cioè con un numero finito di giocatori, ognuno dei quali può scegliere fra un numero finito di mosse – possiede *almeno* un equilibrio di Nash. Come caso particolare, si può anche dimostrare che tutti i giochi di puro conflitto hanno un *unico* equilibrio di Nash, che permette quindi di identificare univocamente la soluzione del gioco. Per esempio, nel gioco della battaglia del Mare di Bismarck (figura 3.9) l'unico equilibrio coincide con la soluzione già trovata ragionando sulla base delle strategie dominanti dei giocatori.

Se tutti i giochi di puro conflitto hanno un unico equilibrio, qual è la soluzione della morra cinese (figura 3.2)? Come abbiamo già osservato alla fine del paragrafo 3.2, essa coincide con la soluzione intuitiva del gioco: scegliere ognuna delle tre mosse "sasso", "carta" e "forbici" con pari probabilità, cioè $1/3$ (cfr. von Neumann 1928). Questo è un esempio di equilibrio di Nash *in strategie miste*, mentre l'equilibrio del gioco in figura 3.10 era *in strategie pure*. Il risultato fondamentale di Nash, per cui tutti i giochi finiti hanno almeno un equilibrio, vale solo se si ammette la possibilità che i giocatori adottino strategie miste: infatti, esistono moltissimi giochi, fra i quali appunto la morra cinese, che non hanno alcun equilibrio in strategie pure.

Se si accettano i postulati TGC1-TGC4 della teoria classica dei giochi, l'equilibrio di Nash appare come un'eccellente concetto di soluzione, soprattutto per i giochi di puro conflitto. Infatti, se le preferenze dei giocatori sono conoscenza comune, come pure il fatto che tutti sono razionalmente e praticamente perfetti, ogni giocatore saprà qual è l'equilibrio del gioco e saprà anche che gli altri giocatori lo sanno; di conseguenza, ciascun giocatore attuerà la strategia che contribuisce alla formazione dell'equilibrio, del tutto fiducioso che anche gli altri si comporteranno nello stesso modo. Identificare la soluzione di ogni gioco con un equilibrio di Nash ha appunto il grande vantaggio che tale concetto cattura il carattere "auto-vincolante" di una "buona" soluzione (cfr. Hargreaves Heap *et al.* 1992, pp. 102-103). Supponiamo infatti che, prima del gioco, i giocatori possano accordarsi su quale strategia sceglierà ognuno di essi. Se, e solo se, i giocatori si accordano sulle strategie che formano un equilibrio di Nash, allora il loro accordo sarà auto-vincolante, nel senso che nessuno avrà alcun incentivo a cambiare la propria strategia e quindi ognuno potrà fidarsi della scelta altrui.

3.5 Due problemi per la teoria classica dei giochi

Il concetto di equilibrio di Nash risponde al problema centrale della teoria classica dei giochi, quello di determinare, sulla base di pochi principi di razionalità generale, la soluzione di qualsiasi possibile gioco. Se le strategie dei giocatori sono in equilibrio, infatti, ognuno ottiene dal gioco il massimo che può aspettarsi date le scelte razionali degli altri giocatori. E dato che, come dimostrato da Nash, ogni gioco finito possiede un punto di equilibrio, tale concetto sembra offrire una soluzione del tutto generale al problema centrale della teoria dei giochi. Inoltre, in alcuni casi, come per esempio quello della battaglia del Mare di Bismarck (figura 3.6), è possibile mostrare che individui reali, impegnati in interazioni rappresentabili in forma astratta come giochi finiti, si comportano effettivamente secondo i dettami della teoria dei giochi, scegliendo strategie che li conducono a un punto di equilibrio.

Tuttavia, la teoria basata sul concetto di equilibrio di Nash deve affrontare almeno due problemi.⁶ In primo luogo, una vastissima classe di casi, in particolare quelli di maggior rilevanza per le scienze sociali, richiede di

6 Si vedano Hargreaves Heap *et al.* (1996) e Hargreaves Heap e Varoufakis (2004, par. 2.5.3 e 3.5) per una rassegna delle principali critiche al concetto di equilibrio di Nash come soluzione generale di un gioco.

considerare giochi che esibiscono una molteplicità di equilibri di Nash. Si propone così il problema di identificare la soluzione di un gioco quando questo ha due o più equilibri di Nash. In secondo luogo, può accadere, e accade di fatto in casi di grande interesse, che un gioco abbia un unico equilibrio di Nash, ma ci siano forti motivi di dubitare che esso rappresenti una “buona soluzione” del gioco stesso. Prima di concludere il capitolo, discutiamo brevemente entrambi questi problemi.

Molteplicità di equilibri di Nash

Consideriamo di nuovo il gioco della guida in figura 3.4. Ognuno dei due automobilisti deve scegliere se accostare sul lato destro (D) o sinistro (S) della strada: non ha alcuna preferenza particolare per una delle due opzioni, ma vuole evitare ad ogni costo lo scontro con l'altro automobilista. Un'occhiata alla matrice del gioco mostra che nessun giocatore dispone di una strategia dominante, ma vi sono due equilibri di Nash, che corrispondono alle celle (D,D) e (S,S). In queste due celle, infatti, la strategia di ogni giocatore è la risposta ottimale a quella dell'altro. In altre parole, se il gioco termina con uno di questi due risultati, entrambi i giocatori hanno ragione di essere soddisfatti della propria mossa, avendo evitato lo scontro.

Un discorso simile vale per il gioco della battaglia dei sessi in figura 3.8. Anche in questo caso, ci sono due equilibri di Nash, nei quali la coppia di fidanzati va assieme al museo (come avrebbe voluto lei) oppure va assieme al cinema (come avrebbe preferito lui). C'è però un'importante differenza fra il gioco della guida e la battaglia dei sessi. Nel primo, una volta raggiunto uno qualsiasi dei due equilibri, entrambi i giocatori sono ugualmente soddisfatti del risultato raggiunto, e lo sarebbero stati anche nell'equilibrio alternativo. Nella battaglia dei sessi, invece, i due equilibri non sono equivalenti dal punto di vista dei giocatori. Infatti, se alla fine la coppia va al museo, lui sarà meno soddisfatto di lei, e viceversa se la coppia va al cinema. In altre parole, i due giocatori hanno preferenze opposte riguardo ai due equilibri del gioco.

In generale, molti giochi non puramente conflittuali, cioè giochi con parziale o totale coincidenza di interessi – come appunto, rispettivamente, la battaglia dei sessi e il gioco della guida – hanno più di un equilibrio. Si noti che questo è il genere di giochi di maggior rilevanza per le scienze sociali, poiché le più interessanti interazioni sociali, compresa (come vedremo nel capitolo 5) la guerra (Schelling 1960, pp. 4-5) sono giochi misti. Questa circostanza mette in evidenza una notevole difficoltà per la teoria classica dei giochi e per la possibilità di una sua feconda applicazione nelle scienze sociali.

Supponiamo infatti di accettare l'idea – che abbiamo chiamato principio di Nash – che la soluzione di un gioco debba essere un equilibrio di Nash. Allora, di fronte a una molteplicità di equilibri e al fatto che, in generale, i giocatori preferiscono equilibri diversi, dobbiamo chiederci *quale* di essi sia la soluzione del gioco; in altre parole, dobbiamo chiederci in che modo i giocatori riusciranno a coordinare le loro strategie così da raggiungere un particolare punto di equilibrio. I tentativi di fornire una risposta adeguata a questo interrogativo – cioè di risolvere il *problema dell'indeterminatezza* – hanno dato origine al cosiddetto *programma di raffinamento*, che, a partire dai contributi dell'economista tedesco Reinhard Selten (nato nel 1930), ha conosciuto un notevole sviluppo soprattutto negli anni Ottanta del secolo scorso.

Parlando di “raffinamento” del concetto di equilibrio di Nash ci riferiamo a una nozione di equilibrio che viene soddisfatta solo da un sottinsieme, di solito molto ristretto, degli equilibri di Nash. Le diverse nozioni “raffinate” di equilibrio proposte nella letteratura sono accompagnate dalla raccomandazione che la soluzione di un gioco debba essere costituita da un equilibrio “raffinato”.⁷ Il concetto più noto di raffinamento è il cosiddetto “equilibrio a mano tremante” proposto da Selten (1975). L'intuizione che sta alla base del raffinamento di Selten è che anche giocatori perfettamente razionali, che conoscono la loro strategia ottimale, talvolta possano adottarne un'altra per sbaglio – cioè per qualche sbadataggine, errore di esecuzione o, per usare l'immagine di Selten, per qualche “tremore” della mano. Se si ammette la possibilità di un errore, si ottiene un gioco leggermente “perturbato” rispetto al gioco originale, in cui talvolta, con una probabilità minima ma diversa da zero, ai giocatori “tremava la mano”. Il gioco perturbato è in equilibrio (a mano tremante) quando la strategia scelta da ogni giocatore è la miglior risposta a quella degli altri giocatori, una volta che tutti riconoscano la possibilità che a ognuno possa talvolta tremare la mano. Dato che non tutti gli equilibri di Nash sono equilibri a mano tremante, tale concetto permette di restringere il campo delle potenziali soluzioni del gioco e, in alcuni casi, quando vi è un solo equilibrio a mano tremante, di individuare la soluzione del gioco.

Vale la pena notare che l'adozione dell'equilibrio a mano tremante proposto da Selten presuppone l'abbandono di quello che abbiamo chiamato il postulato della “perfezione pratica” della teoria classica dei giochi (vedi TGC3), per cui ogni giocatore adotta infallibilmente la propria strategia

7 Per una rassegna delle diverse nozioni di raffinamento proposte nella letteratura si vedano Kohlberg (1990) e Govindan e Wilson (2008).

ottimale. Ciò nonostante, il programma di raffinamento rappresenta il tentativo più avanzato di mantenere intatto l'impianto teorico e concettuale della teoria classica dei giochi, e soprattutto l'idea, compendiata dal postulato dell'universalismo aprioristico (vedi TGC5), per cui pochi principi generali – come il principio di dominanza, il principio di Nash e altri “criteri di ottimalità” basati sull'idea che l'equilibrio di Nash ottimale sia costituito da un equilibrio (opportunamente) raffinato – siano sufficienti a individuare le strategie ottimali dei giocatori. Alcuni buoni motivi per dubitare dell'ottimismo razionalistico di queste proposte sono connessi al fatto che sono stati finora individuati almeno tre dozzine di diversi criteri di ottimalità. Di conseguenza, affinché i giocatori possano coordinarsi su un particolare equilibrio, e quindi individuare la soluzione del gioco, il principio di conoscenza comune dell'altrui razionalità non basta: occorre anche che *tutti* i giocatori identifichino la razionalità con l'adozione di uno *specifico* criterio di ottimalità. Per questo motivo, il programma di raffinamento sembra al momento segnare il passo, a favore di altre proposte che si distaccano in modo più radicale dai presupposti concettuali della teoria classica dei giochi.

Nash contro Pareto

La possibilità che un gioco abbia una molteplicità di equilibri di Nash pone il problema di identificare quale, fra questi, sia la soluzione del gioco. Tuttavia, alcuni difficili problemi sorgono anche in relazione a giochi con un unico equilibrio di Nash. In certi casi, infatti, l'unico equilibrio di un gioco è meno soddisfacente, dal punto di vista di tutti i giocatori, di qualche altro risultato del gioco. L'esempio canonico di questo tipo di problemi è il cosiddetto *dilemma del prigioniero*, un gioco molto noto perché sembra dimostrare l'insufficienza della razionalità individuale nel raggiungere risultati soddisfacenti per *tutti* i partecipanti.

Il dilemma del prigioniero deve il suo nome alla storiella, proposta nel 1950 dal matematico canadese A. W. Tucker (1905-1995), con cui viene solitamente illustrato. Due malviventi vengono arrestati a bordo di un'auto rubata. La polizia può condannarli entrambi a 1 anno per il furto, ma sospetta che i due abbiano commesso una grossa rapina avvenuta recentemente e preferirebbe incriminarli per questo reato, che costerebbe loro 10 anni di prigione. Per ottenere una confessione, il commissario propone a ognuno dei prigionieri la seguente alternativa: o confessare la rapina e tradire il proprio complice, o rimanergli fedele e tacere. Se un solo prigioniero confessa, permettendo al commissario di risolvere il caso, verrà la-

		COLONNA	
		<i>C</i>	<i>D</i>
RIGA	<i>C</i>	-1 0	-10 -5
	<i>D</i>	0 -5	-5 -5

Figura 3.11.
Il dilemma del prigioniero

sciato libero e il complice sconterà l'intera pena, cioè 10 anni. Se entrambi i prigionieri confessano, otterranno entrambi uno sconto di pena di 5 anni per aver collaborato con la polizia, scontando solo i restanti 5. Se entrambi tacciono, la polizia non ha prove sufficienti a incriminarli per la rapina, cosicché entrambi sconteranno solo 1 anno per il furto dell'auto.

La situazione dei due prigionieri – Riga e Colonna – può venir rappresentata come nel gioco in figura 3.11, in cui ciascun giocatore ha due strategie: tacere e non tradire il complice o, viceversa, confessare e tradirlo. Nella terminologia giostica ormai entrata nell'uso, per ogni prigioniero la confessione equivale a *defezionare* nei confronti del compagno, mentre tacere equivale a *cooperare* con lui. Chiameremo quindi *C* (cooperazione) la strategia del silenzio e *D* (defezione) la strategia della confessione. I payoff rappresentano gli anni di galera, e sono quindi negativi o, nel migliore dei casi, nulli.

Dalla matrice del gioco risulta immediatamente chiaro che entrambi i giocatori hanno una strategia dominante, cioè la defezione. Consideriamo, per esempio, Riga. A Riga conviene scegliere *D* indipendentemente da cosa faccia Colonna. Infatti, se Colonna coopera, a Riga conviene defezionare: in questo modo sarà libero (payoff 0) invece di scontare 1 anno di galera (payoff -1). D'altra parte, se Colonna defeziona, a Riga conviene ancora scegliere *D*, in modo da scontare 5 anni di galera (payoff -5) invece che 10 (payoff -10). Colonna ha payoff simmetrici rispetto a Riga, e quindi può fare lo stesso ragionamento; di conseguenza, a entrambi i prigionieri conviene tradire il complice e confessare. Dato che ognuno dei due giocatori deve razionalmente adottare la propria strategia dominante, la soluzione del gioco è (D,D) , cioè l'esito del gioco in cui entrambi i prigionieri scontano 5 anni di carcere. Si noti che (D,D) è l'unico equilibrio di Nash del gioco. Infatti, una volta raggiunto questo risultato, ognuno dei due prigionieri è soddisfatto della propria scelta, data quella altrui: cambia-

re unilateralmente la propria strategia lo porterebbe a scontare 10 anni di galera invece degli attuali 5.

Il “dilemma” risiede nel fatto che il risultato (C,C) è nettamente *migliore*, per entrambi i giocatori, rispetto alla soluzione del gioco (D,D) . Se, infatti, nessuno dei due prigionieri confessasse, ma cooperasse con l’altro, entrambi sconterebbero solo 1 anno di galera invece di 5, raggiungendo un risultato molto più soddisfacente per entrambi rispetto a (D,D) . In altre parole, benché la defezione domini la cooperazione – e sia quindi la soluzione del gioco – quest’ultima garantirebbe risultati migliori a entrambi i giocatori. Ciascun prigioniero si trova quindi alle prese con un dilemma: da un lato sa che l’unica strategia razionale è confessare, ma dall’altro comprende che tacere converrebbe a entrambi.

Nel gioco del prigioniero, il risultato (C,C) è migliore di (D,D) perché, in termini tecnici, (C,C) è *Pareto-superiore* a (D,D) . Il concetto di “Pareto-superiorità” venne introdotto dall’economista italiano Vilfredo Pareto (1848-1923) nell’ambito della cosiddetta economia del benessere, il cui obiettivo era quello di precisare le condizioni in cui un cambiamento sociale aumenta il benessere complessivo di un gruppo di individui, cioè rappresenta un miglioramento dal punto di vista di tutti gli individui coinvolti.⁸ Secondo Pareto, vi è almeno un caso in cui è possibile individuare oggettivamente e senza ambiguità un miglioramento del benessere collettivo. Consideriamo, infatti, due situazioni sociali X e Y e supponiamo che in X nessun individuo stia peggio che in Y e qualche individuo stia addirittura meglio: possiamo allora dire che il “benessere sociale” complessivo è maggiore in X che in Y . In questo caso diciamo quindi che X è Pareto-superiore a Y , o, equivalentemente, che X Pareto-domina Y , o anche che X è un Pareto-miglioramento rispetto a Y . Se X è un Pareto-miglioramento rispetto a Y , e agli individui che si trovano in Y viene chiesto di votare il passaggio da Y a X , alcuni si asterranno, cioè saranno indifferenti fra X e Y , mentre alcuni voteranno a favore: nessuno, comunque, sarà contrario. Questa *unanimità con astensioni* a favore di X mostra che X è considerato da tutti gli individui coinvolti

8 Il problema consisteva nello specificare condizioni di miglioramento che evitino i cosiddetti *confronti interpersonali* fra le utilità dei diversi individui. I confronti interpersonali di utilità sono ritenuti sospetti dalla maggior parte degli economisti e dei filosofi (Hargreaves Heap *et al.*, 1992, pp. 378-380), poiché richiederebbero di confrontare quantitativamente (e quindi di poter misurare) il “benessere” ottenuto, per esempio, da Riga mangiando una mela a quello ottenuto da Colonna mangiando la stessa mela. Diverso il caso dei confronti *intrapersonali* di utilità, che si limitano a un confronto delle preferenze degli individui rispetto a diverse possibili situazioni (preferenze rivelate dal loro comportamento).

un miglioramento rispetto a Y . È importante notare che, date due situazioni X e Y , può accadere che nessuna delle due sia Pareto-superiore all'altra. In gran parte dei casi, infatti, X e Y sono *paretianamente non confrontabili*, nel senso che alcuni individui preferiscono l'una e altri individui preferiscono l'altra, così da rendere impossibile un confronto fra le due situazioni (cfr. Hargreaves Heap *et al.* 1992, p. 435).

Infine, si dice che una situazione X è *Pareto-ottimale*, o *Pareto-efficiente*, quando non vi è nessun'altra situazione Y che sia Pareto-superiore a X . In questo caso, diciamo anche che X costituisce un *ottimo paretiano*. Si noti che una data situazione è un ottimo paretiano semplicemente per l'assenza di alternative (paretianamente) migliori; ciò significa che una situazione paretianamente ottimale non è necessariamente la migliore per tutti gli individui coinvolti. Inoltre, dato un insieme di stati di cose, avremo in generale una molteplicità di ottimi paretiani.

I concetti di Pareto-superiorità e Pareto-ottimalità sono stati utilizzati anche nella teoria dei giochi. In un gioco, infatti, un risultato sarà Pareto-superiore a un altro quando nel primo nessun giocatore avrà payoff peggiori e qualcuno avrà payoff migliori rispetto al secondo; e un risultato sarà Pareto-ottimale quando non è paretianamente dominante da alcun altro risultato. Nel dilemma del prigioniero (si veda la figura 3.11) ci sono tre ottimi paretiani, corrispondenti alle celle (C,C) , (C,D) e (D,C) . L'unico risultato non Pareto-ottimale è proprio la soluzione del gioco, (D,D) , che è paretianamente inferiore a (C,C) . Nel dilemma del prigioniero, e più in generale nei dilemmi sociali, la soluzione del gioco non è un ottimo paretiano.

A partire dagli anni Cinquanta del secolo scorso, il dilemma del prigioniero ha catalizzato l'attenzione non solo dei teorici dei giochi, ma anche degli scienziati sociali e di tutti gli studiosi interessati alla natura, alle possibilità e ai limiti della razionalità umana. La letteratura, sia teorica sia sperimentale, su questo gioco e sulle sue varianti è ormai tanto varia quanto sterminata. Nel prossimo capitolo, vedremo come il dilemma del prigioniero sia stato applicato all'analisi di un problema centrale dell'economia e delle scienze sociali, quello della produzione dei cosiddetti beni pubblici. Nel corso della lunga discussione che ne è seguita, sono venuti alla luce molti dei limiti della teoria classica dei giochi, presentata in questo capitolo, per l'analisi dei fenomeni sociali. Le teorie non classiche dei giochi si propongono di superare almeno alcune di queste limitazioni e il dilemma del prigioniero rappresenta un importante "banco di prova" per ognuna di esse. Il prossimo capitolo passa in rassegna alcune delle "soluzioni" che i sostenitori di queste teorie hanno proposto per uscire dall'*impasse* posta alla teoria classica dal dilemma del prigioniero e da altri "dilemmi sociali".

4.

BENI PUBBLICI E ALTRI DILEMMI SOCIALI

Un problema centrale delle scienze sociali è quello di spiegare la natura, l'emergere e l'evoluzione della cooperazione tra i membri di una società. Per esempio, l'economia studia come molti differenti interessi e valutazioni individuali si coordinino nel mercato e come diverse istituzioni ostacolino o favoriscano tale coordinazione. La cooperazione sociale – cioè la cooperazione fra i diversi individui che compongono la società – è essenziale per raggiungere gli obiettivi universalmente condivisi da tali individui. Fra questi obiettivi, quello principale è la produzione (o fornitura) dei cosiddetti *beni pubblici*.

L'enorme attenzione che economisti e scienziati sociali hanno dedicato al problema della produzione di beni pubblici appare giustificata dall'altrettanto ampia estensione di questo concetto. Essi, infatti, sembrano includere una vastissima classe di beni e servizi: non solo, e in primo luogo, i cosiddetti beni pubblici hobbesiani, come l'ordinamento giuridico, l'amministrazione della giustizia e la gestione della sicurezza e della difesa di una società; ma anche, in secondo luogo, quei beni o servizi ritenuti fondamentali o, comunque, socialmente utili, dalla televisione pubblica all'istruzione, dalla sanità al teatro, fino al mantenimento di un ambiente pulito e alla sopravvivenza di certe rare specie animali.

Comprendere le ragioni che determinano il successo, o il fallimento, di una società nel fornirsi dei beni pubblici di cui necessita è un compito di grande interesse, che gli scienziati sociali hanno affrontato sulla base di svariati approcci, compresa la teoria dei giochi. Come vedremo in questo capitolo, a partire dagli anni Settanta del secolo scorso il problema dei beni pubblici è stato sempre più frequentemente analizzato, in termini giuochistici, come un problema di interazione strategica. In questo senso, si può parlare di diversi *giochi dei beni pubblici* – cioè situazioni in cui ognuno deve decidere, in base alla sua valutazione di quel che faranno gli altri, se – e spesso anche in quale misura – contribuire alla produzione di un determinato bene pubblico.

Le ricerche giochistiche sulla produzione dei beni pubblici hanno ormai raggiunto dimensioni molto ampie. Nell'intento di offrire un'adeguata rappresentazione formale dei giochi dei beni pubblici, sono stati impiegati non meno di una dozzina di modelli di gioco. In questa sede ne considereremo solo tre, cioè il dilemma del prigioniero (già introdotto nel paragrafo 3.5 e che discutiamo nei paragrafi 4.1, 4.2 e 4.3), il gioco del cumulo di neve (paragrafo 4.5) e il gioco della caccia al cervo (paragrafo 4.6), che serviranno anche ad illustrare le conseguenze che un'analisi giochistica ha per la teoria e la filosofia della politica (paragrafo 4.4). Anche se tali giochi riguardano il caso idealizzato di una microsocietà composta da due soli individui, essi possono essere generalizzati al caso in cui i giocatori sono più di due, come accade nelle comuni interazioni sociali connesse alla produzione di beni pubblici.¹

4.1 Il dilemma dei beni pubblici

I beni pubblici vengono tradizionalmente definiti nei termini dei concetti di *non-rivalità* e *non-escludibilità*.² Parlando di non-rivalità di un bene – per esempio, di un ponte – si intende dire che l'uso del bene da parte di un individuo non impedisce che anche altri ne facciano lo stesso uso. Parlando, invece, di non-escludibilità vogliamo dire che è virtualmente impossibile impedire l'uso del bene da parte di individui che non hanno in alcun modo contribuito alla sua produzione.

Una conseguenza della non escludibilità di un bene pubblico è che chiunque preferirebbe usufruirne gratuitamente, nel comodo ruolo dello "scroccone" o, per usare un termine ormai entrato anche nell'uso italiano, del *free rider*. Se, tuttavia, tutti si comportano da *free rider*, nella speranza che siano gli altri a contribuire ai costi di produzione del bene, allora non verrà prodotto alcun bene. Il risultato di un gioco dei beni pubblici dipenderà dalle scelte dei partecipanti, ciascuno dei quali deve decidere se *cooperare* alla produzione del bene (cioè contribuire ai suoi costi di produzione) oppure *defezionare* (cioè non contribuire). Infatti, la produzione

1 I tre giochi sopra menzionati presentano motivi di grande interesse non solo per le scienze sociali, ma anche per la filosofia politica e la sua storia. Infatti, negli ultimi decenni, essi sono stati ampiamente usati per la ricostruzione razionale di alcune fondamentali intuizioni di Hobbes, Hume e Rousseau relative alla produzione di beni pubblici e, più in generale, alla genesi e all'evoluzione dell'ordine sociale.

2 Sulla nozione di bene pubblico si vedano per esempio Hargreaves Heap *et al.* (1992, pp. 374-376 e 193-198) e Taylor (1987, pp. 5 ss.).

del bene potrebbe richiedere un numero minimo di contributori, oppure il superamento di una soglia critica di contributi. Il successo del gioco, cioè l'effettiva produzione del bene è, quindi, tutt'altro che garantito.

Una tesi largamente diffusa tra scienziati sociali e filosofi politici – che potremmo chiamare tesi *convenzionale*, o *hobbesiana* – sostiene che la cooperazione sociale è generalmente impossibile, o comunque molto inefficiente, senza l'intervento di un'autorità in grado di imporre coercitivamente determinati comportamenti a tutti i membri della società. Secondo la tesi convenzionale, la necessità di quell'autorità centralizzata comunemente nota come “stato” dipende proprio dall'impossibilità di ottenere un livello di cooperazione sociale *spontanea* sufficiente alla produzione dei beni pubblici necessari alla società (cfr. Sugden 1989). Come nota Robert Sugden (1992):

[n]ella teoria politica ed economica, è un luogo comune dire che il governo è necessario a causa dell'esistenza dei beni pubblici. I beni pubblici sono beni che ognuno desidera ma di cui nessuno ha interesse a sostenere i costi di produzione, a causa del problema del *free rider*. Questa giustificazione del governo risale ad Hobbes, che sostenne che l'ordine civile è un bene pubblico che solo un sovrano col monopolio dell'uso della forza può fornire. Così, secondo Hobbes, individui razionali in uno “stato di natura” sottoscriverebbero un contratto sociale con il quale ognuno si sottometterebbe volontariamente a un sovrano assoluto, purché ogni altro faccia lo stesso.

Secondo la tesi hobbesiana, a causa del problema del *free rider*, se ogni individuo fosse libero di scegliere o meno se contribuire alla produzione di un bene pubblico nessun bene pubblico verrebbe mai prodotto.

Per questo motivo, la discussione sui beni pubblici è strettamente legata a quella sui cosiddetti dilemmi sociali. Si parla di *dilemma sociale* per indicare una forma di interazione in cui la razionalità individuale, cioè la combinazione delle scelte razionali effettuate dai singoli partecipanti, conduce a un disastro collettivo. In termini un po' meno drammatici (Bicchieri 2006, p. 140), un dilemma sociale è una situazione

in cui ogni membro del gruppo ottiene un risultato migliore se persegue il proprio interesse personale, ma, allo stesso tempo, ciascuno trae beneficio dal fatto che tutti i membri del gruppo assecondino l'interesse comune.

Secondo la tesi convenzionale, la produzione di un bene pubblico, come il ponte nell'esempio sopra, rappresenta un tipico caso di dilemma sociale o, nei termini di Micheal C. Taylor (1987, pp. 3 e 19), di “problema di azione collettiva” (cfr. anche Olson 1965):

[...] la caratteristica definitoria di un problema di azione collettiva [...] risiede grosso modo nel fatto che è inverosimile che egoisti razionali riescano a cooperare per promuovere i propri interessi comuni. [...] Esiste un problema di azione collettiva quando l'azione individuale razionale può portare a un risultato strettamente Pareto-inferiore, cioè a un risultato che è strettamente meno preferito da ogni individuo rispetto ad almeno un altro risultato.

Con riferimento a questa rappresentazione si parla comunemente di *dilemma dei beni pubblici* (cfr. Hargreaves Heap e Varoufakis 2004, pp. 175 ss.).³ Se la produzione di un bene pubblico costituisse *sempre* un dilemma sociale, i membri di una società non sarebbero in grado di produrre nessuno dei beni pubblici di cui vorrebbero disporre. Questa era in effetti, come è noto, la visione di Hobbes, per il quale una società di individui liberi di cooperare, oppure no, piomberà inevitabilmente in uno stato di “guerra di tutti contro tutti”, in cui “la vita dell'uomo è solitaria, povera, sordida, bestiale e corta” (Hobbes 1651, pp. 72-73). Secondo Hobbes, infatti, nella condizione di anarchia disordinata, che è la condizione naturale dell'uomo prima che si instauri un governo, non può sussistere quel tanto di fiducia reciproca necessaria a rendere credibili gli accordi tra due o più individui, accordi che rappresentano il fondamento dell'ordine sociale e, quindi, anche della prosperità economica. L'unica via d'uscita dal dilemma consisterebbe, allora, in un particolare accordo, o contratto, stipulato dai giocatori *prima* di affrontare qualsiasi dilemma sociale. Mediante tale contratto i giocatori darebbero vita, in qualche modo, a un'autorità esterna, dotata di poteri coercitivi, che li costringerebbe a cooperare, così da produrre quei beni pubblici che tutti desiderano. Lo stato, o il governo, sarebbe quindi *necessario* per la soluzione di qualunque dilemma dei beni pubblici.

Come vedremo nel paragrafo 4.4, la tesi hobbesiana è stata criticata da una minoranza di studiosi i quali rifiutano l'idea che lo stato sia una condizione necessaria per la produzione di qualunque bene pubblico. Secondo questi autori, la teoria economica “è eccessivamente pessimistica circa la possibilità dell'ordine spontaneo” e “sottovaluta l'abilità degli individui nel coordinare il loro comportamento al fine di risolvere i loro problemi comu-

3 Per un'utile rassegna delle analisi giostiche del dilemma dei beni pubblici, si veda Kollock (1998); si veda anche Cevolani e Festa (2012b). La cosiddetta “tragedia dei beni comuni”, resa famosa dall'ecologista americano Garrett Hardin (1915-2003), può essere considerata un dilemma dei beni pubblici in cui, per un individuo, cooperare significa astenersi dall'utilizzo eccessivo di una risorsa – per esempio un terreno di pascolo o una zona di pesca – alla quale ogni membro del gruppo ha libero accesso (Hardin 1968). Per una discussione delle differenze tra dilemma dei beni pubblici e tragedia dei beni comuni si vedano Dionisio e Gordo (2006).

ni” (Sugden 1986, p. 4). Che il “pessimismo hobbesiano” della discussione tradizionale sui dilemmi sociali sia eccessivo è in effetti suggerito dagli innumerevoli esempi storici, anche contemporanei, che mostrano come gli individui sono in grado di produrre svariate forme di beni pubblici senza ricorrere all’intervento dello stato.⁴ Inoltre, accanto a queste considerazioni di carattere storico, si possono addurre anche convincenti ragioni teoriche per rifiutare il diffuso pessimismo di economisti e scienziati sociali circa la possibilità e l’efficacia della cooperazione spontanea nella produzione di beni pubblici. Più precisamente, come vedremo nei prossimi paragrafi del capitolo, gli strumenti concettuali della teoria dei giochi consentono di far luce sul perché, in certe condizioni, la produzione di beni pubblici per mezzo della contribuzione volontaria può avere successo.

4.2 Dilemma dei beni pubblici e dilemma del prigioniero

La convinzione che fosse impossibile fornire beni pubblici sulla base della contribuzione volontaria era condivisa anche dagli studiosi che, soprattutto a partire dagli anni Settanta del secolo scorso, intrapresero un’analisi giochistica della produzione di beni pubblici. La maggior parte di loro ritenevano che il dilemma dei beni pubblici non fosse altro che un caso particolare del dilemma del prigioniero. Tale dilemma è rappresentato nella figura 4.1.

La figura mostra che il bene pubblico può essere prodotto anche con il contributo di un solo giocatore. Poiché il bene non è, per ipotesi, escludibile, se un solo giocatore contribuisce ai costi di produzione, il suo concorrente potrà comunque usufruirne nel piacevole ruolo del *free rider*. I payoff di Riga e Colonna nel dilemma dei beni pubblici sono indicati, nella figura 4.1, dalle quattro lettere x , y , w e z (la notazione è quella di Taylor 1976, p. 5). Più precisamente, un giocatore otterrà x se entrambi i giocatori cooperano, y se gioca da *free rider* (cioè se egli defeziona e il suo concor-

4 I beni prodotti sulla base di contributi volontari comprendono, per esempio, il servizio di salvataggio marittimo e le banche del sangue in Gran Bretagna (cfr. Sugden 1986, pp. 3-4). Un’impressionante rassegna di casi storici di produzione privata di beni pubblici – dal diritto della *lex mercatoria* all’istruzione – si trova in Beito *et al.* (2002). L’economista americana Elinor Ostrom (1933-2012), premio Nobel per l’economia nel 2009, ha attirato l’attenzione su diversi casi in cui la tragedia dei beni comuni viene risolta dalla cooperazione spontanea degli individui coinvolti; cfr. Ostrom (1990) e Ostrom *et al.* (1999).

		COLONNA	
		C	D
RIGA	C	bene pubblico nessun <i>free rider</i> x, x	bene pubblico Colonna <i>free rider</i> z, y
	D	bene pubblico Riga <i>free rider</i> y, z	nessun bene pubblico w, w

Figura 4.1.
Il dilemma dei beni pubblici

rente coopera), w se entrambi i giocatori defezionano, e, infine, z se gioca da “babbeo” (*sucker*), cioè se egli coopera e il suo concorrente defeziona.

Il dilemma del prigioniero, e quindi anche il suo caso particolare costituito dal dilemma dei beni pubblici, si fonda sul presupposto che valgano le disuguaglianze $y > x > w > z$ (si veda anche la figura 3.11 nel capitolo precedente). Vale subito la pena notare (come vedremo meglio nel paragrafo 4.4) che tale presupposto è tutt’altro che innocente. Mentre infatti le disuguaglianze $y > x > w$ sono intuitivamente plausibili, la disuguaglianza $w > z$ appare, invece, piuttosto problematica. Infatti, tale disuguaglianza significa che un giocatore preferisce rassegnarsi alla mancata produzione del bene piuttosto che produrlo con i suoi soli sforzi. Tale preferenza appare pienamente plausibile nei casi in cui il bene viene prodotto in quantità variabili, che dipendono dal numero dei contributori. Si può allora immaginare che la quantità del bene prodotta con gli sforzi di un solo contributore sia così modesta da non compensare i costi da lui affrontati. Casi di questo genere sembrano verificarsi piuttosto spesso. Tuttavia, come avremo modo di vedere discutendo il gioco del cumulo di neve (paragrafo 4.5), si presentano frequentemente anche casi in cui $z > w$, cioè casi in cui un giocatore preferisce produrre il bene con i suoi soli sforzi piuttosto che rassegnarsi alla sua mancata produzione (cfr. anche Kollock 1998, figura 2).

Un’importante conseguenza delle disuguaglianze $y > x > w > z$ è che la defezione è la strategia dominante sia di Riga sia di Colonna. Infatti, da $y > x$ segue che la migliore risposta di un giocatore alla cooperazione del concorrente è la defezione; allo stesso modo, da $w > z$ segue che la migliore risposta di un giocatore alla defezione del concorrente è, ancora una volta, la defezione. Così, l’unica soluzione razionale al dilemma dei beni pubblici è la mutua defezione, cioè il risultato (w, w) .

Come si può facilmente immaginare, questa analisi giochistica della produzione di beni pubblici ha immediatamente suscitato forti perplessità.

È infatti evidente che con la mutua defezione non viene prodotto alcun bene pubblico e che entrambi i giocatori ottengono un *payoff* inferiore a quello che avrebbero ottenuto con la mutua cooperazione, che è un ottimo paretiano del gioco. Ciò significa che, se si comportano entrambi razionalmente, i giocatori ottengono un risultato peggiore di quello che otterrebbero comportandosi irrazionalmente. È proprio con riferimento a tale conclusione che il gioco dei beni pubblici – e, più in generale, il gioco del prigioniero – viene considerato un esempio classico di dilemma sociale.

Il dilemma ripetuto del prigioniero

La conclusione paradossale, enunciata qui sopra, per cui individui razionali non dovrebbero mai contribuire alla produzione di beni pubblici che pure desiderano, ha sollevato non pochi dubbi sull'opportunità di utilizzare il gioco del prigioniero per rappresentare situazioni di questo tipo. Fin dalle prime analisi giochistiche del problema, è stato per esempio fatto notare che il dilemma del prigioniero *a un turno* – cioè in cui due o più giocatori interagiscono per la prima e ultima volta – è una rappresentazione irrealistica del dilemma dei beni pubblici. Nelle società reali, infatti, gli individui interagiscono continuamente con altri individui, e prendono decisioni anche sulla base della loro esperienza passata di queste interazioni. In particolare, come dimostrano anche studi sperimentali e ricerche sul campo, la scelta di contribuire o meno alla produzione di determinati beni pubblici è fortemente influenzata dall'osservazione del comportamento altrui in situazioni simili. Per esempio, un individuo potrebbe più facilmente evitare di gettare cartacce per terra in un certo quartiere, se ha notato che le strade sono sempre pulite e ben tenute, perché i vicini si astengono da quel comportamento scorretto.

A partire dalla metà degli anni Settanta alcuni studiosi, per esempio il già citato Taylor (1976, 1987), hanno proposto di analizzare il dilemma dei beni pubblici nei termini del dilemma *ripetuto* del prigioniero. Il gioco del dilemma ripetuto è costituito da un indefinito numero di turni, ciascuno dei quali è una ripetizione del dilemma del prigioniero. I partecipanti non sanno quale sarà l'ultimo turno; conoscono solo la probabilità che, alla fine di ogni turno, il gioco prosegua con un ulteriore turno. Quando il gioco finisce, ogni giocatore ottiene un *payoff* che dipende da quelli ottenuti in ogni turno di gioco. Fra i diversi modi di determinare i *payoff* del dilemma ripetuto, il più semplice consiste nell'identificare il *payoff* "totale" di un giocatore con la somma dei *payoff* "parziali" ottenuti nei diversi turni del dilemma ripetuto (Sugden 1986, pp. 18 ss.). Una proposta più sofisticata si basa sull'idea intuitiva che un uovo oggi è meglio di un uovo domani; in altre parole, si introduce un *fattore di sconto*, in base al quale il *payoff* totale viene calcolato attribuendo un

peso progressivamente decrescente ai payoff parziali ottenuti nei turni futuri (cfr. Taylor 1987, p. 21 e Axelrod 1984, pp. 12 ss.).

Nel dilemma ripetuto del prigioniero si presuppone che, a ogni turno, ciascun giocatore ricordi esattamente non solo le proprie mosse nei turni precedenti, ma anche quelle del concorrente. Ciò significa che, come nella versione classica del dilemma, i giocatori sono perfettamente razionali e infallibili (come richiesto dai postulati TGC2 e TGC3). In particolare, ogni giocatore è in grado di calcolare, prima di iniziare il gioco, la strategia da utilizzare, e di attenersi ad essa ad ogni turno di gioco. Si noti che una strategia per il dilemma ripetuto deve indicare al giocatore la mossa da fare nel primo turno e quelle da fare in ciascun turno successivo, alla luce di quanto è accaduto nei turni precedenti. Come si può facilmente immaginare, le possibili strategie per il dilemma ripetuto sono infinite. Gli esempi forse più semplici di strategie sono la “cooperazione incondizionata” C , che suggerisce di cooperare a ogni turno, e la “defezione incondizionata” D , che suggerisce di defezionare a ogni turno.

L’idea che sta alla base del concetto di equilibrio di Nash – cioè che, in equilibrio, nessun giocatore si deve pentire della strategia adottata – può venir applicata ed estesa anche ai giochi ripetuti. Per esempio, si può facilmente notare che la strategia C di cooperazione incondizionata non è la migliore risposta a se stessa. Infatti, se i due giocatori adottano C , coopereranno a ogni turno di gioco, ottenendo a ogni turno un payoff pari a x (cfr. la figura 4.1). Il loro payoff totale, alla fine del gioco, sarà dunque proporzionale a x (nel caso più semplice sarà x moltiplicato per il numero dei turni di gioco). In ogni caso, tale payoff sarà inferiore al payoff totale che avrebbe ottenuto chi avesse defezionato in almeno un turno di gioco, o, ancor meglio, in tutti (si ricordi, con riferimento alla figura 4.1, che $y > x$). In altre parole, la mutua cooperazione *non* è un equilibrio del dilemma ripetuto del prigioniero. Ciò significa che, a ogni turno di gioco, un giocatore razionale ha un forte incentivo a defezionare se si aspetta che l’altro giocatore cooperi. Al contrario, non è difficile controllare che la mutua defezione, esattamente come nel dilemma a un turno, è un equilibrio, perché D è la miglior risposta a se stessa. Alla fine del dilemma ripetuto, due giocatori che hanno adottato D non avranno motivo di pentirsi della propria scelta.

Fino a qui il dilemma ripetuto assomiglia al dilemma a un turno. Ma le somiglianze non vanno oltre. Infatti, mentre nel dilemma a un turno la mutua defezione è l’*unico* equilibrio, nel dilemma ripetuto, è solo uno tra i molti possibili equilibri del gioco. Vi sono, infatti, diverse altre strategie di equilibrio, cioè strategie che, quando vengono adottate da entrambi i giocatori, danno luogo a un equilibrio. Tra queste la più semplice è la strategia

T – dove “ T ” sta per *tit for tat* (colpo su colpo) – la quale suggerisce di cooperare al primo turno e, a ogni turno successivo, di copiare la scelta del concorrente nel turno immediatamente precedente. Ciò significa che T è una strategia di *reciprocità*, con la quale un giocatore ripaga il concorrente della sua stessa moneta, premiando la cooperazione con la cooperazione e punendo la defezione con la defezione. David Hume (1739, pp. 552-553) illustrava l’idea della reciprocità in questi termini:

È così che imparo a prestare un servizio a un altro senza provare per lui una vera benevolenza; infatti io prevedo che egli mi renderà il servizio attendendosi un altro dello stesso tipo, per conservare la medesima reciprocità di buoni uffici con me o con altri. Di conseguenza, dopo che gli ho prestato i miei servizi e che egli è in possesso del vantaggio che nasce dalla mia azione, è spinto a fare la sua parte prevedendo le conseguenze di un suo rifiuto.

Si può mostrare che – date certe condizioni, non troppo restrittive, sulla probabilità che ciascun turno del dilemma ripetuto sia seguito da un ulteriore turno e sul modo di determinare i payoff totali dei possibili risultati del gioco – T è una strategia di equilibrio (cfr. Taylor 1987, pp. 67 ss.; Sugden 1986, pp. 114 ss.). Inoltre, è interessante notare che T è una *strategia cooperativa*, nel senso che due “reciprocatori” che adottano T si comporteranno, di fatto, esattamente come due operatori incondizionati che adottano C . Infatti, i due giocatori cooperano al primo turno e poi copiano le rispettive mosse al turno successivo, continuando così a cooperare a ogni turno. In tal modo essi otterranno un payoff totale (proporzionale a x) maggiore di quello (proporzionale a w) ottenuto da due defezionatori incondizionati che adottano D . Di conseguenza, entrambi i giocatori preferiranno l’equilibrio (T, T) , determinato dalla mutua reciprocità, all’equilibrio (D, D) , determinato dalla mutua defezione.

Se (D, D) e (T, T) fossero gli unici due equilibri del dilemma ripetuto, allora i giocatori avrebbero buone ragioni per adottare T . Sfortunatamente, tuttavia, (D, D) e (T, T) *non* sono gli unici due equilibri del gioco. Si è infatti scoperto che, oltre a T , vi sono molte altre strategie di reciprocità e che tutte queste strategie sono strategie di equilibrio (cfr. ancora Taylor 1987, pp. 73 ss. e Sugden, *ivi*). Per esempio, un’altra strategia di reciprocità è T' , che coopera al primo turno, risponde alla cooperazione con la cooperazione nel turno successivo, ma punisce la defezione con la defezione nei due turni successivi. Si pone, quindi, il problema di stabilire quale sia, tra le molte strategie di equilibrio del dilemma ripetuto, la strategia ottimale, cioè quella che dovrebbe essere adottata da qualunque giocatore razionale. In altre parole, si pone il problema di stabilire quale, tra i molti equilibri

del gioco, ne rappresenti la soluzione. Come si ricorderà, nel caso di un gioco con un unico equilibrio, il principio di Nash impone di considerare quell'equilibrio come la soluzione del gioco. D'altra parte, nel caso di un gioco con molti equilibri, tale principio non ci è di grande aiuto; infatti, ci dice solo che uno di questi equilibri è la soluzione, senza dirci quale. Come abbiamo visto nel paragrafo 3.5, molti importanti sviluppi della teoria classica dei giochi sono stati ispirati proprio dall'intento di "raffinare" il concetto di equilibrio di Nash, formulando opportuni criteri per identificare, fra i molteplici equilibri di un gioco, l'equilibrio ottimale, cioè quello che un giocatore perfettamente razionale considererebbe la soluzione del gioco. Tuttavia, sfortunatamente, questi sviluppi non hanno finora condotto a conclusioni ampiamente condivise. In conclusione, l'analisi del dilemma ripetuto del prigioniero mostra solo come la spontanea cooperazione reciproca sia in effetti possibile, ma non come i giocatori possano di fatto arrivare ad adottare efficaci strategie di reciprocità.

4.3 Beni pubblici nelle teorie non classiche dei giochi

Come abbiamo già osservato, il pessimismo hobbesiano circa la possibilità di produrre beni pubblici tramite la cooperazione spontanea degli individui coinvolti caratterizza buona parte delle analisi classiche di questo problema. Negli ultimi decenni, tuttavia, diversi studiosi hanno messo in discussione questa visione pessimistica, raggiungendo conclusioni molto più sfumate, e nel complesso ottimistiche, circa le potenzialità della cooperazione spontanea in problemi di azione collettiva. In particolare, le analisi giochistiche del problema dei beni pubblici basate sulle versioni non classiche della teoria dei giochi hanno progressivamente chiarito i meccanismi strategici, cognitivi ed evolutivi che favoriscono l'emergere di soluzioni cooperative spontanee a diversi tipi di dilemma sociale. Nel seguito, ci concentriamo su alcuni aspetti essenziali delle teorie comportamentali, cognitive, evolucionistiche e adattive dei giochi, che gettano nuova luce sul dilemma del prigioniero come caso paradigmatico di dilemma sociale.

Preferenze sociali nelle teorie comportamentali

Una letteratura ormai sterminata, basata su studi sperimentali e ricerche sul campo, ha dimostrato che gli individui alle prese con interazioni che presentano la forma del dilemma del prigioniero manifestano – a dispetto

dei principi della teoria classica dei giochi – una spiccata tendenza alla cooperazione (cfr. Bowles 2004, p. 19, Camerer 2003, Guala 2005 e Smith 2008). Si noti che tali ricerche riguardano il classico dilemma del prigioniero a un turno, cioè il caso in cui due individui affrontano il gioco una volta sola, senza alcuna prospettiva di incontrarsi di nuovo, come invece accade nel dilemma ripetuto del prigioniero. Ciò significa che la scelta di cooperare, da parte di un giocatore, non può dipendere dalla speranza che il concorrente gli renderà il favore in un successivo incontro.

Come si è visto, la teoria classica dei giochi non è in grado giustificare la razionalità della cooperazione nel dilemma del prigioniero a un turno. Sulla scorta di questa constatazione, alcuni autori hanno concluso che le scelte cooperative in questo genere di gioco sono semplicemente una delle tante forme di irrazionalità che affliggono il comportamento degli esseri umani. Questa conclusione, però, è tutt'altro che inevitabile. Per esempio, l'idea che il comportamento cooperativo nel dilemma a un turno sia irrazionale è stata respinta con forza dai sostenitori di quella che abbiamo chiamato teoria comportamentale dei giochi. Il lettore ricorderà che questa teoria, fondata sullo studio sperimentale del comportamento umano nelle interazioni strategiche, è caratterizzata dal rifiuto del presupposto TGC1 della teoria classica dei giochi, secondo il quale i giocatori hanno sempre preferenze autointeressate. Secondo la teoria comportamentale, gli esseri umani hanno, normalmente, preferenze sociali o eterointeressate, basate sia sull'altruismo – cioè sulla considerazione delle preferenze degli altri giocatori – sia sulla reciprocità – cioè sulla considerazione delle strategie adottate dagli altri, e quindi del modo in cui viene condotto il gioco.

Una più precisa caratterizzazione di questi due tipi di preferenze sociali può venire tracciata nei termini delle nozioni di payoff “materiali”, “immateriali” e “globali”. Parleremo di *payoff materiali* per indicare la rappresentazione quantitativa di preferenze autointeressate. Per esempio i payoff indicati nella matrice della figura 4.1 rappresentano le preferenze autointeressate di giocatori alle prese con un dilemma dei beni pubblici. Parleremo, invece, di *payoff globali* per indicare la rappresentazione quantitativa di preferenze sociali. Il payoff globale che un giocatore attribuisce a un determinato risultato è la combinazione di due ingredienti: il primo è il payoff materiale da lui ottenuto con quel risultato, il secondo è un *payoff immateriale*, che può dipendere dal payoff materiale acquisito dal concorrente oppure dalla specifica combinazione di strategie adottate dai due giocatori, cioè dal “processo intenzionale” che ha prodotto quel risultato. Se il payoff immateriale del giocatore viene determinato nel primo modo, diremo che le sue preferenze sono *altruistiche*; se viene determinato

nel secondo modo, diremo che sono *basate sulla reciprocità*. L'espressione "basate sulla reciprocità" sta qui a indicare il fatto che un giocatore preferisce le combinazioni di strategie caratterizzate da reciprocità, cioè quelle in cui egli ripaga il concorrente con la sua stessa moneta, rispondendo con strategie "gentili" alle strategie gentili e con strategie "meschine" a quelle meschine.

Se i partecipanti a un dilemma dei beni pubblici hanno qualche genere di preferenze sociali – non importa se altruistiche o basate sulla reciprocità – allora affronteranno il gioco in maniera diversa da quella di giocatori con preferenze autointeressate. Infatti, considereranno i payoff materiali indicati nella figura 4.1 solo come un ingrediente dei loro payoff globali, che in alcuni casi potrebbero quindi risultare molto differenti da quelli materiali. Di conseguenza, i giocatori potrebbero trovarsi ad affrontare un gioco con una struttura ben diversa da quella dell'originario dilemma dei beni pubblici. In particolare, potrà accadere che la cooperazione venga identificata da entrambi come la strategia ottimale, cosicché il gioco avrà come risultato la produzione spontanea del bene pubblico (cfr. Smith 2008, p. 48 e capitoli 10-12).

Punti focali e imperfezione pratica nelle teorie cognitive

Le difficoltà in cui si è imbattuta la teoria classica nell'analisi dei giochi con molti equilibri hanno aperto la strada a quelle che abbiamo chiamato teorie cognitive dei giochi. Come si ricorderà, tali teorie abbandonano il presupposto TGC5 della teoria classica, secondo il quale la soluzione di un gioco può sempre venir identificata grazie a principi di razionalità universali e determinati sulla base di considerazioni a priori. In conflitto con questa idea, i fautori delle teorie cognitive ritengono che, in generale, la soluzione di un gioco con molteplici equilibri possa venire identificata solo impiegando, assieme ai principi di razionalità della teoria classica, anche appropriati principi contestuali, formulati con riferimento agli aspetti essenziali del "contesto cognitivo" entro il quale si svolge il gioco.

A questo riguardo, Schelling, che può essere ritenuto uno fra i primi e più illustri sostenitori delle teorie cognitive dei giochi, ha sostenuto che la convergenza dei giocatori su un particolare equilibrio, da identificare come la soluzione di un gioco, non dipende solo dal suo valore utilitaristico, cioè dai payoff che i giocatori ottengono con quell'equilibrio, ma anche (e, spesso, soprattutto) dalla sua *prominenza* o *rilevanza*, cioè dalla sua peculiare "riconoscibilità" nel contesto cognitivo del gioco. Sulla base di diversi esperimenti informali, effettuati in classe, con gli studenti nella veste di soggetti sperimentali, Schelling (1960, cap. 3 e 4) mostrò che,

in una grande varietà di interazioni sociali, gli individui sono in grado di coordinarsi su uno degli esiti del gioco, che si distingue per particolari caratteristiche dipendenti dal contesto e spesso arbitrarie, ma più o meno immediatamente riconoscibili da tutti i giocatori coinvolti. Tali caratteristiche contraddistinguono quello che Schelling (1960, p. 67) chiama un *punto focale* del gioco, cioè un punto di riferimento “per le aspettative di ogni persona circa quello che l’altro si aspetta che lui creda che ci si aspetti da lui”. In questo senso, un punto focale può essere descritto come un “punto di convergenza delle attese” reciproche dei giocatori (Sugden e Zamarrón 2006, p. 610). Per esempio, secondo la maggior parte dei soggetti intervistati da Schelling (1960, p. 66), due persone che non si conoscono e devono incontrarsi a New York, senza precedenti accordi e senza la possibilità di comunicare, dovrebbero farsi trovare a mezzogiorno in punto all’ufficio informazioni della Grand Central Station, che identifica il punto focale di questo gioco di coordinazione.

L’analisi di Schelling suggerisce che, nella maggior parte dei giochi misti o di pura coordinazione con più di un equilibrio di Nash – come il gioco della guida (figura 3.4) o la battaglia dei sessi (figura 3.8) introdotti nel capitolo 3 o il gioco del cumulo di neve che considereremo nel paragrafo 4.4 – vi sono punti focali in grado di far convergere spontaneamente le scelte dei giocatori coinvolti. In molti casi, l’identificazione di un punto focale dipende “dall’immaginazione più che dalla logica” (Schelling 1960, p. 67) e da indizi impliciti nel contesto del gioco; in altri casi, può dipendere da indicazioni esplicite, eventualmente fornite da una terza parte “fuori” dal gioco (cfr. McAdams e Nadler 2005). Per esempio, può ancora accadere di osservare, dipinta sul muro di vecchi edifici all’ingresso di alcuni paesi, la scritta “mantenere la destra”. Anche in assenza di un codice della strada che disciplini il senso di marcia prevedendo sanzioni per chi non lo rispetta, indicazioni di questo genere sono presumibilmente sufficienti a risolvere il gioco della guida, spingendo entrambi i giocatori ad adottare una delle loro due strategie di equilibrio. Per tornare al problema centrale di questo capitolo, non è difficile immaginare che, in molti casi assimilabili a un dilemma ripetuto del prigioniero, gli equilibri che derivano dall’adozione di strategie di reciprocità come T o T' abbiano, agli occhi dei giocatori, una rilevanza particolare che li spinga ad adottare spontaneamente quelle strategie, e quindi a risolvere il relativo dilemma dei beni pubblici.

Un altro importante aspetto del contesto cognitivo delle interazioni strategiche tra esseri umani è dato dalla circostanza che i giocatori sono, di norma, consapevoli della propria e altrui imperfezione pratica. Per esempio, sanno che possono commettere errori nell’esecuzione del-

la mossa prescelta. Tenere conto di questo aspetto dei giochi significa rinunciare ai presupposti TGC3 e TGC4 della teoria classica – secondo i quali i giocatori sono praticamente perfetti e condividono una conoscenza comune della loro perfezione – e sostituirli con adeguati modelli teorici dell'imperfezione pratica dei giocatori in carne e ossa. Seguendo questa strada, i teorici cognitivi hanno fatto luce su alcuni meccanismi, precedentemente ignorati, del dilemma ripetuto. Per esempio, hanno compreso che, diversamente da quanto si era ipotizzato tra gli anni Settanta e Ottanta del secolo scorso, la strategia *T non* è la strategia ottimale per giocatori praticamente imperfetti.

Supponiamo, infatti, che entrambi i partecipanti al dilemma ripetuto adottino *T* e, quindi, rispondano alla cooperazione con la cooperazione. Tuttavia, se i giocatori possono commettere errori nell'esecuzione di una mossa, allora, molto probabilmente, prima o poi uno di loro defezionerà per errore. Di conseguenza, il suo concorrente, in accordo con *T*, risponderà con la defezione, e al turno successivo il primo giocatore, sempre in accordo con *T*, defezionerà a sua volta. In questo modo, per effetto di un singolo errore di esecuzione, i giocatori cadranno nella trappola della mutua defezione, da cui usciranno soltanto quando uno di loro commetterà un ulteriore errore di esecuzione, effettuando involontariamente una mossa cooperativa. L'estrema vulnerabilità di *T* agli errori di esecuzione ne sconsiglia l'adozione da parte di giocatori in carne ed ossa, cioè di giocatori praticamente imperfetti. Occorre, quindi, individuare una strategia di reciprocità che permetta di evitare la trappola della mutua defezione. In questa sede, non possiamo prendere in esame i risultati ottenuti dai teorici cognitivi nelle loro indagini su questo e altri problemi relativi al dilemma ripetuto. Ci preme, tuttavia, sottolineare che l'approccio cognitivo ha condotto a una comprensione ormai piuttosto approfondita del modo in cui gli esseri umani affrontano il dilemma ripetuto e di come, in molti casi, sono in grado di scegliere efficaci strategie di reciprocità.

Adattamento e apprendimento nelle teorie evoluzionistiche

Un contributo essenziale all'analisi del dilemma del prigioniero e di altre forme di interazione connesse ai giochi dei beni pubblici è venuto, già a partire dagli anni Settanta, dalla teoria evoluzionistica dei giochi. Come si ricorderà, tale teoria nasce dall'applicazione della teoria dei giochi alla biologia e all'etologia, cioè allo studio del comportamento animale. Lo studio delle interazioni animali ha rivelato che, in molti casi, la loro struttura è identica a quella di ben note forme di interazione umana. In particolare,

alcune interazioni animali condividono le caratteristiche fondamentali dei giochi dei beni pubblici (cfr. Hauert *et al.* 2006). Si pensi, per esempio, alle grida d'allarme che un uccello può emettere per segnalare al proprio gruppo l'avvicinarsi di un predatore. Così facendo, l'uccello produce un bene pubblico – che consiste nell'accresciuta sicurezza del gruppo – ma lo fa a sue spese, poiché deve affrontare un costo rappresentato dal rischio di attirare su di sé l'attenzione del predatore.

La teoria evuzionistica fornisce appropriati modelli formali per l'analisi di molti interessanti problemi relativi all'evoluzione delle strategie animali. Un'illustrazione intuitiva di questi problemi viene fornita dal seguente esempio. Supponiamo che le interazioni fra i membri di una popolazione animale abbiano la forma del dilemma ripetuto del prigioniero e che valgano le seguenti condizioni: (1) le interazioni sono del tutto casuali, nel senso che, in qualunque momento, un individuo ha la stessa probabilità di interagire con qualunque altro membro della popolazione; (2) un individuo usa la stessa strategia in qualsiasi interazione, cosicché ogni possibile strategia è rappresentata da una determinata percentuale, eventualmente nulla, di membri della popolazione – per esempio, la strategia di cooperazione incondizionata *C* potrebbe essere adottata dal 10% dei membri, la strategia di defezione incondizionata *D* dal 20%, la strategia di reciprocità *T* dal 45%, e così via. Allora i payoff ottenuti da un individuo in una lunga serie di interazioni non dipenderanno solo dalla sua strategia, ma anche da quelle degli altri membri della popolazione.

Poiché si presuppone che un individuo trasmette la propria strategia alla prole, e che payoff più elevati si traducono in maggiori vantaggi riproduttivi, cioè in una prole più numerosa, a ogni turno del gioco (che rappresenta una generazione nello sviluppo della popolazione) osserveremo una *riproduzione differenziale* delle strategie rappresentate nella popolazione, a vantaggio di quelle che hanno garantito payoff più elevati agli individui che le hanno adottate. Questo significa che, con l'arrivo della seconda generazione, le strategie di maggiore successo verranno adottate da una crescente percentuale di membri della popolazione. In tal modo, con il passare delle generazioni, può accadere che una determinata strategia diventi sempre più frequente e, in certi casi, finisca con l'invadere la popolazione, nel senso di venire adottata da tutti, o quasi tutti, i suoi membri. Ci si può chiedere, allora, a quali condizioni determinate strategie potranno diffondersi in una popolazione, e persino invaderla, mentre altre si avvieranno all'estinzione. Come vedremo nel prossimo capitolo, un compito fondamentale della teoria evuzionistica dei giochi consiste, appunto, nell'offrire risposte plausibili a interrogativi di questo tipo e, più in generale, nel far luce sui mec-

canismi che governano l'evoluzione strategica delle popolazioni animali. In quest'ultimo quarantennio tale compito è stato affrontato con notevole successo da una folta schiera di studiosi, i quali si sono avvalsi di svariati metodi di indagine, dall'osservazione sul campo alle tecniche di simulazione computerizzata.

Anche l'altra branca delle teorie evoluzionistiche, cioè quelle che abbiamo chiamato *teorie adattive* (cfr. figura 2.2), ha contribuito alla nostra comprensione dei meccanismi strategici e cognitivi che favoriscono la cooperazione nei giochi dei beni pubblici e in altre interazioni sociali. Le teorie adattive dei giochi si concentrano sull'evoluzione culturale, piuttosto che su quella biologica, di una popolazione, studiando come diversi *schemi di comportamento* emergono e si diffondono al suo interno. Il presupposto è che gli individui che formano la popolazione non siano astratti agenti iperrazionali ma agenti che, come gli esseri umani in carne ed ossa, hanno significative ma limitate capacità di ragionamento e di previsione. Ciò significa che ogni individuo coinvolto in una determinata interazione sociale, per esempio un dilemma ripetuto del prigioniero, adatterà una certa strategia basandosi su una rapida valutazione intuitiva della situazione, sull'abitudine e su altri fattori più o meno arbitrari – non, in ogni caso, su un attento calcolo costi-benefici come ipotizzato dalla teoria classica dei giochi.

Naturalmente, può benissimo darsi che la scelta strategica così operata sia tutt'altro che ottimale. Tuttavia, secondo le teorie adattive, gli esseri umani possono migliorare, nel corso del tempo, le loro strategie, attraverso un processo di *apprendimento dall'esperienza* che li mette in grado di ottenere un crescente *adattamento* all'ambiente, cioè alla popolazione di individui con cui sono soliti interagire. Più precisamente, le teorie adattive ipotizzano che, dopo un certo numero di interazioni, il giocatore si guarderà intorno e confronterà i propri *payoff* con quelli ottenuti dai membri della popolazione che hanno adottato altre strategie. L'esito di questo confronto potrà indurlo a sostituire la sua strategia con quella che, a quanto gli risulta, ha ottenuto il maggiore successo fino a quel momento; e, in linea di principio, questa procedura di perfezionamento strategico può venire attuata per un numero indefinito di volte.

Secondo i sostenitori delle teorie adattive, il processo di ripetuta *imitazione* delle strategie di maggiore successo svolge un ruolo molto importante nell'evoluzione strategica di una popolazione umana. Come si può facilmente intuire, tale processo di imitazione conduce alla riproduzione differenziale delle strategie di maggiore successo svolgendo, in tal modo, un ruolo strettamente simile a quello della trasmissione genetica nell'evolu-

luzione delle popolazioni animali. La stretta somiglianza tra i processi di riproduzione differenziale delle strategie di maggiore successo nelle popolazioni animali e umane apre la strada all'impiego sistematico dell'apparato formale delle teorie evoluzionistiche nell'analisi dell'evoluzione strategica delle popolazioni umane.

Tornando al tema di questo capitolo, converrà ricordare che le teorie adattive dei giochi hanno permesso di studiare le condizioni che favoriscono la diffusione delle strategie di reciprocità nelle popolazioni umane alle prese con il dilemma ripetuto del prigioniero. Per esempio, Robert Axelrod, nel suo famoso volume *The Evolution of Cooperation* (1984), ha impiegato ingegnosi metodi di simulazione computerizzata per analizzare l'evoluzione strategica di popolazioni alle prese con il dilemma ripetuto del prigioniero, giungendo alla conclusione che, in un'ampia varietà di condizioni iniziali, la strategia di reciprocità *T* finisce con l'invadere la popolazione. Anche se le ricerche condotte successivamente da diversi autori hanno in parte indebolito la tesi di Axelrod sulla quasi inevitabile tendenza di *T* ad imporsi, hanno comunque confermato la tesi, molto meno restrittiva, che l'evoluzione strategica favorisce l'emergere di qualche strategia di reciprocità (cfr. Sugden 1986, cap. 6).

4.4 Beni pubblici, teoria dei giochi e filosofia della politica

I contributi offerti dalle teorie non classiche dei giochi all'analisi del dilemma dei beni pubblici sembrano suggerire una visione più ottimistica, rispetto alla convenzionale tesi hobbesiana, della possibilità della cooperazione spontanea alla produzione di tali beni. A questo proposito, vale la pena notare che economisti e teorici della politica hanno espresso una grande varietà di vedute sia sul problema dei beni pubblici, sia sull'approccio giochistico a tale problema, sia, infine, sulle implicazioni di tali questioni per la filosofia politica. In particolare, diversi studiosi hanno criticato l'idea che il problema dei beni pubblici costituisca un dilemma sociale e, sempre nell'ambito dell'approccio giochistico, hanno proposto altri giochi, diversi dal dilemma del prigioniero, per rappresentare questo problema. Una proposta di questo tipo, che prenderemo in esame nel seguito, è quella del filosofo e teorico della politica ungherese Anthony de Jasay (nato nel 1925). Conviene quindi – prima di procedere alla discussione dettagliata (nei paragrafi 4.5 e 4.6) di queste proposte alternative – passare velocemente in rassegna la grande varietà di posizioni teoriche e metodologiche sul problema dei beni pubblici.

Beni pubblici e teoria dei giochi

Buona parte della discussione in quest'ambito ruota, in particolare, attorno ai seguenti quattro interrogativi – dei quali i primi tre sono di carattere sostanziale, mentre l'ultimo è di natura metodologica:

(1) Quali sono i beni pubblici o, in altre parole, come li si può adeguatamente definire?

(2) Data un'adeguata definizione di bene pubblico, è giustificata la credenza nell'esistenza di beni pubblici?

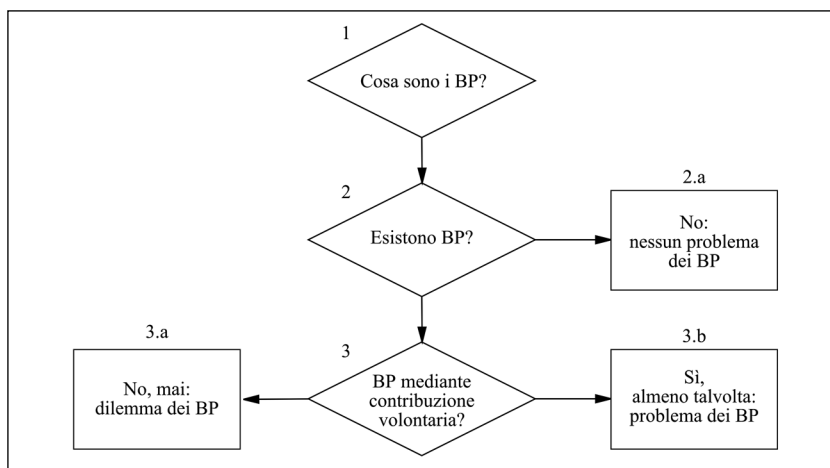
(3) Supponendo che i beni pubblici esistano, la loro produzione richiede necessariamente l'intervento dello stato, o può essere realizzata, almeno in certi casi, attraverso varie forme di cooperazione spontanea?

(4) Nell'analisi del problema dei beni pubblici dovremmo, oppure no, fare uso della teoria dei giochi?

Le relazioni che intercorrono tra gli interrogativi (1)-(3) vengono rappresentate nel "diagramma di flusso" nella figura 4.2, dove l'acronimo "BP" sta per "beni pubblici".

L'interrogativo che appare nel primo nodo del diagramma, cioè nel rombo (1), concerne la reale *natura* dei beni pubblici o, in termini sostanzialmente equivalenti, il modo migliore per definire tali beni. Data un'adeguata definizione di bene pubblico, possiamo affrontare l'interrogativo che appare nel rombo (2), concernente l'*esistenza* dei beni pubblici. Una

Figura 4.2. Posizioni teoriche sui beni pubblici



possibile risposta a questo interrogativo, suggerita dai teorici cosiddetti anarco-capitalisti, consiste nel negare l'esistenza dei beni pubblici – vedi rettangolo (2.a) – e nell'affermare, invece, che qualunque bene è privato, cioè che qualunque bene può essere fornito dal mercato (Rothbard 1973, cap. 10-14; 1981; Hoppe 1989a; 1989b, cap. 10). Se si accetta la risposta anarco-capitalista all'interrogativo (2), allora non dobbiamo neppure porci l'interrogativo (3) concernente la *fornitura* dei beni pubblici e, più precisamente, la possibilità che tali beni siano forniti grazie a varie forme di cooperazione spontanea. Ciò significa che la risposta anarco-capitalista all'interrogativo (2) rappresenta una *dissoluzione* del problema dei beni pubblici, formulato nel rombo (3).

Se invece si ammette l'esistenza di almeno alcuni beni pubblici, allora affrontare l'interrogativo (3) diventa non solo legittimo, ma inevitabile. Come si è visto sopra, la risposta convenzionale a questo interrogativo – che appare nel rettangolo (3.a) – consiste nel concepire il problema dei beni pubblici come un dilemma, cioè nell'escludere la possibilità che tali beni possano essere prodotti senza l'intervento coercitivo dello stato, in base alla semplice cooperazione spontanea tra i membri della società. Tuttavia, una minoranza di studiosi ha difeso, in contrasto con la risposta convenzionale (3.a), la possibilità che almeno alcuni beni pubblici vengano prodotti su base volontaria e non coercitiva – vedi rettangolo (3.b). Come vedremo fra poco, alcuni fra i più solidi argomenti addotti in difesa della risposta (3.b) sono stati elaborati da de Jasay, il quale si è impegnato nell'analisi sistematica delle condizioni che rendono possibile la produzione di beni pubblici sulla base di varie forme di contribuzione volontaria.⁵

Venendo ora all'interrogativo metodologico (4) – concernente l'opportunità di usare gli strumenti concettuali offerti dalla teoria dei giochi nell'analisi del problema dei beni pubblici – vale la pena notare che anche su questo punto si è manifestata una grande eterogeneità di posizioni tra gli studiosi di scienze sociali e filosofia politica. Possiamo osservare, per esempio, che tanto nell'ambito delle tradizioni di pensiero liberale e anarchica, quanto al di fuori di esse, incontriamo studiosi che applicano, in misura più o meno ampia, la teoria dei giochi e altri che la ignorano o, talvolta, manifestano un'aperta avversione al suo impiego. Per esempio, fra i sostenitori di un approccio giochistico in ambito liberale si possono citare almeno Rawls (1971), Gauthier (1986), Nozick (1974), Boudon (1979) e lo stesso de Jasay (1985). Come vedremo nel capitolo 7, i pensatori che si

5 Tra gli studiosi che concordano con la risposta (3.b) occorre anche ricordare, almeno, i nomi di Taylor (1987) e Axelrod (1984, 1997), dell'economista e anarchico David Friedman (1989, capitoli 34, 36 e 39) e di Vernon Smith (2008).

ispirano alla Scuola Austriaca di economia, fra i quali molti anarco-capitalisti come Rothbard (1973), tendono a ignorare o rifiutare esplicitamente la teoria dei giochi quale valido strumento d'analisi; un esempio di questa posizione è Bertrand Lemennicier (2006).

Se si considerano le tre variabili teoriche costituite da “liberalismo”, “anarchia” e “(impiego della) teoria dei giochi” possiamo avere, in linea di principio, otto possibili posizioni teoriche, a seconda che ciascuna di queste variabili sia presente, oppure no. Tali posizioni sono raffigurate nella tabella 1, dove le lettere “L”, “A” e “G” in cima alle prime tre colonne indicano, rispettivamente, il liberalismo, l'anarchia e l'impiego della teoria dei giochi nell'analisi del problema dei beni pubblici (e, più in generale, dell'ordine politico-sociale), mentre l'occorrenza dei simboli “+” e “-” in ogni riga della tabella indica, rispettivamente, la presenza o l'assenza della variabile teorica indicata in cima alla corrispondente colonna. Infine, nell'ultima colonna si indicano alcuni studiosi o movimenti significativi che sostengono le posizioni teoriche rappresentate in ciascuna delle otto righe.

Tabella 1.

	L	A	G	<i>Esempi:</i>
1.	-	-	-	Marxisti classici
2.	-	-	+	Marxisti analitici
3.	-	+	-	Anarco-collettivisti
4.	-	+	+	Michael J. Taylor
5.	+	-	-	Mises, Hayek, Milton Friedman
6.	+	-	+	Nozick, Gauthier
7.	+	+	-	Rothbard, David Friedman
8.	+	+	+	de Jasay

Come si può immediatamente osservare, nessuna riga dell'ultima colonna è stata lasciata vuota. Ciò significa che tutte le possibili combinazioni delle variabili “liberalismo”, “anarchia” e “teoria dei giochi” sono state effettivamente sostenute da qualche significativo studioso o movimento. In particolare, l'ultima riga ci mostra il nome di Anthony de Jasay, il quale può essere considerato come uno fra i più importanti teorici anarco-liberali e, fino ad oggi, anche come l'unico fra loro che abbia fatto un uso ampio e sistematico della teoria dei giochi.⁶

6 Benché minoritaria, la teoria anarchica e anche anarco-liberale annovera ormai un certo numero di sostenitori (si veda Boettke 2005 per una classificazione di diversi tipi di anarchismo da un punto di vista liberale). Questa teoria ha conosciu-

Beni pubblici e teoria dei giochi nell'anarco-liberalismo di de Jasay

Quello dei beni pubblici è un tema ricorrente nelle opere di de Jasay e, in modo particolare, nel volume *Social Contract, Free Ride* (1989).⁷ Come anticipato, de Jasay (1989) sostiene che la produzione di un bene pubblico non costituisce un genuino *dilemma* sociale, ma semplicemente un *problema* il quale ammette, almeno in linea di principio, una soluzione “anarchica”, basata sulla contribuzione volontaria. Gli argomenti da lui addotti a favore di questa tesi spaziano dalla critica della definizione convenzionale di bene pubblico all’analisi, in termini giochistici, della produzione dei beni pubblici. Per motivi di spazio, ci limiteremo qui a una concisa illustrazione di quest’ultimo punto.

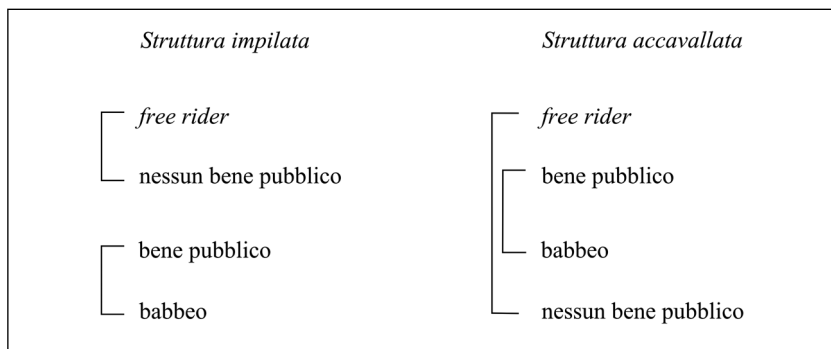
Secondo la tesi convenzionale, la cooperazione spontanea nella produzione di un bene pubblico non è possibile poiché nessun individuo razionale contribuirà volontariamente ai costi necessari. Tale affermazione si fonda su uno specifico presupposto relativo alle preferenze individuali tra i quattro risultati del gioco rappresentato nella figura 4.1. Si presume, infatti, non solo che (1) tutti preferiscano usufruire del bene pubblico da *free rider* piuttosto che da onesti contributori, ma anche che (2) tutti preferiscano che nessun bene venga prodotto piuttosto che produrlo da babbei, cioè sostenendone unilateralmente i costi. De Jasay (1989, pp. 146 ss.) ritiene che questo presupposto non sia universalmente valido e che, al contrario, esso descriva solo *una* delle possibili “strutture di incentivi” del gioco, cioè solo uno dei possibili ordinamenti delle preferenze individuali tra i risultati del gioco.

Secondo de Jasay occorre considerare attentamente anche altre strutture di incentivi che possono caratterizzare il problema dei beni pubblici. Come si può vedere dalla figura 4.3, le scelte di ciascun giocatore – e, di conseguenza, l’effettivo risultato del gioco – sono determinate dalle sue preferenze tra i quattro risultati possibili: quello in cui il bene pubblico

to recentemente un notevole sviluppo grazie all’analisi economica dell’anarchia nell’ambito della scuola di *Public Choice* (cfr. Tullock 1972; Buchanan 1975): si vedano Powell e Stringham (2009) per una rassegna e Stringham (2007) per una raccolta di alcuni dei principali contributi. Altri contributi interessanti provengono dal campo dell’economia sperimentale (Powell e Wilson 2008; Smith 2010); a questo proposito, si veda anche Cevolani (2011).

7 Sul problema dei beni pubblici si vedano anche de Jasay (1991, pp. 206-210; 1997, specialmente capitoli 1 e 6; 1998; 2002, capitoli 2 e 3; 2010, pp. 28-31). Per una presentazione più estesa di questo e di altri aspetti dell’opera di de Jasay, rimandiamo a Cevolani (2008) e Cevolani e Festa (2011).

Figura 4.3. Due strutture di incentivi per il problema dei beni pubblici



viene prodotto congiuntamente, quello in cui gioca da *free rider*, quello in cui gioca da babbeo e, infine, quello in cui nessun bene pubblico viene prodotto.

Nella parte sinistra della figura 4.3 viene rappresentata una delle possibili strutture di incentivi – precisamente, quella che de Jasay (1989, p. 148) chiama struttura “impilata” (*stacked*), poiché la coppia di risultati (*free rider* e nessun bene pubblico) determinati dalla scelta di non contribuire è preferita rispetto alla coppia di risultati (bene pubblico e babbeo) corrispondenti alla scelta di contribuire. La struttura impilata offre una chiara rappresentazione grafica del presupposto centrale della tesi convenzionale, che può venire così espresso: il peggior risultato possibile nel caso in cui l’individuo decida di non contribuire (nessun bene pubblico) è comunque migliore del peggior risultato determinato dalla scelta di contribuire (babbeo). Tale caratteristica è invece assente nel caso della struttura “accavallata” (*straddle*), rappresentata nella parte destra della figura 4.3. Possiamo vedere, infatti, che in questo ordinamento di preferenze la coppia di risultati corrispondenti alla non contribuzione sta, per così dire, “a cavalcioni” della coppia corrispondente alla contribuzione. In particolare, mentre la scelta di giocare da *free rider* è ancora preferibile a quella di contribuire alla produzione del bene pubblico, in questo caso giocare da babbeo è meglio che trovarsi senza alcun bene pubblico.

De Jasay (1989, in particolare capitoli 6 e 7) dimostra che una struttura accavallata è condizione necessaria e sufficiente affinché il problema dei beni pubblici sia risolvibile anche in assenza di coercizione, sulla base della contribuzione volontaria di almeno alcuni giocatori. Inoltre, suggerisce che, in molti casi, la produzione di beni pubblici esibisce proprio una strut-

tura di incentivi accavallata. In tali casi, la fornitura di un bene pubblico non costituisce più un dilemma del prigioniero, ma può venir rappresentata dal gioco “del cumulo di neve”, che analizzeremo nel prossimo paragrafo. La soluzione del problema dei beni pubblici offerta da de Jasay consiste, appunto, nel rifiutare l’idea che il dilemma del prigioniero offra una rappresentazione adeguata del problema dei beni pubblici e nel sostenere, invece, che tale problema debba venire rappresentato come un gioco del falco e della colomba. Vale la pena segnalare due tra le conseguenze della soluzione proposta da de Jasay.

I Se il gioco in figura 4.1 è caratterizzato da una struttura accavallata, allora un giocatore potrebbe decidere – sulla base delle sue attese circa le decisioni altrui – sia di contribuire sia di non contribuire. Ciò significa che in questo caso – diversamente da quanto accade nel dilemma del prigioniero – la scelta di non contribuire *non* è la strategia dominante.

II La scelta di contribuire non è semplicemente possibile, ma risulta anche, in diverse situazioni, la scelta migliore. Per esempio, de Jasay mostra che contribuire è la strategia migliore nel caso in cui i benefici attesi da tale scelta sono significativamente superiori alle perdite che ci si attende di subire nel caso in cui il bene pubblico non venga prodotto. In tal caso la possibilità di fallimento dell’impresa può spingere un individuo a contribuire, anche correndo il rischio di fare la parte del babbeo.

La possibilità dell’anarchia ordinata

Un esame dettagliato dell’analisi del problema dei beni pubblici proposta da de Jasay – e dei suoi aspetti problematici⁸ – esula dagli obiettivi di questo libro. La posizione teorica di de Jasay è tuttavia interessante per una valutazione della discussione contemporanea sulla possibilità di qualche forma di “anarchia ordinata”, cioè di un ordine sociale spontaneo che emerge in assenza di un’autorità centrale come lo stato. Come si è visto, un tratto distintivo dell’approccio di de Jasay è l’ampio ricorso alla teoria dei giochi nell’analisi dell’ordine sociale, approccio che colloca questo autore in una posizione peculiare rispetto ad altre correnti del liberalismo. Peculiari appaiono anche gli esiti della riflessione di de Jasay, che lo allontanano sia dalla tesi convenzionale sull’impossibilità dell’anarchia sia dalle posizioni degli anarco-capitalisti. Infatti, a differenza di questi ultimi, de Jasay ammette esplicitamente l’esistenza dei beni pubblici e si impegna

8 La soluzione di de Jasay è stata criticata da Sugden (1992) e da Kliemt e Lahno (1992); si veda anche la replica di de Jasay (1992) a Sugden.

nell'esplorazione sistematica dei modi possibili – e dei modi migliori – per la loro produzione.

Occorre anche rilevare che la visione anarco-liberale di de Jasay rimane decisamente problematica, con tratti pessimistici raramente rintracciabili nelle concezioni degli anarco-capitalisti “ortodossi”. Sotto i colpi della sua critica “scettica” (Buchanan 2007, p. 4), infatti, non cadono solo le costruzioni teoriche dei fautori dello stato, “minimo” o “massimo” che sia, ma anche la “robustezza” dell'ideale anarchico. Infatti, in diverse sue opere, de Jasay (1989, 1997) insiste sul quasi inevitabile emergere dello stato e sulla sua quasi irresistibile tendenza a espandersi. Un esempio delle tendenze espansionistiche dello stato – anch'esse analizzate da de Jasay con ampio ricorso alla teoria dei giochi – è costituito proprio dall'infinito ampliamento della sfera dei beni pubblici e dalla parallela restrizione di quella delle scelte individuali. A questo proposito, de Jasay non nasconde il suo scetticismo nei riguardi degli espedienti suggeriti da molti liberali per frenare la crescita incontrollata dello stato – a partire dalle sofisticate “architetture costituzionali” escogitate da Hayek (1973-1979). Le costituzioni vorrebbero imprigionare la politica in una sorta di cintura di castità. Tuttavia – osserva un po' maliziosamente de Jasay (1985, p. 205) – si tratta di una cintura di castità “con la chiave sempre a portata di mano” che “potrà al massimo ritardare il momento in cui la natura farà il suo corso”. Come vedremo nei capitoli 6 e 7, l'approccio giochistico alla genesi e all'evoluzione delle norme sociali può risultare estremamente utile per valutare questa e altre posizioni sulla possibilità di un ordine anarchico spontaneo.

4.5 Oltre il dilemma del prigioniero

Nel paragrafo 4.3, abbiamo visto come lo sviluppo delle diverse versioni della teoria dei giochi ha contribuito in maniera essenziale a comprendere come gli individui cooperino spontaneamente nei dilemmi sociali, e a mitigare il pessimismo della tesi convenzionale, per cui la cooperazione spontanea sarebbe impossibile o altamente improbabile. In particolare, l'analisi teorica e sperimentale del dilemma del prigioniero, sia a un turno sia ripetuto, ha chiarito i meccanismi strategici, cognitivi ed evolutivi grazie ai quali i giocatori possono raggiungere una soluzione cooperativa in questo genere di interazioni sociali. D'altra parte, il dilemma (ripetuto) del prigioniero, pur essendo il più famoso, non è l'unico gioco adeguato per rappresentare il problema della produzione di beni pubblici e altri dilemmi sociali. Anzi, come abbiamo osservato nel precedente paragrafo, il dilem-

ma dei beni pubblici, rappresentato nella figura 4.1, è caratterizzato da un presupposto piuttosto specifico e non del tutto plausibile, per cui ciascun giocatore preferisce rassegnarsi alla mancata produzione del bene piuttosto che produrlo da babbeo, cioè con i suoi soli sforzi (questo è il significato della disuguaglianza $w > z$ che i payoff del dilemma devono soddisfare).

A partire dagli anni Ottanta del secolo scorso, la tesi che *qualunque* problema dei beni pubblici rappresenti sempre un dilemma, cioè sia adeguatamente descritto come un caso particolare di dilemma del prigioniero, è stata rifiutata da un numero crescente di studiosi, fra i quali i già citati Sugden (1986), Taylor (1987) e de Jasay (1989). Essi hanno sostenuto che vi sono molti casi in cui ciascun giocatore preferisce produrre il bene pubblico con i suoi soli sforzi piuttosto che rassegnarsi alla sua mancata produzione. Per esempio, secondo Sugden (1986), la struttura di diversi giochi dei beni pubblici è essenzialmente identica a quella del cosiddetto (gioco del) cumulo di neve (*snowdrift game*), che deve il nome alla seguente storiella (*ibidem*, par. 7.3, p. 132):

Supponi di guidare la tua auto d'inverno su una strada isolata e di restare bloccato, assieme a un'altra auto, da un cumulo di neve. Sia tu sia l'altro viaggiatore siete stati previdenti e avete a bordo delle pale. È quindi chiaro che dovrete cominciare a spalare. O no? Dopotutto, l'altro non può aprirsi una via d'uscita senza, al tempo stesso, liberare la strada anche per te. Se pensi che sia in grado di fare il lavoro da solo, perché darti la pena di aiutarlo?

Il gioco del cumulo di neve è rappresentato in figura 4.4, dove i termini "strada libera" e "strada occupata" indicano, rispettivamente, il bene pubblico di cui i viaggiatori vorrebbero disporre e la mancata produzione di tale bene, mentre i termini "cooperare" e "defezionare" indicano, rispettivamente, le azioni di spalare e non spalare.

Il cumulo di neve è caratterizzato dalla disuguaglianza $z > w$, cioè dal fatto che ciascun giocatore preferisce liberare la strada spalando da solo piuttosto che rimanere bloccato.⁹ Tale disuguaglianza è l'aspetto distintivo della struttura dei payoff del cumulo di neve (in cui abbiamo che $y > x > z$

9 Come nota Sugden (1986/2004, p. 133), il gioco del cumulo di neve è equivalente al (cioè ha la stessa struttura di incentivi del) gioco "del pollo" o "del falco e della colomba", originariamente discusso da Maynard Smith (1982) in relazione al conflitto fra animali. Tali giochi sono stati ampiamente usati nell'analisi del problema dei beni pubblici: si vedano, per esempio, Taylor e Ward (1982), Taylor (1987, capitolo 2 e in particolare pp. 36 ss.) e de Jasay (1989, p. 187; 2002, p. 48). Per un confronto dettagliato fra questi contributi, si veda l'Appendice al capitolo 7 di de Jasay (1989).

		COLONNA	
		C	D
RIGA	C	strada libera nessun <i>free rider</i> x, x	strada libera Colonna <i>free rider</i> z, y
	D	strada libera Riga <i>free rider</i> y, z	strada bloccata w, w

Figura 4.4.
Il gioco del cumulo di neve

$> w$), struttura che, per il resto, è identica a quella del dilemma del prigioniero (in cui $y > x > w > z$).

Il modo migliore di mettere a confronto i giochi del dilemma del prigioniero e del cumulo di neve, rappresentati nelle figure 4.1 e 4.3, è interpretare esplicitamente i payoff x, y, w e z dei giocatori nei termini dei costi e dei benefici che un giocatore ottiene nei diversi esiti del gioco. In particolare, possiamo identificare tali payoff con la differenza $(v - c)$ tra il *valore* v che ogni giocatore attribuisce all'eventuale disponibilità del bene e il *costo* c che deve eventualmente pagare per la sua produzione (cfr. Sugden 1986, p. 133). Supponiamo, inoltre, che i valori di v e c vengano determinati in accordo con tre criteri intuitivamente molto plausibili:

- (1) poniamo $v = 0$ se il bene non è disponibile e $v > 0$ se è disponibile;
- (2) poniamo $c = 0$ se il giocatore non contribuisce alla produzione del bene;
- (3) poniamo $c_1 > c_2 > 0$, dove c_1 e c_2 sono, rispettivamente, il costo affrontato da un giocatore che contribuisce da solo alla produzione del bene e quello affrontato da un giocatore che lo produce cooperando con il concorrente.

Se adottiamo i criteri (1)-(3), allora i payoff x, y, w e z , attribuiti ai quattro possibili risultati del gioco, sono dati dalle seguenti uguaglianze:

$$\begin{aligned}
 x &= v - c_2 \\
 y &= v \\
 w &= 0 \\
 z &= v - c_1
 \end{aligned}$$

Si comprende facilmente che, se $v < c_1$, cioè se il valore attribuito alla disponibilità del bene è così basso che non vale la pena produrlo da soli, allora $w > z$; se, al contrario, $v > c_1$, allora $z > w$. Nel primo caso abbiamo a che fare con un dilemma del prigioniero, nel secondo con un gioco del cumulo di neve: la matrice in figura 4.5 può quindi rappresentare entrambi i giochi, a seconda dei valori relativi di v e c_1 .

Si può dimostrare che, diversamente da quanto accade nel dilemma del prigioniero, dove la strategia ottimale è la defezione, nel cumulo di neve la strategia ottimale non coincide con nessuna strategia pura a disposizione dei giocatori, bensì con una particolare strategia mista. Per motivi di semplicità espositiva, concentreremo l'attenzione sul caso particolare del gioco in cui il valore della strada libera e i suoi costi di produzione siano, rispettivamente, $v = 4$, $c_1 = 3$ e $c_2 = 1$, e quindi i payoff dei viaggiatori sono $x = 3$, $y = 4$, $w = 0$ e $z = 1$. Un'attenta considerazione di questo caso, rappresentato nella figura 4.6, basterà, infatti, a mettere in luce alcune importanti caratteristiche *generali* del cumulo di neve.

		COLONNA	
		C	D
RIGA	C	$(v - c_2)$ $(v - c_2)$	v $(v - c_1)$
	D	$(v - c_2)$ v	0 0

Figura 4.5.
Il gioco dei beni pubblici in forma generale: se $v < c_1$ la matrice è quella di un dilemma del prigioniero, se $v > c_1$ è quella di un gioco del cumulo di neve

		COLONNA	
		C	D
RIGA	C	3 3	4 1
	D	1 4	0 0

Figura 4.6.
Gioco del cumulo di neve con payoff $x = 3$, $y = 4$, $w = 0$ e $z = 1$

Anzitutto, la figura 4.6 mostra che la defezione è la migliore risposta alla cooperazione, mentre la cooperazione è la migliore risposta alla defezione. Ciò implica che nessuna strategia pura del gioco domina l'altra e che nessuna di esse è una strategia di equilibrio, dato che i risultati (C,C) e (D,D) non sono equilibri. La figura 4.6 mostra anche che il cumulo di neve presenta due equilibri, costituiti dai risultati (C,D) e (D,C) , cioè dalle combinazioni di strategie in cui uno dei viaggiatori coopera e l'altro defeziona. Tali caratteristiche del gioco indicano chiaramente che la strategia ottimale per il cumulo di neve – cioè la strategia che dovrebbe essere adottata da qualunque giocatore razionale – non coincide con nessuna strategia pura. Secondo la teoria classica dei giochi la strategia ottimale è una strategia mista $(p, 1 - p)$, dove il valore di p dipende dai valori assegnati ai payoff y, x, z, w . Per esempio, nel caso particolare rappresentato nella figura 4.6, la strategia ottimale è $(1/2, 1/2)$, cioè la strategia che consiste nella decisione di cooperare con una probabilità del 50%. Essa è infatti l'unica strategia di equilibrio del gioco, cioè l'unica strategia che rappresenta la migliore risposta a se stessa.

Dato il carattere probabilistico di $(1/2, 1/2)$, non possiamo stabilire con certezza quali azioni verranno attuate da due giocatori che adottano questa strategia e neppure quali payoff otterranno. Possiamo, però, determinare la probabilità delle loro azioni e il cosiddetto valore atteso dei loro payoff. Vediamo così che: (1) vi è una probabilità del 75% che almeno un giocatore cooperi e che, di conseguenza, il bene pubblico venga prodotto; (2) il valore atteso del payoff di ciascun giocatore è pari a 2.¹⁰ Ciò significa che, diversamente da quanto accade nel dilemma del prigioniero, nel cumulo di neve la possibilità che due giocatori razionali possano produrre un bene pubblico in base alla contribuzione volontaria è tutt'altro che evanescente.

Occorre notare, tuttavia, che il valore atteso, pari a 2, dei payoff di giocatori che adottano la strategia ottimale $(1/2, 1/2)$ risulta inferiore al payoff, pari a 3, che essi otterrebbero con la mutua cooperazione. Dobbiamo quindi ammettere che il gioco rappresentato nella figura 4.6 costituisce un dilemma sociale – anche se, in confronto al dilemma del prigioniero, si tratta di una forma, per così dire, attenuata di dilemma sociale. Si noti che questa conclusione non vale solo per il gioco rappresentato nella figura 4.6 ma, più in generale, per qualunque caso del cumulo di neve.

10 Tale valore atteso è approssimativamente uguale alla media dei *payoff* che, con ogni probabilità, un giocatore che adotta $(1/2, 1/2)$ otterrebbe in una lunga serie di incontri con un concorrente che adotta la stessa strategia.

Il lettore non di stupirà troppo nell'apprendere che, nella vita reale, i giocatori alle prese con il cumulo di neve cooperano molto più frequentemente di quanto previsto dalla teoria classica. I meccanismi che stanno alla base della loro accentuata tendenza alla cooperazione sono molto simili a quelli illustrati nel paragrafo 4.3 con riferimento al dilemma del prigioniero: infatti, tali tendenze sono alimentate dalle preferenze sociali di molti giocatori e, nel caso della versione ripetuta del cumulo di neve, dalla diffusa adozione di strategie di reciprocità.

Un ulteriore meccanismo che favorisce l'emergere della cooperazione è dato dalla possibilità che, nell'ambito di una popolazione di giocatori alle prese con il cumulo di neve, emergano determinate *convenzioni* in grado di rompere la simmetria del gioco (cfr. Sugden 1986, pp. 134 ss.). Parlando di simmetria ci riferiamo qui al fatto che, nel cumulo di neve, i due giocatori Riga e Colonna sono perfettamente intercambiabili, poiché dispongono di identiche strategie e hanno gli stessi payoff. Possiamo immaginare, tuttavia, che la simmetria venga infranta da una convenzione che permetta di distinguere senza ambiguità i due giocatori, sulla base di alcune loro caratteristiche prominenti, e di decidere chi dei due deve cooperare e chi invece può defezionare, giocando il ruolo di *free rider autorizzato*. Per esempio, la convenzione potrebbe prescrivere che, se Riga arriva per primo a ridosso del cumulo, allora deve spalare da solo, mentre Colonna, che è arrivato per secondo, può tranquillamente attendere, senza muovere un dito, che Riga liberi la strada. Se tutti i membri della popolazione seguono la convenzione, allora ogni incontro tra due giocatori Riga e Colonna si concluderà con uno dei risultati (C,D) e (D,C) ; in particolare, ogni giocatore otterrà un payoff pari a 4 se gli viene assegnato il ruolo del *free rider autorizzato* e un payoff pari a 1 se gli viene assegnato il ruolo di chi deve contribuire da solo alla produzione del bene.

Ci possiamo ora chiedere se a un individuo conviene appartenere a una popolazione dove vige una convenzione di questo tipo, piuttosto che a una popolazione dove non ne vige alcuna e i giocatori adottano la strategia ottimale $(1/2, 1/2)$. La risposta a questo interrogativo dipende dalla frequenza con cui, negli incontri con altri membri della popolazione, ogni individuo riceve i due ruoli di chi è autorizzato a defezionare e di chi deve cooperare unilateralmente. Consideriamo, per esempio, il caso in cui, in una lunga serie di incontri, ciascun ruolo viene attribuito a qualunque membro della popolazione con una frequenza pari al 50%. Un caso del genere è rappresentato dalla convenzione che attribuisce l'intero lavoro di spalatura al viaggiatore che arriva per primo a ridosso del cumulo, dato che si può ragionevolmente ritenere che ogni membro della popolazione riceverà

entrambi i ruoli con una frequenza che, in una lunga serie di incontri, è approssimativamente pari al 50%. Se vige una convenzione del genere, ognuno riceverà un payoff pari a 4 nel 50% degli incontri e un payoff pari a 1 nell'altro 50%; di conseguenza, il valore medio dei payoff di un giocatore, nell'insieme degli incontri, sarà pari a 2,5. Poiché tale valore è superiore a 2, cioè al valore medio dei *payoff* ottenuti dai membri di una popolazione di giocatori razionali che applicano costantemente la strategia ottimale (1/2, 1/2), un individuo è sicuramente avvantaggiato dall'appartenere a una popolazione in cui vige questa convenzione.

4.6 Dilemmi della sfiducia

Gli studiosi che si sono occupati delle origini evolutive del comportamento umano hanno suggerito che la cattura di grossi animali, effettuata da piccoli gruppi di cacciatori-raccoglitori, è uno dei primi esempi di produzione di beni pubblici della storia dell'umanità (cfr. Ridley 1996, p. 108 e Camerer 2003, pp. 376 ss.). Un semplice modello delle interazioni strategiche tra i partecipanti a una battuta di caccia è il gioco della caccia al cervo, che deve il suo nome a un passo di Jean-Jacques Rousseau nel *Discorso sull'origine dell'ineguaglianza* (1755, p. 75, traduzione modificata):

Se si dovesse cacciare un cervo, ognuno comprenderebbe che deve restare fermo e fiducioso al suo posto; ma se una lepre dovesse passare nei paraggi di uno dei cacciatori, non possiamo dubitare che egli lascerebbe il proprio posto per catturarla, senza alcuno scrupolo per gli altri, che perderebbero la loro preda.

In questo passaggio, Rousseau descrive una tipica situazione in cui la produzione di un bene pubblico richiede la cooperazione di *tutti* i membri del gruppo, ciascuno dei quali, tuttavia, potrebbe facilmente cedere alla tentazione di perseguire un piccolo beneficio immediato, senza curarsi del fatto che il suo cedimento farebbe fallire l'impresa comune.

Molti studiosi ritengono che situazioni di questo genere possano venire adeguatamente rappresentate dalla matrice in figura 4.7 – dove “cooperare” indica l'azione di tenere la propria posizione nella battuta di caccia e “defezionare” quella di abbandonarla per catturare una lepre (per un'analisi approfondita della caccia al cervo e delle sue applicazioni nelle scienze sociali, si veda Skyrms 2004).

La figura 4.7 mostra che il bene pubblico costituito dalla cattura del cervo viene prodotto *solo* nel caso in cui *entrambi* i giocatori cooperano:

		COLONNA	
		<i>C</i>	<i>D</i>
RIGA	<i>C</i>	cervo 10, 10	Colonna lepre 0, 2
	<i>D</i>	Riga lepre 2, 0	Riga lepre Colonna lepre 2, 2

Figura 4.7.
Il gioco della caccia al cervo

nessun giocatore può, quindi, sperare di giocare nell'allettante ruolo di *free rider*. Questa caratteristica della caccia al cervo rappresenta una notevole differenza con i giochi del dilemma del prigioniero e del cumulo di neve, nei quali il contributo di un solo giocatore basta a produrre il bene pubblico, consentendo così all'altro giocatore di comportarsi da *free rider*.

Dalla matrice del gioco, vediamo inoltre che la cooperazione è la migliore risposta alla cooperazione, mentre la defezione è la migliore risposta alla defezione. Ciò significa che nessuna strategia pura del gioco domina l'altra e che vi sono due equilibri, vale a dire la mutua cooperazione e la mutua defezione. Poiché cooperazione e defezione sono entrambe strategie di equilibrio, possiamo chiederci quale di esse verrebbe adottata da due giocatori ideali, del genere ipotizzato dalla teoria classica dei giochi. Uno sguardo alla figura 4.7 mostra che, con la mutua cooperazione, entrambi i giocatori ottengono un payoff, pari a 10, di gran lunga superiore a quello, pari a 2, associato alla mutua defezione. Ciò significa che l'equilibrio (C,C) Pareto-domina l'equilibrio (D,D) . Si noti, inoltre, che l'equilibrio di mutua cooperazione non è dominato da nessun altro equilibrio del gioco – in altre parole, è l'unico equilibrio Pareto ottimale del gioco. Non vi è, quindi, alcun dubbio che due giocatori ideali sceglierebbero la cooperazione.

Se la razionalità individuale viene definita nei termini della teoria classica dei giochi, allora si può affermare che la caccia al cervo *non* è un dilemma sociale. Infatti, la soluzione del gioco – determinata dalla combinazione delle scelte razionali dei partecipanti – non conduce affatto a un disastro collettivo, bensì al raggiungimento dell'interesse comune, dato dalla cattura del cervo. D'altra parte, non appena abbandoniamo le idealizzazioni della teoria classica per riflettere sull'interazione tra due cacciatori in carne e ossa, la certezza che essi sceglierebbero inevitabilmente la cooperazione comincia a sgretolarsi.

Infatti, diversamente da quanto ipotizzato dal presupposto TGC3 della teoria classica, può accadere che un cacciatore non sia praticamente perfetto e che, di conseguenza, ceda alla tentazione di defezionare, avventandosi su una lepre di passaggio. Ciò non si verificherebbe per una scelta consapevole del cacciatore, ma per effetto della sua debolezza di volontà. D'altra parte, la possibilità di defezioni di questo genere è il punto di partenza di un processo cognitivo che può condurre entrambi i giocatori alla decisione razionale di defezionare. Si supponga, infatti, che entrambi i cacciatori attribuiscono una probabilità non trascurabile all'ipotesi che il concorrente defezioni. Poiché la defezione è la migliore risposta alla defezione, potrebbe allora accadere che entrambi i cacciatori *decidano* di defezionare, in risposta alla probabile defezione altrui. Tale defezione non sarebbe determinata dall'imperfezione dei cacciatori ma dalla loro preoccupazione per l'imperfezione altrui – cioè dalla loro *sfiducia* circa la forza di volontà o le capacità cognitive del concorrente.

Inoltre, la “trappola della sfiducia” potrebbe manifestarsi anche in modi più insidiosi di quelli appena descritti. Potrebbe accadere, per esempio, che i giocatori Riga e Colonna – per quanto dotati di notevole forza di volontà e grande fiducia nell'altrui forza di volontà – condividano una sfiducia “di secondo livello”, nel senso che sono convinti che il concorrente abbia scarsa fiducia nella loro forza di volontà e che, quindi, tema la loro defezione. In una situazione di questo genere è probabile che entrambi decidano di defezionare. Infatti, se Riga ritiene molto probabile che Colonna ritenga molto probabile che Riga defezioni, allora Riga potrebbe decidere di defezionare per rispondere alla probabile defezione di Colonna, indotta dal suo probabile timore della defezione di Riga. Lo stesso ragionamento si applica, naturalmente, anche a Colonna. Il lettore ci perdonerà se sospendiamo qui la nostra esplorazione dei livelli di sfiducia reciproca, di numero virtualmente infinito, che possono affliggere i partecipanti alla caccia al cervo. Ci limitiamo a osservare che l'effetto dei diversi livelli di sfiducia ha carattere cumulativo. Non deve quindi sorprendere che, in molti casi, la combinazione di diversi livelli di sfiducia porti entrambi i giocatori ad attribuire un'elevata probabilità all'ipotesi che il concorrente defezioni, così da indurli a rispondere con la defezione alla sua probabile defezione. Queste considerazioni suggeriscono che, nella prospettiva delle teorie cognitive dei giochi, la caccia al cervo va considerata come un dilemma sociale. Infatti, il disastro collettivo associato alla mancata cattura del cervo viene determinato dalla combinazione delle scelte razionali dei due giocatori e, più precisamente, dalla loro decisione di defezionare presa sulla base all'elevata probabilità che attribuiscono all'ipotesi che il concorrente defezioni.

Le ricerche condotte nell'ambito delle teorie comportamentali e cognitive dei giochi hanno esplorato in profondità i delicati meccanismi della fiducia che governano le interazioni strategiche simili alla caccia al cervo. Tali ricerche suggeriscono che spesso i partecipanti alla caccia al cervo attribuiscono un'elevata probabilità alla defezione del concorrente e hanno, quindi, buone ragioni per defezionare. Infatti, la defezione fornisce una ragionevole *assicurazione* contro i rischi della defezione altrui.¹¹ Per esempio, nel gioco della figura 4.7, la defezione garantisce al giocatore un payoff, pari a 2, che non dipende in alcun modo dal fatto che il concorrente cooperi, oppure defezioni.

Un'interessante generalizzazione della caccia al cervo al caso di molti giocatori è costituita dal cosiddetto dilemma di Wolf, ampiamente studiato nell'ambito della teoria comportamentale dei giochi.¹² In questo gioco, una ventina di persone viene sistemata in laboratorio, ognuna davanti a un pulsante, con un cronometro che segna lo scorrere del tempo. Il gioco dura 10 minuti, durante i quali ogni soggetto deve decidere se premere o meno il pulsante. Mentre l'esperimento è in corso, nessuno può sapere se qualcun altro ha premuto il pulsante. Se, allo scadere del tempo, nessuno avrà premuto il pulsante, ciascun partecipante riceverà 10 euro. Se, invece, qualcuno lo ha premuto, tutti coloro che l'hanno premuto riceveranno 2 euro, mentre chi non l'ha premuto non riceverà nulla. Il lettore può controllare che la caccia al cervo rappresentato dalla figura 4.7 non è altro che la versione a due giocatori del dilemma di Wolf. Nella prospettiva della teoria classica dei giochi, la strategia ottimale di ciascun giocatore è la cooperazione, che consiste nel non premere il pulsante. Tuttavia gli esperimenti finora effettuati mostrano che la maggioranza dei soggetti, temendo la defezione altrui, preferisce invece defezionare, cosicché preme il pulsante prima dello scadere del tempo (cfr. Camerer 2003, cap. 7).

Quanto detto finora non deve, tuttavia, spingerci a nutrire un eccessivo pessimismo circa l'esito delle innumerevoli cacce al cervo praticate nella vita sociale. Dopo tutto il fatto che oggi siamo qui a discutere amabilmente di caccia al cervo, dipende anche dal successo che, nella lunga storia dell'umanità, ha coronato molte battute di caccia al cervo. Ciò significa che i nostri antenati hanno scoperto espedienti efficaci per disinnescare la trappola della sfiducia reciproca. Le attuali ricerche sulla caccia al cervo

11 Per questa ragione, la caccia al cervo va spesso sotto il nome di "gioco della garanzia" (*assurance game*); cfr. Sen (1967, p. 115).

12 Questo gioco è stato introdotto da Douglas Hofstadter (1985, p. 753), che ne attribuisce l'idea al matematico Robert Wolf.

mirano, per l'appunto, a far luce sulle condizioni che favoriscono il consolidarsi della fiducia, e quindi della cooperazione, tra coloro che partecipano a questo gioco. Sembra ormai chiaro, per esempio, che un elemento importante è costituito dal numero dei partecipanti: tanto più elevato è questo numero, tanto più difficile è l'instaurarsi della fiducia reciproca. L'importanza del numero dei giocatori per l'emergere della cooperazione nei giochi dei beni pubblici era già stata notata da Hume (1739, p. 570, traduzione modificata):

Due vicini possono mettersi d'accordo per prosciugare un canale, che posseggono in comune; poiché è facile per ognuno di loro conoscere le inclinazioni dell'altro, e ciascuno capisce che la conseguenza immediata di non fare la sua parte è l'abbandono dell'intero progetto. Ma è difficile, e invero impossibile, che un migliaio di persone si accordino per un'impresa del genere, poiché è difficile concertare un piano così complicato e ancora più difficile eseguirlo, dato che ciascuno cercherà un pretesto per evitare il fastidio e la spesa dell'impresa, in modo da scaricare tutto il fardello sugli altri.

Come si è visto, le teorie comportamentali e cognitive dei giochi suggeriscono che i partecipanti alla caccia al cervo cooperano molto meno di quanto ci si attenderebbe dalla prospettiva della teoria classica. Tuttavia, occorre ricordare che, nella vita sociale, il gioco della caccia al cervo si presenta molto spesso nella sua versione ripetuta, che in questi ultimi anni è stata approfonditamente esplorata nell'ambito delle teorie evoluzionistiche e adattive dei giochi (si veda Skyrms 2004). Le ricerche sull'argomento, condotte con un ampio impiego di metodi di simulazione computerizzata, hanno mostrato che l'evoluzione di una popolazione di giocatori che si affrontano ripetutamente nella caccia al cervo conduce, in un'ampia varietà di condizioni, all'emergere di strategie di reciprocità e convenzioni che favoriscono la mutua cooperazione. La genesi e l'evoluzione di convenzioni di questo tipo è l'argomento del prossimo capitolo.

5. COMUNICAZIONE, DETERRENZA E STRATEGIA DEL CONFLITTO

Storicamente, un forte impulso allo sviluppo della teoria dei giochi venne dalle ricerche sulla strategia militare e sulle relazioni internazionali promosse negli Stati Uniti a partire dalla fine della seconda guerra mondiale e per tutto il corso della guerra fredda. Un ruolo del tutto particolare in queste ricerche ha giocato, e gioca tuttora, la RAND Corporation (*Research ANd Development*), un *think tank* fondato alla fine degli anni Quaranta del secolo scorso per fornire servizi di analisi e ricerca strategica alle forze armate statunitensi. L'elenco dei ricercatori e dei collaboratori della RAND – che comprende ad oggi una trentina di premi Nobel, soprattutto matematici, fisici ed economisti – include i nomi di molti fra i maggiori teorici dei giochi: basti citare von Neumann, Nash, Schelling e Robert Aumann. Anche i matematici Merrill Flood (1908-1991) and Melvin Dresher (1911-1992) lavoravano alla RAND quando, nel 1950, cominciarono a studiare il gioco che poi Tucker avrebbe reso noto come “dilemma del prigioniero”. Un altro ricercatore della RAND, il teorico dei sistemi, esperto di strategia militare e futurologo Herman Kahn (1922-1983), descrive come segue uno dei problemi tipici che venivano discussi all'epoca (1962, p. 59)

Un esperimento di pensiero di cui ho fatto uso molte volte, e con molte varianti, negli ultimi venticinque o trent'anni comincia con la domanda: “Supponiamo che il presidente degli Stati Uniti venga informato che una bomba da molti megaton è stata lasciata cadere su New York. Cosa pensate che farebbe?” Nella metà degli anni Cinquanta, quando questa domanda fu posta per la prima volta, la risposta usuale era “Schiaccerebbe tutti i bottoni per lanciare gli ordigni nucleari e andrebbe a casa”. Il dialogo tra me e l'uditorio continuava più o meno così:

KAHN: E poi cosa accade?

UDITORIO: I sovietici fanno la stessa cosa.

KAHN: E dopo cosa succede?

UDITORIO: Nulla. Entrambe le parti sono state distrutte.

KAHN: Perché allora il Presidente ha agito così?

Ne seguiva un ripensamento generale del problema, e l'uditorio concludeva che, forse, il Presidente non avrebbe dovuto lanciare un immediato attacco totale di rappresaglia.

La situazione discussa da Kahn è un tipico esempio di interazione strategica, in cui due giocatori – in questo caso i presidenti degli Stati Uniti e dell'URSS – devono prendere decisioni che dipendono dalla scelte altrui. Ma ciò che l'esperimento mentale di Kahn vuole mettere in luce è il ruolo che una minaccia – in questo caso quella di una reciproca distruzione atomica – può giocare nel determinare l'esito dell'interazione. A partire da problemi di questo tipo, la teoria dei giochi è stata applicata all'analisi delle interazioni in cui la comunicazione di minacce (e promesse) influisce in modo fondamentale sulle scelte strategiche dei giocatori. Come vedremo meglio nel seguito, al di là delle sue origini "belliche", una teoria di questo tipo ha applicazione di più ampia portata, applicandosi all'analisi strategica di qualsiasi forma di comunicazione fra i giocatori. Così, uno dei nostri obiettivi è quello di mettere in luce la possibilità, ancora troppo spesso trascurata, di feconde interazioni tra le analisi strategiche dei conflitti internazionali e l'analisi filosofica della razionalità pratica: da un lato, gli strumenti concettuali propri dell'analisi filosofica potrebbero svolgere un ruolo rilevante per l'indagine sui fondamenti degli studi strategici; dall'altro, i problemi centrali di questi studi potrebbero costituire un importante banco di prova per lo sviluppo della teoria filosofica della razionalità pratica (cfr. Festa 2003, 2004).

In questo capitolo, ci concentriamo sul ruolo della *deterrenza* nella strategia atomica e, più in generale, nella cosiddetta teoria strategica (paragrafo 5.1). Sulla base della teoria del conflitto elaborata da Schelling (1960), che presentiamo nel paragrafo 5.2, analizziamo due importanti aspetti della strategia della deterrenza, cioè la logica della deterrenza (paragrafo 5.3) e la metodologia della deterrenza (paragrafo 5.4), per prendere infine in esame alcuni interessanti "paradossi della deterrenza" (paragrafo 5.5).

5.1. Teoria strategica e deterrenza

Sono in molti a ritenere che la fine dell'Unione Sovietica abbia scongiurato, almeno per il prossimo futuro, la possibilità di una terza guerra mondiale combattuta con armi atomiche. Sfidando questa opinione, alcuni studiosi sostengono che un terzo conflitto mondiale di carattere atomico è già stato combattuto. Questo conflitto si sarebbe concluso nel 1991 con la

resa dell'URSS, seguita dalla sua scomparsa come stato. La fine dell'Unione Sovietica non andrebbe quindi intesa come l'evento che ha impedito un conflitto atomico, bensì come l'esito di un conflitto atomico, cioè di un conflitto condotto anche, e soprattutto, attraverso un prolungato scambio di sofisticate *minacce atomiche*. Indipendentemente da ogni valutazione sulla sua plausibilità storica, questa tesi ha il merito di insistere sul ruolo fondamentale delle minacce e di altre strategie dissuasive nei conflitti tra stati, soprattutto se dotati di armi atomiche. Le ricerche sull'impiego delle strategie dissuasive nei conflitti internazionali vanno talvolta sotto il nome di *strategia della dissuasione*, o della *deterrenza*.¹³ Ci occuperemo qui delle basi teoriche della strategia della deterrenza, con particolare riferimento alla deterrenza atomica.

Natura e fini della teoria strategica

Il termine "strategia" viene impiegato in molte accezioni che si riferiscono, in un modo o nell'altro, a un identico oggetto, cioè alla "condotta e [alle] conseguenze delle relazioni umane nel contesto di un effettivo o possibile conflitto armato" (Luttwak, 1987, p. 22). Qui, però, siamo interessati soltanto alla strategia intesa come "teoria strategica" (*ivi*). Vi sono diversi modi di vedere la *natura* e i *fini* della teoria strategica. È opportuno ricordare che, anche per lo studio del comportamento strategico, vale una distinzione applicabile all'analisi di ogni forma di comportamento umano: quella tra teorie descrittive e normative. Le teorie strategiche *descrittive*, sviluppate soprattutto dagli studiosi di storia e sociologia militare, mirano alla spiegazione delle effettive strategie perseguite nei conflitti bellici. Pur riconoscendo il grande rilievo di tali teorie, qui ci occuperemo solo dell'approccio *normativo* all'analisi strategica, cioè delle teorie volte a stabilire i principi che governano la razionalità strategica.

L'idea, oggi molto diffusa, che la strategia possa avere carattere normativo risale almeno a Clausewitz (cfr. Luttwak 1987, p. 336). Vi sono state ampie discussioni sul fondamento e il livello di generalità delle pretese normative della teoria strategica. Nella seconda metà del Novecento, teorie strategiche caratterizzate da un elevato grado di generalità sono state sviluppate da alcuni studiosi che si occupavano di strategia atomica. Fra que-

13 L'anglismo "deterrenza" è entrato nell'uso comune con un significato molto aderente alla sua radice etimologica latina: la deterrenza, infatti, mira a *distogliere* l'avversario dal suo eventuale proposito di scatenare un'aggressione, sfruttando il suo *terrore* del contrattacco nucleare dell'agredito.

sti, un posto di rilievo va attribuito al generale André Beaufre, padre della forza atomica francese, nota anche come *force de frappe*. Beaufre (1963, p. 17) ritiene che l'analisi dei conflitti tra potenze atomiche imponga di superare l'antico concetto di strategia militare, vista come l'"arte di impiegare le forze militari per raggiungere i risultati determinati dalla politica". Questa definizione, infatti, è "troppo ristretta in quanto si riferisce solamente alle forze militari" (*ivi*); occorre, invece, includere nel dominio della strategia anche l'impiego di "forze materiali" diverse da quelli militari – per esempio, la forza economica – e di forze non-materiali che, in mancanza di termini migliori, possiamo definire "psicologiche".

Appare, quindi, più appropriato intendere la strategia come "l'arte di fare concorrere la forza per raggiungere gli scopi della politica" (*ivi*). Riprendendo un'intuizione del maresciallo Ferdinand Foch, Beaufre (*ibidem*, p. 18) definisce, più compiutamente, la strategia come

[...] l'arte che consente, a prescindere da qualsiasi tecnica, di dominare i problemi che ogni duello pone in sé, e di impiegare le tecniche con la massima efficacia; è quindi l'arte della dialettica delle forze o ancora, più precisamente, *l'arte della dialettica delle volontà che usano la forza per risolvere il loro conflitto*.

Poiché il conflitto tra stati è un caso peculiare di "dialettica delle volontà" – dove ciascuno dei contendenti cerca di far prevalere la propria volontà su quella altrui – non deve sorprendere che i mezzi utilizzabili a questo scopo possano includere le forze psicologiche. Infatti, nella "dialettica delle volontà" che contrappone due stati, ciascuno cerca "di ottenere dall'avversario un certo tipo di decisione". Questa "decisione è un avvenimento di carattere psicologico che si vuol produrre nell'avversario: convincerlo, appunto, che impegnare la lotta o proseguirla è perfettamente inutile" (*ivi*).

Non è difficile cogliere la distanza che separa la visione strategica di Beaufre dalla concezione "tradizionale" – spesso attribuita a Clausewitz – per la quale la strategia si occupa di come far terminare la guerra attraverso una battaglia vittoriosa. Il carattere angusto di tale concezione emerge chiaramente in tutti quei casi in cui gli scopi della politica *non* possono venire raggiunti attraverso una guerra aperta, come, per esempio, quando lo scopo politico di uno stato consiste nel semplice mantenimento dello *status quo*.

La strategia della deterrenza

Grazie al suo livello di generalità, la concezione strategica di Beaufre è applicabile a tutta la grande varietà di obiettivi che possono venire stabiliti dalla politica. In particolare, tale concezione fornisce un'appropriata corni-

ce teorica per l'analisi dei problemi della dissuasione militare che sorgono nell'ambito dei conflitti tra potenze atomiche. In una guerra combattuta attraverso lo scambio di colpi atomici le forze armate tradizionali non sono in grado di proteggere il territorio di uno stato dalla distruzione fisica e dalla contaminazione nucleare. Quale dovrà essere, allora, la funzione di tali forze nei conflitti tra potenze atomiche? E, soprattutto, con quali altri mezzi ci si potrà difendere da un attacco atomico?

Il tentativo di rispondere a domande di questo tipo ha dato origine alla cosiddetta *strategia atomica*, che rappresenta una parte fondamentale delle moderne teorie strategiche. In linea di principio, il rischio di un attacco atomico può venire affrontato in quattro modi diversi, non necessariamente incompatibili, vale a dire: "1) *La distruzione preventiva* delle armi avversarie [...]; 2) *L'intercettazione delle armi atomiche* [...]; 3) *La protezione fisica* contro gli effetti delle esplosioni [...]; 4) *La minaccia di rappresaglia*" (Beaufre, *ivi*, p. 54). Attraverso quest'ultima si tenta di dissuadere l'avversario dall'effettuare un "primo colpo" atomico, minacciandolo di rispondere con un "secondo colpo" che gli infliggerà danni intollerabili.

Si noti che le modalità di impiego dei mezzi menzionati ai punti 1-3 dipendono, in larga misura, dalle specifiche tecnologie disponibili in un determinato momento. Al contrario, la peculiare natura della minaccia di rappresaglia la rende relativamente indipendente dagli sviluppi tecnologici. Anche se, naturalmente, l'efficacia di tale minaccia si basa su un ovvio presupposto tecnologico – costituito dall'effettiva disponibilità di una *force de frappe* e dalla sua potenza – va rilevato che i problemi fondamentali della minaccia di rappresaglia non sono di carattere tecnologico, bensì "psicologico". Tali problemi, che riguardano la dinamica dell'interazione tra gli attori del conflitto, hanno a che fare con interrogativi di questo genere: in che modo si può convincere l'avversario della credibilità di una minaccia? Come si possono combinare in un unico schema di comportamento minacce e promesse? Qual è la strategia dissuasiva ottimale tra quelle a disposizione?

La strategia della deterrenza cerca di rispondere a interrogativi di questo tenore, suggerendo i metodi più efficaci per "raggiungere direttamente la volontà dell'avversario senza dover passare attraverso una prova di forza" (Beaufre, *ibidem*, p. 57). Mentre quella che potremmo chiamare *strategia applicativa* si propone di identificare le migliori *modalità d'uso* dei sistemi d'arma, la strategia della deterrenza "non concerne l'*applicazione* efficiente della forza, ma l'*utilizzo di una forza potenziale*" (Schelling, 1960, p. 5): così, per esempio, la strategia della deterrenza atomica non ha a che fare con l'*effettivo* uso delle armi atomiche, ma con i modi più efficaci per

sfruttare la “forza potenziale” derivante dalla *possibilità* di usare tali armi. Anche se la distinzione tra strategia applicativa e strategia della deterrenza può applicarsi all’intera teoria strategica, tale distinzione si è precisata solo con lo sviluppo della strategia atomica, nel cui ambito il ruolo della strategia della deterrenza è apparso ben presto fondamentale.

Inoltre, vale la pena notare che la nozione di sfruttamento della forza potenziale - che sta alla base della strategia della deterrenza - non riguarda soltanto la *gestione* delle forze militari dispiegate dai contendenti in un dato momento, ma si applica a *tutte* le fasi del conflitto. Strategie dissuasive di vario genere possono venire utilizzate, infatti, anche nelle fasi precedenti e successive al dispiegamento delle forze militari, vale a dire nella *preparazione* dello scontro e nella sua *conduzione*. Per quanto riguarda la preparazione della guerra, basti ricordare che, prima che i sistemi d’arma siano dispiegati per il combattimento, occorre che vengano fatti uscire da caserme, porti e arsenali. Prima ancora, però, occorre studiarne la fattibilità, progettarli, costruirne i prototipi, collaudarli, avviarne la produzione su larga scala, addestrare i reparti operativi al loro impiego, e così via. È evidente che, in relazione a ciascuna di queste operazioni, si possono mettere in atto minacce, promesse e altre strategie dissuasive. Per esempio, un governo potrebbe minacciare di dispiegare le proprie forze nucleari in modo tale da renderle pronte all’uso con un limitato preavviso, o minacciare di investire un’enorme quantità di fondi nella progettazione di un efficiente sistema anti-missili, o promettere di non produrre per i prossimi cinque anni un determinato sistema d’arma.

Per quanto riguarda la fase della conduzione dello scontro bellico, occorre notare che una guerra può essere combattuta con diversi gradi di intensità. Prima ancora di individuare i più efficienti modi di impiego dei sistemi d’arma a disposizione, occorre decidere a quale livello di intensità condurre lo scontro. Mentre il primo compito è proprio della strategia applicativa, è evidente che le decisioni relative al livello di intensità dello scontro riguardano da vicino la strategia della deterrenza. Infatti, persino nel corso dei conflitti più violenti, i contendenti possono tacitamente accordarsi di non superare certe soglie di intensità o, al contrario, minacciare di farlo. La storia militare ci mostra che promesse e minacce di questo genere funzionano piuttosto bene, al punto che le guerre non vengono quasi mai combattute al massimo grado di intensità, e che virtualmente ogni guerra può venire considerata come una *guerra limitata*.¹⁴

14 Il mancato impiego dei gas nella seconda guerra mondiale costituisce un esempio classico del carattere limitato di quasi tutti i conflitti militari, inclusi quelli più violenti ed estesi.

Poiché i problemi relativi all'efficiente applicazione delle strategie dissuasive, nelle diverse fasi di un conflitto tra stati, riguardano più i meccanismi generali dell'interazione umana che le caratteristiche tecnologiche dei sistemi d'arma, è lecito supporre che le competenze richieste per lo sviluppo di una strategia della deterrenza siano diverse da quelle proprie dei militari di professione. Questa è in effetti l'opinione di Schelling (1960, p. 10):

possiamo fare un'utile distinzione tra l'*applicazione* della forza e la *minaccia* della forza. La deterrenza riguarda l'utilizzo di una forza potenziale. Essa consiste nel persuadere un potenziale nemico che nel suo stesso interesse dovrebbe evitare certi tipi di attività. C'è un'importante differenza fra le qualità intellettuali necessarie per portare a compimento una missione militare e quelle necessarie per utilizzare una capacità militare *potenziale* per perseguire gli obiettivi di una nazione. Una teoria della deterrenza è, di fatto, una teoria dell'abile *non-uso* della forza militare e a questo fine la deterrenza richiede capacità ben più ampie di quelle strettamente militari.

Come emerge chiaramente dalle ricerche condotte da Schelling e da altri studiosi, tali competenze comprendono la capacità di analizzare il comportamento razionale nei contesti di interazione tra due o più agenti, avvalendosi anche di strumenti concettuali molto raffinati come, per esempio, la teoria dei giochi.¹⁵

5.2. Strategie dissuasive e teoria del conflitto

Molte interazioni sociali presentano una struttura facilmente riconducibile ai modelli formali dalla teoria dei giochi. Molte altre, però, presentano elementi che, nonostante la loro presumibile rilevanza per il processo decisionale, restano fuori dal quadro concettuale della teoria classica dei giochi. Uno di questi elementi consiste nell'impiego di minacce, promesse e altre *strategie dissuasive*.

Nel tentativo di comprendere la natura, il funzionamento e il carattere razionale delle strategie dissuasive, il quadro concettuale della teoria dei

15 Per questo motivo, dovrebbe risultare del tutto naturale l'idea che i filosofi e gli studiosi interessati all'analisi della razionalità e del comportamento umano possano contribuire allo *sviluppo* della teoria strategica. Finora, tuttavia, i filosofi si sono impegnati prevalentemente nella critica dei presupposti ideologici delle teorie strategiche e nella denuncia dell'inaccettabilità etico-politica dell'idea stessa che un intellettuale possa occuparsi dello sviluppo di una teoria strategica adeguata all'età atomica.

giochi è stato esteso in diverse direzioni. Le ricerche su questi problemi hanno dato origine a un campo di studi interdisciplinari – al quale ci riferiremo con il termine “teoria del conflitto” – che ha trovato la sua prima organica trattazione in *The Strategy of Conflict* (1960) di Schelling.¹⁶

I conflitti tra individui – o altri tipi di agenti, come partiti o stati – possono venire analizzati da diversi punti di vista (cfr. Arielli e Scotto 1998). Per esempio, un’analisi sociologica dei conflitti non mancherebbe di descriverne anche gli aspetti irrazionali. La teoria del conflitto, invece, si occupa solo degli aspetti razionali e coscienti del conflitto, visto come un contesto in cui i partecipanti cercano di “vincere” attraverso un “comportamento conflittuale consapevole, intelligente e sofisticato” (Schelling 1960, p. 3). Alla pari della teoria dei giochi, con la quale condivide il carattere normativo, la teoria del conflitto si propone di identificare i principi che governano il comportamento di un agente razionale coinvolto in un conflitto con altri agenti razionali (*ibidem*, pp. 16-17). Si noti che i conflitti più interessanti e diffusi, anche nelle relazioni internazionali, possono venire rappresentati come giochi misti, piuttosto che come giochi a somma zero. Ciò significa che, come nota Schelling (*ibidem*, pp. 4-5), il concetto di vittoria non coincide necessariamente con la rovina dell’avversario:

Il conflitto puro, in cui gli interessi dei due antagonisti sono completamente opposti, è un caso eccezionale; potrebbe darsi in una guerra di sterminio totale, ma altrimenti neppure in guerra. Per questa ragione, “vincere” in un conflitto non ha un significato strettamente competitivo; non vuol dire vincere rispetto ai propri avversari. Significa piuttosto guadagnare relativamente al proprio sistema di valori; e questo si può ottenere contrattando, attraverso un reciproco accomodamento, ed evitando comportamenti reciprocamente dannosi.

In altre parole, può accadere che, in un conflitto rappresentato da un gioco misto, alcuni dei possibili risultati siano *congiuntamente indesiderabili*, nel senso entrambi i giocatori ricevono un payoff molto basso. Il comune interesse a evitare che il risultato del conflitto sia congiuntamente indesiderabile non basta, tuttavia, a mettere d’accordo i giocatori, dato che le loro preferenze tra gli altri risultati saranno diverse: potrebbe persino accadere che qualche risultato sia rovinoso per un giocatore ed estremamente desiderabile per l’altro. Come abbiamo visto nei capitoli precedenti, i principi proposti dalla teoria dei giochi per “risolvere” i giochi a somma variabile

16 Questo campo di studi è stato variamente denominato come “teoria della negoziazione”, “teoria della strategia”, “strategia del conflitto” e “teoria del conflitto”: cfr. Schelling (1960, pp. xiv e 3).

non sempre garantiscono a entrambi i giocatori la possibilità di evitare risultati altamente indesiderabili. Fortunatamente, nella vita reale, gli agenti riescono abbastanza spesso a scongiurare tali risultati ricorrendo a svariate forme di comunicazione, tacita o esplicita, attraverso le quali veicolano minacce, scambiano promesse o, più in generale, conducono qualche forma di negoziazione. Come vedremo, la teoria del conflitto cerca di spiegare il funzionamento – e il carattere razionale – di svariate strategie dissuasive che danno prova di grande efficacia nella vita reale.

5.3. Logica della deterrenza

Consideriamo un gioco misto a due giocatori. Poiché la mossa del nostro concorrente dipende dalle sue attese circa le nostre scelte, è naturale pensare che potremmo modificare a nostro vantaggio le sue scelte, modificando le sue attese circa le nostre. Qualcosa del genere accade effettivamente in molte interazioni sociali, quando almeno uno dei giocatori cerca di modificare le attese – e, quindi, le scelte – dell'avversario attraverso minacce, promesse e altre strategie dissuasive.

Mosse strategiche

Le strategie dissuasive sono ben diverse dalle “mosse di base” del gioco, consistenti nell'effettuare determinate azioni rappresentate, a seconda dei casi, da una riga o una colonna della matrice di gioco. Tali strategie corrispondono, invece, alle *mosse strategiche*, così caratterizzate da Schelling (1960, p. 186):¹⁷

Se l'essenza di un gioco di strategia è la dipendenza della corretta scelta d'azione di ogni persona dipende da ciò si aspetta faccia l'altra, potrebbe essere utile definire una “mossa strategica” come segue: una mossa strategica è una mossa che influenza la scelta dell'altra persona, in maniera a noi favorevole, influenzando le aspettative dell'altra persona sul modo in cui noi ci comporteremo. In questo modo, si vincola la scelta che farà il partner, vincolando il proprio comportamento. L'obiettivo è quello di [...] comunicare in maniera convincente all'altro giocatore una modalità di comportamento (includere le risposte condizionate al comportamento dell'altro) che lascerà all'altro un semplice problema di massimizzazione la cui soluzione sarà un ottimo per noi, e distruggerà la capacità dell'altro di fare la stessa cosa.

17 Sul concetto di mossa strategica si vedano anche Ullmann-Margalit (1977, pp. 164-167) e Bicchieri (1993, p. 190).

Le mosse strategiche non sono semplicemente diverse dalle mosse di base, ma si pongono, per così dire, a un livello superiore rispetto a queste ultime. Infatti, effettuare una mossa strategica significa assumere, prima ancora di cominciare il gioco, certi impegni circa le mosse di base che si attueranno nel corso del gioco.

Quando un giocatore effettua una mossa strategica si impegna in anticipo ad attuare una linea d'azione che preferirebbe evitare, se non fosse vincolato a mantenere l'impegno. L'impegno ad attuare una determinata linea d'azione non può, quindi, consistere semplicemente nell'annunciare alla controparte che la si attuerà: esso consiste, piuttosto, nel vincolarsi visibilmente alla sua attuazione. Quando una mossa strategica viene effettuata nella maniera appropriata, ricorrendo a vincoli visibili, parleremo di *applicazione* della mossa.¹⁸

Vi sono tanti generi di mosse strategiche, o strategie dissuasive, quanti sono i generi di impegni che, prima di cominciare a giocare, possiamo assumere in merito alla nostra futura condotta di gioco. Il giocatore che deve "muovere" per primo può assumere solo impegni non-condizionali, vincolandosi alla scelta di una determinata mossa di base. Se, invece, deve muovere per secondo, può assumere anche diversi tipi di impegni condizionali, può cioè impegnarsi a rispondere in un certo modo alle possibili scelte del concorrente. Per esempio, se i due giocatori hanno solo due mosse a disposizione, Riga potrebbe impegnarsi a scegliere la prima delle proprie mosse, se Colonna sceglierà la seconda, o a scegliere la seconda se Colonna sceglierà la prima. Potrebbe anche assumere entrambi gli impegni e, più in generale – se la matrice del gioco include molte righe e colonne – potrebbe assumere sofisticati impegni condizionali, preannunciando uno "schema di risposta" per ogni possibile mossa dell'avversario.

Nella vita di ogni giorno, risulta piuttosto semplice distinguere tra minacce, promesse e diversi altri impegni. Tuttavia, non è immediatamente chiaro se, e come, queste distinzioni possano venire formulate rigorosamente sulla base della matrice di gioco e, più specificamente, della struttura dei payoff. Il primo compito della teoria del conflitto è proprio quello di analizzare la

18 Usiamo "applicazione" nel senso del termine inglese *enforcement*. Questo vocabolo, ampiamente usato in contesti giuridici, può essere tradotto, a seconda dei casi, con "esecuzione", "sanzione" o "applicazione": per esempio, "*enforcement of judgments (orders)*" indica l'esecuzione di una sentenza o di un provvedimento, mentre "*enforcement of the law*" indica l'applicazione della legge. Il concetto di *enforcement* viene usato anche nel linguaggio comune, con riferimento a minacce, promesse e altri tipi di impegni: per questo genere di uso, il termine semi-tecnico "applicazione" ci sembra preferibile ad altre possibili traduzioni.

struttura formale delle diverse strategie dissuasive. In mancanza di un termine standard, ci riferiremo a tale analisi – che rappresenta, per così dire, la componente pura della teoria del conflitto – con l'espressione *logica della deterrenza*. Oltre alla logica della deterrenza, la teoria del conflitto comprende anche una componente applicata – che potremmo denominare *metodologia della deterrenza* – il cui obiettivo è quello di indagare i metodi di applicazione delle mosse strategiche, cioè gli accorgimenti cui un giocatore può ricorrere per vincolarsi visibilmente a una linea d'azione.

Avvertimenti

La più semplice strategia dissuasiva è costituita da un tipo di impegno non-condizionale che possiamo chiamare *avvertimento*. La struttura formale di un avvertimento può venire illustrata con riferimento al gioco rappresentato nella figura 5.1.

In base al principio di dominanza, Riga dovrà scegliere la sua strategia dominante R1. Poiché questa circostanza è nota a Colonna, quest'ultimo potrà massimizzare il suo payoff scegliendo C1. La soluzione del gioco sarà quindi (R1,C1), che assegna a Riga e Colonna, rispettivamente, i payoff 1 e 7. Come si può vedere dalla matrice, tale risultato non è particolarmente desiderabile per Riga, dato che la sua funzione dei payoff include valori di gran lunga superiori a 1.

Mentre, in assenza di comunicazione, Riga deve rassegnarsi al payoff 1, se gli viene consentito di comunicare con Colonna potrà sfruttare questa opportunità per migliorare la propria sorte a danno di Colonna. Infatti, Riga potrebbe avvertire Colonna che attuerà comunque la mossa R2. Se Colonna ritiene credibile l'avvertimento di Riga, allora la sua scelta si riduce a quella tra le celle (R2,C1) e (R2,C2): poiché il suo payoff in (R2,C2) è più elevato

		COLONNA	
		C1	C2
RIGA	R1	1 7	9 3
	R2	0 0	8 2

Figura 5.1.
Riga avverte Colonna che sceglierà R2

di quello in (R2,C1), Colonna farà bene a scegliere R2, consentendo in tal modo a Riga di ottenere un payoff pari a 8. Si noti che il successo di Riga viene ottenuto a spese di Colonna, che deve ora accontentarsi del payoff 2.

Minacce

Un ruolo di preminenza fra le strategie dissuasive utilizzate nella vita di ogni giorno spetta, presumibilmente, alle *minacce*. Nell'analisi della struttura formale di minacce, promesse e altre strategie dissuasive, si rivela molto utile ricorrere a un'ingegnosa rappresentazione grafica dei giochi a due introdotta da Schelling (1960, pp. 55 ss.).

Per esempio, il grafico della figura 5.2 (Schelling 1960, p. 57, Fig. 2) rappresenta un gioco nel quale i due giocatori – che chiameremo Nord ed Est – hanno una coppia di alternative ciascuno: Nord può scegliere fra le mosse A e a , mentre Est può scegliere fra le mosse B e b . I quattro risultati del gioco, associati alle possibili combinazioni delle mosse dei due giocatori, corrispondono ai punti AB , Ab , aB e ab . Per ciascun risultato, il payoff di Est viene misurato orizzontalmente sull'asse delle ascisse, mentre quello di Nord viene misurato verticalmente sull'asse delle ordinate. Questa rappresentazione permette di distinguere molte varianti qualitativamente diverse di giochi a due, in base alle posizioni relative di AB , Ab , aB e ab .

A livello informale, la minaccia può venire descritta come una strategia dissuasiva in forma condizionale, il cui tratto distintivo consiste nel fatto che chi applica la minaccia – per dissuadere il concorrente dall'attuazione di una determinata mossa – non avrebbe alcun interesse a metterla davvero in atto, se si trovasse di fronte all'eventualità di dover castigare il trasgressore. Il successo della minaccia non consiste, quindi, nella sua attuazione: una minaccia riuscita è, invece, quella che non occorre attuare, perché ha effettivamente raggiunto lo scopo di distogliere la controparte da una determinata linea d'azione.¹⁹

Un'adeguata analisi della struttura formale della minaccia, compatibile con la caratterizzazione appena suggerita, dovrebbe rispondere ai seguenti interrogativi: (1) quali caratteristiche deve avere la matrice di

19 Luce e Raiffa (1957, p. 111) identificano le minacce con gli impegni espressi attraverso locuzioni del tipo “Questo farà più danno a te che a me”. Nella sua nitida caratterizzazione del concetto intuitivo di minaccia, Schelling (1960, p. 42) osserva che “qualcuno minaccia di fare qualcosa che non avrebbe incentivo a fare, ma che ha funzione di deterrente con la promessa di un danno reciproco [...] Il carattere distintivo [della] minaccia è che colui che minaccia non ha incentivi a realizzare quanto dice né prima né dopo l'evento critico.”

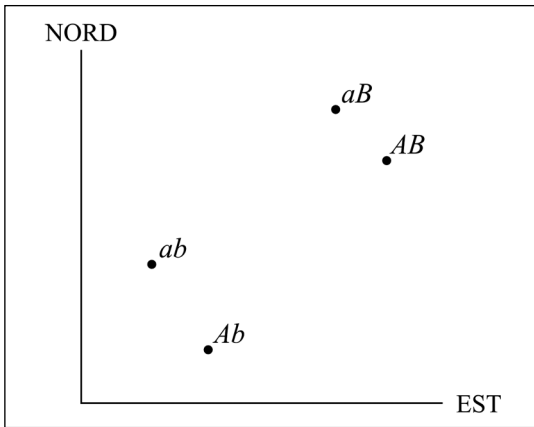


Figura 5.2.
Est minaccia Nord di
rispondere con b ad a

gioco affinché, in primo luogo, una minaccia *possa* riuscire e, in secondo luogo, l'autore della minaccia non abbia alcun interesse ad attuarla, nel caso di insuccesso? (2) se chi rivolge una minaccia non ha interesse ad attuarla, come può convincere il concorrente che un'eventuale trasgressione verrà punita? Rivolgeremo ora la nostra attenzione al primo interrogativo, che riguarda la logica della deterrenza, rinviando l'esame del secondo – che concerne, invece, la metodologia della deterrenza – al prossimo paragrafo.

Un esempio di struttura dei payoff che rende possibile l'impiego di una minaccia è quella associata al gioco della figura 5.2, dove supponiamo che Nord debba muovere per primo (cfr. Schelling 1960, pp. 56-57). Poiché sia Nord sia Est hanno a disposizione strategie dominanti costituite, rispettivamente, da a e B , la soluzione del gioco è aB . Si noti che tale risultato – mentre è preferibile a ogni altro dal punto di vista di Nord – si trova solo al secondo posto, dopo AB , nell'ordine di preferenza di Est.

In assenza di comunicazione, Est deve rassegnarsi ad aB , ma le sue prospettive cambiano se gli viene consentito di comunicare con Nord. In questo caso, infatti, Est può migliorare le sue prospettive ai danni di Nord attraverso un'appropriata minaccia volta a dissuadere Nord dall'attuazione di a . Più precisamente, Est può minacciare Nord di rispondere con b alla mossa a : se, in risposta all'eventuale trasgressione di Nord, tale minaccia venisse attuata, il gioco si concluderebbe con ab , cioè con un risultato congiuntamente indesiderabile. La minaccia di Est può avere due diversi esiti:

Esito 1. La minaccia ha successo. Ciò significa che Nord crede alla minaccia di Est. La sua scelta si restringe, quindi, alle opzioni ab e AB : poiché Nord preferisce la seconda, rinuncerà ad a in favore di A . Attuando la sua strategia dominante B , Est ottiene ora AB , cioè un risultato preferibile ad Ab e a ogni altro risultato.

Esito 2. La minaccia non ha successo. Ciò significa che, non credendo alla minaccia di Est, Nord attua la sua strategia dominante a . La scelta di Est si restringe allora a due alternative: (i) punire la trasgressione di Nord attuando b , così da ottenere ab ; (ii) rinunciare alla punizione attuando B , così da ottenere aB . Si noti che, dal punto di vista di Est, il risultato aB della rinuncia è nettamente preferibile al risultato ab della punizione.

Nel gioco della figura 5.2, alcune caratteristiche della struttura dei payoff – vale a dire le relazioni tra ab , AB e Ab , indicate nella descrizione dell'Esito 1 – corrispondono al fatto che la minaccia può riuscire. Altre caratteristiche della struttura dei payoff – vale a dire le relazioni tra ab e aB , indicate nella descrizione dell'Esito 2 – corrispondono invece al “carattere distintivo” della minaccia, cioè al fatto che chi minaccia non ha incentivi a punire il trasgressore. Ciò significa che le relazioni tra AB , Ab , aB e ab indicate nella descrizione degli Esiti 1 e 2, consentono di classificare come una minaccia l'impegno di Est a rispondere con b all'eventuale mossa a di Nord.

Dall'analisi delle alternative (i) e (ii) di Est nell'Esito 2 emerge che, se la minaccia non riesce, Est non ha alcun interesse a punire l'eventuale trasgressione di Nord. Ne segue che, se Est potesse decidere liberamente e razionalmente sulla sola base dei suoi payoff, non attuerebbe la propria minaccia. Più in generale, se i giocatori fossero sempre in grado di scegliere liberamente e razionalmente – e questa circostanza fosse universalmente nota – nessuna minaccia verrebbe mai attuata e, quindi, nessuna minaccia avrebbe successo. Questo implica che la credibilità e il successo di una minaccia dipendono dalla capacità del suo autore di sbarazzarsi della propria libertà di scegliere se infliggere o no la punizione e, più precisamente, dalla sua capacità di vincolarsi visibilmente all'attuazione della minaccia, in risposta a un'eventuale trasgressione.²⁰

La struttura formale di una minaccia può venire analizzata anche sulla base della consueta rappresentazione di un gioco in forma di matrice. Si consideri, per esempio, il gioco della figura 5.3, dove si suppone che Riga muova per secondo (cfr. Schelling 1960, p. 147, matrice di sinistra della

20 L'analisi degli accorgimenti necessari a vincolare visibilmente il proprio comportamento rientra, come si è detto, nel campo della metodologia della deterrenza, di cui ci occuperemo nel prossimo paragrafo.

		COLONNA	
		C1	C2
RIGA	R1	1 2	2 1
	R2	0 0	0 0

Figura 5.3.
Riga minaccia Colonna di rispondere con R2 a C1

Fig. 9). Poiché Riga e Colonna dispongono di strategie dominanti date, rispettivamente, da R1 e C1, la soluzione del gioco è (R1,C1).

Come si vede dalla matrice, (R1,C1) si trova al primo posto nell'ordine di preferenze di Colonna, ma solo al secondo in quello di Riga. In assenza di comunicazione, Riga deve rassegnarsi a ottenere (R1,C1); se invece gli si permette di comunicare con Colonna, può migliorare le sue prospettive minacciando Colonna di rispondere con R2 alla sua eventuale mossa C1. Se Colonna crede alla minaccia di Riga, allora la sua scelta si riduce alle opzioni (R2,C1) e (R1,C2): poiché Colonna preferisce la seconda, rinuncerà a C1 in favore di C2, consentendo così a Riga di ottenere un payoff 2, invece del payoff 1 che avrebbe ottenuto senza ricorrere a minacce.

Promesse

Anche se le promesse costituiscono una strategia dissuasiva di ampio impiego, non è affatto ovvio come si possa dare una definizione rigorosa del concetto di promessa. Sembra chiaro che un giocatore *X*, nel fare una promessa al proprio concorrente *Y*, si impegna ad attuare una mossa che, oltre ad essere vantaggiosa per lui, verrebbe favorevolmente accolta da *Y*. Questa formulazione esprime l'idea intuitiva che una promessa è il contrario di una minaccia: infatti, mentre si minaccia qualcosa di congiuntamente indesiderabile, si promette qualcosa di congiuntamente desiderabile. Ciò significa che una promessa riuscita può venire intesa come uno "scambio di favori" tra i giocatori.

Conviene distinguere tra due tipi di promesse, che si differenziano anche in relazione al modo in cui lo scambio di favori viene ottenuto (Schelling 1960, pp. 153 ss.):

(1) la *promessa* è un impegno condizionale *unilaterale* preso dal giocatore che muove per secondo, allo scopo di incentivare la controparte a fare una scelta vantaggiosa per entrambi;

(2) lo *scambio di promesse* è un impegno non-condizionale *bilaterale*, in forma di *quid pro quo*, preso da giocatori che muovono simultaneamente.

Possiamo illustrare la struttura formale di una promessa unilaterale con riferimento al gioco della figura 5.4, dove si suppone che Nord debba muovere per primo (Schelling 1960, p. 58, Fig. 3).

La soluzione del gioco può venire facilmente determinata. Infatti, in base al principio di dominanza, Nord dovrà scegliere la sua strategia dominante a : poiché si assume che questa circostanza sia nota a Est, la strategia ottimale di quest'ultimo, cioè la strategia che gli consentirà di massimizzare il proprio payoff, sarà B : ne segue che la soluzione del gioco è aB . Tuttavia, come si vede immediatamente dal grafico, aB è un risultato congiuntamente indesiderabile; infatti, entrambi i giocatori preferirebbero di gran lunga Ab ad aB .

In assenza di comunicazione, i giocatori devono rassegnarsi ad aB , ma le loro prospettive migliorano nettamente se la comunicazione viene consentita. In tal caso, infatti, Est può promettere a Nord di rispondere con b alla sua eventuale mossa A . Se Nord si fidasse della promessa di Est, e se Est mantenesse la propria promessa, il gioco si concluderebbe con Ab , cioè con un risultato congiuntamente desiderabile. La promessa di Est può avere due diversi esiti:

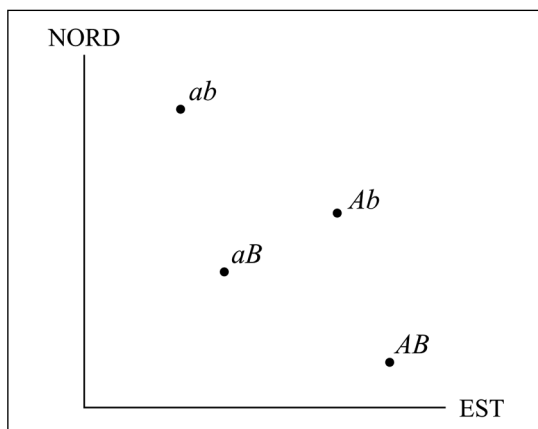


Figura 5.4.
Est promette a Nord di rispondere con b ad A

Esito 1. La promessa ha successo. Ciò significa che, facendo affidamento sulla promessa di Est, Nord attua A . La scelta di Est si restringe allora a due alternative: (i) mantenere la promessa attuando b , così da ottenere il risultato congiuntamente desiderabile Ab ; (ii) violare la promessa attuando B , così da ottenere AB . Si noti che, dal punto di vista di Est, il risultato Ab in cui mantiene la promessa è nettamente peggiore del risultato AB in cui la viola.

Esito 2. La promessa non ha successo. Non fidandosi di Est, Nord sceglie a . A questo punto Est attua la propria strategia ottimale B , e il gioco si conclude con il risultato congiuntamente indesiderabile aB .

L'analisi della struttura formale della promessa mette in luce problemi del tutto analoghi a quelli che si presentano in relazione alle minacce. Si ricorderà che, quando una minaccia non riesce, colui che l'ha effettuata non ha alcun interesse a punire il trasgressore. Analogamente, dalla descrizione delle alternative (i) e (ii) di Est nell'Esito 1, emerge che, quando una promessa riesce, il suo autore non ha alcun interesse a premiare il concorrente che si è fidato di lui. Ne segue che, se i giocatori fossero sempre in grado di scegliere liberamente, nessuna promessa verrebbe mai mantenuta e, quindi, nessuna promessa potrebbe avere successo. Ciò significa che – proprio come accade nel caso della minaccia – il successo della promessa dipende dalla capacità del suo autore di sbarazzarsi della propria libertà di scelta e, più precisamente, dalla sua capacità di vincolarsi visibilmente al mantenimento della promessa.

Ci resta da illustrare la struttura formale di uno scambio di promesse. A questo scopo, possiamo considerare il gioco rappresentato nella matrice della figura 5.5, che costituisce un esempio del dilemma del prigioniero:

		COLONNA	
		C1	C2
RIGA	R1	5	6
	R2	-4	-3

Figura 5.5.
Riga e Colonna
promettono di scegliere,
rispettivamente, R1 e C1

la soluzione del gioco è data dal risultato congiuntamente indesiderabile (R2,C1) che assegna a entrambi un payoff -3 (cfr. figura 3.11). Tuttavia, se si consente la comunicazione tra i giocatori, questi possono migliorare le loro prospettive, effettuando un appropriato scambio di promesse. Infatti, possono impegnarsi a scegliere le loro strategie non-dominanti R1 e C1: se entrambi manterranno la promessa, otterranno il risultato congiuntamente desiderabile (R1,C1) che assegna a entrambi un payoff 5.

Altri tipi di mosse strategiche

Avvertimenti, minacce e promesse non esauriscono il ricco arsenale delle strategie dissuasive. Tali strategie includono, tra l'altro, svariate combinazioni di minacce e promesse. Una strategia di questo tipo può venire illustrata con riferimento al gioco della figura 5.6, dove si suppone che Nord debba muovere per primo (cfr. Schelling 1960, p. 58, Fig. 4).

Poiché Nord sceglierà in ogni caso la sua strategia dominante a , il payoff di Est verrà massimizzato da b : la soluzione del gioco è quindi ab . Il grafico mostra, però, che ab – pur essendo preferibile a ogni altro risultato dal punto di vista di Nord – non risulta molto attraente per Est. In assenza di comunicazione, Est deve rassegnarsi ad ab , ma le sue prospettive cambiano se la comunicazione viene consentita. In tal caso, infatti, Est può migliorare la propria sorte a danno di Nord, applicando una strategia dissuasiva che combina una promessa e una minaccia. Più precisamente, Est minaccerà Nord di attuare B in risposta alla sua eventuale scelta di a , così da infliggergli il risultato congiuntamente indesiderabile aB ; nel contempo, gli prometterà di

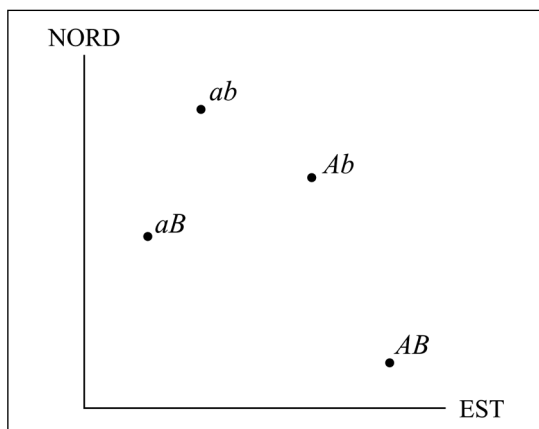


Figura 5.6.
Est minaccia Nord di rispondere con B ad a e gli promette di rispondere con b ad A

attuare b in risposta alla sua eventuale scelta di A , consentendogli così di evitare il pessimo risultato AB e di ottenere, invece, il risultato, congiuntamente desiderabile, Ab . Si noti che la sola minaccia non sarebbe sufficiente a distogliere Nord da a , dato che la scelta di A potrebbe condurre Nord al pessimo risultato AB , nel caso in cui Est non prometta di astenersi da B .

Questo esempio mostra come, in generale, la comunicazione fra giocatori combini minacce, promesse e altri tipi di strategie dissuasive. A questo proposito, vale la pena notare che abbiamo finora discusso la struttura formale di un certo numero di strategie dissuasive con riferimento a giochi in cui ciascun giocatore deve scegliere tra due sole alternative. Nei giochi in cui dispongono di più di due alternative, i giocatori possono escogitare una varietà ancora maggiore di strategie dissuasive, costituite da diversi *schemi di risposta* alle possibili mosse della controparte.

Supergiochi

Poiché un conflitto è un gioco in cui, oltre all'insieme delle mosse di base, si possono effettuare anche determinate mosse strategiche, o strategie dissuasive, è naturale ipotizzare che esso possa venire rappresentato come un particolare *supergioco*, ottenuto espandendo la matrice dal gioco originario, così da includere tra le mosse a disposizione dei giocatori anche un certo numero di strategie dissuasive. La strategia dissuasiva ottimale potrebbe poi venire individuata applicando le usuali regole di scelta della teoria dei giochi.

La possibilità appena descritta è stata presa seriamente in considerazione da Schelling (1960, pp. 175 ss.). La sua analisi parte dal rilievo che, poiché un giocatore applica una strategia dissuasiva quando si vincola visibilmente a una determinata mossa di base, tale applicazione equivale, di fatto, alla visibile e irreversibile riduzione, nella matrice del gioco originario, di alcuni dei payoff del giocatore. Più precisamente, l'applicazione di una mossa strategica è matematicamente equivalente a invocare una penalità nel caso non si mantenga l'impegno, cioè a "sottrarre l'ammontare della pena dai propri payoff in tutte le celle che non corrispondono alla strategia così selezionata" (*ibidem*, pp. 175-176).

Questo punto può venire chiarito con un esempio. Supponiamo che, nel gioco della figura 5.7.a, Riga si impegni ad attuare la mossa R2.

L'impegno di Riga equivale a trasformare R2 nella sua strategia dominante, cioè a decurtare i payoff della riga R1 applicando a ciascuno di questi una penalità pari, per esempio, a -5: in tal modo il gioco della figura 5.7.a viene trasformato in quello della figura 5.7.b Analogamente, impe-

Figura 5.7. Riga trasforma il primo gioco (a) nel secondo (b) impegnandosi su R2 e nel terzo (c) impegnandosi su R1

		COLONNA				COLONNA				COLONNA	
		C1	C2			C1	C2			C1	C2
RIGA	R1	2	5	0		2	5	0		2	5
	R2	0	1	2		-3	-4	2		-5	0
		a		b		c					

gnarsi ad attuare la mossa R1 equivale ad applicare ai payoff nella riga R2 la penalità -5 , cioè a trasformare il gioco della figura 5.7.a in quello della figura 5.7.c.

Possiamo ora vedere come l'interpretazione appena descritta delle mosse strategiche consenta di rappresentare il conflitto come un particolare supergioco. Si consideri il gioco della figura 5.7.a, dove si suppone che Riga debba muovere per secondo. In questo gioco Colonna, muovendo per primo, ha a disposizione due sole strategie: attuare C1 o attuare C2. Muovendo per secondo, Riga può scegliere, invece, fra quattro strategie:

- (i) attuare in ogni caso R1 (strategia C1-R1, C2-R1);
- (ii) attuare in ogni caso R2 (strategia C1-R2, C2-R2);
- (iii) attuare R1 in risposta a C1 e R2 in risposta a C2 (strategia C1-R1, C2-R2);
- (iv) attuare R1 in risposta a C2 e R2 in risposta a C1 (strategia C1-R2, C2-R1).

La soluzione di questo gioco può venire determinata senza troppe difficoltà. Basta, infatti, osservare che, muovendo per secondo, Riga sceglierà la mossa che gli consentirà di massimizzare il proprio payoff; ciò significa che, se Colonna sceglierà C1, allora Riga risponderà con R1, e se Colonna sceglierà C2, risponderà con R2. Nel primo caso Colonna otterrà un payoff pari a 5 e nel secondo un payoff pari 2: di conseguenza Colonna sceglierà C1 e, come si è detto, Riga risponderà con R1. Ciò significa che la soluzione del gioco è (R1,C1). Si noti che tale risultato si trova in cima all'ordine di preferenza di Colonna – al quale assicurerà un payoff pari a 5 – ma non è del tutto soddisfacente per Riga, che dovrà accontentarsi di un payoff di 2.

Supponiamo ora che Riga, e lui soltanto, abbia il potere di applicare determinate mosse strategiche, costituite da impegni non-condizionali sulla sua mossa. È evidente che, nel supergioco che incorpora le mosse strategiche di Riga, entrambi i giocatori avranno a disposizione un maggior numero di strategie rispetto al gioco originario. Infatti, Riga può ora scegliere anche fra tre diverse possibilità di impegno:

- (i) non impegnarsi affatto (scelta 0);
- (ii) impegnarsi su R1 (scelta 1);
- (iii) impegnarsi su R2 (scelta 2).

Di conseguenza, le *strategie estese* a disposizione di Riga nel supergioco saranno le $12 = 3 \times 4$ strategie ottenute combinando le tre scelte 0, 1 e 2 con le quattro strategie (C1-R1,C2-R1), (C1-R2,C2-R2), (C1-R1,C2-R2) e (C1-R2,C2-R1) del gioco originario. Per esempio, Riga può scegliere di non impegnarsi prima di cominciare a giocare, e poi affrontare il gioco sulla base della seconda strategia del gioco originario (strategia 0, C1-R2,C2-R2); oppure può impegnarsi a scegliere R1, e poi adottare la terza strategia del gioco originario (strategia 1, C1-R1,C2-R2); e così via. Anche le strategie estese a disposizione di Colonna saranno più numerose delle strategie di cui disponeva nel gioco originario. Infatti, Colonna disporrà ora di $8 = 2^3$ strategie estese, che vengono ottenute scegliendo con quale delle due mosse di base C1 e C2 rispondere a ciascuna delle tre scelte 0, 1 e 2 di Riga. Ciascuna di queste strategie viene denotata da una tripla di coppie di simboli: per esempio, “0-C1, 1-C2, 2-C1” significa “Scegli C1 se Riga non si impegna, C2 se si impegna su R1, e C1 se si impegna su R2”.

La combinazione delle dodici strategie di Riga con le otto di Colonna darà luogo alla matrice di $96 = 12 \times 8$ celle del supergioco rappresentato nella figura 5.8. Si noti che i payoff del supergioco sono determinati sulla base dei payoff del gioco originario in figura 5.7. In particolare, per ogni cella della matrice, i payoff di Colonna dipenderanno solo dalle mosse di base “incorporate” nelle strategie estese di Riga e Colonna corrispondenti a quella cella: più precisamente i payoff di Colonna saranno identici a quelli che si trovano nella figura 5.7.a, in corrispondenza a tali mosse di base. I payoff di Riga, invece, saranno identici a quelli della corrispondente cella della figura 5.7.a, se non si impegna; a quelli della figura 5.7.c, se si impegna su R1; e a quelli della figura 5.7.b, se si impegna su R2.

Schelling (1960, pp. 154-156) dimostra, sulla base degli usuali principi di scelta della teoria dei giochi, che la soluzione di questo supergioco è costituita da una qualunque delle quattro celle con i payoff in grassetto:

Figura 5.8. Il supergioco determinato dal gioco in figura 5.7

	C.I	C.II	C.III	C.IV	C.V	C.VI	C.VII	C.VIII
	0-C1	0-C1	0-C1	0-C1	0-C2	0-C2	0-C2	0-C2
	1-C1	1-C1	1-C2	1-C2	1-C1	1-C1	1-C2	1-C2
	2-C1	2-C2	2-C1	2-C2	2-C1	2-C2	2-C1	2-C2
R.I: 0, C1-R1, C2-R1	2	5	2	5	1	0	1	0
R.II: 0, C1-R2, C2-R2	0	1	0	1	5	2	5	2
R.III: 0, C1-R1, C2-R2	2	5	2	5	5	2	5	2
R.IV: 0, C1-R2, C2-R1	0	1	0	1	1	0	1	0
R.V: 1, C1-R1, C2-R1	2	5	1	0	2	5	1	0
R.VI: 1, C1-R2, C2-R2	-5	1	0	2	-5	1	0	2
R.VII: 1, C1-R1, C2-R2	2	5	0	2	2	5	0	2
R.VIII: 1, C1-R2, C2-R1	-5	1	1	0	-5	1	1	0
R.IX: 2, C1-R1, C2-R1	-3	0	-3	5	-3	0	-3	5
R.X: 2, C1-R2, C2-R2	0	2	0	1	0	2	0	1
R.XI: 2, C1-R1, C2-R2	-3	5	-3	2	5	2	-3	5
R.XII: 2, C1-R2, C2-R1	0	1	0	0	0	1	0	0

tali celle si trovano sull'intersezione della riga R.X con le colonne C.II, C.IV, C.VI e C.VIII. Si noti che la riga x corrisponde alla strategia (2, C1-R2, C2-R2) di Riga – che consiste nell'impegnarsi su R2, e nell'attuare poi in ogni caso R2 – mentre le colonne C.II, C.IV, C.VI e C.VIII corrispondono alle quattro strategie di Colonna che prevedono la risposta C2 alla scelta iniziale 2 di Riga. Come si può osservare dalla matrice, scegliendo le loro strategie ottimali, Riga e Colonna otterranno, rispettivamente, i payoff 5 e 2. Ciò significa che Riga può sfruttare la possibilità di impegnarsi su R2 per ribaltare l'esito del gioco originario, accrescendo il proprio payoff da 2 a 5, a scapito di Colonna. Il significato intuitivo della soluzione del supergioco individuata da Schelling può venire illustrato con riferimento alle matrici della figura 5.7. Possiamo notare, infatti, che la strategia ottimale (2, C1-R2, C2-R2) di Riga equivale alla scelta –

fra le tre possibilità costituite dai giochi della figura 5.7 – di giocare al gioco (b). In questo gioco, attuando la sua strategia dominante R2, Riga può realizzare un payoff pari a 5; invece, scegliendo di giocare al gioco originario (a) (che equivale alla scelta di non impegnarsi) o al gioco (c) (che equivale alla scelta di impegnarsi su R1) Riga otterrebbe soltanto un payoff pari a 2. In risposta alla scelta di Riga di giocare al gioco (b), Colonna non può fare altro che attuare – in accordo con una qualunque delle strategie C.II, C.IV, C.VI e C.VIII – la mossa di base C2, che gli consente di massimizzare il proprio payoff nel gioco (b).

Oltre a mostrare concretamente la possibilità di rappresentare i conflitti come supergiochi, l'esempio appena descritto fornisce una suggestiva indicazione della complessità di tale compito. Infatti, basta mettere a disposizione del solo giocatore che muove per secondo due semplici mosse strategiche per passare dalla matrice di quattro celle del gioco originario a quella di 96 celle del corrispondente supergioco. Inoltre, occorre osservare che le dimensioni della matrice del supergioco crescono in modo più che esponenziale al crescere sia delle dimensioni della matrice del gioco originario sia del numero di mosse strategiche consentite ai giocatori. Per esempio, partendo da un gioco con una matrice 3×3 – dove Colonna muove per primo e Riga per secondo – e consentendo a entrambi i giocatori solo le mosse strategiche fondamentali, il corrispondente supergioco avrà una sbalorditiva matrice con più di un *googol* – cioè 1 seguito da un centinaio di zero – di colonne (Schelling 1960, p. 183, n. 33).

Pur riconoscendo l'indubbio interesse teorico della dimostrazione che i conflitti possono essere rappresentati come supergiochi, di fronte all'enorme complessità di tali rappresentazioni occorre chiedersi se esse possano essere di aiuto nell'analisi di specifici conflitti, cosa di cui lo stesso Schelling sembra dubitare. Infatti, il problema fondamentale nell'analisi di un conflitto non è quello di individuare la mossa strategica ottimale, nell'ambito dell'insieme, quasi sempre enorme, di mosse strategiche logicamente possibili. Il vero problema è, piuttosto, quello di identificare la classe, di solito molto piccola, delle mosse strategiche praticamente possibili, cioè delle mosse strategiche applicabili nelle condizioni di uno specifico conflitto. Dopo avere risolto questo problema di metodologia della deterrenza, la mossa strategica ottimale potrà venire individuata – senza bisogno di ricorrere ai supergiochi – mediante l'analisi della struttura formale delle poche mosse strategiche effettivamente applicabili.

5.4. La metodologia della deterrenza

La logica della deterrenza permette di individuare le mosse strategiche logicamente possibili in un determinato conflitto. Tuttavia, la possibilità *logica* di applicare una mossa strategica non basta a garantirne la possibilità *pratica*, vale a dire l'effettiva applicabilità. Quest'ultima, infatti, non dipende soltanto dalla struttura formale del conflitto, ma anche dalla sua *struttura materiale*: tale struttura è costituita dalle condizioni fisiche, tecnologiche, culturali, e comunicative entro le quali esso si svolge, cioè da elementi molto diversi da quelli rappresentati nella matrice di gioco.

Se si ammette che l'applicabilità delle mosse strategiche dipende dalla struttura materiale del conflitto, occorre anche ammettere che la metodologia della deterrenza non può comprendere principi universalmente validi, determinabili sulla base di considerazioni puramente a priori. A questo proposito, Schelling (1960, p. 190) osserva che “non possiamo derivare da considerazioni a priori, con strumenti puramente analitici, i principi rilevanti per un *esito positivo* del gioco, i principi *strategici* e le proposizioni di una teoria *normativa*”. Si deve, al contrario, riconoscere che la metodologia della deterrenza necessita di una solida base empirica, cioè di affidabili conoscenze empiriche relative alle svariate sfere di attività – comprese in uno spettro che va dalla storia militare ai conflitti sociali, dalle interazioni della vita quotidiana alle relazioni internazionali – in cui possono sorgere i conflitti. Anche se la fatica di scoprire – o inventare – i più efficaci *metodi di applicazione* delle mosse strategiche graverà sempre sulle spalle di chi prende parte a un conflitto, la metodologia della deterrenza può affrontare il compito, più limitato e realistico, di classificare i metodi di applicazione delle mosse strategiche e analizzarne l'efficacia. Nel seguito di questo paragrafo illustreremo alcuni di questi metodi.

Contratti e punizioni

Gli impegni bilaterali – e, in particolare, lo scambio di promesse – possono venire applicati attraverso il *metodo del contratto esplicito*. Con questo metodo gli impegni vengono trasformati in contratti che prevedono il ricorso a un arbitro in grado di amministrare una punizione a chi viola l'impegno sancito dal contratto. Ciò significa che l'irrevocabilità e l'obbligatorietà degli accordi viene garantita alterando, attraverso un sistema di punizioni, le funzioni dei payoff dei giocatori così da penalizzare i risultati delle mosse del gioco che violano l'impegno. Il ricorso a contratti espliciti garantiti da un arbitro non è infrequente nelle relazioni internazionali: si pensi, per esempio,

alle dispute tra piccoli stati che ricorrono all'arbitrato di potenze maggiori. Tuttavia, è assai dubbio che questo metodo sia applicabile ai conflitti tra potenze atomiche evolute, dato che in questo caso è virtualmente impossibile individuare un arbitro *super partes*, che possa somministrare una punizione senza dover temere le minacce dei contendenti.

In mancanza di un arbitro, due giocatori possono garantire l'obbligatorietà di un impegno bilaterale mediante accorgimenti in grado di far scattare in modo quasi automatico il sistema di punizioni, in risposta a una violazione dell'impegno. Per esempio, due stati che si scambiano la promessa di non dare inizio a uno scontro militare, possono applicare il loro impegno ricorrendo al "sistema degli ostaggi": in tal modo uno scontro militare produrrebbe quasi automaticamente l'uccisione degli ostaggi che, si presume, rappresenta una notevole penalità per entrambi i contendenti. Diversamente da quanto si potrebbe pensare, il sistema degli ostaggi non è un puro ricordo del passato, ma rappresenta, in una sua versione aggiornata, un elemento costitutivo della strategia atomica. Infatti, la bilancia del terrore atomico – fondata sulla quasi completa rinuncia alla costruzione di sistemi antimissile e di altre strutture difensive – equivale a "uno scambio completo di tutti gli ostaggi immaginabili", poiché "evitare la difesa, in effetti, equivale a fare dell'intera popolazione un ostaggio, senza preoccuparsi di metterla fisicamente nella mani dell'altro" (Schelling, 1960, p. 159).

Mettere in gioco la propria reputazione

Mentre il ricorso a contratti espliciti appare particolarmente indicato nello scambio di promesse, minacce e altri impegni di carattere unilaterale richiedono differenti metodi di applicazione. Tali impegni, infatti, non prevedono alcun accordo con la controparte, che anzi potrebbe essere danneggiata dall'eventuale attuazione dell'impegno preso unilateralmente da un giocatore. Anche in questo caso, tuttavia, un giocatore può fare ricorso a metodi di applicazione basati sull'alterazione della propria funzione dei payoff e, più specificamente, sulla penalizzazione dei risultati di tutte le mosse del gioco diverse dal mantenimento dell'impegno.

Per esempio, un giocatore può alterare la sua funzione dei payoff ricorrendo al "metodo della reputazione", che consiste nel mettere in gioco la propria reputazione di uomo, o donna, d'onore, condannandosi a "perdere la faccia" nel caso in cui non si mantenga l'impegno preso. Il metodo della reputazione è afflitto, però, da una notevole limitazione legata al fatto che il danno subito da chi perde la faccia si manifesta solo dopo la fine del gioco:

tale danno, infatti, consiste essenzialmente nella perdita di credibilità del giocatore nell'ambito delle successive interazioni nelle quali sarà coinvolto. Ciò significa che perdere la faccia è una penalità significativa solo per quei giocatori sui quali si proietta l'"ombra del futuro" (Axelrod 1984, p. 18), cioè l'ombra dei futuri svantaggi strategici connessi alla perdita della propria reputazione. Il metodo della reputazione appare, quindi, del tutto inefficace quando viene utilizzato da un agente per il quale il gioco attuale costituisce, per così dire, l'ultimo gioco, cioè da un agente che in futuro non dovrà presumibilmente misurarsi con altri giocatori.

La limitazione appena descritta rappresenta un ostacolo decisivo all'impiego del metodo della reputazione nella strategia della deterrenza atomica, dato che il conflitto atomico ha tutti i caratteri dell'ultimo gioco, quello che può distruggere il giocatore, la controparte, e la possibilità stessa di essere impegnati in futuri conflitti.

Bruciarsi i ponti alle spalle

I contratti espliciti, il sistema degli ostaggi, e la messa in gioco della propria reputazione, sono metodi attraverso i quali le funzioni dei payoff dei giocatori vengono modificate al fine di rendere desiderabile – e quindi credibile – la scelta della mossa sulla quale si sono impegnati. Se non si trova nessun sistema per modificare la funzione dei payoff, si può comunque rendere credibile l'impegno ad attuare una certa mossa ricorrendo ad accorgimenti in grado di distruggere interamente la libertà di attuare una condotta alternativa.

Così, per esempio, l'esercito che si brucia i ponti alle spalle si priva volontariamente e irreversibilmente di ogni via di fuga, rendendo in tal modo assolutamente credibile la propria minaccia di combattere fino all'ultimo uomo. Il "metodo dei ponti bruciati" include una grande varietà di sistemi e tattiche che operano attraverso "qualche volontario ma irreversibile sacrificio della libertà di scelta" (Schelling 1960, p. 22). Il modo migliore per bruciarsi i ponti alle spalle dipenderà, naturalmente, dalla struttura materiale del conflitto. Nel caso, per esempio, di un conflitto tra potenze atomiche, un efficace metodo di applicazione della minaccia di rappresaglia atomica potrebbe consistere nell'impiego di qualche meccanismo di risposta automatica in grado di reagire a un attacco atomico nemico scatenando una rappresaglia massiccia indipendentemente dalla – e anche contro la – volontà di chi ha attivato il meccanismo.

Contromosse

Poiché l'applicabilità delle mosse strategiche dipende dalla struttura materiale del conflitto, un giocatore razionale, in grado di comprendere i vantaggi che la controparte potrebbe trarre dall'applicazione di una determinata mossa strategica, dovrebbe considerare attentamente la possibilità di inibire l'applicazione di quella mossa attraverso opportune contromosse, volte a modificare la struttura materiale del conflitto.

Poiché l'applicazione di una mossa strategica consiste nel vincolarsi visibilmente a una certa linea d'azione, una contromossa può basarsi su tattiche in grado di danneggiare il sistema di comunicazione attraverso il quale l'avversario cerca di rendere visibile il suo vincolo. A questo proposito, occorre ricordare che ogni forma di comunicazione presuppone alcuni elementi di base, come l'identificazione del destinatario della comunicazione, e il ricorso a qualche canale di comunicazione attraverso il quale fare giungere il messaggio al destinatario. Ciò implica che si può danneggiare o distruggere la comunicazione rendendo impraticabili almeno uno dei suoi elementi di base: per esempio, ci si può sottrarre all'identificazione attraverso opportune forme di mimetizzazione, oppure si possono distruggere fisicamente tutti i possibili canali di comunicazione che la controparte potrebbe utilizzare per comunicarci i suoi impegni strategici.²¹ Si noti che le contromosse del tipo appena descritto sono di natura ben diversa dalle mosse strategiche: una contromossa, infatti, *non* è un tentativo di contrastare la mossa strategica della controparte con un'altra mossa strategica, bensì una manovra volta a *inibire* l'applicazione della mossa strategica della controparte.

Il quadro della razionalità strategica fin qui tratteggiato indica che gli agenti coinvolti in un conflitto si trovano di fronte ad almeno quattro diversi problemi di scelta: (1) Qual è la mossa di base ottimale del gioco? (2) Qual è la mossa strategica ottimale tra quelle applicabili? (3) Qual è il migliore metodo di applicazione per tale mossa? (4) Quali sono le contromosse ottimali per anticipare e inibire le mosse strategiche della controparte? Mentre la teoria dei giochi affronta solo il primo quesito, la teoria del conflitto si propone di affrontare sistematicamente anche gli altri tre.

21 Su questo punto si veda Schelling (1960, pp. 170 ss.). Per quanto riguarda la mimetizzazione, si può osservare che questa tattica viene spesso adottata dai gruppi terroristici che, proprio per questo, sembrano difficilmente esposti a minacce: si veda il profetico saggio di Schelling "Thinking about Nuclear Terrorism", pubblicato nel 1982, e ristampato in Schelling (1984, cap. 14).

5.5. Paradossi della deterrenza

La storia militare offre numerosi elementi a conferma dell'affermazione che (Luttwak, 1987, p. 23)

*l'intero regno della strategia è pervaso da una logica paradossale tutta sua, in contrasto con la logica lineare ordinaria, in base alla quale viviamo in tutte le altre sfere della nostra esistenza.*²²

Qualunque cosa si pensi circa il carattere paradossale dell'“intero regno della strategia”, è chiaro che almeno la strategia della deterrenza esibisce aspetti altamente paradossali, che un'adeguata teoria del conflitto non può trascurare. In questo paragrafo illustreremo brevemente alcuni di questi paradossi.

Il paradosso del vantaggio strategico

Quando pensiamo all'esito che potrà avere un conflitto, tendiamo ad attribuire un vantaggio strategico al giocatore “più forte”, dove la forza sembra dipendere da elementi quali la libertà di manovra di un giocatore o la presenza, nella sua funzione dei payoff, di molti valori elevati. I conflitti che si svolgono nel mondo reale ci mostrano, però, che talvolta la debolezza può trasformarsi in forza e che il giocatore “più debole” può trovarsi in una condizione di vantaggio strategico. Parlando di paradosso del vantaggio strategico ci riferiamo, appunto, alla possibilità che si verifichino casi di questo genere. Per esempio, il metodo dei ponti bruciati, illustrato nel paragrafo 5.4, mostra come un apparente indebolimento di un giocatore – costituito dalla severa autolimitazione della sua libertà di manovra – possa rafforzarne la posizione strategica. Vedremo ora come la posizione strategica di un giocatore possa venire rafforzata anche dalla diminuzione di alcuni dei suoi payoff.

Un caso di questo genere può venire illustrato con riferimento al gioco della figura 5.2, considerato nel paragrafo 5.3 (cfr. Schelling 1960, p. 57). In questo gioco la minaccia di Est restringe l'ambito di scelta di Nord alle opzioni *ab* e *AB*: preferendo la seconda opzione, Nord dovrebbe rinunciare

22 Schelling (1960, p. 21) osserva che “i risultati raggiunti da un'analisi teorica del comportamento strategico sono spesso qualcosa di paradossale, in quanto spesso contraddicono il senso comune o le regole vigenti”. L'insistenza sul carattere paradossale dei principi strategici è un motivo ricorrente negli studi sull'argomento: si vedano, per esempio, Aron (1963, cap. VI, “Logica e paradossi della teoria strategica”) e Luttwak (1987, cap. I, “L'uso cosciente del paradosso in guerra”).

alla sua strategia ottimale a in favore di A . Poiché la riuscita della minaccia dipende dal fatto che Nord preferisce AB ad ab , un sistema efficace per bloccare la minaccia è quello di diminuire il payoff di Nord in AB , “trascinando” verticalmente AB sotto il livello di ab . Si vede bene che questo indebolimento di Nord è solo apparente e che, in realtà, la posizione strategica di Nord ne viene rafforzata, dato che la diminuzione del payoff di Nord in AB rende logicamente impossibile la minaccia di Est. Al riparo da questa minaccia, Nord può quindi scegliere a , così da ottenere il risultato aB , che gli garantirà un payoff superiore anche a quello inizialmente associato ad AB . Nel seguito, illustriamo brevemente altre due forme nelle quali può presentarsi il paradosso del vantaggio strategico, vale a dire l’uso dell’irrazionalità per scopi razionali e l’applicazione di minacce che lasciano qualcosa al caso.

L’uso dell’irrazionalità per scopi razionali

Il possesso di un alto grado di razionalità rappresenta, generalmente, un elemento di forza per la conduzione dei conflitti. Talvolta, però, la razionalità può trasformarsi in uno svantaggio, mentre l’irrazionalità può venire usata per scopi razionali, cioè per rafforzare la posizione strategica di un giocatore (cfr. Schelling 1960, p. 21, e Parfit 1984, pp. 17-19). La storia della guerra fredda offre svariati esempi di uso strategico dell’irrazionalità (Beaufre, 1963, p. 60):

Se si ha a che fare con un pazzo, non bisogna irritarlo! La fermezza di Dulles, le collere e le scarpe di Krusciov, l’ostinazione fredda di De Gaulle corrispondono a questo gioco psicologico, la cui influenza può superare tutti i calcoli ricavati da fattori materiali. In realtà, l’elemento decisivo riposa sulla volontà di scatenare il cataclisma. Far credere che si ha questa volontà è più importante di tutto il resto. Naturalmente ciascuno bara; ma fino a che punto?

La tattica di simulare follia, stupidità e altre forme di irrazionalità – in aggiunta a quella di cui si è naturalmente dotati – è vecchia quanto il mondo. Anche se offrire una giustificazione teorica dell’efficacia di questa tattica è tutt’altro che semplice, ci sembra che l’uso strategico dell’irrazionalità potrebbe fondarsi sul seguente meccanismo.

Supponiamo di interagire con un avversario irrazionale, o comunque molto lontano dal modello ideale della teoria classica dei giochi, cioè di un agente che opera le sue scelte a mente fredda e con lucida intelligenza. Nell’interazione con questo avversario, la nostra “mossa migliore” non coinciderà necessariamente con la strategia ottimale, poiché quest’ultima viene definita in base all’assunzione che la controparte sia un agente ra-

zionale, capace di riconoscere e attuare la *sua* strategia ottimale. Nel caso, invece, di un avversario dotato di razionalità imperfetta, occorre presumere che, molto probabilmente, egli attuerà qualche mossa “sub-ottimale”. Se conoscessimo bene il genere e il grado della sua irrazionalità, potremmo prevedere con una certa sicurezza la sua mossa e scegliere la nostra mossa sulla base di tale previsione. In realtà, non possiamo mai essere completamente certi dell’irrazionalità del nostro avversario e dobbiamo, quindi, accontentarci di scegliere la nostra mossa in base a previsioni altamente incerte circa le sue scelte. È proprio l’impossibilità di sapere se la controparte è davvero irrazionale ad aprire la strada all’uso strategico dell’irrazionalità. Il nostro avversario, infatti, potrebbe simulare stupidità, irascibilità o follia, allo scopo di indurci ad attuare una strategia sub-ottimale, scelta in base alla nostra erronea previsione che egli attuerà una determinata mossa sub-ottimale. A quel punto, il nostro avversario potrà approfittare del nostro errore, attuando la sua strategia ottimale, oppure una strategia ancora più astuta, basata sulla sua corretta previsione della nostra mossa sub-ottimale.

Se si accetta questa spiegazione intuitiva dell’uso strategico dell’irrazionalità, occorre chiedersi se, e come, essa possa venire incorporata nella teoria del conflitto. Ci sembra che, a tale scopo, occorra estendere l’apparato concettuale della teoria, per lasciare spazio anche alle interazioni in cui non si può applicare l’assunzione di razionalità, secondo la quale tutti i giocatori sono individui razionali, e a conoscenza della razionalità dei loro avversari. Più specificamente, si dovrebbero formulare criteri di decisione applicabili alle interazioni in cui uno o entrambi i giocatori possono prendere seriamente in considerazione sia la possibilità che l’avversario sia irrazionale sia quella che simuli di esserlo. Lo sviluppo, nel corso degli ultimi decenni, delle teorie cognitive dei giochi, basate sul rifiuto del postulato della perfetta razionalità dei giocatori, promette di gettare nuova luce su criteri di questo tipo, e di offrire nuovi strumenti concettuali all’analisi del conflitto.

Minacce che lasciano qualcosa al caso

Abbiamo finora considerato solo mosse strategiche di tipo “deterministico”. In particolare, il carattere deterministico di una minaccia può venire enfatizzato esprimendola in questo modo: “Se farai x , allora *certamente* risponderò con y ”. La forza di una minaccia deterministica sembra attenuarsi, o venir meno, quando la minaccia viene “indebolita” e sostituita da minacce probabilistiche come: “Se farai x , allora probabilmente risponderò con y ”, oppure “Se farai x , allora vi è una probabilità p che io risponda con y ”. Il paradosso delle minacce probabilistiche consiste nel fatto che, in

certi casi, esse possono funzionare egregiamente bene e, talvolta, persino meglio delle corrispondenti minacce deterministiche.

A prima vista l'idea stessa di minaccia probabilistica appare problematica. Minacciare qualcuno significa, infatti, vincolarsi visibilmente all'attuazione di una particolare mossa, così da convincere la controparte di avere perso la libertà di non attuarla. Nel caso di minacce probabilistiche, tuttavia, non è immediatamente chiaro in cosa possa consistere tale vincolo. Si potrebbe anzi pensare che la stessa forma probabilistica della minaccia impedisca di vincolarsi in modo visibile: se annuncio alla controparte che, in risposta a y , *potrei* rispondere con x , quest'ultima potrà credere che ho conservato la mia libertà di scelta e che, di conseguenza, non attuerò la minaccia. I dubbi circa l'applicabilità di una minaccia probabilistica possono venire fugati riflettendo sulla peculiare natura della sua probabilità di attuazione: tale probabilità, infatti, *non* è un effetto della libertà di chi minaccia, bensì di qualche elemento casuale sottratto, almeno parzialmente, alla sua libera scelta. La minaccia probabilistica deve quindi venire intesa – per usare l'efficace espressione di Schelling (1960, p. 188) – come una “minaccia che lascia qualcosa al caso”.

In linea di principio, l'elemento casuale utilizzato per vincolarsi all'attuazione di minacce probabilistiche potrebbe essere qualche processo randomizzato – come il lancio di un dado – il cui esito determina, mediante opportuni automatismi, l'attuazione della minaccia. Di solito, però, tale elemento consiste nella creazione di situazioni di tensione, confusione e rischio, in cui la razionalità della scelta viene appannata dalla fretta, dal nervosismo e dalle difficoltà di comunicazione, aprendo così la via a decisioni che “lasciano qualcosa al caso”. Le minacce probabilistiche sembrano svolgere un ruolo molto importante nelle relazioni internazionali e, in particolare, nella strategia della deterrenza militare. Si pensi, per esempio, alla minaccia probabilistica di scatenare “involontariamente” una guerra di grandi proporzioni, e persino una guerra atomica: tale minaccia può venire applicata, per esempio, dando inizio a una guerra limitata, utilizzata come un generatore di rischi. In modo simile, se una guerra limitata è già in corso, la minaccia probabilistica di scatenare un conflitto di vaste proporzioni può venire applicata mettendo in opera un comportamento rischioso, temerario e avventuristico nella condotta delle operazioni.

La macchina dell'apocalisse

Il metodo dei ponti bruciati – che consiste nello sbarazzarsi della propria libertà di scelta – potrebbe venire applicato affidando a qualche automati-

smo la nostra risposta a determinate mosse dell'avversario. Immaginiamo, per esempio, che la risposta a un attacco atomico venga affidata alla macchina dell'apocalisse, costruita dai sovietici nel film di Stanley Kubrick, *Dr. Strangelove, or How I Learned to Stop Worrying and Love the Bomb* (1963). In risposta a un attacco, la macchina scatena automaticamente un'esplosione così potente da distruggere la vita su tutto il pianeta. Inoltre, la macchina scatena l'apocalisse anche nel caso in cui si tenti in qualunque modo di manometterla. Naturalmente tale caratteristica serve a proteggere la macchina dai possibili ripensamenti dei suoi costruttori che, in caso di attacco, avrebbero tutto l'interesse a disattivarla per non subirne gli effetti totalmente catastrofici: è proprio l'impossibilità di disattivare la macchina a garantirne il potere deterrente.

Discutendo alcuni problemi connessi all'uso della macchina dell'apocalisse, Skyrms (1996, pp. 22-25 e 38-42) pone questo interrogativo: se, dopo aver subito un attacco, ci si accorgesse che la macchina dell'apocalisse si è rotta – e si è quindi recuperata la propria libertà di scelta – si dovrebbe ugualmente scatenare l'apocalisse, oppure sarebbe meglio astenersene? È evidente che, se la macchina dell'apocalisse si rompe, chi ha subito l'attacco si trova in una situazione identica a quella del presidente degli Stati Uniti nell'esperimento mentale di Kahn riportato all'inizio di questo capitolo. In quel caso il presidente, informato di un attacco atomico su New York, deve decidere se attuare, oppure no, la propria minaccia di rispondere con una rappresaglia massiccia a qualunque attacco. Poiché attuare questa minaccia significa applicare la dottrina strategica della mutua distruzione assicurata – diventata famosa con il suggestivo acronimo MAD (*Mutual Assured Destruction*) – l'interrogativo di Skyrms equivale alla domanda se sia razionale, oppure no, applicare MAD, nel caso in cui si fosse chiamati a fare questa scelta.

Skyrms (1996, p. 24) sostiene che MAD è una strategia irrazionale, che non dovrebbe venire in alcun caso attuata. Questa risposta è in pieno accordo con la morale popolare secondo la quale “una strategia che include una minaccia che non è nell'interesse dell'agente attuare se fosse chiamato a farlo, e avesse l'opzione di non attuarla, è una strategia difettosa” (*ivi*). A sua volta, tale morale sembra avere un solido fondamento teorico, costituito dal seguente fondamentale principio di razionalità pratica, che va talvolta sotto il nome di “principio di razionalità modulare” (*ibidem*, p. 24):²³

23 Un piano credibile per future contingenze esibisce *razionalità modulare* nel senso che esso è costituito da *moduli* per ognuno dei quali viene specificata una scelta razionale. In conflitto con il principio di razionalità modulare, alcuni autori difendono la concezione – che Skyrms (1996, p. 38) denomina ironicamente “teologia dell'impegno” – secondo la quale occorrerebbe attenersi a massime del tipo

In un piano credibile per future contingenze, in una situazione in cui un agente si trova di fronte a una sequenza di scelte, il suo piano dovrebbe specificare una scelta *razionale* in ciascun punto della sequenza, definita in relazione alla sua situazione in quel punto di scelta.

Il principio di razionalità modulare implica che non si dovrebbe adottare MAD come un “piano per future contingenze” e che, comunque, non si dovrebbero attuare le minacce che ne stanno alla base. In effetti, l’idea che MAD non rappresenti una plausibile dottrina strategica si è ampiamente diffusa nel pensiero strategico occidentale già a partire dall’inizio degli anni Sessanta, aprendo così la via a un vasto dibattito strategico orientato alla ricerca di più efficaci schemi di risposta a eventuali attacchi atomici. Si pensi, per esempio, alla dottrina dell’*escalation*, che si occupa della conduzione di conflitti atomici a intensità variabile, tra due contendenti che, in ogni momento, devono decidere se varcare, oppure no, una determinata soglia di intensità (cfr. Kahn 1963 e Freedman 1986).

First strike: l’autoavverarsi di profezie catastrofiche

In alcuni conflitti, che potremmo chiamare *giochi del terrore*, le mosse a disposizione dei giocatori includono la possibilità di attuare un attacco a sorpresa, cioè un “primo colpo” in grado di eliminare l’avversario o, almeno, di annullarne la capacità di combattimento. In relazione alla convenienza e alla probabilità dell’effettivo uso di un primo colpo, i giochi del terrore possono disporsi in un continuum, che ha per estremi il Gioco del Proiettile e quello della Freccia. Nel Gioco del Proiettile, i giocatori sono pistoleri con un mira perfetta, muniti di pistole precise e potenti: chi spara per primo ucciderà all’istante l’avversario. Nel Gioco della Freccia, i giocatori sono arcieri muniti di frecce avvelenate: chi scocca la prima freccia ucciderà certamente l’avversario, ma il veleno manifesterà i suoi effetti abbastanza lentamente da consentirgli di eliminare l’aggressore.

Pur essendo caratterizzati dal perfetto equilibrio della “bilancia del terrore” – cioè da una perfetta parità di forze tra i giocatori – i due giochi differiscono grandemente per quanto riguarda la stabilità dell’equilibrio. Nel Gioco della Freccia la bilancia è perfettamente stabile: poiché il primo colpo non può distruggere la capacità reattiva dell’avversario, nessun giocatore avrà interesse ad attuare una scelta che condurrebbe all’eliminazione di entrambi. Nel Gioco del Proiettile, al contrario, la bilancia del terrore

“Mantieni in ogni caso i tuoi impegni”. Tra i difensori di questa concezione un posto di rilievo spetta ai Gauthier (1984) e McClellan (1990).

è completamente instabile: poiché entrambi i giocatori hanno a disposizione un primo colpo immediatamente letale, ciascuno di loro avrà un forte incentivo a colpire per primo, allo scopo di fugare la possibilità, sia pure remota, di essere eliminato dall'eventuale primo colpo del rivale.

È interessante notare che, nel Gioco del Proiettile, l'incentivo ad eliminare l'avversario sussiste anche quando nessun giocatore riporta un vantaggio diretto da tale eliminazione, ed entrambi sono a conoscenza di questa circostanza. In tal caso, nessuno dei due desidererebbe colpire per primo, se fosse certo che neppure l'avversario lo farà. Supponendo, però, che entrambi i giocatori abbiano un piccolissimo timore che l'altro colpisca per primo, ci si può chiedere se, e quanto, questo timore renda probabile che uno dei due decida di anticipare con il proprio primo colpo la temuta possibilità di un primo colpo altrui. Sulla base di una penetrante analisi condotta con l'aiuto di un sofisticato modello matematico Schelling (1960, cap. 9) suggerisce una risposta piuttosto sorprendente a questo interrogativo. Anche un piccolissimo timore iniziale che il nostro l'avversario colpisca per primo è destinato ad accrescersi progressivamente – attraverso un particolare “effetto moltiplicatore” – così da indurci ad anticipare il suo primo colpo con il nostro. Ciò significa che le profezie catastrofiche dei partecipanti al Gioco del Proiettile sono destinate ad autoavverarsi.

L'effetto moltiplicatore sopra menzionato può venire informalmente descritto come segue. Anche se il mio timore che l'avversario colpisca spontaneamente per primo è piccolissimo, occorre aggiungervi un secondo timore: quello che lui colpisca per primo a causa del suo timore iniziale che io colpisca spontaneamente per primo. Naturalmente anche il suo timore iniziale viene accresciuto da un secondo timore: quello che io colpisca per primo per timore che lui colpisca spontaneamente per primo. A questo punto occorrerà aggiungere al timore totale di ciascuno di noi anche un terzo timore: quello che l'avversario colpisca per primo sulla base del suo timore totale, costituito dalla somma del suo timore iniziale e del suo secondo timore. Come si può immaginare, questo processo di amplificazione, generato dalla riflessione sul timore altrui, è destinato a proseguire con l'aggiunta di un terzo e un quarto timore, e così via, all'infinito. Anche se i timori che via via si aggiungono al timore iniziale sono sempre più piccoli, il timore totale di entrambi diventerà ben presto abbastanza grande da rendere praticamente certo che uno dei due anticipi con il suo primo colpo il temuto primo colpo altrui.

Schelling riesce a mostrare in modo piuttosto convincente che, nel Gioco del Proiettile, la sfiducia reciproca – cioè il timore del timore altrui – conduce quasi inevitabilmente a un esito catastrofico, a dispetto del fat-

to che i contendenti hanno tutto l'interesse e la volontà di evitarlo. Se si accetta l'analisi di Schelling, è difficile sottrarsi alla conclusione che solo una bilancia del terrore vicina alla condizione di perfetta stabilità – tipica del Gioco della Freccia – può dare sufficienti garanzie di non destabilizzarsi nel corso del tempo. È del tutto chiaro che i partecipanti al Gioco del Proiettile non hanno alcun motivo di compiacersi della loro possibilità di far partire un primo colpo immediatamente letale, e sarebbero ben lieti di trasformare i loro proiettili in frecce avvelenate. Più in generale, è evidente che i partecipanti a un gioco del terrore hanno tutto l'interesse a modificare il gioco così da stabilizzare la bilancia del terrore. Che i giocatori siano in grado, oppure no, di attuare questo compito, dipenderà, ovviamente, dalle specifiche caratteristiche del gioco. Si pensi, per esempio, ad alcuni accorgimenti che potrebbero venire utilizzati per stabilizzare il Gioco del Proiettile: si potrebbe diminuire la precisione o la potenza della pistola, così da rendere improbabile che il proiettile colpisca l'avversario, o che lo uccida all'istante; oppure si potrebbero utilizzare protezioni antiproiettile per diminuire la letalità del primo colpo.

Impegnandosi nel tentativo di modificare un gioco del terrore, i giocatori esercitano una raffinata forma di razionalità strategica. Infatti, essi si impegnano in una sorta di “metagiuoco” le cui mosse – che potremmo chiamare “metamosse” – consistono nell'accordarsi per modificare, a vantaggio di entrambi, la natura del gioco originario. Ci sembra, dunque, del tutto plausibile che un'adeguata analisi della razionalità strategica debba considerare non solo le mosse di base, le mosse strategiche e le contromosse dei partecipanti a un conflitto, ma anche le loro possibili metamosse. I giochi atomici tra potenze rivali sono esempi paradigmatici di giochi del terrore. In questi giochi il primo colpo – il cosiddetto *first strike* – consiste nel lancio di un attacco atomico volto alla completa distruzione delle forze atomiche dell'avversario, così da lasciarlo alla mercé dell'aggressore.

Per stabilire in quale punto del continuum compreso tra il Gioco del Proiettile e quello della Freccia si trovi uno specifico gioco atomico, occorre un esame molto approfondito delle tecnologie offensive e difensive dei contendenti. Se entrambi hanno a disposizione un *first strike* estremamente preciso e potente, il gioco risulta molto simile al Gioco del Proiettile e – a dispetto delle buone intenzioni di entrambi – la catastrofe finisce con il diventare virtualmente inevitabile. Se, invece, una cospicua parte delle forze atomiche dei contendenti è invulnerabile al *first strike* avversario, il gioco si avvicina a quello delle Freccia, e la buona volontà dei contendenti potrà bastare a evitare la catastrofe. Tra gli anni Sessanta e gli anni Ottanta del secolo scorso, gli strateghi nucleari hanno condotto un'enorme quantità di

ricerche per individuare le metamosse – cioè gli “schemi di disarmo” o, per meglio dire, di controllo degli armamenti – che USA e URSS avrebbero potuto attuare per stabilizzare la “delicata bilancia del terrore”.²⁴ Ci preme rilevare che la letteratura sull’argomento non riguarda solo gli strateghi e gli storici, ma può offrire numerosi spunti di riflessione anche agli studiosi interessati alla teoria del conflitto e, più in generale, all’analisi filosofica della razionalità pratica.

24 L’ampia diffusione dell’espressione “delicata bilancia del terrore” si deve a un influente articolo di Wohlstetter (1959). Per una dettagliata analisi dei problemi relativi all’individuazione dei più appropriati schemi di disarmo si veda la parte conclusiva del volume di Schelling (1960, parte quarta, in particolare il cap. 9). Si vedano anche Schelling (1966) e Schelling e Halperin (1961).

6. NATURA, GENESI ED EVOLUZIONE DELLE NORME SOCIALI

Il comportamento di un individuo all'interno di un gruppo o dell'intera società è solitamente governato da diverse regole di condotta, le cosiddette *norme sociali*. Esse includono sia gli usi e i costumi di una società, sia le sue convenzioni, le sue leggi e le sue istituzioni fondamentali. In effetti, le norme sociali sono talmente pervasive e onnipresenti che è difficile addirittura pensare a una qualsiasi forma di interazione sociale che non sia governata, almeno in parte, da una di tali regole (cfr. Bicchieri e Muldoon 2011; Young 2008). Per questo motivo, l'analisi della natura, dell'origine e dell'evoluzione delle norme sociali – cioè di come tali regole fanno la loro comparsa in un gruppo di individui, si sviluppano al suo interno o si diffondono verso l'esterno, e infine, eventualmente, spariscono – è un problema tradizionale delle scienze sociali. Antropologi, sociologi, economisti e teorici del diritto e della politica hanno a lungo studiato i diversi aspetti delle norme e il loro funzionamento all'interno di diversi gruppi sociali. Già Hume (1739) aveva analizzato in profondità il ruolo che le norme giocano nel fare emergere e prendere forma a diversi tipi di ordinamento sociale.

Più recentemente, la teoria dei giochi, nelle sue diverse varianti, ha offerto agli studiosi un potente strumento d'analisi per affrontare il problema delle norme sociali in un quadro concettuale rigoroso e unitario, che promette di diventare il “paradigma” condiviso da tutti gli scienziati sociali (Gintis 2009).¹ In questo capitolo, presentiamo gli elementi essenziali dell'approccio giochistico alle norme sociali (paragrafo 6.1), mostrando come esso sia stato applicato all'analisi delle origini biologiche ed evoluti-

1 Per alcuni dei principali contributi all'approccio giochistico all'analisi delle norme sociali, si vedano Lewis (1969), Ullman-Margalit (1977), Sugden (1986), Bicchieri (1993, 2006), Binmore (1994, 1998, 2005), Vanderschraaf (1999), Alexander (2000, 2007) e Gintis (2007, 2009). Per una rassegna, si vedano Young (2008) e Bicchieri e Muldoon (2011) e, per un confronto critico fra le diverse definizioni giochistiche di norma sociale, Paternotte e Grose (2012). Si veda anche Guala (2012) per una discussione dell'adeguatezza empirica della concezione lewisiana delle norme.

ve delle norme (paragrafo 6.2) e della loro evoluzione nelle società animali, umane e artificiali (paragrafi 6.3 e 6.4).

6.1 *L'approccio giochistico alle norme sociali*

Parlando genericamente di “norma” sociale, ci si riferisce a concetti e fenomeni spesso molto diversi fra loro (cfr. Paternotte e Grose 2012). Per esempio, si dice che è “buona norma” salutare incontrando qualcuno, tenere aperta la porta per farlo passare, non mangiare rumorosamente, vestirsi in modo adeguato alla circostanza; in generale, mantenere comportamenti educati e civili rispetto all’ambiente sociale in cui ci si trova ad agire. In casi come questi, spesso si parla non tanto di norme, quanto di *convenzioni* sociali, o di “usi e costumi” di una società. In questo senso, *norme* vere e proprie sarebbero invece, per esempio, guidare sul lato giusto della strada, non superare il limite di velocità, non rubare, non uccidere; tanto che esistono apposite “normative” che stabiliscono esplicitamente il comportamento corretto da tenere in questi casi. D’altra parte, spesso gli studiosi usano il termine “norma” non solo per riferirsi alle leggi, civili o morali, di una società, ma anche per indicare alcune *istituzioni* fondamentali, essenziali al suo funzionamento, che regolano il comportamento reciproco degli individui: il diritto, la moneta, la lingua, e così via.

Cos’è una norma sociale

In questo libro, in mancanza di una terminologia stabilita, useremo il termine “norma” in senso molto esteso, per riferirci sia alle convenzioni, sia ai precetti morali, sia alle leggi (positive e consuetudinarie), sia infine alle istituzioni sociali (cfr. Young 2008). Ciò non toglie che esistano notevoli differenze fra questi concetti. In particolare, alcuni studiosi preferiscono distinguere nettamente fra norme e convenzioni, per esempio sulla base del fatto che le prime, ma non le seconde, prevedono esplicite *sanzioni* in caso di violazione (Sugden 1986; Elster 1989; Bicchieri 2006). Se vengo sorpreso a guidare troppo velocemente o a rubare, rischio una multa o anche la galera; mentre non corro simili rischi se mastico a bocca aperta o mi presento a un matrimonio in calzoncini e infradito. D’altra parte, questa distinzione non è così netta e si può sostenere che, in realtà, infrangere una norma sociale comporta sempre una sanzione più o meno esplicita, formalizzata o gravosa. Per esempio, se “salto la coda” all’ufficio postale, è probabile che le persone in fila esprimano la loro indignazione con mormorii e occhiate ed è anche

possibile che qualcuno mi riprenda apertamente e a voce alta, svergognandomi davanti a tutti. Se continuo a comportarmi maleducatamente a tavola o in diverse occasioni pubbliche, è probabile che la gente smetta di invitarmi a cena o ad altri incontri simili; e così via. In altre parole, esistono infiniti gradi o livelli di gravità della punizione in cui un individuo rischia di incorrere per aver violato una norma. A un estremo della scala ci sono reazioni informali e spontanee come la disapprovazione o l'ostracismo, all'altro estremo l'esplicita sanzione legale di un reato. Inoltre, il livello di gravità può cambiare nel tempo: per esempio, in Italia fumare al ristorante è stato per decenni un comportamento del tutto comune, poi è diventato "socialmente indesiderabile" o maleducato e, infine, è stato trasformato in un'infrazione esplicitamente sanzionabile con una multa.

Un'altra distinzione rilevante fra diversi tipi di norma, che si sovrappone in parte alla precedente e complica il tentativo di classificarli, è basata sulla maggiore o minore facilità di *applicazione* di una norma. Alcune norme sono "auto-applicative" (*self enforcing*) nel senso che a un individuo basta sapere che sono rispettate all'interno di un certo gruppo per adeguarsi spontaneamente. Per esempio, guidare sul lato destro della strada (in Italia) o stringere con la destra la mano di chi si saluta sono casi di norme auto-applicative. Altre norme richiedono invece una sanzione esplicita per assicurarne l'applicazione. Anche in questo caso, c'è una gradazione pressoché infinita nel tipo di sanzione necessaria per sostenere il rispetto di una norma, dalla "pressione sociale" informale (*peer pressure*) fino a un'esplicita legislazione. Per esempio, è probabile che la maggior parte delle persone si astenga spontaneamente dal rubare, e non solo né soprattutto per il timore delle sanzioni previste dalla legge; ma è altrettanto probabile che la maggior parte delle persone paghi le tasse solo perché vi è costretto.

Gli esempi appena considerati mostrano che esistono tanti diversi tipi di norma, e tanti diversi "livelli di normatività" nelle regole che governano il comportamento degli individui in società. Tuttavia, questi esempi sono accomunati da alcune caratteristiche essenziali, che definiscono in modo sufficientemente preciso il concetto di norma. In particolare, le norme non sono semplici *abitudini* o regole di comportamento largamente, o anche universalmente, accettate e seguite nella società (cfr. Bicchieri e Muldoon 2011). Per esempio, lavarsi i denti tutti i giorni, non barare mentre si gioca a solitario o pregare prima di addormentarsi non sono norme, almeno nel senso qui discusso: sono infatti regole di comportamento che un individuo segue del tutto indipendentemente dal resto della società. Quando invece un individuo rispetta o viola una norma *sociale*, il suo comportamento è influenzato da ciò che egli si aspetta facciano gli altri e da ciò che pensa

che gli altri si aspettino da lui: in altre parole, il comportamento guidato da norme ha carattere *strategico* (cfr. Paternotte e Grose 2012). Analizzare il funzionamento di una norma sociale significa quindi analizzare il comportamento di individui coinvolti in un'interazione strategica e comprendere come la norma influisca su di esso sulla base delle aspettative che ogni individuo nutre sul comportamento e sulle aspettative altrui.

Norme ed equilibri

Per questo motivo, era del tutto naturale attendersi che i filosofi e gli scienziati sociali interessati all'analisi delle norme tentassero di applicare la teoria dei giochi al proprio campo d'indagine. In effetti, a partire dai primi contributi di Schelling (1966), David Lewis (1969) e Edna Ullmann-Margalit (1977), alcuni filosofi ed economisti – in particolare Sugden (1986), Cristina Bicchieri (1993, 2006), Kenneth Binmore (1994, 1998) e Herbert Gintis (2007, 2009) – hanno proposto un'analisi giochistica delle norme sociali basata sul concetto di equilibrio di Nash. Tale analisi chiarisce come le norme sociali dipendono dalle credenze e dalle preferenze di coloro che vi si conformano, definendo come segue una norma:

NORMA SOCIALE. Data una regolarità N del comportamento individuale in una determinata popolazione, diremo che N è una *norma sociale* nel caso in cui:

- (1) Quasi tutti *preferiscono* conformarsi a N a condizione (e solo a condizione) che quasi tutti si conformino a loro volta.
- (2) Quasi tutti *credono* che quasi tutti si conformino a N .

La condizione (1) equivale ad affermare che una norma sociale N è un equilibrio di Nash, cioè una combinazione di strategie in cui la strategia di ognuno è la risposta ottimale alle strategie altrui. La condizione (2) esprime la credenza di quasi tutti i membri della popolazione che, fra gli equilibri possibili, quasi tutti sceglieranno N .

La definizione appena vista ha il merito di illustrare, in modo semplice e rigoroso, il ruolo giocato dalle norme sociali nel regolare il comportamento di un individuo: in particolare, le norme permettono a un individuo di coordinare le proprie azioni con quelle altrui, adottando un comportamento appropriato alla specifica situazione sociale, cioè appropriato rispetto alle attese reciproche di ogni individuo sulle azioni di ogni altro. Inoltre, la definizione di norma come equilibrio di Nash spiega anche la “forza

normativa” che distingue questo concetto da quello di semplice abitudine o regolarità comportamentale: in equilibrio, ogni individuo si aspetta che tutti gli altri si conformino alla norma, e quindi è incentivato conformarsi a sua volta (Young 2008).

Naturalmente, la definizione giochistica appena illustrata spiega la *natura* delle norme sociali, ma non la loro *genesì* né la loro *evoluzione*. Sostenere che una norma è un equilibrio di Nash, infatti, non significa ancora spiegare come, nei diversi tipi di gioco, un particolare equilibrio venga identificato come quello rilevante. Gli studiosi che accettano tale definizione hanno in effetti offerto diverse analisi del *funzionamento* di una norma sociale, cioè di come gli individui giungano a coordinare le proprie azioni facendo emergere uno specifico equilibrio del gioco (cfr. Paternotte e Grose 2012).

Approcci contrattualisti, funzionalisti ed evolucionisti

Prima di discutere, nei prossimi paragrafi, il problema della genesi e dell'evoluzione delle norme sociali, conviene illustrare alcuni aspetti generali dei diversi approcci all'analisi giochistica delle norme che sono emersi nel corso degli ultimi decenni. A questo proposito, è interessante osservare come il primo tentativo di applicare la teoria dei giochi allo studio di questioni etico-politiche, inclusa la natura e il funzionamento delle norme sociali, risale già alla metà degli anni Cinquanta dello scorso secolo – cioè agli anni in cui Nash dà forma definitiva alla teoria classica. In una conferenza intitolata *Theory of Games as a Tool for the Moral Philosopher* (1955), il filosofo inglese Richard Braithwaite (1990-1990) propose infatti di affrontare il problema della giustizia distributiva in termini giochistici. A partire da questo pionieristico contributo, le applicazioni etico-politiche della teoria dei giochi hanno dato origine a tre diversi approcci distinti approcci: quello contrattualista, funzionalista ed evolucionista (cfr. Verbeek e Morris 2010).

L'*approccio contrattualista* si rifà direttamente alle ricerche di Braithwaite (1955) ed è ispirato dall'intento di riformulare e fondare su basi giochistiche la tradizionale teoria del contratto sociale. Tale approccio è stato fatto proprio da diversi studiosi, tra i quali John Harsanyi (1955), Brian Barry (1965), John Rawls (1971) e David Gauthier (1967, 1987). Per esempio, i cosiddetti contrattualisti della scelta razionale hanno usato gli strumenti concettuali della teoria della contrattazione – un'importante area della teoria dei giochi sviluppata da Nash – per giustificare le norme e le istituzioni politiche, viste come il risultato di una contrattazione tra agenti razionali. I sostenitori dell'approccio contrattualista tendono a identificare le norme morali con determinati “principi di giustizia” posti alla base di un sistema di

norme e istituzioni politiche; ciò significa che un sistema politico “giusto” andrebbe difeso con riferimento a una “comunità morale” formata da agenti razionali che condividono alcuni fondamentali principi morali.

L'*approccio funzionalista* mira invece a identificare la funzione delle norme sociali o morali, o, equivalentemente, a identificare i problemi che insorgono in assenza di tali norme (Ullmann-Margalit 1978). Un fondamentale contributo a questo approccio si deve alla filosofa israeliana Edna Ullmann-Margalit (1946-2010), che per prima ha utilizzato l'apparato formale della teoria dei giochi per esplorare la funzione delle norme morali. La Ullmann-Margalit (1977) riesce, fra l'altro, a mostrare che l'adesione alle norme morali consente agli attori impegnati in diversi tipi di interazioni strategiche di ottenere risultati molto migliori di quelli che sarebbero consentiti dal loro puro interesse egoistico, perseguito in conformità con i principi di scelta della teoria classica dei giochi.

Infine, l'*approccio evoluzionista* concepisce una norma sociale come l'effetto non intenzionale delle interazioni dei membri di una società. Tenta quindi – in accordo con i principi fondamentali delle teorie evoluzionistiche dei giochi – di spiegare l'esistenza delle diverse norme sociali mettendo in luce le ragioni che stanno alla base del loro successo evolutivo. Nelle loro analisi dei processi di formazione e sviluppo delle norme, gli “evoluzionisti” fanno ampio ricorso a *simulazioni computerizzate* – o *società artificiali* – in grado di riprodurre le interazioni tra agenti (cfr. Harms e Skyrms 2008). Va inoltre osservato che, soprattutto a partire dagli anni Novanta, i sostenitori dell'approccio evoluzionista hanno fatto ampio ricorso non solo alle teorie evoluzionistiche dei giochi ma anche a quelle cognitive (Bicchieri 1993, 2006).

Vale la pena mettere in rilievo alcune importanti somiglianze e differenze fra i tre approcci sopra menzionati. In primo luogo, mentre l'approccio contrattualista ha fini essenzialmente precrittivi, gli approcci evoluzionista e funzionalista hanno intenti apertamente descrittivi. Per questa ragione, i contrattualisti si avvalgono quasi esclusivamente della teoria classica dei giochi, mentre i funzionalisti e gli evoluzionisti fanno ampio uso anche delle teorie evoluzionistiche e cognitive. In secondo luogo, i funzionalisti e gli evoluzionisti tendono a trascurare il problema di valutare lo specifico contenuto delle norme e di giustificarne le pretese normative. Sono invece accomunati dall'interesse per la descrizione di quelle norme o pratiche morali che vengono comunemente considerate buone. Si può anzi affermare che la maggior parte delle indagini effettuate nell'ambito degli approcci funzionalista ed evoluzionista si basano, più o meno esplicitamente, sull'identificazione delle norme morali con le cosiddette *norme prosociali*, cioè

con le norme che risultano utili al funzionamento di una società e al benessere dei suoi membri.

Infine, va osservato che le differenze tra gli approcci funzionalista ed evolucionista non sono meno importanti delle loro somiglianze. In particolare, i funzionalisti tendono a considerare le norme come variabili “esogene” – cioè come fattori che regolano, per così dire, dall’esterno il comportamento individuale – piuttosto che come variabili “endogene”, cioè prodotti non intenzionali dell’interazione fra individui (Bicchieri e Muldoon 2011). Per esempio, i sostenitori dell’approccio evolucionista ritengono che identificare la funzione delle norme non basti a spiegarne l’emergere, la persistenza e l’evoluzione. Infatti, per spiegare come una norma possa di fatto emergere e diffondersi, non è sufficiente mostrare che, a posteriori, essa svolge una funzione utile all’interno della società (Elster 1989). Inoltre, spiegare la genesi delle norme sociali è un problema distinto da quello di spiegare l’evoluzione di tali norme, una volta che esse siano, in qualche modo, già emerse in una popolazione. Come osserva Bicchieri (1993, p. 242), la comprensione dei motivi per cui le norme sociali persistono nel tempo “non getta alcuna luce sul processo di formazione delle norme”; infatti “il problema di *come* emergono le norme è diverso dal problema del perché esse emergono e si stabilizzano” (*ivi*). Per questo motivo, gli evolucionisti ritengono che l’approccio funzionalista si precluda la possibilità di spiegare la genesi e la persistenza anche di norme disfunzionali o *anti-sociali*, cioè che non svolgono alcuna funzione positiva nei riguardi della società nel suo complesso.

Genesi ed evoluzione delle norme sociali

Come abbiamo appena notato, il problema della genesi di una norma sociale e quello della sua evoluzione sono due problemi ben distinti, che richiedono, in generale, strumenti di analisi diversi. Il primo richiede di capire come una norma sociale emerge all’interno di un certo gruppo di individui che precedentemente non riconoscevano tale norma. I due giocatori che si trovano per la prima volta ad affrontare il gioco della guida (figura 3.4) del capitolo 3, cioè a decidere su quale lato della strada accostare per incrociarsi senza incidenti, costituiscono un gruppo di questo tipo. Il problema dell’evoluzione di una norma è invece quello di capire come una norma, una volta emersa all’interno di gruppo, si stabilisca permanentemente nel gruppo, si diffonda all’esterno di esso “contagiando” altri gruppi o l’intera società o, invece, perda terreno e infine sparisca.

Diversi studiosi hanno esteso l'analisi giochistica delle norme sociali anche alla loro genesi ed evoluzione. Il primo tentativo di affrontare sistematicamente, in termini giochistici, il problema della *genesì* delle norme sociali è stato proposto dalla Ullmann-Margalit in *The Emergence of Norms* (1977). La sua tesi fondamentale è che la genesi delle norme può venire ricondotta alla circostanza che “certi tipi di norme sono soluzioni possibili ai problemi posti da certi tipi di situazioni di interazione sociale” (*ibidem*, p. vi). Prima di tutto, l'autrice mostra in modo convincente che un grande numero di situazioni della vita appartengono a tre forme paradigmatiche di interazione sociale, tra le quali il dilemma del prigioniero, e che esistono effettivamente determinate norme che risolvono i problemi posti da queste tre forme di interazione. A suo giudizio “la dimostrazione di questa correlazione” – tra determinate classi di problemi e le norme che ne facilitano la soluzione – può venire intesa come “una spiegazione dell'emergere di queste norme” (*ibidem*, p. 6).

Quest'ultima affermazione appare, però, assai problematica. Infatti, sebbene possiamo accettare l'idea che la capacità di risolvere i problemi connessi a una determinata forma di interazione sociale sia una condizione *necessaria* affinché un determinato schema di comportamento “emerga” fino ad acquisire lo status di norma sociale, non sembrano esserci buoni motivi per ritenere che tale condizione sia anche *sufficiente* a generare una norma. Al contrario, non si può escludere la possibilità che il genere di norme adatto a risolvere i problemi posti da certe situazioni di interazione sociale non venga “scoperto” da nessuno o, comunque, stenti ad emergere. Una buona spiegazione dell'emergere delle norme sociali deve quindi indicare almeno alcune condizioni sufficienti per il loro emergere. Ciò significa che, per spiegare adeguatamente la genesi delle norme, occorre risolvere un problema che, a dispetto del titolo, viene completamente eluso nel volume di Ullman-Margalit: *come*, cioè attraverso quale processo, potrebbero venire scoperte ed emergere norme potenzialmente utili in determinati tipi di interazione sociale?²

Un'interessante risposta a questo interrogativo si deve a Cristina Bicchieri (1993, cap. 6; 2006). Secondo la Bicchieri il processo di sviluppo delle norme sociali si svolge attraverso due stadi: nel primo, una norma emerge all'interno di piccoli gruppi, nel contesto di ripetute interazioni strategiche governate dalla razionalità individuale e da specifici processi di apprendimento; nel secondo, la norma può estendersi all'intera popolazione tramite appropriati meccanismi evolutivi. L'analisi del processo di formazione delle norme richiede, quindi, l'applicazione sia delle teorie cognitive dei giochi sia

2 Questa “omissione” viene rilevata, per esempio, da Taylor (1987, pp. 29-30). Si veda anche Bicchieri (1993, p. 242).

della teorie evoluzionistiche: le prime verranno impiegate per studiare l'emergere delle norme all'interno di piccoli gruppi, le seconde per analizzare la loro propagazione ed evoluzione nell'intera popolazione.

La spiegazione della genesi delle norme proposta dalla Bicchieri si basa sull'idea che gli individui siano in grado di apprendere, attraverso interazioni ripetute, con quale genere di concorrente hanno a che fare e, quindi, ad affrontarlo in modo adeguato. A partire da questo genere di *apprendimento*, derivante dall'interazione all'interno di *piccoli gruppi*, si possono creare regolarità di comportamento suscettibili di diffondersi, attraverso determinate dinamiche evolutive, in tutta la popolazione, fino a divenire norme sociali. Il fondamentale ruolo dei piccoli gruppi nella nascita delle norme dipende dal fatto che, nell'interazione ripetuta tra due persone o in un piccolo gruppo, ciascun agente può usare le proprie azioni come "segnali" e può anche sperimentare comportamenti diversi per imparare qualcosa sulle caratteristiche del partner con cui sta interagendo. Questo genere di apprendimento non sarebbe possibile in un grande gruppo dove le azioni di ogni particolare agente tendono ad essere ignorate.

Il processo di apprendimento mediante il quale una regola di comportamento emerge all'interno di un piccolo gruppo è qualcosa di molto diverso dalla dinamica evolutiva attraverso la quale, in condizioni favorevoli, tale regola può estendersi all'intera popolazione, fino a diventare una norma sociale. Rappresentare le norme sociali come il risultato di un processo di apprendimento in un contesto di interazione strategica significa accettare la tesi che queste "dipendono dalle scelte e, in ultima analisi, dalle preferenze e credenze individuali" (Bicchieri 1993, p. 242). A sua volta, ciò significa che "[l]a genesi delle norme è un esempio di ordine spontaneo, una forma di coordinazione che ha luogo senza l'intervento di un'autorità centrale e che non presuppone precedenti accordi o "contratti" tra le parti" (*ibidem*, p. 265).

Nel seguito del capitolo, passiamo in rassegna tre diversi approcci all'analisi della genesi e dell'evoluzione delle norme sociali (cfr. Bicchieri e Muldoon 2011, par. 8). Il primo è stato proposto in biologia e spiega l'emergere e la diffusione di una norma sulla base della selezione naturale di comportamenti prosociali e cooperativi. Il secondo è basato sul concetto di strategia di reciprocità, già introdotto nel paragrafo 4.2, e aggiunge all'approccio puramente biologico un importante ingrediente cognitivo, prevedendo che i giocatori siano in grado di distinguere chi tende a rispettare o a violare una norma. Il terzo approccio, sviluppato soprattutto da teorici dei giochi sulla base di simulazioni al computer, combina la teoria cognitiva e la teoria evoluzionistica dei giochi per arrivare ad un'analisi adeguata dell'emergere e dell'evolvere di una norma sociale.

6.2 Norme, altruismo e origini evolutive della cooperazione

Fra i diversi tipi di norma, giocano un ruolo di primo piano le cosiddette norme prosociali, cioè le norme che garantiscono la cooperazione e la coordinazione fra i membri di un gruppo e quindi il mantenimento e il buon funzionamento dell'ordinamento sociale (Bicchieri e Muldoon 2011, par. 1). Norme di questo tipo prescrivono per esempio l'altruismo, la lealtà, l'onestà, il mantenere le promesse fatte, il ricambiare i favori ricevuti e così via. In generale, possiamo chiamare "cooperativi" comportamenti di questo tipo, e "norme di cooperazione" le norme che li regolano e incentivano. Molti studiosi, in particolare quelli che adottano un approccio biologico all'analisi delle norme sociali, si sono concentrati sulle norme di cooperazione, tanto che, in larga parte della letteratura, l'analisi della genesi e dell'evoluzione delle norme coincide *tout court* con quella della genesi e dell'evoluzione della *cooperazione* (cfr. Cevolani e Festa 2012a).

Che l'uomo fosse un animale sociale era già stato osservato da Aristotele. A più di due millenni di distanza, le moderne scienze umane hanno messo in luce che l'uomo è un animale sociale *sui generis*, caratterizzato da una spiccata e persistente tendenza alla cooperazione. Poiché tale tendenza è chiaramente rintracciabile nell'intera storia umana, dal Pleistocene fino alle moderne società industriali, si può affermare che l'*homo sapiens* è "una specie cooperativa" (Bowles e Gintis 2011). Il carattere universale della propensione umana alla cooperazione fa anzi sospettare che essa sia un tratto fondamentale della natura umana. Un suo aspetto distintivo consiste nella tendenza degli esseri umani – a esclusione di una piccola percentuale di sociopatici – a interagire pacificamente con un gran numero di individui, ben al di là della ristretta cerchia dei famigliari, dei parenti e degli amici. La natura generalmente cooperativa dei rapporti che gli esseri umani intrattengono con un gran numero di estranei costituisce un tratto peculiare della cooperazione umana, che differenzia l'*homo sapiens* da tutte le altre specie animali. Le innumerevoli interazioni sulle quali si basa il funzionamento delle moderne economie di mercato forniscono un impressionante esempio dell'estensione, efficacia e stabilità della cooperazione tra estranei.

Mentre il carattere pervasivo della cooperazione umana è attestato dall'esperienza quotidiana di ciascuno di noi e ampiamente documentato dalle scienze sociali, la spiegazione di questo fenomeno appare tutt'altro che semplice. Con riferimento ai molti aspetti ancora oscuri della cooperazione umana, non sembra inappropriato parlare di *enigma della cooperazione*. Non sorprende, quindi, che la formulazione di una soddisfacente

teoria della cooperazione – cioè, di una teoria in grado di dirci come, perché e a quali condizioni gli esseri umani cooperano – venga comunemente considerata un obiettivo fondamentale delle scienze sociali.³

Lungi dal dissipare l'enigma della cooperazione, le ricerche sull'argomento condotte nel secolo scorso hanno alimentato i dubbi, conducendo a conclusioni paradossali. Infatti, alcune fra le più accreditate teorie del comportamento umano sembrano implicare la conclusione che l'uomo *non* dovrebbe cooperare – o, almeno, che non dovrebbe farlo nella misura in cui lo fa. Tale conclusione rappresenta quello che potremmo chiamare il *paradosso della cooperazione* umana. Si noti, inoltre, che possiamo osservare varie forme di cooperazione altruistica non solo nelle società umane, ma anche nelle popolazioni animali (cfr. Okasha 2009). Per esempio, molti animali che vivono in gruppo hanno l'abitudine di lanciare un segnale di allarme all'avvicinarsi di un predatore: in tal modo accrescono la sicurezza del gruppo, correndo però il rischio non trascurabile di attirare su di sé l'attenzione del predatore.

La ricerca di una buona spiegazione dell'altruismo animale ha costituito una sfida importante per la biologia evuzionistica fin dai suoi albori. Già Charles Darwin, infatti, aveva notato che un individuo altruista, sacrificandosi a favore degli altri, tende a ridurre il proprio vantaggio riproduttivo, cioè le proprie possibilità di lasciare una discendenza numerosa. Questo sembra implicare che gli altruisti dovrebbero estinguersi, a tutto vantaggio degli egoisti. D'altra parte, tale conclusione rende incomprensibile la persistente diffusione dell'altruismo in molte specie animali, compresa quella umana. Com'è possibile, infatti, che gli schemi di comportamento altruistico siano stati selezionati nel corso dall'evoluzione biologica, così da essere ancora oggi molto diffusi?

Darwin proponeva di risolvere questo problema facendo appello a un meccanismo evolutivo che sarebbe poi stato chiamato “selezione di gruppo”:

Una tribù che ha molti membri [...] che si aiutano l'un l'altro e si sacrificano per il bene comune, riuscirà vittoriosa su molte altre tribù; e questa sarebbe selezione naturale (Darwin 1971, cap. 5, p. 173).

3 A questo proposito, ricordiamo che la lista dei 25 grandi problemi aperti della scienza contemporanea, pubblicata nel 2005 su *Science*, in occasione del 125° anniversario della rivista, include il problema dell'origine e dell'evoluzione della cooperazione umana (Kenedy e Norman 2005, p. 75). I principali scritti sull'argomento pubblicati negli ultimi decenni sono discussi in Bowles e Gintis (2011, pp. 6-7), Rubin (2002, pp. 39-40) e Friedman (2008, pp. 355 ss.). Per una rassegna, si veda Cevolani e Festa (2012a).

Darwin sosteneva, sia pure con qualche perplessità, che l'ipotesi della selezione di gruppo, cioè l'idea che la selezione agisse anche al livello del gruppo, avrebbe permesso di spiegare la diffusione dell'altruismo, tramite la prevalenza di gruppi con un'alta percentuale di altruisti (Okasha 2009, par. 1). L'ipotesi darwiniana della selezione di gruppo venne adottata dalla maggior parte degli evoluzionisti ed era ancora ampiamente accettata negli anni Trenta e Quaranta dello scorso secolo quando – grazie all'opera di studiosi come Julian Huxley (1887-1975), Ronald Fisher (1890-1962) e John B. S. Haldane (1892-1964) – venne elaborata la cosiddetta sintesi neodarwiniana, che pose le basi della moderna biologia evoluzionistica.⁴

Addomesticato dall'ipotesi della selezione di gruppo, il problema dell'altruismo ritornò però prepotentemente alla ribalta negli anni Sessanta, quando alcuni studiosi – come George Williams (1926-2010), William Hamilton (1936-2000) e Maynard Smith – promossero quella che molti considerano “un'autentica rivoluzione nella biologia” (Ridley 1996, p. 61). Il primo risultato di questa rivoluzione fu l'abbandono generale dell'ipotesi della selezione di gruppo. Si ripropose così lo spinoso problema di offrire una buona spiegazione del comportamento altruistico negli animali.

La prima spiegazione di questo genere venne formulata dagli stessi Williams e Hamilton, nell'ambito della cosiddetta teoria del *gene egoista* emersa da questa rivoluzione. L'espressione, coniata da Richard Dawkins (1976), si riferisce all'ipotesi che la selezione naturale non agisca al livello degli individui, né tantomeno a quello dei gruppi, ma piuttosto a quello dei geni, da cui dipendono le caratteristiche fisiche e comportamentali degli individui. Ciò implica, fra l'altro, che ogni gene egoista guida il comportamento del suo portatore cercando di massimizzare il numero di proprie

4 Fra gli studiosi che accettarono l'idea darwiniana della selezione di gruppo figura in una posizione preminente il principe russo Pëtr Kropotkin (1842-1921), autore di un pionieristico quanto influente studio sull'altruismo negli animali (*Mutual Aid*, 1902). All'infaticabile attività di teorico e organizzatore del movimento anarchico internazionale, Kropotkin univa robusti interessi scientifici, specie nei campi della geografia, della geologia e della zoologia. I capitoli di *Mutual Aid* riproducono una serie di articoli pubblicati tra il 1890 e il 1896 sulla rivista *Nineteenth Century*, in polemica con Thomas Henry Huxley (1825-1895), il celebre “mastino di Darwin” le cui posizioni aprirono la strada al darwinismo sociale. Il volume di Kropotkin continua, anche ai nostri giorni, a sollevare l'interesse di biologi evoluzionisti e teorici dei giochi; si vedano, per esempio, Skyrms (1996, pp. 45, 46, 51, 62; 2004, p. 6) e Bowles e Gintis (2011, p. 7). Il biologo Lee Alan Dugatkin (2011) offre un suggestivo ritratto della vita e dell'opera di Kropotkin; per quanto riguarda, in particolare, l'influenza di tale opera sull'approccio evoluzionistico si vedano anche Gould (1991, cap. 22) e Dugatkin (2006, cap. 1-2).

copie nella generazione successiva. Se il comportamento di un individuo è vantaggioso, nel senso che tende ad aumentare la sua discendenza, i geni di quell'individuo tenderanno a diffondersi nella popolazione; in caso contrario, tenderanno a sparire.

La teoria del gene egoista permette di spiegare l'altruismo individuale facendo ricorso al meccanismo della *selezione parentale*. Se, come spesso accade fra gli animali, uomo compreso, il comportamento altruistico di un individuo è rivolto ai suoi consanguinei, che condividono con lui buona parte del patrimonio genetico, allora tale comportamento risulta vantaggioso per i geni dell'altruista, poiché incrementa le probabilità che essi si riproducano nella generazione successiva. In questo modo, un comportamento altruista, anche se svantaggioso per l'individuo, è vantaggioso per il gene. In questa prospettiva "genocentrica", l'individuo non è altro che "il giocattolo e lo strumento con cui si gingilla un comitato di geni egoisti" (Ridley 1996, p. 63).

A partire dagli anni Settanta del secolo scorso, la spiegazione dell'altruismo e della cooperazione negli animali e nell'uomo ha avuto un posto di primo piano nell'agenda dei biologi evuzionisti e degli scienziati sociali. Un forte impulso alle ricerche su questo argomento venne dalla pubblicazione del volume *Sociobiology*, con il quale il biologo evuzionista statunitense Edward Wilson (1975) stabilì l'omonima disciplina. L'obiettivo fondamentale della sociobiologia è quello di analizzare il comportamento umano alla luce della biologia evuzionistica, partendo dal presupposto che la maggior parte dei nostri comportamenti sono adattamenti selezionati durante lo sviluppo evolutivo della nostra specie, e possono quindi venir spiegati in termini di vantaggio riproduttivo. L'approccio sociobiologico è stato aspramente criticato, fin dal suo apparire, anche da eminenti biologi evuzionisti, come Stephen Jay Gould e Richard Lewontin. Ciò non ne ha impedito la rapida diffusione, come dimostra l'impetuoso sviluppo, sul tronco della sociobiologia, di un nuovo programma di ricerca che va sotto il nome di *psicologia evuzionistica* (Barkow, Cosmides e Tooby 1992). Gli psicologi evuzionisti cercano di far luce sui meccanismi che stanno alla base del comportamento umano analizzando tali meccanismi come adattamenti evolutivi all'ambiente in cui l'uomo si è evoluto, il cosiddetto *ambiente dell'adattamento evolutivo*.

Il concetto di ambiente dell'adattamento evolutivo (AAE) è stato introdotto, negli anni Sessanta del secolo scorso, dallo psichiatra inglese John Bowlby (1969). L'AAE corrisponde approssimativamente al periodo del Pleistocene (sulla scala geologica) o del Paleolitico (in archeologia), cioè al periodo compreso fra 2,6 milioni e 12 000 anni fa. Il genere *Homo* sa-

rebbe apparso circa 2,5 milioni di anni fa, la specie *Sapiens* (la nostra) circa 200 000 anni fa e il “moderno” *Homo Sapiens* (essenzialmente identico agli attuali esseri umani anche dal punto di vista intellettuale e comportamentale) circa 50 000 anni fa. L’evoluzione dell’uomo sarebbe quindi avvenuta quasi interamente in un contesto profondamente diverso da quello attuale. In particolare, dato che la *rivoluzione neolitica*, che segna la nascita e la diffusione dell’agricoltura e dell’allevamento, è datata a circa 10 000 anni fa, il moderno *Homo Sapiens* avrebbe vissuto quattro quinti della sua esistenza da nomade cacciatore-raccogliitore, cioè non diversamente dai suoi antenati *Homo*. Dato che l’AAE aveva caratteristiche molto diverse dalle società attuali, gli psicologi evuzionisti ipotizzano che le preferenze e le tendenze comportamentali acquisite nell’AAE non siano necessariamente vantaggiose nelle società moderne (cfr. Rubín 2002, pp. 60 ss.; Friedman 2008; Ridley 1996).

Reciprocità e cooperazione con gli estranei

A mezzo secolo di distanza dalla rivoluzione biologica degli anni Sessanta, molti aspetti della cooperazione e dell’altruismo negli animali restano ancora oscuri. Per esempio, il meccanismo della selezione parentale non può, per sua stessa natura, spiegare i comportamenti altruistici nei riguardi di chi non appartiene al gruppo ristretto dei consanguinei. Si tratta di una limitazione niente affatto secondaria, se si pensa al fatto che comportamenti di questo genere sono del tutto comuni tra gli esseri umani e vengono spesso osservati anche nel resto del regno animale.

Una spiegazione evuzionistica della cooperazione con gli estranei, proposta da Robert Trivers (1971), si basa sulle cosiddette strategie di *reciprocità*. Una strategia di questo genere consiste nell’aiutare un altro individuo contando sul fatto che, alla prossima occasione, questo ricambierà il favore. Un esempio è quello dei pipistrelli vampiro che condividono regolarmente parte del sangue succhiato alle vittime con i loro compagni meno fortunati nella caccia.⁵ Lungi dall’essere indiscriminata, la generosità di questi pipistrelli è regolata dalla reciprocità, cioè dalla loro capacità di riconoscere i membri del gruppo e dalla tendenza a condividere il cibo solo con coloro che in precedenza si sono comportati generosamente.

Secondo diversi studiosi, la reciprocità ha svolto un ruolo fondamentale nell’evoluzione della capacità umana di cooperare con gli estranei. Il suo successo in termini di adattamento ne avrebbe anzi fatto un istin-

5 Ci riferiamo al *Desmodus rotundus*, o vampiro comune, diffuso in Sud e Centro America (cfr. Wilkinson 1984, pp. 181-184).

to insopprimibile della natura umana. Alcune convinzioni molto diffuse, come “occhio per occhio, dente per dente”, potrebbero derivare proprio dalle pressioni selettive che hanno favorito la diffusione delle strategie di reciprocità nel moderno *Homo Sapiens*. La stessa origine avrebbe quella “particolare inclinazione della natura umana [...] a trafficare, barattare e a scambiare una cosa con l'altra” da cui, secondo Adam Smith (1776, p. 72), emergono gli scambi di mercato e la divisione del lavoro (si veda anche Ridley 1996, p. 99).

Nelle società umane la cooperazione altruistica con estranei è un fenomeno non meno diffuso della cooperazione mutualistica. La nostra esperienza quotidiana ci mostra innumerevoli casi in cui un individuo aiuta un estraneo senza aver alcuna ragionevole aspettativa di venir ricambiato in futuro – e, spesso, sperando di non aver mai bisogno di esser ricambiato. Si pensi, per esempio, a tutti coloro che fanno l'elemosina, si impegnano nelle organizzazioni di volontariato, offrono aiuto alle vittime di incidenti stradali o, più semplicemente, prestano il cellulare a qualcuno che lo ha smarrito e deve avvertire la moglie che arriverà a cena in ritardo.

Spiegare la tendenza degli esseri umani a comportarsi altruisticamente con gli estranei è una sfida di grande importanza per l'approccio evolutivistico. Alcuni studiosi pensano che tali comportamenti altruistici potrebbero essere il lascito evolutivo di certe forme di cooperazione diffuse nelle società primitive (Ridley 1996; Rubin 2002; Friedman 2008, cap. 1). Sembra infatti che nell'AAE la società umana fosse composta da piccoli gruppi di individui (fra i 25 e i 150), spesso imparentati fra loro. In gruppi di questo genere, il comportamento altruistico avrebbe potuto evolvere sia grazie alla selezione parentale sia grazie alla reciprocità. L'ipotesi della psicologia evolutivistica, per cui i nostri istinti comportamentali si sono evoluti in questo periodo e non sono più cambiati in modo sostanziale, spiegherebbe perché ancora oggi tendiamo ad adottare comportamenti altruistici. Infatti, anche se nelle nostre società, molto più estese, complesse e anonime rispetto all'AAE, i comportamenti altruistici non portano più alcun vantaggio riproduttivo – dato che di solito i beneficiari non sono né parenti né persone da cui possiamo aspettarci di essere ricambiati – ci comportiamo spesso altruisticamente poiché abbiamo ereditato questa tendenza dai nostri antenati che vivevano nell'AAE, dove essa risultava vantaggiosa.

Selezione di gruppo ed evoluzione culturale

Le difficoltà incontrate nel tentativo di formulare una soddisfacente spiegazione generale dell'altruismo animale e umano sulla sola base dei

meccanismi della selezione parentale e della reciprocità hanno indotto molti studiosi – tra i quali il biologo evolucionista David Wilson, assieme al filosofo Elliott Sober e, più recentemente, al padre della sociobiologia, Edward Wilson – a riconsiderare la vecchia ipotesi della selezione di gruppo, riproponendone alcune versioni opportunamente aggiornate (cfr. Sober e Wilson 1999, Wilson 2005, Wilson e Wilson 2007). Per esempio, secondo la teoria della *selezione multilivello* (SML), il processo evolutivo opera su una gerarchia di diversi livelli di complessità biologica che, dal basso verso l'alto, includono, almeno, i livelli del gene, della cellula, dell'individuo e del gruppo. Secondo SML, il fatto che un comportamento sia svantaggioso a un certo livello di selezione, per esempio a quello individuale, non significa che lo sia anche per gli altri livelli; anzi, considerato nel suo complesso, quel comportamento potrebbe risultare vantaggioso. In particolare, comportamenti come l'altruismo, che risultano vantaggiosi per il gruppo ma svantaggiosi per i suoi membri, potrebbero diffondersi in tutti quei casi in cui la “pressione selettiva” a livello di gruppo è più forte di quella a livello individuale.⁶

Le ricerche sull'*evoluzione culturale* delle società umane – cioè sull'origine e lo sviluppo di credenze, norme, tradizioni e istituzioni – sono un tema classico delle scienze sociali, il cui studio sistematico risale almeno a David Hume. Una delle ragioni che, nell'ultima metà del ventesimo secolo, avevano spinto i biologi evolucionisti a ignorare tali ricerche era che l'evoluzione culturale veniva spesso analizzata nei termini dell'allora screditata idea di selezione di gruppo.⁷ La recente diffusione di SML e di altre versioni aggiornate della teoria della selezione di gruppo ha riaperto l'interesse dei biologi evolucionisti per l'evoluzione culturale e i suoi rapporti con

6 Le indagini condotte di recente su diversi animali sembrano confermare la SML; si veda, per esempio, Wade *et al.* (2010).

7 L'esempio più famoso del ricorso all'idea di selezione di gruppo, che considereremo ancora nel capitolo 7, è rappresentato dalla concezione evolucionistica dell'ordine sociale elaborata da Hayek (1973-1979, par. I.1, in particolare pp. 25-28). Hayek era un grande ammiratore degli sviluppi della biologia e dell'etologia, a cui s'era avvicinato anche grazie alla conoscenza di prima mano delle ricerche effettuate da Konrad Lorenz e Ludwig von Bertalanffy, entrambi suoi amici, coetanei e compatrioti. D'altra parte, Hayek criticava aspramente «gli errori della sociobiologia» (Epilogo, p. 528) che, a suo giudizio, avrebbe del tutto trascurato la dimensione specificamente culturale dell'evoluzione umana (Epilogo, in particolare pp. 528-31, 538-40 e 548-49). Sulla teoria hayekiana dell'evoluzione culturale e sulle sue relazioni concettuali con i recenti sviluppi dell'approccio evolucionistico, si vedano Koppl (2004, in particolare il contributo di Rubin e Gick e la parte II), Zywicki (2000), Caldwell (2002) e Marciano (2009).

l'evoluzione biologica. I frutti di questo rinnovato interesse non si sono fatti attendere.⁸ In particolare, per quanto riguarda il tema che in questa sede più ci interessa, occorre ricordare che le ricerche sull'interazione tra evoluzione biologica e culturale condotte in questi ultimi anni hanno consentito di comprendere che alcuni limiti delle spiegazioni puramente biologiche della cooperazione e dell'altruismo possono venir superati facendo ricorso anche ai meccanismi dell'evoluzione culturale.

6.3 *L'evoluzione della cooperazione e delle norme di giustizia*

Mentre i biologi interessanti alla genesi e all'evoluzione del comportamento altruistico e cooperativo si sono concentrati soprattutto sui meccanismi della selezione naturale (individuale o di gruppo) operanti all'interno di popolazioni animali, i filosofi e i teorici dei giochi hanno adottato un approccio più generale a questo problema, basato sia sulle teorie cognitive sia sulle diverse versioni della teoria evoluzionistica dei giochi. Abbiamo già visto all'opera questo approccio nel capitolo 4, discutendo il problema della cooperazione alla produzione di un bene pubblico. Tale problema rappresenta un caso particolare di quello, più generale, dell'emergere e del diffondersi di strategie cooperative. In effetti, come mostrato da Axelrod (1984), l'analisi del dilemma ripetuto del prigioniero permette di studiare non solo possibili soluzioni a diversi dilemmi sociali, ma anche la genesi e l'evoluzione di qualsiasi strategia cooperativa, e quindi di illuminare il problema dell'origine e dello sviluppo delle norme prosociali.

L'evoluzione della cooperazione

In effetti, lo scopo dell'influente volume *The Evolution of Cooperation* di Axelrod (1984) è proprio quello di analizzare i meccanismi attraverso i quali determinate strategie basate sulla cooperazione possono affermarsi spontaneamente anche in una società di egoisti razionali. Axelrod propone di studiare questo problema cercando di identificare le condizioni necessarie e sufficienti per l'affermarsi di "norme di cooperazione" in un gioco ripetuto del prigioniero. Come si ricorderà, Axelrod impiegò una si-

8 Si pensi, per esempio, ai recenti sviluppi della cosiddetta teoria della coevoluzione gene-cultura, le cui basi sono state poste a partire dagli anni Settanta dello scorso secolo; si vedano Cavalli-Sforza e Feldman (1981), Lumsden e Wilson (1981), Richerson e Boyd (1985), Fehr e Fischbacher (2003), Boyd e Richerson (2005); e, per una rassegna, Lewens (2008).

mulazione al computer per confrontare le prestazioni di diverse strategie, cooperative e non, per il dilemma ripetuto; da tale confronto, emerse come vincitrice la strategia del “colpo su colpo” che abbiamo chiamato T , che coopera alla prima mossa e in ogni mossa successiva riproduce semplicemente la scelta effettuata dal concorrente nella mossa immediatamente precedente. In altre parole, T è caratterizzata dal massimo grado di reciprocità e, quindi, dalla massima disponibilità a cooperare con altre strategie che si dimostrino inclini alla cooperazione. Axelrod organizzò anche un secondo torneo computerizzato, al quale ammise sessantadue strategie proposte da studiosi a conoscenza dei risultati del primo torneo. Nonostante alcune di queste strategie fossero ispirate dallo sforzo consapevole di perfezionare T , la versione originale di T vinse anche il secondo girone.

Il successo di T nel secondo girone, contro una così grande varietà di avversari, suggerisce l'ipotesi che T sia una strategia talmente “robusta” da comportarsi bene nelle più svariate situazioni, e persino nell'ambito di popolazioni nelle quali possa essere rappresentata qualsiasi strategia. Per determinare il “grado di robustezza” di T – e di strategie dello stesso genere, fondate su criteri di reciprocità e disponibilità a cooperare – Axelrod effettuò svariati esperimenti di simulazione basati sui modelli dinamici della teoria evuzionistica dei giochi. Questi esperimenti consistono in una lunga serie di gare nelle quali le strategie di maggior successo sono ammesse a partecipare a un numero maggiore di incontri, mentre viene limitato il numero delle strategie meno fortunate. Ogni gara rappresenta una generazione di individui ciascuno dei quali impiega sempre una stessa strategia nelle sue interazioni con gli altri membri della popolazione. Il successo di una strategia in ogni singola gara determina la numerosità della sua “prole”, e quindi la misura in cui tale strategia sarà rappresentata nella generazione successiva. A partire da una determinata distribuzione iniziale delle strategie nella popolazione – e dal valore di alcuni altri parametri rilevanti, è possibile stabilire, attraverso ben congegnati esperimenti di simulazione, come evolveranno le diverse strategie e quale sarà la loro distribuzione finale nella popolazione.

Gli esperimenti condotti da Axelrod fanno luce sulle ragioni del successo di TFT – e, più in generale, delle strategie “cooperative” – nelle più svariate condizioni iniziali, e anche sui meccanismi attraverso i quali tali strategie possono penetrare in un “mondo di cattivi” (dove domina la strategia D “Defeziona sempre!”), diffondersi in questo mondo fino a invaderlo e resistere, poi, al tentativo di invasione da parte di altre strategie. Axelrod mostra che il meccanismo fondamentale attraverso il quale le strategie cooperative possono prosperare nel gioco iterativo del prigioniero è quello

della correlazione, per cui un individuo che usa una certa strategia ha una probabilità sufficientemente alta di incontrarne un altro con la sua stessa strategia. Tale meccanismo è essenzialmente identico a quello studiato da Skyrms, che ha proseguito le ricerche di Axelrod concentrandosi sull'evoluzione delle norme di equità e giustizia.

L'evoluzione delle norme di giustizia

In *The Evolution of Social Contract* (1996), il filosofo della scienza americano Brian Skyrms mostra in modo piuttosto convincente che i modelli dinamici della teoria evuzionistica dei giochi possono venire utilmente impiegati per spiegare l'evoluzione di un'ampia gamma di norme sociali. Molta attenzione viene dedicata, in particolare, al "problema della giustizia". Rinunciando al tentativo – che sarebbe, del resto, estraneo alla sua prospettiva teorica – di formulare una teoria generale della giustizia o della società giusta, l'autore esamina l'evoluzione culturale della giustizia distributiva con l'obiettivo di spiegare come potrebbero essersi sviluppate alcune norme di comportamento improntate a criteri di equità distributiva. A tale scopo, nei primi due capitoli del volume, Skyrms analizza una serie di semplici giochi di contrattazione che vale la pena di illustrare in qualche dettaglio.

Immaginiamo che io e te dobbiamo accordarci su come suddividere una torta di cioccolato miracolosamente caduta dal cielo, e che, in caso di mancato accordo, la torta si deteriori, così da lasciare entrambi a mani vuote. Supponendo che questa contrattazione venga effettuata in una situazione di perfetta simmetria, in cui nessuno può rivendicare particolari diritti nei confronti dell'altro, sembra ovvio che ciò che dovremmo fare, in base a qualunque immaginabile principio di equità, è dividere la torta in parti uguali.

Tuttavia, questo non ci impedisce di immaginare comportamenti diversi. Per esempio, potrebbe esserci qualche mercanteggiamento preliminare del tipo: "Cosa ne diresti di lasciarne 2/3 a me e tenerne 1/3 per te? No, io ne prendo il 60% e ne lascio il 40% a te!". Questa forma di contrattazione può venir rappresentata come un gioco in cui ciascuno di noi scrive privatamente la sua richiesta finale di una percentuale di torta – corrispondente al valore minimo che sarebbe disposto ad accettare – su un pezzo di carta e lo consegna a un immaginario arbitro del gioco. Se il totale delle richieste supera il 100% l'arbitro tiene la torta per sé. Altrimenti otteniamo entrambi quello che chiediamo. Diversi esperimenti psicologici hanno mostrato che di solito i soggetti che partecipano a questo gioco si comportano come ci attenderemmo, cioè che agiscono con equità, dividendosi la torta a metà. Potremmo anche cercare di dimostrare, sulla base di qualche teoria della

scelta razionale, che in questo caso la decisione moralmente giusta coincide con quella più conveniente per entrambi i soggetti.

In un modello di gioco leggermente più complicato, la suddivisione della torta deve essere effettuata sotto un cosiddetto “velo d’ignoranza”: io e te dobbiamo accordarci su come dividere la torta tra gli individui *A* e *B*, e un arbitro deciderà poi se tu sei *A* e io *B*, o viceversa. Anche in questo caso potremmo tentare di dimostrare che la scelta moralmente giusta coincide con quella più conveniente per entrambi i soggetti. Infatti, proprio partendo dall’analisi di modelli di gioco non troppo diversi da questo, autori come Harsanyi e Rawls sono arrivati a interpretare le regole giuste come quelle regole che sarebbero prodotte da scelte razionali effettuate dietro veli d’ignoranza opportunamente specificati.

La possibile coincidenza tra scelta razionale e giustizia lascia, però, il posto a un evidente conflitto in altri tipi di gioco, come il *gioco dell’ultimatum*, di cui ci occuperemo tra poco. Una vasta letteratura psicologica indica chiaramente che i soggetti sperimentali che partecipano a questo gioco rivelano una stabile inclinazione per un comportamento che può venire descritto come giusto, cooperativo, o altruistico – e che si trova in conflitto con le norme basate sulle teorie della scelta razionale e, in particolare, sulla teoria “classica” dei giochi. Un obiettivo teorico fondamentale di Skyrms è quello di spiegare l’alta frequenza di comportamenti di questo genere in base alla teoria evoluzionistica dei giochi.

Nel gioco dell’ultimatum occorre suddividere un bene, per esempio una somma di 10 euro, tra due giocatori. A differenza del gioco della torta, spetta ora a uno dei due giocatori, che chiameremo *Proponente*, dare l’ultimatum, cioè fare una proposta all’altro giocatore, che chiameremo *Controparte*. Quest’ultimo può solo accettare o rifiutare la proposta: se la accetta il Proponente ottiene quello che chiede e la Controparte quello che resta, se la rifiuta nessun giocatore ottiene nulla. Si noti che un perfetto egoista illuminato dovrebbe, quando impersona la Controparte, accettare qualsiasi offerta positiva del Proponente. Gli esperimenti hanno mostrato che molti soggetti affrontano il gioco sulla base di *principi di equità* che includono non solo norme per effettuare proposte eque quando ci si trova nel ruolo di Proponente, ma anche norme per “punire” offerte inique quando ci si trova in quello di Controparte. Tali soggetti, per esempio, sono disposti a rinunciare a 1 o 2 euro per punire un Proponente ingordo che chiede di avere 8 o 9 dollari su 10. Ciò significa che essi agiscono sulla base di norme che – in conflitto con il postulato TGC1 – violano l’imperativo di massimizzare il guadagno monetario atteso.

Skyrms applica la teoria evoluzionistica dei giochi a una versione semplificata del gioco dell’ultimatum in cui occorre spartire una torta divisa in

dieci fette e ciascun giocatore ha soltanto due opzioni: il Proponente può chiedere per sé cinque pezzi oppure nove, e la Controparte può accettare o rifiutare la proposta. Si noti, tuttavia, che mentre il Proponente ha solo due strategie a propria disposizione – “Chiedi 9!”, “Chiedi 5!” – la Controparte ne ha quattro: “Accetta tutto!”, “Rifiuta tutto!”, “Accetta se il Proponente chiede 5, e rifiuta se chiede 9!”, “Accetta se il Proponente chiede 9, e rifiuta se chiede 5!”. Immaginiamo una popolazione in cui gli individui giocano ripetutamente tra loro, talvolta nel ruolo di Proponente e talvolta in quello di Controparte, e supponiamo che ciascun individuo adotti una strategia fissa, che gli dice cosa fare in ciascun ruolo. Vi saranno otto strategie da considerare – corrispondenti alle otto possibili combinazioni delle strategie del Proponente e della Controparte – che caratterizzano otto diversi tipi di individui. Ad alcuni di questi, particolarmente interessanti nel contesto della nostra discussione, conviene attribuire un nome: lo Spregiudicato (*Gamesman*) chiede 9 nel ruolo di Proponente e accetta tutto nel ruolo di Controparte; l’Imparziale (*Fairman*) chiede 5 e accetta la richiesta di 5, ma rifiuta quella di 9; il Cane Arrabbiato (*Mad Dog*) chiede 9 e accetta la richiesta di 9, ma rifiuta quella di 5; il Disinteressato (*Easy Rider*) chiede 5 e accetta tutto.

In uno dei modelli dinamici per il gioco dell’ultimatum esplorati da Skyrms, si assume che in ogni turno di gioco le coppie siano formate in modo casuale, e che sia determinato nello stesso modo anche il ruolo assunto da ogni individuo. Inoltre, si suppone che le fette di torta guadagnate nel gioco rappresentino qualche vantaggio evolutivo, in grado di influenzare la numerosità della prole, a cui viene trasmessa la stessa strategia del genitore. Sulla base di queste assunzioni, che corrispondono alla dinamica dei riproduttori della teoria evuzionistica dei giochi, si può programmare un computer allo scopo di simulare questa dinamica e osservare come evolveranno popolazioni in cui sono rappresentate proporzioni variabili delle otto strategie.

Simulazioni di questo genere conducono a conclusioni illuminanti, e spesso inattese. Per esempio, se nella popolazione iniziale sono rappresentate proporzioni uguali di ciascuna strategia, si trova – come è naturale aspettarsi – che il processo dinamico conduce, dopo un certo punto, all’estinzione degli Imparziali. Tuttavia, piuttosto sorprendentemente, si scopre anche che gli Spregiudicati non giungeranno a dominare l’intera popolazione, che evolverà invece verso uno stato polimorfico con circa l’87% di Spregiudicati e il 13% di Cani Arrabbiati. Ciò che sorprende è la persistenza della stravagante strategia dei Cani Arrabbiati che rifiutano le offerte eque e accettano quelle inique. In base alla teoria classica dei giochi, la strategia degli Cani Arrabbiati è irragionevole in quanto dominata

da altre strategie. Un aspetto interessante della teoria evuzionistica dei giochi, basata sulla dinamica dei riproduttori, consiste proprio nel fatto di consentire, in determinate condizioni, la sopravvivenza di strategie “anomale” come questa.

Ci possiamo ora chiedere se esistono condizioni in cui un'altra strategia anomala, vale a dire quella degli Imparziali – che si basa su una semplice norma di giustizia distributiva – possa sopravvivere. A questo proposito, Skyrms dimostra che, per evitare l'estinzione degli Imparziali, è sufficiente che si attribuiscono valori appropriati alle proporzioni con cui le diverse strategie sono rappresentate nella popolazione iniziale. Per esempio, se si parte con una popolazione composta per il 30% da Imparziali – e il resto equamente suddiviso tra le altre sette strategie – allora la dinamica evuzionistica condurrà la popolazione a uno stato polimorfico con circa il 64% di Imparziali e il 34% di Disinteressati. Ciò significa che la strategia anomala degli Imparziali può non solo sopravvivere, ma anche prosperare.

Un'interessante estensione del modello dinamico sopra illustrato si basa sull'assunzione che gli accoppiamenti non siano effettuati in modo completamente casuale, bensì in base a qualche *meccanismo di correlazione*, tale che – come sembra accadere in natura – individui che usano la stessa strategia tendono a giocare tra di loro più spesso che con individui che usano una strategia diversa. Utilizzando questo modello dinamico “correlato”, Skyrms riesce a dimostrare che, in condizioni favorevoli di correlazione, possono ottenere un straordinario successo non solo i comportamenti imparziali, ma anche quelli *altruistici*, in cui un soggetto accresce il guadagno degli altri giocatori a spese del proprio. La maggior parte delle applicazioni del modello correlato discusse da Skyrms sono di carattere etologico. Per esempio, il modello viene utilizzato per spiegare l'evoluzione del senso di possesso e dei sistemi di segnalazione nel mondo animale, e anche il fatto che molto spesso in natura vediamo funzionare la “legge di aiuto reciproco” ipotizzata da Kropotkin (1902). D'altra parte, Skyrms ipotizza che anche l'evoluzione culturale – e, in particolare, l'evoluzione dei principi che regolano il contratto sociale implicito – sia governata, almeno in parte, dai meccanismi del modello correlato. Per esempio, questo modello potrebbe spiegare sia lo sviluppo delle norme che regolano il possesso e la proprietà sia un certo numero di tendenze cooperative e altruistiche. In altre parole, ci sono buone ragioni per ritenere che il modello correlato possa essere utilizzato anche per spiegare l'evoluzione di determinate norme sociali, e quindi per studiare il processo dell'evoluzione culturale umana.⁹

9 Cfr. Skyrms (*ibidem*, p. 61) e D'Arms *et al.* (1998, pp. 93-97).

Per evitare possibili equivoci, occorre sottolineare che, come vedremo meglio nel prossimo capitolo, l'obiettivo principale di Skyrms è quello, consapevolmente limitato, di elaborare un modello concettuale che ci permetta di formulare *possibili spiegazioni* dell'evoluzione di strategie, norme e comportamenti che stanno alla base del contratto sociale. Di conseguenza, Skyrms non spiega "come effettivamente" si siano sviluppati certi comportamenti, ma si limita a mostrare "come sia possibile" che questi comportamenti abbiano potuto sopravvivere e prosperare, date determinate *condizioni iniziali* dei processi dinamici ipotizzati. Se, poi, la spiegazione possibile sia anche una *spiegazione effettiva*, o almeno *plausibile*, dipenderà dalla validità delle ipotesi circa le condizioni iniziali.

Evoluzione, algoritmi genetici e metanorme

Una grave limitazione degli esperimenti di simulazione di Axelrod (1984) e Skyrms (1996) è costituito dalla loro "impostazione ecologica", cioè dalla circostanza che – mentre si consente che nelle generazioni successive possano variare le frequenze con cui le diverse strategie sono inizialmente rappresentate nella popolazione – non si prevede la possibilità di introdurre *nuove* strategie. Sarebbe naturalmente di grande interesse passare da un'impostazione puramente ecologica a un'ottica evolutivista, che consentirebbe di introdurre nuove strategie nella popolazione (Axelrod, 1984, p. 49). In effetti, l'evoluzione biologica è in grado di fare proprio questo – attraverso la mutazione e altri meccanismi che forniscono "fonti di variabilità" per le strategie rappresentate nella popolazione. Viene naturale chiedersi se si possano escogitare esperimenti di simulazione in grado di rappresentare adeguatamente sia i meccanismi di riproduzione differenziale sia quelli volti ad assicurare fonti di variabilità nelle strategie impiegate.

Dopo il 1984 Axelrod ha escogitato un metodo per generare automaticamente nuove strategie, senza alcuna progettazione umana, e lo ha utilizzato per effettuare diversi esperimenti di simulazione, descritti in *The Complexity of Cooperation* (1997). Il metodo escogitato da Axelrod è costituito da un'appropriata versione dell'*algoritmo genetico* sviluppato da John Holland (1975).¹⁰ Gli algoritmi genetici sono stati sviluppati – in un'area di ricerche sull'intelligenza artificiale nota come *teoria della complessità* – con l'intento di riprodurre i meccanismi sui quali si basa il grande successo

10 Per un'esposizione di carattere introduttivo agli algoritmi genetici si vedano Holland (1992) e Riolo (1992).

dell'evoluzione biologica nello “scoprire” metodi complessi ed efficaci per l'adattamento a una grande varietà di situazioni ambientali. Come si è detto, l'evoluzione biologica utilizza sia meccanismi genetici di riproduzione differenziale – che consentono alla prole di ereditare le caratteristiche o strategie più riuscite dei genitori – sia meccanismi per la “scoperta” di nuove caratteristiche o strategie. Questi ultimi includono la riproduzione sessuale, attraverso la quale la prole eredita parte del materiale genetico da un genitore e parte dall'altro, e la mutazione. Gli algoritmi genetici devono il proprio nome al fatto di imitare la doppia funzione dei cromosomi nell'evoluzione biologica: da un lato, forniscono una rappresentazione di ciò che l'organismo diverrà; dall'altro, il loro materiale genetico è suscettibile di subire trasformazioni, così da trasmettere alla generazione successiva un nuovo materiale.

In particolare, l'algoritmo genetico utilizzato da Axelrod ha due tratti salienti: (a) è in grado di rappresentare qualsiasi strategia come una stringa di geni su un cromosoma che può subire trasformazioni genetiche; (b) può modellare tali trasformazioni mediante due “operatori genetici” – incrocio (o riproduzione sessuale) e mutazione – le cui caratteristiche sono formalmente identiche a quelle dei loro corrispondenti biologici. A partire da una popolazione iniziale ottenuta pescando a caso tra tutte le strategie possibili, un algoritmo genetico conduce, generazione dopo generazione, a popolazioni finali molto diverse da quella iniziale, non solo per la diversa frequenza delle strategie sopravvissute ma anche per la comparsa di nuove strategie. Dopo molte generazioni si ottiene spesso, come risultato della simulazione, una popolazione composta da individui che, nel dato ambiente strategico, ottengono mediamente un successo maggiore di quelli della popolazione iniziale.

Applicando per la prima volta gli algoritmi genetici alla teoria evuzionistica dei giochi, Axelrod (1997, cap. 1) effettua un esperimento di simulazione con il gioco ripetuto del prigioniero. A partire da una popolazione di strategie scelte in modo del tutto casuale entro l'insieme di tutte le possibili strategie – così da escludere ogni sbilanciamento iniziale in direzione della cooperazione o della reciprocità – la simulazione continua lasciando che la popolazione evolva in base all'algoritmo genetico. Sulla base dell'analisi di molti esperimenti di questo tipo, si giunge all'interessante conclusione che vi è una spiccata tendenza delle popolazioni a evolvere velocemente, da un punto di partenza del tutto casuale, verso strategie sostanzialmente simili a *T*. Questa conclusione mostra che la robustezza di *T* – e, più in generale, delle strategie improntate a reciprocità e tendenza alla cooperazione – è ancora maggiore di quanto si poteva sospettare.

Altri interessanti risultati vengono ottenuti da Axelrod (1997, cap. 3) analizzando la versione a n persone del dilemma del prigioniero. Una ben nota difficoltà di questa versione del dilemma è che, quando n è abbastanza grande, non si comprende bene in che modo la cooperazione possa svilupparsi sulla base di norme basate sulla reciprocità, dato che i giocatori che cooperano non hanno la possibilità di punire adeguatamente quelli che defezionano. Axelrod supera questa difficoltà esplorando un nuovo tipo di gioco a n persone dove, oltre alla cooperazione e alla defezione, viene data ai giocatori l'opzione aggiuntiva di individuare e punire sia coloro che sono stati visti defezionare sia chiunque altro eviti di punire i violatori. Si noti che, per evitare di rendere il problema banale, Axelrod assume che l'identificazione dell'impostore non è automatica e che la punizione ha un costo anche per chi la attua. Questo meccanismo addizionale viene denominato "metanorma": cioè "una norma per cui occorre punire coloro che non puniscono la defezione" dalla norma principale (Axelrod, 1997, p. 56).¹¹ Non riguarda, quindi, semplicemente l'imposizione della norma "primaria", ma considera questa stessa imposizione alla stregua di una norma e ne punisce la violazione. Secondo Axelrod, se la violazione di una norma viene percepita come particolarmente grave da un certo gruppo di individui, questi considereranno proporzionalmente grave anche il fatto che quella violazione venga tollerata. In altre parole, il grado di "vendicatività" (*vengefulness*) relativo alla violazione di una norma e quello relativo alla tolleranza di una violazione sarebbero correlati. Ammesso questo presupposto, Axelrod è in grado di mostrare che, dato un certo "grado di vendicatività" dei membri di una società, alcune norme sociali verranno "spontaneamente" rispettate e garantite, cioè che appropriate metanorme possono favorire l'insorgenza e la stabilizzazione di strategie di cooperazione.

6.4 Società artificiali, teorie ABM e complessità della cooperazione

Le simulazioni computerizzate di Axelrod e Skyrms permettono di scoprire il "risultato aggregato" di alcuni semplici schemi di comportamento individuale, rappresentati dalle strategie utilizzate da ciascuno dei membri di un determinata popolazione. Di conseguenza, permettono di spiegare l'origine, il funzionamento, l'evoluzione e l'eventuale declino di norme

11 La nozione di metanorma è molto simile a quella, introdotta e discussa da de Jasay (1991), di "convenzione satellite". Su questo punto, cfr. Cevolani (2008).

e istituzioni sociali sulla base delle interazioni fra individui e delle conseguenze, intenzionali e non, di tali interazioni.

L'idea che alcune semplici regolarità nelle interazioni tra i membri di un sistema sociale possano produrre complesse regolarità nel comportamento dell'intero sistema è tutt'altro che recente. Tuttavia, solo nell'ultimo trentennio questa idea è stata opportunamente formalizzata e incorporata nelle cosiddette *società artificiali*, cioè in programmi per la simulazione computerizzata dei sistemi sociali. Queste tecniche di simulazione sono spesso note come *agent-based models*, o con l'acronimo ABM; parleremo quindi di teorie ABM per riferirci alle teorie della complessità sociale formulate nei termini di *agent-based models*. Si noti che anche le teorie evoluzionistiche dei giochi basate su simulazioni computerizzate rientrano a pieno titolo nella più vasta classe delle teorie ABM. Le popolazioni studiate da Axelrod e Skyrms formano infatti società artificiali delle quali è possibile studiare diversi aspetti sulla base del comportamento dei singoli membri.

Uno dei primi tentativi di formulare una teoria ABM si deve a Schelling (1978), che impiegò i nuovi metodi di simulazione per affrontare il vecchio problema delle relazioni tra “micromotivazioni” individuali e “macrocomportamento” del sistema sociale. Schelling osserva che in molte situazioni gli obiettivi di un individuo sono strettamente dipendenti dal comportamento di altri individui che perseguono i loro obiettivi, o dalle caratteristiche di un ambiente formato da altri individui che perseguono i loro obiettivi: in situazioni di questo genere gli agenti esibiranno un *comportamento contingente*, cioè un comportamento che dipende da quello altrui. Di conseguenza ciascun agente, reagendo secondo certi schemi di comportamento alle caratteristiche del sistema cui appartiene, contribuirà a modificarlo. Questo meccanismo, ben noto agli economisti, condurrà in molti casi a “risultati aggregati” che nessun agente desidera e che nessun agente, a eccezione forse di qualche scienziato sociale, comprende di aver contribuito a determinare.

L'applicazione forse più nota delle teorie ABM è una simulazione ideata da Schelling (1978, pp. 137-155) per spiegare, sulla base del comportamento individuale contingente, il processo di segregazione a livello sociale. Il termine “segregazione” indica qualunque processo collettivo innescato da un comportamento individuale discriminatorio, cioè dalla tendenza a prendere alcune decisioni – dove abitare, dove sedersi, che lavoro scegliere o evitare, e così via – sulla base di determinate caratteristiche degli altri individui, come il sesso, l'età, la religione o il colore della pelle. Il modello di segregazione di Schelling è stato applicato, in primo luogo, alla segregazione spaziale, cioè ai processi che portano alla formazione di quartieri

abitati soltanto, o quasi, da individui dello stesso gruppo etnico. Quello che segue è un semplice esempio, tratto da Axelrod (1997, p. 215), del modo in cui tale modello viene applicato.

Possiamo rappresentare lo spazio residenziale come una scacchiera con sessantaquattro quadrati sui quali vengono collocati casualmente un certo numero di agenti Bianchi e Neri, avendo cura che ogni quadrato ospiti non più di un agente e che alcuni quadrati siano lasciati vuoti. I *vicini immediati* di un agente sono gli occupanti degli otto quadrati adiacenti al suo. Immaginiamo ora che gli agenti possano migrare sulla scacchiera in base a questa semplice regola di comportamento: un agente è soddisfatto della sua collocazione sulla scacchiera, e resta quindi nel suo quadrato, se almeno tre dei suoi vicini immediati sono del suo stesso colore; altrimenti l'agente non è soddisfatto e si sposta nel più vicino tra i quadrati liberi che gli permette di essere soddisfatto. Possiamo ora distribuire casualmente sulla scacchiera una quarantina di agenti Bianchi e Neri, applicare a tutti la regola di comportamento appena descritta e ripetere il processo una cinquantina di volte. Vedremo allora molto distintamente il manifestarsi della segregazione spaziale, cioè della tendenza degli agenti ad aggregarsi in zone cromaticamente omogenee. Le piccole dimensioni di questo specifico modello ci consentono di effettuare la simulazione anche "a mano". Tuttavia, non appena rendiamo il modello più realistico, considerando scacchiere con alcune migliaia di caselle e agenti, e numerose ripetizioni del processo, diventa indispensabile ricorrere a tecniche di simulazione computerizzata. Tali tecniche ci permetteranno di rispondere a interessanti interrogativi sulla struttura del processo di segregazione; possiamo chiederci, per esempio, se la tendenza alla segregazione sia più o meno veloce nel caso in cui uno dei due colori prevalga nettamente nella popolazione o se vi sia una più spiccata tendenza delle minoranze a compattarsi in zone cromaticamente omogenee.

Il modello di Schelling mostra con estrema chiarezza come un macrofenomeno sociale – cioè la segregazione – emerga spontaneamente a partire da alcune semplici strategie di comportamento degli individui coinvolti, senza che nessuno di essi lo abbia previsto o programmato. Questo genere di spiegazione è quello richiesto, nelle scienze sociali, dal cosiddetto principio dell'individualismo metodologico: cioè mostrare come le azioni individuali possano dar luogo a diversi tipi di macrofenomeni sociali che non erano stati progettati, voluti o previsti da nessuno, e che si presentano quindi come "ordini spontanei". I principi dell'individualismo metodologico sono certamente soddisfatti dalle teorie ABM e dalle diverse versioni delle teorie dei giochi che abbiamo considerato finora. Le ipotesi fondamentali di questi teorie, infatti, riguardano le credenze, le preferenze, le azioni e le

interazioni individuali; tali ipotesi vengono poi usate per spiegare svariati tipi di macrofenomeni sociali. Come vedremo più ampiamente nel prossimo capitolo, spiegazioni di questo genere sollevano interessanti interrogativi metodologici, cioè relativi al modo in cui le scienze sociali studiano il comportamento strategico di individui in società e le sue conseguenze.

7. TEORIA DEI GIOCHI, SCUOLA AUSTRIACA E METODO DELLE SCIENZE SOCIALI

L'approccio giochistico all'analisi del comportamento umano – illustrato nei capitoli precedenti in relazione ad alcuni dei principali problemi delle scienze sociali – solleva diversi interessanti interrogativi metodologici, cioè relativi ai presupposti, gli strumenti e i fini coi quali lo studioso della società affronta il proprio oggetto d'indagine. La teoria dei giochi, soprattutto nelle sue varianti più recenti, sembra poter fornire un solido quadro teorico per tutte le scienze sociali (Gintis 2009; Harsanyi 1968), che permette di integrare ed applicare coerentemente i risultati dei più tradizionali studi storici, antropologici ed evolucionistici sull'origine e lo sviluppo delle preferenze e delle tendenze comportamentali dell'uomo (Bowles e Gintis 2011; Ridley 1996; Rubin 2002; Friedman 2008). Una descrizione e valutazione dettagliata degli aspetti metodologici dell'approccio giochistico allo studio della società esula dai fini del presente lavoro – e sarebbe peraltro prematura, considerato lo stato ancora magmatico e in piena evoluzione delle ricerche. Ci sembra però possibile offrire un primo bilancio dei principali risvolti metodologici dell'applicazione della teoria dei giochi alle scienze sociali, che costituisce il contenuto di questo capitolo.

Piuttosto sorprendentemente, da questo bilancio emerge una notevole affinità metodologica fra l'approccio giochistico e uno dei più noti, anche se meno frequentati, approcci teorici alle scienze sociali, quello della cosiddetta Scuola Austriaca di economia.¹ Un confronto fra questi due approcci appare quanto meno azzardato alla luce del fatto che gli "austriaci" – dal fondatore della Scuola, Carl Menger (1840-1921), ai suoi più importanti esponenti nel Novecento, come Ludwig von Mises (1881-1973), Friedrich August von Hayek (1899-1992) e Israel Kirzner (nato nel 1930) – si sono sempre distinti per la loro avversione sia all'approccio matematico e formale ai problemi delle scienze sociali, sia all'uso di metodi empirici o

¹ Per una introduzione ai problemi metodologici ed epistemologici delle scienze sociali da una prospettiva austriaca, si vedano Di Nuoscio (2006) e Infantino (1995).

sperimentali. In particolare, secondo la Scuola Austriaca, l'economia non è una scienza sperimentale e le teorie economiche non sono empiricamente controllabili alla stregua di quelle delle scienze naturali. Per questo motivo, gli austriaci difendono una visione rigorosamente anti-positivista e anti-riduzionista delle scienze sociali, rifiutando quello che Hayek (1952) ha chiamato “scientismo”, cioè l'applicazione dei metodi delle scienze naturali all'analisi dei fenomeni sociali.²

A prima vista, l'approccio giochistico – basato su una teoria matematica e sempre più integrato con lo studio empirico e sperimentale della scelta e del comportamento umano – sembrerebbe quanto mai lontano dall'approccio austriaco. Nel seguito, mostriamo invece come questo divario sia più apparente che reale, e come questi due approcci siano accomunati da profonde affinità che rendono non solo compatibili ma potenzialmente integrabili le analisi di diversi fenomeni sociali condotte nell'ambito dei due programmi di ricerca. Non a caso, recentemente diversi studiosi – compreso per esempio il premio Nobel e “padre” della cosiddetta economia sperimentale Vernon Smith (nato nel 1927) – hanno sottolineato la rilevanza dei contributi, non solo metodologici, degli austriaci per gli ultimi sviluppi delle scienze sociali.³ Dopo aver messo in luce le origini “austriache” della teoria dei giochi (paragrafo 7.1), discutiamo gli importanti punti di contatto fra l'approccio giochistico e la concezione hayekiana dell'evoluzione culturale (paragrafo 7.2). Tale confronto permette di identificare alcuni essenziali aspetti metodologici del tipo di spiegazione offerta dall'approccio giochistico dei fenomeni sociali complessi (paragrafo 7.3) e di illustrare alcune possibili applicazioni di tale approccio alla filosofia politica (paragrafo 7.4).

2 Dal punto di vista della teoria economica “ortodossa” (o *mainstream*), e in particolare neoclassica, questi presupposti metodologici confinano la Scuola Austriaca fra gli approcci cosiddetti “eterodossi”, o almeno alla periferia della disciplina (cfr. Hausman 2012, par. 1.2). Per una breve rassegna aggiornata della storia e dei principali risultati della Scuola Austriaca si veda White (2008).

3 Si vedano in particolare le riflessioni di Smith (1999, 2005) dedicate all'opera di Mises e di Hayek, e l'impianto interamente hayekiano del suo ultimo libro (Smith 2008). Sui rapporti fra la metodologia austriaca, l'approccio giochistico e altri approcci più o meno “eterodossi” all'economia e alle scienze sociali (come l'economia sperimentale e comportamentale) si vedano Foss (2000), Festa (1999, 2001, 2007, 2012), Holcombe (2009), Oprea e Powell (2010), Seagren (2011) e Cevolani (2011). Sulla metodologia dell'economia sperimentale, si veda Guala (2005).

7.1 Le origini “austriache” della teoria dei giochi

L'influenza della Scuola Austriaca sulla genesi della teoria dei giochi – anche se poco appariscente e largamente ignorata – è stata tutt'altro che trascurabile. A questo riguardo, Kelly (2009) ha ricostruito l'interessante legame fra i primi teorici dei giochi e l'insegnamento degli austriaci e, in particolare, di Mises. In particolare, Mises esercitò una notevole influenza su Morgenstern, che apparteneva alla stessa generazione di Hayek. Morgenstern era stato studente di Mises e ne aveva frequentato il seminario privato nel 1924, succedendo poi ad Hayek, nel 1931, come direttore dell'Istituto per le Ricerche sul Ciclo Economico di Vienna. Nel volume *The Limits of Economics* (1937), scritto mentre svolgeva questa funzione, Morgenstern abbracciò una concezione misesiana dell'economia. Tale concezione dovette esercitare una certa influenza sulle prime formulazioni della teoria dei giochi. Infatti, mentre il merito dei risultati matematici contenuti in *Theory of Games and Economic Behaviour* (1944) va attribuito a von Neumann, fu certamente Morgenstern a fornire la teoria economica usata per la formulazione dei problemi affrontati nel volume. Un altro esempio dell'influenza austriaca sullo sviluppo della teoria dei giochi – ancora ricostruito da Kelly (*ibidem*, p. 38-40) – è la connessione tra Mises e Nash, attraverso Bert Hozelitz (1913-1995), uno dei primi studenti di Mises. Infatti, nel 1947, Hozelitz tenne un corso di Economia internazionale al Carnegie Tech. Uno dei suoi studenti era appunto Nash; in effetti, quello di Hozelitz fu l'unico corso di economia frequentato da Nash. Dopo la vittoria del Nobel, Nash riconobbe che il processo di pensiero che lo condusse a formulare la nozione di equilibrio era stato probabilmente influenzato dalla Scuola Austriaca, attraverso l'insegnamento di Hozelitz.

Come vedremo nel corso di questo capitolo, diversi programmi di ricerca sviluppati nell'ambito della Scuola Austriaca presentano notevoli affinità concettuali con la teoria dei giochi. Per esempio, alcuni recenti sviluppi della teoria sono ispirati da idee che appaiono molto vicine a quelle elaborate dagli austriaci – e, in particolare, da Hayek – nelle loro ricerche sulla coordinazione delle credenze e dei piani d'azione individuali nell'ambito dei processi di mercato. Come osserva Foss (2000), questi problemi furono affrontati anche da Morgenstern in diversi lavori pubblicati prima del suo incontro con von Neumann. In uno di questi lavori, Morgenstern (1935, pp. 58-59) illustra il problema della coordinazione dei piani d'azione individuali con la parabola, divenuta poi famosa, di Holmes e Moriarty:

Quando Sherlock Holmes è inseguito dal suo nemico Moriarty, parte da Londra per Dover con un treno che ferma a una stazione intermedia e, anziché

continuare il suo viaggio fino a Dover, scende alla stazione intermedia. Egli ha visto Moriarty alla stazione, lo ritiene molto intelligente e si aspetta che Moriarty prenda un treno speciale più veloce per attenderlo a Dover. Questa supposizione di Holmes risulta esatta. Ma che cosa sarebbe avvenuto se Moriarty fosse stato ancor più intelligente, avesse valutato in modo superiore le capacità intellettuali di Holmes e ne avesse in tal modo previsto l'azione? Sarebbe palesemente andato alla stazione intermedia. Holmes avrebbe a sua volta potuto calcolare ciò e avrebbe quindi dovuto decidersi per Dover. Alla qual cosa Moriarty avrebbe di nuovo "reagito" diversamente. Con maggior riflessione essi non avrebbero fatto assolutamente nulla oppure quello intellettualmente meno dotato avrebbe dovuto consegnarsi all'altro già alla Victoria Station: l'intera fuga sarebbe infatti divenuta inutile.

La parabola offre una vivida raffigurazione delle enormi difficoltà che, in molte interazioni sociali, ostacolano il tentativo di raggiungere un equilibrio fondato sulle previsioni del comportamento altrui. Morgenstern (1935) mostra piena consapevolezza che interazioni non molto dissimili da quella in cui si trovarono coinvolti Holmes e Moriarty – cioè interazioni caratterizzate dalla differenza delle credenze e dall'imperfezione delle previsioni degli agenti – si presentano frequentemente nelle attività economiche. Occorre ammettere che, dell'enfasi di Morgenstern sulla difficoltà di raggiungere equilibri fondati sulla coordinazione di credenze, resta ben poco nel volume del 1944 in collaborazione con von Neumann; qui, infatti, gli autori si occupano quasi esclusivamente di situazioni statiche, caratterizzate da credenze omogenee e previsioni perfette (cfr. Foss, 2000, p. 44). Tuttavia, piuttosto sorprendentemente, la parabola "austriaca" di Holmes e Moriarty anticipa i problemi – affrontati recentemente nell'ambito delle teorie cognitive dei giochi – relativi ai paradossi della conoscenza comune e, più in generale, al ruolo delle credenze degli agenti nelle interazioni sociali. Sembra quindi ragionevole riconoscere con Foss (2000, p. 42) che "le frontiere della teoria dei giochi contemporanea si sono considerevolmente estese all'analisi di problemi che sono da lungo tempo centrali per l'economia austriaca", e condividere, quindi, la sua conclusione che "una più stretta interazione tra gli economisti giochistici e gli austriaci farebbe probabilmente bene non solo agli austriaci, ma anche agli studiosi di teoria dei giochi" (Foss, 2000, p. 53).

L'interazione auspicata da Foss fra approccio austriaco e approccio giochistico non si è finora tradotta in pratica. Fin dalle origini, in effetti, e nonostante i punti di contatto che abbiamo appena discusso, il volume di von Neumann e Morgenstern sulla teoria dei giochi suscitò scarso interesse negli austriaci. Una delle poche osservazioni, di tono fortemente critico, sulla

proposta di applicare la teoria dei giochi all'analisi dei fenomeni sociali ed economici, si deve a Ludwig von Mises (1949, pp. 112-113):

Lo scopo immediato di un gioco è di sconfiggere l'avversario secondo le regole del gioco stesso. Questo è un caso speciale e peculiare dell'azione. La maggior parte delle azioni non tendono a sconfiggere o a far perdere nessuno. Esse tendono a migliorare le condizioni. Può accadere che questo miglioramento sia ottenuto a spese di altri. Ma certamente non è sempre il caso. Esso non è, per dirlo in termini piani, certamente il caso nell'ambito del funzionamento regolare di un sistema sociale basato sulla divisione del lavoro. Non c'è la minima analogia tra il gioco e la condotta degli affari entro una società di mercato. [...] Colui che desidera interpretare la condotta degli affari come una mariuoleria è su una cattiva strada.

Tratto caratteristico dei giochi è l'antagonismo tra due giocatori o gruppi di giocatori. Tratto caratteristico degli affari entro la società, cioè entro un ordine basato sulla divisione del lavoro, è la concordia degli sforzi dei suoi membri. [...] Nell'ambito di un'economia di mercato la concorrenza non implica antagonismo nel senso in cui questo termine è applicato all'urto ostile di interessi incompatibili. [Si deve quindi concludere] che è erroneo applicare la terminologia del mutuo sterminio ai problemi della mutua cooperazione nell'ambito della società. I termini militari sono inappropriati alla descrizione delle funzioni economiche.

Successivamente Mises (1962, pp. 79-81) ribadì, in termini ancora più netti, la sua ostilità all'uso della teoria dei giochi nelle scienze sociali:

Sarebbe difficile fraintendere in modo più radicale l'essenza della cooperazione sociale e lo sforzo economico dell'umanità civilizzata di quanto si farebbe considerandola come un combattimento o come la duplicazione ludica di un combattimento, cioè come un gioco. Nella cooperazione sociale ognuno, nel servire i propri interessi, serve gli interessi del suo prossimo. Spinto dal desiderio di migliorare le proprie condizioni, egli migliora anche le condizioni degli altri. Il fornaio non danneggia quelli per cui sforna il pane; li aiuta. Tutti sarebbero danneggiati se il fornaio smettesse di produrre il pane e il medico non curasse più i malati. Il produttore di scarpe non ricorre a una "strategia" al fine di sconfiggere i suoi clienti rifornendoli di scarpe. La competizione sul mercato non deve essere confusa con la spietata competizione biologica prevalente tra animali e piante o con le guerre tuttora combattute tra nazioni – sfortunatamente non ancora completamente – civilizzate.

[Alla luce della fallace visione degli affari economici suggerita dalla teoria dei giochi] la preoccupazione del produttore è quella di danneggiare il pubblico. La sua abilità è la strategia, per così dire, l'arte di infliggere quanto male possibile al nemico. Gli avversari di cui egli pianifica la rovina sono i suoi potenziali consumatori così come i suoi competitori [...]. Il metodo più appropriato di investigare scientificamente le attività commerciali e il processo

di mercato, si dice, è di analizzare il comportamento e la strategia delle persone impegnati nei giochi. [...] In un gioco ci sono vincitori e perdenti. Ma lo scambio commerciale è sempre vantaggioso per entrambe le parti. Se sia il venditore che il compratore non considerassero la transazione come l'azione più vantaggiosa che essi possano scegliere nelle condizioni date, non porterebbero a termine l'affare.

All'epoca in cui venne espresso, lo scetticismo di Mises sull'opportunità di applicare la teoria dei giochi nelle scienze economiche e sociali era del tutto legittimo. In quel periodo, infatti, l'attenzione dei teorici dei giochi era prevalentemente diretta all'analisi dei giochi a somma zero, caratterizzati da un radicale conflitto di interessi tra i giocatori. Successivamente, a partire dalla fine degli anni cinquanta, i teorici dei giochi si occuparono sempre più approfonditamente dei giochi cooperativi e a motivazione mista, i quali sembrano in grado di fornire un adeguato modello per molti tipi di interazioni economiche e sociali. Così, Mises ha tutte le ragioni per rifiutare l'idea che lo scambio tra venditore e acquirente sia un gioco a somma zero, nel quale uno dei due guadagna quello che l'altro perde. Tuttavia, l'affermazione misesiana che entrambe le parti traggono beneficio da una transazione può essere adeguatamente espressa in termini giochistici, considerando la transazione come un caso particolare di gioco cooperativo.

Nonostante gli sviluppi recenti della teoria dei giochi abbiano mostrato sempre maggiori affinità con alcune posizioni della Scuola austriaca, piuttosto sorprendentemente l'ostilità degli austriaci verso la teoria dei giochi è rimasta sostanzialmente inalterata. Ci pare che tale ostilità sia ormai priva di solide motivazioni e che, al contrario, vi siano buone ragioni per indurre gli studiosi che operano nella tradizione austriaca a interessarsi alla teoria dei giochi e impegnarsi nel tentativo di applicare i suoi potenti strumenti formali nello sviluppo dei loro programmi di ricerca. Nei prossimi paragrafi, mostriamo come una feconda interazione fra l'approccio giochistico e quello austriaco possa fondarsi su alcune solide affinità metodologiche fra questi due programmi di ricerca.

7.2 Teoria dei giochi e concezione hayekiana dell'evoluzione culturale

Nel capitolo 6, abbiamo visto come i teorici dei giochi hanno affrontato il problema della genesi e dell'evoluzione delle norme sociali, concentrandosi su quelle norme di coordinazione e cooperazione che contribuiscono alla formazione di una grande varietà di ordini sociali spontanei. Un approccio molto noto all'analisi delle norme sociali e degli ordini spontanei è

la concezione dell'evoluzione culturale proposta da Hayek. Tra l'approccio hayekiano e quello giochistico vi sono tali differenze di stile di ricerca e linguaggio da far apparire azzardato ogni tentativo di confronto. Tuttavia, a dispetto delle apparenze, ci sembra che i due approcci presentino significativi punti di contatto e che interpretare e sviluppare le tesi di Hayek sull'evoluzione culturale alla luce della teoria dei giochi possa dischiudere feconde direzioni di ricerca. In questo paragrafo, richiamiamo alcune tesi teoriche e metodologiche che caratterizzano la concezione hayekiana e l'analisi giochistica dell'evoluzione culturale, per poi sottolineare alcuni rilevanti punti di contatto tra questi due approcci.

Norme sociali e ordine spontaneo nella concezione hayekiana dell'evoluzione culturale

Il nocciolo della concezione hayekiana dell'evoluzione culturale ci pare costituito da tre tesi *teoriche* circa la natura e l'evoluzione delle norme sociali – tesi che verranno indicate con “TH.1-TH.3”, dove “T” e “H” stanno per “teorico” e “Hayek”. In stretta connessione con le ricerche effettuate in qualità di scienziato sociale, Hayek affronta anche i problemi metodologici posti da tali ricerche, sviluppando un orientamento che può venire compendiato in tre tesi *metodologiche*, che verranno indicate con “MH.1-MH.3”, dove “M” sta per “metodologico”.⁴

Le tesi teoriche TH.1-TH.3 possono venire così formulate.

TH.1. CARATTERE TACITO E INCONSAPEVOLE DELLE NORME SOCIALI. La maggior parte delle *regole di condotta* individuale sono date da norme sociali, cioè da regole di condotta la cui violazione viene sanzionata socialmente. In genere gli agenti non sono in grado di articolare verbalmente le norme alle quali si attengono e, molto spesso, non sono neppure pienamente consapevoli di attenersi a tali norme. A tale proposito Hayek (1973-1979, p. 534; cfr. anche pp. 27, 96, 100 e 208) osserva che

l'uomo ha certamente più spesso imparato a fare le cose giuste senza capire perché lo erano, ed è ancora oggi meglio servito dalle sue abitudini che dalla sua capacità di comprensione. [...] Fu un repertorio di regole apprese che disse

4 Sulla concezione hayekiana dell'evoluzione culturale si vedano Hayek (1942-1943; 1967, specialmente capp. 1, 2 e 4; 1973-1979, specialmente vol. 1). Queste ricerche hanno un importante punto d'avvio in *Scientism and the Study of Society* (1942-1943) e trovano una magistrale sintesi nel primo dei tre volumi di cui si compone *Law, Legislation and Liberty* (1973-1979), significativamente intitolato *Rules and Order*.

all'uomo quale era il modo giusto o sbagliato di agire in diverse circostanze, e che gli diede una crescente capacità di adattarsi a condizioni mutevoli – e in particolare di cooperare con gli altri membri del suo gruppo. [...] *Vi era allora probabilmente molta più "intelligenza" incorporata nel sistema di regole di comportamento, di quanta ve ne fosse nei pensieri dell'uomo su quanto lo circondava.*

Ne consegue che, di solito, il comportamento conforme a norme sociali non richiede particolari calcoli o riflessioni circa le loro possibili conseguenze. Ciò significa che, contrariamente a un'opinione diffusa, le regole della razionalità strumentale – cioè i principi generali applicabili al raggiungimento di un fine, sulla base delle preferenze e delle credenze individuali – svolgono un ruolo piuttosto modesto nelle interazioni sociali.⁵

TH.2. EVOLUZIONE DELLE NORME SOCIALI. La maggior parte delle norme sociali sono il risultato di un processo evolutivo fondato su meccanismi di selezione di gruppo che favoriscono i gruppi di individui che usano tali norme.⁶

TH.3. AZIONI INDIVIDUALI, NORME SOCIALI E ORDINI SPONTANEI. Alcuni sistemi di semplici norme sociali possono dare luogo a ordini spontanei caratterizzati da un elevato grado di complessità. A loro volta, tali ordini spontanei forniscono l'ambiente per l'evoluzione delle norme sociali in uso e l'emergere di nuove norme, o regole di condotta. Il successo evolutivo di una norma dipende dall'efficienza dell'ordine spontaneo che questa – in combinazione con l'insieme delle altre norme adottate in una determinata società – contribuisce a creare.⁷ Secondo Hayek (1967, p. 150), infatti, si può mostrare che

i sistemi di regole di condotta si sviluppano come degli insiemi, ovvero che il processo di selezione evolutivistica opererà sull'ordine inteso come un tutto; e il fatto che una nuova regola, combinata con tutte le altre regole del gruppo, nell'ambiente particolare in cui si trova, aumenti o diminuisca l'efficienza dell'intero gruppo, dipenderà dall'ordine cui porta la condotta individuale. Una conseguenza di ciò è che una nuova regola di condotta individuale, che in una situazione può essere di detrimento, in un'altra può dimostrarsi benefica.

5 Sulla natura delle regole di condotta abitualmente seguite nell'interazione sociale si veda Hayek (1973-1979, specialmente pp. 17-18 e 98-104).

6 Cfr. Hayek (1967, specialmente cap. 4, significativamente intitolato, "Note sulla evoluzione dei sistemi delle regole di condotta (L'interazione tra regole di condotta individuale e l'ordine sociale delle azioni)". Si veda anche Hayek (1973-1979, specialmente pp. 15, 25-28, 32-37, 96, 104 e 190).

7 Cfr. Hayek (1967, pp. 146, 153 e 160-161; 1973-1979, pp. 51-52, 58-61 e 97).

Veniamo ora alle tesi metodologiche MH.1-MH.3.

MH.1. INDIVIDUALISMO METODOLOGICO E SPIEGAZIONE DELL'ORDINE SOCIALE. Le scienze sociali debbono spiegare i processi attraverso i quali nascono, evolvono e decadono le norme sociali a partire dall'esame delle azioni degli individui che compongono una determinata società e delle conseguenze, intenzionali e non, di tali azioni.⁸ In altre parole, il fine delle scienze sociali è quello di "spiegare gli effetti involontari e imprevisi" delle azioni umane (Hayek, 1952, p. 26) e, in particolare, gli *ordini spontanei* che frequentemente emergono come risultato non intenzionale delle azioni di molti individui che interagiscono tra loro sulla base di determinate regole di condotta. A tale proposito Hayek (1952, p. 43) osserva che "[è] solo nella misura in cui un certo tipo di ordine emerge come risultato dell'azione dei singoli, ma senza essere stato da alcuno di essi consciamente perseguito, che si pone il problema di una [...] spiegazione teorica [dei fenomeni sociali]".

MH.2. PECULIARITÀ METODOLOGICA DELLE SCIENZE SOCIALI: SPIEGAZIONI DI PRINCIPIO. Le scienze sociali non possono usare gli stessi metodi delle scienze naturali. Infatti, esse "hanno tratti metodologici specifici di carattere permanente" (Hayek, 1967, p. 67), adeguati al peculiare carattere del loro oggetto e, più precisamente, alla grande complessità degli ordini spontanei che derivano dalle interazioni di individui operanti in base a determinati sistemi di regole di condotta (Cfr. Hayek 1967, p. 151). In particolare, le teorie delle scienze sociali non mirano a "predizioni specifiche di eventi particolari", ma solo a scoprire "se i fattori che abbiamo scelto siano realmente presenti nei fenomeni particolari che vogliamo spiegare, e se siano rilevanti e sufficienti a spiegare ciò che osserviamo" (*ibidem*, p. 54), almeno nei suoi aspetti qualitativi. Ciò significa che la spiegazione di un fenomeno sociale potrà dirci soltanto "quali *tipi* di eventi aspettarci e quali no" (*ibidem*, p. 53). Spiegazioni di questo genere rappresentano "storie congetturali" (*ibidem*, p. 157) e vanno intese come "spiegazioni di principio", ottenute attraverso "la riproduzione di un principio su modelli grandemente semplificati" (*ibidem*, p. 70). Lo scopo di una storia congetturale è di carattere essenzialmente teorico, piuttosto che storico; infatti "la storia congetturale è la ricostruzione di un tipo ipotetico di processo, che potrebbe non essere mai stato osservato, ma che, se fosse avvenuto, avrebbe prodotto fenomeni del tipo che osserviamo" (Hayek, 1967, p. 157). Poiché "il valore principale di *ogni* scienza consiste nel dirci quali sarebbero le conseguenze se le cose fossero, per qualche aspetto, rese differenti da quel-

8 Per un'analisi dell'individualismo metodologico di Hayek e della Scuola Austriaca, si veda Nozick (1997, cap. 4). Si veda anche Antiseri e Pellicani (1992).

le che sono” (Hayek, 1973-1979, p. 24), i teorici sociali devono accettare l’idea che “[u]na scienza sociale feconda deve essere in grandissima parte uno studio di ciò che *non* è: una costruzione di modelli ipotetici di mondi possibili che potrebbero esistere se mutassero alcune delle caratteristiche modificabili” (*ivi*). Vale la pena notare, inoltre, che secondo Hayek le spiegazioni di principio sono appropriate non solo per gli ordini sociali, bensì per tutti i fenomeni complessi, compresi quelli naturali. In particolare, egli ritiene che le spiegazioni di principio della teoria sociale abbiano caratteri estremamente simili a quelle elaborate nella teoria dell’evoluzione (cfr. Hayek, 1967, pp. 54-55 e 85-87).

MH.3. PRESUPPOSTI TEORICI E DEDUZIONE NELLE SPIEGAZIONI DI PRINCIPIO. La spiegazione di principio di un determinato fenomeno sociale – ottenuta riproducendo determinati principi, o presupposti teorici, su un modello altamente semplificato – consiste nella *deduzione*, a partire da quei presupposti teorici, degli aspetti qualitativi del fenomeno. Ciò significa che occorre dedurre, a partire da certi presupposti teorici circa le regole di condotta individuale seguite dai membri di una società, le caratteristiche qualitative dell’ordine spontaneo che si sviluppa sulla base di tali regole. Tale compito comporta notevoli difficoltà: per esempio, quella di identificare il “giusto” insieme di presupposti teorici che potrebbero consentirci di dedurre gli aspetti qualitativi di un fenomeno sociale. Qui “la questione importante non è se le ipotesi o le leggi usate per la spiegazione dei fenomeni siano vere, bensì se abbiamo selezionato, dal nostro bagaglio di teorie accettate, le ipotesi appropriate e se le abbiamo combinate nel modo giusto” (Hayek, 1967, pp. 47-48). A proposito della difficoltà di fornire adeguate spiegazioni deduttive dei fenomeni sociali Hayek (1967, p. 47) osserva che

[n]essuno può [...] trarre tutte le conseguenze estraibili dalla nostra conoscenza, o estraibili addirittura da alcune delle più banali e incontestate proposizioni che usiamo nella vita quotidiana; è spesso sempre un compito eccessivamente difficile decidere quanto di ciò che osserviamo può essere spiegato con leggi già conosciute, o potrebbe essere spiegato se possedessimo tutti i dati rilevanti. Trarre più conclusioni significative possibili da tutto ciò che già sappiamo non è certo un compito puramente deduttivo: nella scelta dei problemi, esso deve essere guidato dall’osservazione. Ma, nonostante sia l’osservazione a sollevare i problemi, la risposta riposa solo nella deduzione.

Norme sociali e ordine spontaneo nelle teorie dei giochi

I risultati ottenuti, da Axelrod, Skyrms e altri, nell’ambito delle teorie evuzionistiche e adattive dei giochi suggeriscono tre tesi *teoriche* che

sottendono l'analisi giochistica delle norme sociali, che indicheremo con "TG.1-TG.3", dove "T" e "G" stanno per "teorico" e "giochistico". Tali tesi possono venire così formulate.

TG.1. CARATTERE TACITO E INCONSAPEVOLE DELLE NORME SOCIALI. Le norme sociali possono venire intese come schemi di comportamento adottati da determinati gruppi della popolazione. Di solito tali schemi di comportamento sono privi di attributi quali la razionalità, l'intenzionalità, o anche solo la consapevolezza.

TG.2. EVOLUZIONE DELLE NORME SOCIALI. L'evoluzione di determinate norme sociali può venire spiegata ipotizzando che il gruppo che ha adottato tali norme abbia avuto buoni risultati in confronto a quelli ottenuti dalla media della popolazione e che tali norme si siano poi diffuse attraverso l'imitazione da parte di membri della popolazione che avevano adottato strategie di minore successo.

TG.3. AZIONI INDIVIDUALI, NORME SOCIALI E ORDINI SPONTANEI. Le azioni dei membri di una determinata società, effettuate in conformità a certe norme sociali, portano alla formazione di ordini spontanei che non erano stati né previsti né progettati da nessuno (cfr. Bicchieri (1993, p. 7) e Bicchieri e Muldoon 2011).

Veniamo ora alle quattro tesi che compendiano l'orientamento metodologico emerso in stretta connessione con l'analisi giochistica dell'evoluzione culturale – tesi che verranno indicate con "MG.1-MG.4", dove "M" e "G" stanno per "metodologico" e "giochistico". Tali tesi possono venire formulate come segue.

MG.1. INDIVIDUALISMO METODOLOGICO E SPIEGAZIONE DELL'ORDINE SOCIALE. L'approccio giochistico alle norme sociali soddisfa il principio dell'individualismo metodologico, secondo il quale "tutte le teorie vere della scienza sociale sono *riducibili* a teorie dell'azione umana individuale, con l'aggiunta di condizioni di contorno che specificino le condizioni in cui una persona agisce" (Nozick, 1997, p. 132). In particolare, l'analisi giochistica dell'evoluzione culturale spiega la formazione di ordini spontanei a partire dalle azioni individuali dei membri di una determinata società, cioè "come azioni indipendenti di singoli individui possono generare un "ordine" che non era stato previsto né progettato" e "come le azioni individuali si coordinano e producono esiti collettivi che non sono né programmati né intenzionalmente perseguiti dagli individui" (Bicchieri, 1993, p. 7). Diversi studiosi impegnati nello sviluppo dell'approccio giochistico hanno messo in rilievo la sua conformità al principio dell'individualismo metodologico. Per esempio, Axelrod (1984, p. 13) afferma che il suo approccio consiste nel "partire da talune ipotesi circa le motivazioni individuali per dedurne

le conseguenze sul comportamento dell'intero sistema". Si vedano anche Bicchieri (1993, pp. 15-23 e cap. 6.7) e Ullmann-Margalit (1977, p. 11).

MG.2. PECULIARITÀ METODOLOGICA DELLE SCIENZE SOCIALI: SPIEGAZIONI SPECULATIVE. Di solito le analisi giochistiche dell'evoluzione culturale non sono il frutto di ricerche empiriche, ma vengono effettuate nell'ambito della cosiddetta "sociologia speculativa" (Ullmann-Margalit, 1977, p. vii; 1978). La *spiegazione speculativa* di un evento sociale può essere intesa come "una descrizione degli aspetti essenziali delle situazioni in cui un *tale* evento *potrebbe* verificarsi: è una storia di come qualcosa potrebbe accadere [...] non di ciò che effettivamente è accaduto" (*ibidem*, p. 1). La forma di conoscenza che possiamo attenderci da questo genere di spiegazioni è quella che, nella tradizione del neoempirismo, va sotto i nomi di *esplicazione* o *ricostruzione razionale* (Carnap (1962, capitolo 1 e pp. 576-577). Oltre a questa fondamentale differenza tra le ricostruzioni razionali degli eventi sociali e le teorie della sociologia empirica, un'altra differenza è che le ricostruzioni razionali sono motivate dall'intento di "fornire una migliore comprensione di un corpo di credenze ricevute" (*ivi*) e, a tale scopo, fanno uso *solo* di assunzioni ampiamente accettate, o di senso comune.

MG.3. PRESUPPOSTI TEORICI E SIMULAZIONE NELLE SPIEGAZIONI SPECULATIVE. Nella ricerca di adeguate spiegazioni speculative dei fatti sociali il ricorso ai tradizionali strumenti della chiarificazione concettuale, della deduzione e dell'induzione si rivela insufficiente. Infatti, alcune spiegazioni speculative dell'evoluzione delle norme sociali possono essere elaborate solo sulla base di appropriati metodi di simulazione computerizzata, cioè i metodi ABM. Axelrod (1997, p.3) osserva che ABM è "un terzo modo di fare scienza" che si differenzia sia dalla deduzione che dall'induzione (cfr. Grüne-Yanoff e Weirich, 2010). ABM condivide con la deduzione il fatto di partire da un insieme di espliciti presupposti teorici. Tuttavia, nonostante la grande semplicità dei presupposti adottati in molte applicazioni di ABM, risulta in generale impossibile dedurre conseguenze interessanti. Ciò dipende dal fatto che tali presupposti riguardano le caratteristiche degli agenti e delle loro interazioni, mentre le conseguenze alle quali siamo interessati vertono sugli effetti di larga scala delle interazioni. Si tratta di effetti quasi sempre sorprendenti e difficili da anticipare, poiché di solito è impossibile dedurli, o calcolarli, mediante gli abituali procedimenti logico-matematici.

L'uso di metodi di simulazione appare, quindi, non semplicemente utile, bensì indispensabile. Come osserva Axelrod (*ibidem*, p. 3), "il fine di ABM è di aiutare l'intuizione" a comprendere alcuni fondamentali *aspetti qualitativi* di certi tipi di processi sociali. Fornire un'accurata rappresentazione quantitativa di particolari fenomeni sociali non rientra, invece, tra gli scopi delle spiegazioni speculative, e può persino rivelarsi controproducente. Per

esempio, l'utilità delle simulazioni con il gioco ripetuto del prigioniero non dipende dall'accuratezza del modello, ma dall'estrema semplicità dei suoi presupposti che, proprio per questo, riescono a catturare fondamentali aspetti qualitativi di una grande varietà di interazioni sociali. Le conseguenze dei presupposti di partenza, trovate attraverso ABM, non vengono espresse da teoremi provati deduttivamente, ma sono costituite da "dati simulati", i quali vengono poi analizzati con le consuete procedure statistico-induttive. Questo significa che ABM include un fondamentale ingrediente induttivo. D'altra parte, tale ingrediente è ben diverso dai tradizionali metodi induttivi usati nella sociologia empirica per il fatto che le procedure induttive vengono applicate a dati ottenuti attraverso la simulazione piuttosto che attraverso l'osservazione diretta del mondo reale.

MG.4. ROBUSTEZZA DELLE SPIEGAZIONI SPECULATIVE. Per valutare l'adeguatezza delle spiegazioni speculative elaborate nell'ambito dell'approccio giochistico sono stati proposti alcuni criteri quali la *robustezza* (cfr. D'Arms *et al.* 1998). Si considerino, per esempio, le spiegazioni speculative proposte da Axelrod e Skyrms. Tali spiegazioni non mirano a spiegare come si siano *effettivamente* sviluppate determinate norme di cooperazione e di equità. Il loro scopo, invece, è quello di mostrare *come sia possibile* che tali norme abbiano potuto sopravvivere e prosperare, dati determinati tipi di condizioni iniziali. In altre parole, le spiegazioni speculative intendono fornire una *spiegazione possibile*, o *potenziale*, piuttosto che una *spiegazione effettiva*, dell'evoluzione delle norme di cooperazione ed equità. Poiché, di solito, non si riesce a stabilire se le condizioni iniziali ipotizzate nelle spiegazioni speculative si sono effettivamente verificate, occorre richiedere che una spiegazione speculativa sia *robusta*, cioè che valga in una grande varietà di condizioni iniziali o, ancor meglio, in *quasi tutti* i tipi di condizioni iniziali. In base al criterio di robustezza possiamo affermare, per esempio, che le spiegazioni speculative proposte da Axelrod e Skyrms sono sufficientemente – o, in certi casi, straordinariamente – robuste. La robustezza conferisce alle spiegazioni speculative una generalità che permette di analizzare tutti quei casi, invero molto frequenti, in cui la determinazione dei dettagli risulta difficile o impossibile. Infatti, una spiegazione robusta dell'evoluzione di determinate norme sociali si applica a una grande varietà di condizioni iniziali, cioè a quasi tutti i valori che possono venire attribuiti ai parametri rilevanti.

Un confronto fra approccio hayekiano e approccio giochistico

Alla luce di quanto si è visto, sembra del tutto ragionevole affermare che le principali tesi teoriche e metodologiche di Hayek sull'evoluzione cultu-

rale hanno strette relazioni concettuali con le corrispondenti tesi elaborate nell'ambito dell'approccio giochistico. Per quanto riguarda, in particolare, le tesi teoriche, tre rilevanti punti di contatto tra Hayek e l'approccio giochistico sono i seguenti.

T.1. CARATTERE TACITO E INCONSAPEVOLE DELLE NORME SOCIALI. Le tesi TH.1 e TG.1 sono accomunate dall'idea che le norme sociali sono di solito costituite da schemi di comportamento adottati in maniera tacita e inconsapevole.

T.2. EVOLUZIONE DELLE NORME SOCIALI. Le tesi TH.2 e TG.2 sono accomunate dall'idea che la maggior parte delle norme sociali sono il risultato di un processo evolutivo fondato su determinati meccanismi di trasmissione culturale, quali la capacità degli individui di riconoscere e imitare le norme che sembrano avere portato buoni risultati ai gruppi che le hanno adottate.

T.3. AZIONI INDIVIDUALI, NORME SOCIALI E ORDINI SPONTANEI. Le tesi TH.3 e TG.3 sono accomunate dall'idea che le azioni individuali basate su determinate regole di condotta possono condurre alla formazione di ordini spontanei caratterizzati da elevata complessità e stabilità.

Per quanto riguarda le tesi metodologiche, tre rilevanti punti di contatto tra Hayek e l'approccio giochistico sono i seguenti.

M.1. INDIVIDUALISMO METODOLOGICO E SPIEGAZIONE DELL'ORDINE SOCIALE. La tesi MH.1 e MG.1 sono accomunate dall'idea che la spiegazione dell'evoluzione delle norme sociali e della formazione di ordini spontanei deve essere effettuata sulla base di presupposti teorici relativi alle azioni individuali.

M.2. PECULIARITÀ METODOLOGICA DELLE SCIENZE SOCIALI. La tesi MH.2, da un lato, e le tesi MG.2 e MG.4, dall'altro, sono accomunate dall'idea della peculiarità metodologica delle scienze sociali. Per esempio, sembra chiaro che le spiegazioni speculative dell'evoluzione culturale, elaborate nell'ambito dell'approccio giochistico, condividono i tratti fondamentali delle spiegazioni di principio hayekiane. In particolare, entrambi i tipi di spiegazione si basano su presupposti, relativamente ovvi, che riguardano determinati aspetti qualitativi dell'evoluzione delle norme sociali. Anche la nozione di robustezza – nella tesi MG.4 – viene usata per valutare gli aspetti qualitativi delle spiegazioni speculative: per esempio, si valuta la robustezza delle spiegazioni dell'emergere della strategia *T* (colpo su colpo) o di altre strategie *qualitativamente simili*, nei loro aspetti fondamentali, a *T*.

M.3. PRESUPPOSTI TEORICI, DEDUZIONE E SIMULAZIONE. Le tesi MH.3 e MG.3 mettono in rilievo la necessità di scoprire le conseguenze che semplici regole di condotta possono avere sull'ordine sociale. Tuttavia, le due tesi suggeriscono procedimenti piuttosto diversi per la scoperta di tali conseguenze. Infatti, Hayek sembra credere che, se si identificano i “giusti” presupposti teorici, sia poi *sufficiente* fare uso delle procedure deduttive

per raggiungere conclusioni interessanti. D'altra parte, i risultati ottenuti nell'ambito dell'approccio giochistico mostrano che non sempre la deduzione basta a far luce sui meccanismi dell'evoluzione culturale. Per esempio, non si riesce a immaginare come i risultati ottenuti da Axelrod e Skyrms attraverso ben congegnati metodi di simulazione avrebbero potuto essere ottenuti sulla base della sola deduzione. A quanto ci risulta, Hayek non ha mai fatto riferimento alla possibilità di usare metodi di simulazione nelle scienze sociali. Tuttavia, non crediamo si possa parlare, a tale proposito, di un'autentica differenza di prospettive metodologiche tra Hayek e i sostenitori dell'approccio giochistico. Infatti, ai tempi in cui lo studioso austriaco operava, l'applicazione dei metodi di simulazione nelle scienze sociali era ancora in una fase embrionale.

Non ci sembra, quindi, del tutto arbitrario ipotizzare che Hayek avrebbe apprezzato metodi di simulazione quali ABM, che si ispirano al tentativo – sotteso anche alle sue spiegazioni di principio – di scoprire le conseguenze che certe semplici regole di condotta possono avere sull'ordine sociale che si sviluppa sulla loro base. Questa ipotesi viene suffragata dall'interesse che, in diverse occasioni, Hayek mostra per la possibile applicazione di certi sofisticati modelli formali nello sviluppo della teoria sociale. Per esempio, egli osserva (Hayek, 1967, p. 67; cfr. anche Hayek 1973-1979, pp. 536-537) che, probabilmente,

quando l'avanzamento delle scienze toccherà fenomeni sempre più complessi, le teorie che forniscono esclusivamente spiegazioni di principio, o che descrivono solamente una gamma di fenomeni che certi tipi di strutture sono in grado di produrre, diventeranno la regola piuttosto che l'eccezione. Certi recenti sviluppi, come la cibernetica, la teoria degli automi o la teoria dei sistemi e forse anche la teoria della comunicazione, sembrano appartenere a questo tipo.

7.3 Spiegazione e complessità nelle scienze sociali

I rilievi di Hayek sul probabile sviluppo di una vera e propria “scienza della complessità” si sono rivelati, in gran parte, profetici. In effetti, nella seconda metà dello scorso secolo sono state sviluppate almeno due dozzine di teorie matematiche, fisiche e biologiche accomunate dall'uso del termine “complessità”. Anzi, fino a una ventina d'anni fa si parlava di complessità quasi esclusivamente con riferimento a queste teorie. Successivamente l'idea di complessità ha avuto una straordinaria diffusione nelle scienze sociali, dove è stata spesso accompagnata dalla fiducia che le teorie della complessità sviluppate nelle scienze naturali avrebbero innescato una

feconda svolta nello studio del mondo sociale. Come vedremo, da quanto si è visto finora, questa fiducia era del tutto infondata. Nel contempo, però,osterremo che la distinzione “semplice/complesso” svolge un ruolo di primo piano nell’analisi del metodo scientifico: infatti, la scienza può venire identificata con la ricerca della semplicità, cioè con la ricerca di semplici teorie in grado di spiegare adeguatamente la straordinaria complessità del mondo. Da questo punto di vista, l’approccio giochistico all’emergere degli ordini spontanei, e in particolare alla formazione ed evoluzione delle norme sociali, assieme alle simulazioni sviluppate dalle teorie ABM, forniscono una rappresentazione semplice, esplicativamente adeguata e probabilmente vera della complessità sociale.

Teorie semplici per un mondo complesso

Non è del tutto inappropriato vedere l’attività scientifica come un’incessante ricerca della semplicità. Ciò significa che gli scienziati sono sempre alla ricerca di teorie in grado di ricondurre la straordinaria complessità del mondo visibile ad alcune semplici regolarità che sfuggono allo sguardo del profano. Per esempio, gli astronomi hanno spiegato il bizzarro moto visibile dei pianeti interpretandolo come il risultato di moti circolari o ellittici, o come l’effetto della legge di gravitazione. Allo stesso modo, i matematici hanno cercato di ricondurre la variopinta complessità che caratterizza il “lato visibile” del mondo delle figure e dei numeri a poche e semplici regolarità; per esempio, la moderna teoria dei numeri naturali ha mostrato che le svariate proprietà dei numeri naturali, molte delle quali note fin dall’antichità, sono derivabili da pochi e semplici assiomi.

Se si vuole sviluppare e giustificare l’idea intuitiva, oggi ampiamente accettata, che la ricerca della semplicità sia un obiettivo fondamentale dell’indagine scientifica, occorre precisare cosa si intende esattamente parlando di semplicità delle teorie e complessità del mondo. In altre parole, si devono affrontare sia i problemi logico-epistemologici consistenti nel definire la semplicità delle teorie e giustificare la preferibilità delle teorie più semplici, sia il problema ontologico consistente nel fornire una soddisfacente caratterizzazione della complessità del mondo.

Risolvere il problema logico della semplicità significa definire un preciso criterio comparativo per stabilire quale, fra due teorie in competizione, è più semplice. Sulla base di un criterio di questo genere possiamo affrontare il problema epistemologico della semplicità, possiamo cioè chiederci se vi siano buone ragioni per seguire la regola metodologica che consiglia di accettare, *ceteris paribus*, la più semplice tra le teorie in competizione.

Qui la clausola *ceteris paribus* indica che la regola può essere applicata solo “a parità di condizioni”, cioè solo nella scelta fra teorie con la stessa capacità di spiegare i fenomeni. Sfortunatamente, sembra esserci una tensione sistematica tra le virtù teoriche della semplicità e dell’adeguatezza esplicativa; ciò significa che, date due teorie rivali *A* e *B*, accade spesso che *A* è più semplice di *B*, ma è dotata di minore adeguatezza esplicativa. In questo caso occorrerà applicare una versione più sofisticata della regola di semplicità che consenta di “pesare” opportunamente, nella scelta fra teorie rivali, semplicità e adeguatezza esplicativa.

Molti studiosi hanno giustificato la regola di semplicità sulla base di considerazioni pragmatiche: una teoria più semplice può essere più facilmente utilizzata per spiegare, prevedere e manipolare i fenomeni. Altri hanno invece fatto appello a considerazioni estetiche: poiché la semplicità di una teoria è uno dei fattori che spingono a considerarla bella ed elegante, la preferenza degli scienziati per la semplicità andrebbe intesa come una manifestazione della loro tendenza a operare scelte teoriche sulla base di valutazioni estetiche. Difficilmente si può negare il ruolo delle considerazioni pragmatiche ed estetiche nelle scelte teoriche degli scienziati: resta però aperto il problema normativo di stabilire se considerazioni di questo genere siano sufficienti a giustificare la regola di semplicità. Nell’ambito del realismo scientifico, cioè della concezione secondo la quale il fine fondamentale dell’indagine scientifica è costituito dalla ricerca della verità, è stata proposta una diversa e, a nostro giudizio, più convincente giustificazione della regola di semplicità (cfr. McAllister 1996). I realisti sostengono che le teorie più semplici sono dotate di maggiore plausibilità, cioè che hanno una maggiore probabilità di essere (almeno approssimativamente) vere. Nella prospettiva dei realisti la semplicità è una guida alla verità; in altre parole, la ricerca della semplicità va intesa come un obiettivo strumentale, funzionale al perseguimento della verità.

Se si accetta la concezione realista secondo la quale le teorie scientifiche dovrebbero fornire una rappresentazione vera della struttura del mondo, occorrerà ammettere che la ricerca di teorie vere dotate di grande semplicità viene inevitabilmente limitata dalla reale complessità del mondo, cioè dall’effettivo grado di complessità del sistema di oggetti sottoposti ad indagine. Piuttosto sorprendentemente, la nozione generale di complessità di un sistema di oggetti, nonostante la sua ovvia importanza per l’analisi del metodo scientifico, non ha finora ricevuto molta attenzione da parte dei filosofi della scienza. Dobbiamo a Nicholas Rescher (1998) una della poche ricerche sistematiche sul concetto filosofico di complessità. Rescher sostiene che la complessità ontologica del mondo è letteralmente inesauribile:

ne segue che la scienza può soltanto darcene un'immagine approssimata e imperfetta. Rescher (1998, cap. 2) distingue sei tipi di complessità ontologica, raggruppati in tre coppie che corrispondono ai tre aspetti principali della complessità di un sistema di oggetti:

- | | |
|-------------------------------|-------------------------------|
| A. COMPLESSITÀ COMPOSIZIONALE | 1. Complessità costituzionale |
| B. COMPLESSITÀ STRUTTURALE | 2. Complessità tassonomica |
| C. COMPLESSITÀ FUNZIONALE | 3. Complessità organizzativa |
| | 4. Complessità gerarchica |
| | 5. Complessità operativa |
| | 6. Complessità nomica |

La complessità compositiva di un sistema dipende dal numero dei suoi componenti (complessità costituzionale) e dalla loro varietà, cioè dal numero dei differenti tipi di componenti (complessità tassonomica). La complessità strutturale dipende, invece, dal modo in cui gli elementi del sistema interagiscono fra loro (complessità organizzativa) e sono collegati l'uno con l'altro (complessità gerarchica): per esempio, possono essere collegati da relazioni gerarchiche di subordinazione, con diversi livelli di sottosistemi, come nel caso di un esercito o un ministero, oppure da relazioni di coordinazione, come nei sistemi di libero mercato. È interessante notare che un'estrema semplicità organizzativa può coesistere con un'elevata complessità gerarchica: è il caso delle strutture frattali nelle quali un unico modo di organizzazione viene ripetuto *ad infinitum* a diversi livelli di scala dando così luogo a un'infinita complessità gerarchica. Infine, la complessità funzionale di un sistema dipende dalla varietà dei suoi modi di operare e svilupparsi nel corso del tempo (complessità operativa) e dal carattere, più o meno elaborato, delle leggi che governano le relazioni o interazioni tra i suoi elementi (complessità nomica).

L'imprevedibilità di alcuni sistemi è prodotta proprio dalla loro complessità nomica, come accade nei sistemi caotici. Mentre nel linguaggio comune il termine "caos" indica spesso l'assenza di ogni ordine e legge, nelle moderne teorie dei sistemi caotici il caos rappresenta un caso di estrema complessità nomica. Secondo le leggi di un sistema caotico, piccolissimi cambiamenti nelle condizioni del sistema in un determinato istante bastano a produrre, dopo un periodo di tempo sufficientemente lungo, cambiamenti di larga scala; quindi, dato che non possiamo osservare i dettagli più minuti delle attuali condizioni del sistema, non saremo in grado di prevederne con precisione lo stato futuro. Come si può scorgere da questa breve illustrazione, il concetto di complessità ontologica integra elementi distinti; non

possiamo quindi limitarci a parlare di sistemi più o meno complessi, o di diversi gradi di complessità, ma dobbiamo anche distinguere tra numerosi tipi di complessità.

Conviene ora tornare brevemente sulla questione del rapporto tra semplicità delle teorie e complessità del mondo. Abbiamo già osservato che il grado di complessità di un sistema di oggetti pone un limite alla semplicità delle nostre migliori teorie intorno a quel sistema. In molti casi, tuttavia, la scarsa semplicità delle nostre teorie dipende, più che dall'effettiva complessità del sistema, dalla nostra incompetenza cognitiva, che ci porta ad attribuire al sistema un grado di complessità maggiore di quello davvero esistente. In casi di questo genere possiamo dire che le nostre teorie sono "complicate", cioè inutilmente complesse (cfr. Rescher 1998, p. 16). La complicazione teorica viene normalmente considerata come un difetto che occorre eliminare o almeno ridurre. Questa idea ha ispirato svariate versioni e giustificazioni della regola di semplicità. In particolare, molti fra i sostenitori della concezione realista della scienza sostengono le seguenti tesi: (1) la scelta fra due o più teorie relative a un determinato sistema di oggetti deve basarsi sulla valutazione della loro semplicità e della loro capacità di spiegare adeguatamente i fenomeni, cioè gli aspetti accertati del sistema; (2) la teoria scelta sulla base di questi due fattori ha una più elevata probabilità di essere vera, o approssimativamente vera; ciò implica fra l'altro, che tale teoria ha una più elevata probabilità di attribuire al sistema il "giusto" tipo e grado di complessità.

Scienza versus retorica della complessità

Poiché qualunque teoria scientifica verte su determinati sistemi di oggetti e qualunque sistema è caratterizzato da un certo tipo e grado di complessità, è legittimo dire che qualunque teoria scientifica è una teoria di sistemi (complessi) o anche una teoria della complessità (di sistemi). Tuttavia questa affermazione può apparire provocatoria dato che, nell'ultimo secolo, il termine "complessità" è stato adottato, e quasi confiscato, da diverse discipline scientifiche, nel cui ambito ha acquisito significati tecnici piuttosto distanti dall'uso che se ne fa abitualmente nel linguaggio comune e in quello filosofico. Nella lista che segue, fornita a titolo puramente illustrativo e senza nessuna pretesa di completezza, elenchiamo alcune teorie della complessità elaborate in diversi ambiti del sapere scientifico, accompagnandole con i nomi degli studiosi che hanno maggiormente contribuito al loro sviluppo.

A. LOGICA, MATEMATICA, SCIENZA DEI CALCOLATORI E DEI SISTEMI ARTIFICIALI

- Teorie della complessità computazionale e dell'informazione algoritmica (Andrej Kolmogorov, Ray Solomonoff e Gregory Chaitin)
- Teorie della complessità logica e descrittiva (Ronald Fagin)
- Teorie degli automi cellulari (John von Neumann e Stephen Wolfram)
- Cibernetica (Norbert Wiener)

B. SCIENZE FISICHE

- Teorie dei sistemi caotici (David Ruelle e Edward Lorenz)
- Teoria dei sistemi dissipativi (Ilya Prigogine)
- Teoria delle catastrofi (René Thom)
- Teoria dei frattali (Benoit Mandelbrot)

C. SCIENZE BIOLOGICHE

- Teoria generale dei sistemi (Ludwig von Bertalanffy)
- Teorie dell'auto-organizzazione e dei sistemi adattivi complessi, algoritmi genetici e vita artificiale (John Conway, Stuart Kauffman, John Holland e Mitchel Resnick)

D. NEUROSCIENZE, SCIENZE COGNITIVE E PSICANALISI

- Teorie dell'autopoiesi (Humberto Maturana e Francisco Varela)
- Epistemologia genetica e teorie dei sistemi simbolici (Jean Piaget, Gregory Bateson e Edgar Morin)
- Teoria psicanalitica degli insiemi infiniti (Ignacio Matte Blanco)

E. SCIENZE SOCIALI

- Teorie funzionaliste, strutturaliste, comunicative e critiche dei sistemi sociali (Bronislaw Malinowski, Talcott Parsons, Claude Lévi-Strauss, Niklas Luhman e Jürgen Habermas)
- Teorie evoluzionistiche dell'ordine sociale (Friedrich August von Hayek)
- Teoria delle organizzazioni complesse (Herbert Simon)

Può essere interessante osservare che alcune fra le teorie della complessità originariamente sviluppate nei campi disciplinari A-D della precedente lista sono state poi applicate anche alle scienze sociali. Inoltre si è sostenuto che le nozioni di complessità utilizzate nell'ambito delle diverse teorie della complessità esibiscono una notevole "somiglianza di famiglia". Queste circostanze hanno ispirato la convinzione che le teorie della complessità proposte nell'ambito di svariate discipline scientifiche possano venire, in qualche modo, unificate; da tale unificazione nascerebbero una "scienza" e

persino una “logica” della complessità pressoché universalmente applicabili e, in ogni caso, applicabili all’analisi dei sistemi sociali. A quanto ci risulta, questa convinzione non ha finora condotto a risultati scientifici degni di nota; tuttavia il suo successo di pubblico, in filosofia e nelle scienze umane, è innegabile. Tra le conseguenze di questo successo va segnalata la diffusione di una *retorica della complessità*, caratterizzata da un suo peculiare gergo, preso a prestito dal linguaggio delle filosofie postmoderne e afflitto dalla stessa oscurità e ampollosa vaghezza. Fortunatamente, negli ultimi anni alcune voci critiche si sono alzate per denunciare gli usi impropri delle nozioni di complessità tratte dalla matematica e dalle scienze naturali; per esempio, nel settimo capitolo del volume di Sokal e Bricmont (1999) si può trovare una gustosa rassegna delle vere e proprie imposture intellettuali, basate sull’impiego disinvolto dei concetti di caos e complessità, di cui si sono resi responsabili alcuni fra i più famosi teorici postmoderni.

Robustezza, rappresentatività e flessibilità delle spiegazioni giochistiche

Come abbiamo visto, le teorie ABM e le teorie dei giochi soddisfano pienamente i principi dell’individualismo metodologico, per cui occorre spiegare l’origine, il funzionamento, l’evoluzione e l’eventuale sparizione delle istituzioni sociali sulla base delle interazioni fra individui e delle conseguenze, intenzionali e non, di tali interazioni. Possiamo quindi affermare che, se l’individualismo metodologico fornisce una buona caratterizzazione dell’adeguatezza esplicativa delle teorie formulate nelle scienze sociali, allora le teorie ABM e le teorie dei giochi sono esplicativamente adeguate: infatti, sono in grado di spiegare una grande varietà di macrofenomeni sociali in accordo con i principi individualistici. Ci sembra, inoltre, difficile negare la notevole semplicità delle teorie ABM e delle teorie dei giochi, semplicità che si riflette, per esempio, nel piccolo numero di concetti e ipotesi fondamentali utilizzati nella spiegazione dei fenomeni sociali. Per i sostenitori di una concezione realista delle teorie scientifiche, del genere descritto nel primo paragrafo, la circostanza che le teorie ABM e le teorie dei giochi siano semplici ed esplicativamente adeguate è un’ottima ragione per ritenere che tali teorie siano anche almeno approssimativamente vere e forniscano, quindi, un’immagine sostanzialmente corretta della complessità sociale.

Recentemente, i filosofi della scienza si sono interrogati su come si possa valutare l’adeguatezza esplicativa delle teorie sviluppate nell’ambito dell’approccio giochistico e ABM (cfr. Grüne-Yanoff e Weirich 2010). Per esempio, D’Arms *et al.* (1998) hanno formulato alcuni plausibili criteri di

valutazione – vale a dire *robustezza*, *rappresentatività* e *flessibilità* – al fine di valutare le spiegazioni proposte da Skyrms (1996) della genesi e dell'evoluzione delle istituzioni sociali. Vale la pena illustrare più nei dettagli questi criteri, dato che possono venire applicati, in generale, per valutare qualsiasi analisi dei fenomeni sociali sviluppata sulla base delle teorie evuzionistiche e adattive dei giochi e delle teorie ABM.

Nel capitolo 6 abbiamo utilizzato il concetto di “robustezza” per indicare la capacità di alcune strategie – come la strategia *T* del “colpo su colpo” studiata da Axelrod – di affermarsi in una grande varietà di situazioni. Tuttavia, come abbiamo visto nel paragrafo 7.2, nell’accezione di D’Arms *et. al.* (1998) si può parlare di robustezza non solo con riferimento a particolari strategie, ma anche a determinate spiegazioni generali, come il modello di Skyrms dell’evoluzione delle norme di giustizia. In questo modello il successo di una strategia dipende dal contesto: per esempio dal modo in cui le diverse strategie sono inizialmente distribuite nella popolazione. Poiché la distribuzione iniziale di strategie in una data popolazione è normalmente ignota, come possiamo sapere che il modello spiega davvero il prevalere di *quelle* specifiche strategie in *quella* popolazione? Potremmo rispondere a questo interrogativo osservando che il modello di Skyrms non mira a spiegare come effettivamente si siano sviluppate certe norme di giustizia. Il suo scopo è, invece, quello di mostrare come sia possibile che queste norme abbiano potuto sopravvivere e prosperare, dati determinati *tipi di condizioni iniziali*, nei processi dinamici ipotizzati. D’altra parte, questa risposta non risolve la difficoltà legata al fatto che, nella maggior parte dei casi, non si riuscirà mai a stabilire se le condizioni iniziali ipotizzate nel modello giochistico si sono effettivamente verificate.

Tale difficoltà può venire, in qualche modo, superata richiedendo che un modello adeguato sia *robusto*, cioè che valga in una grande varietà di condizioni iniziali o, ancor meglio, in *quasi tutti* i tipi di condizioni iniziali. In base al criterio di robustezza possiamo affermare, per esempio, che la spiegazione di Skyrms dell’evoluzione delle norme di giustizia è sufficientemente robusta e che la spiegazione di Axelrod dell’evolvere della cooperazione è straordinariamente robusta. La robustezza conferisce alle spiegazioni giochistiche una generalità che permette di analizzare tutti quei casi, invero molto frequenti, in cui la determinazione dei dettagli risulta difficile o impossibile. Ciò significa che un modello giochistico sufficientemente robusto ci consentirà di elaborare spiegazioni plausibili dell’evoluzione di certe norme di comportamento anche quando i dettagli circa i “meccanismi di basso livello” – biologici o sociali – che potrebbero avere determinato quelle norme sono ignoti.

Sebbene la robustezza sia una caratteristica molto importante, è perfettamente possibile immaginare un modello giochistico estremamente robusto e, al tempo stesso, del tutto inadeguato. Tale risulterebbe, infatti, un modello robusto che non fosse in grado di offrire, almeno a livello qualitativo, una *rappresentazione accurata* delle circostanze entro le quali emergono ed evolvono alcune interessanti forme di interazione. Il *criterio di rappresentatività* consiste, appunto, nella richiesta di mostrare, attraverso un sufficiente numero di esempi, che un modello possiede tale capacità di rappresentazione. Per affermare che un determinato modello giochistico è rappresentativo dobbiamo mostrare, per esempio, che la struttura delle retribuzioni del gioco su cui esso si basa corrisponde alle circostanze effettive di un determinato tipo di interazioni umane. Per esempio, Axelrod (1984, capitolo 4) ha cercato di mostrare che il suo modello giochistico dell'evoluzione di T è applicabile alla guerra di trincea (cfr. D'Arms *et. al.*, 1998, p. 89). D'altra parte, non tutti i modelli presentati nella letteratura sembrano possedere un alto grado di rappresentatività; per esempio, "è difficile pensare che la famiglia di giochi di divisione della torta esplorata da Skyrms sia rappresentativa di circostanze frequenti della nostra vita" (*ivi*).

Infine, un modello è *flessibile* quando è in grado di rappresentare una grande varietà di meccanismi e processi (D'Arms *et. al.*, 1998, p. 91). Per esempio, la spiegazione di Skyrms dell'evoluzione delle norme di equa contrattazione rivela buone doti di flessibilità, dato che la dinamica dei replicatori potrebbe modellare tanto l'evoluzione delle propensioni genetiche di certi meccanismi psicologici quanto la scelta di strategie da parte di agenti razionali. Il criterio di flessibilità è ispirato da motivazioni non molto dissimili da quelle che stanno alla base del criterio di robustezza; infatti, ci esenta dalla necessità di identificare certi importanti dettagli, in questo caso i "dettagli causali" relativi agli specifici meccanismi e processi causali che determinano l'evoluzione di particolari tipi di norme.

7.4 Teoria dei giochi, individualismo e filosofia della politica

Prima di concludere la nostra discussione, riteniamo utile almeno accennare a un interessante problema messo in luce dal confronto fra l'approccio austriaco – e specialmente hayekiano – e quello giochistico, cioè lo stretto nesso che esiste fra scienze sociali e filosofia della politica. Infatti, anche se l'ideale di una teoria sociale abbastanza potente da implicare una filosofia politica in grado di dirci come organizzare o dirigere la società è ormai largamente screditato, è difficile immaginare una teoria sociale – cioè un

corpus di ipotesi circa i meccanismi di funzionamento della società – del tutto priva di implicazioni per la filosofia politica. Ciò dipende dal fatto che la comprensione dei “meccanismi sociali” fornisce inevitabilmente almeno alcune indicazioni sulle possibilità di una “tecnologia sociale”, cioè sui modi in cui un’ autorità centrale potrebbe influenzare i (macro)fenomeni sociali. Sembra, quindi, perfettamente legittimo chiedersi quali siano le possibili implicazioni delle teorie giochistiche della genesi e dell’evoluzione delle istituzioni sociali per la filosofia politica.⁹

A questo proposito, occorre almeno notare come le teorie giochistiche dell’interazione sociale si rivelino molto efficaci per fare luce sui meccanismi attraverso i quali, in assenza di qualunque autorità centrale, una popolazione di egoisti razionali tende a sviluppare un insieme di norme di comportamento improntate a criteri di equità, giustizia, cooperazione, e persino altruismo. Come abbiamo visto (soprattutto nei capitoli 4 e 6), nei modelli giochistici e ABM svariati “principi di buon comportamento” – cioè norme e regolarità di comportamento ritenute desiderabili da diverse filosofie etico-politiche – possono diffondersi e prevalere in una popolazione senza l’intervento di alcuna autorità centrale. Inoltre, questo approccio ci permette anche di comprendere perché tale intervento potrebbe produrre una cascata di indesiderabili conseguenze non intenzionali sull’ordine spontaneo prodotto da determinati sistemi di norme sociali. In altre parole, i risultati ottenuti nell’ambito dell’approccio giochistico contribuiscono significativamente a minare l’idea che un intervento politico centrale sia *necessario* ad assicurare la prevalenza di norme ritenute moralmente, socialmente o politicamente auspicabili.

9 Recentemente, diversi autori hanno indagato le conseguenze dell’approccio evolucionistico (e anche giochistico) al comportamento umano per la (filosofia) politica: si vedano, ai due estremi dello spettro politico, Singer (2000) e Arnhart (2005), ma soprattutto i contributi di Rubin (2002) e Friedman (2008). Come nota Gilberto Corbellini (2009) nella prefazione all’edizione italiana del libro di Rubin, il tentativo di applicare l’approccio evolucionistico alla (filosofia) politica rischia di venire frainteso. Il fraintendimento più comune e pericoloso deriva dalla mancata distinzione tra il “darwinismo politicizzato” e la “politica darwiniana”: il primo è costituito, come è accaduto nel caso del darwinismo sociale di fine Ottocento, dall’“interpretazione strumentale e priva di coerenza logica o di basi scientifiche delle idee darwiniane per difendere qualche particolare ideologia politica”; la seconda, invece, consiste nell’“uso delle conoscenze evolucionistiche sulla natura umana per meglio comprendere le origini delle preferenze politiche individuali, la loro distribuzione sociale e le dissonanze tra gli adattamenti ancestrali e l’ambiente attuale” (*ibidem*, p. 9).

I risultati appena menzionati si rivelano di particolare interesse per quelle filosofie della politica che – come nel caso di Hayek, ma anche di de Jasay – sottolineano i limiti e i pericoli dell'interventismo statale e dell'"ingegneria sociale". A questo proposito ci sembrano significativi alcuni rilievi di Skyrms (1996, pp. 108-109, corsivo nostro):

La filosofia politica è lo studio di come le società potrebbero essere organizzate. Se la possibilità è interpretata in modo generoso abbiamo la teoria utopistica. Quelli che vorrebbero trattare con "gli uomini così come sono" hanno bisogno di lavorare con un senso più ristretto di possibilità. L'interesse per la dinamica interattiva dell'evoluzione biologica, dell'evoluzione culturale e dell'apprendimento fornisce alcuni interessanti vincoli.

In conflitto con il rozzo determinismo dei darwinisti sociali del diciannovesimo secolo, e con Hegel e Marx, la [teoria evoluzionistica dei giochi] mostra che il caso tipico è quello in cui non c'è un unico risultato predeterminato, ma piuttosto una profusione di possibili risultati di equilibrio. La teoria prevede quello che gli antropologi hanno sempre saputo: che sono possibili molti stili alternativi di vita sociale. [...] Gli equilibri [sociali] hanno una stabilità variabile. Alcuni sono più facili da rovesciare. Altri sono più robusti. Quelli instabili possono essere sensibili ad alcuni tipi di perturbazione, ma non ad altri. Anche coloro che vogliono cambiare il mondo devono prima imparare come descriverlo.

I rilievi di Skyrms richiamano la nostra attenzione su come gli sviluppi delle scienze sociali possono influenzare la filosofia della politica. Quest'ultima, infatti, può venire intesa come il tentativo di rispondere alla domanda: *qual è il miglior ordine politico possibile?* Come nota Cubeddu (1997, p. 18):

[s]e i protagonisti della teoria politica sono gli individui, con le loro azioni e le relative conseguenze per gli altri individui, il problema [fondamentale della teoria politica] può essere riformulato come la ricerca del migliore ordine politico, ricerca che a sua volta presuppone una teoria sulla formazione dell'ordine sociale o, più brevemente, una teoria sociale.

In altre parole, un'adeguata teoria sociale è il necessario fondamento di qualsiasi filosofia politica. Ciò è particolarmente evidente in Hayek, che sottolinea a più riprese come la sua teoria sociale individualistica ed evoluzionistica sia un presupposto fondamentale del suo liberalismo politico. A suo avviso, infatti, una delle concezioni fondamentali su cui si basa "il mantenimento di una società di uomini liberi" è quella che "un ordine autogenerantesi o spontaneo, e la struttura ordinata di un'organizzazione, sono due cose molto diverse, e che tale differenza dipende dai due generi molto diversi di regole o leggi che prevalgono al loro interno" (von Hayek, 1973-

1979, p. 7). Mentre il costruttivismo – cioè l’idea che le norme e le istituzioni sociali vengano esplicitamente progettate e applicate da un’autorità centrale – “ci dà la sensazione di avere un potere illimitato nel realizzare i nostri desideri”, la teoria sociale evolutzionistica “ci porta a considerare che vi sono limiti a ciò che possiamo tentare, e a riconoscere che alcune delle nostre speranze sono delle illusioni” (*ibidem*, p. 13). Inoltre, un’adeguata teoria evolutzionistica illumina i meccanismi attraverso i quali gli interventi di un’autorità centrale possono interferire, talvolta disastrosamente, con l’ordine spontaneo generato dalle interazioni tra individui che seguono determinati sistemi di regole di condotta. Sono proprio queste le ragioni per cui “il liberalismo restringe il controllo deliberato sull’intero ordine sociale all’implementazione di quelle sole regole generali che sono necessarie per la formazione di un ordine spontaneo i cui dettagli non siamo in grado di prevedere” (*ibidem*, pp. 45-46).

Il contrasto fra l’approccio costruttivista e quello “ecologico” (Smith 2008), basato sul concetto di ordine spontaneo, alla razionalità umana e alle istituzioni sociali è illustrato, per esempio, dalla discussione che Hayek e Skyrms offrono delle norme di giustizia. Per entrambi, il problema della giustizia non è quello di immaginare e “calare dall’alto” il modello corretto di “società giusto”, quanto quello di mostrare come “regole giuste” emergono “dal basso”, cioè dall’interazione individuale. Il secondo volume di Hayek (1973) è una devastante critica del “miraggio della giustizia sociale”, cioè delle pretese esplicative e normative solitamente associate all’idea di giustizia sociale. Per Hayek questo concetto risulta privo di un contenuto determinabile; a suo avviso, infatti, la giustizia è un attributo della condotta umana e non di una società nel suo complesso:

Propriamente parlando, soltanto il comportamento umano può essere considerato giusto o ingiusto. Se si applicano questi termini a uno stato di cose, hanno senso soltanto se si ritiene qualcuno responsabile del suo avvento o dell’averlo permesso. Un semplice fatto, o stato di cose che nessuno può mutare, può essere buono o cattivo, ma non giusto o ingiusto (p. 219).

Di conseguenza, secondo von Hayek, cercare di spiegare l’evoluzione della giustizia significa essenzialmente cercare di spiegare l’emergere e l’evoltersi di regole di comportamento basate su criteri di giustizia, e di analizzare in quali condizioni tali regole possono più facilmente prosperare. Come si può vedere, è la stessa prospettiva che sottende l’analisi del “problema della giustizia” condotta da Skyrms, anche se il suo approccio è motivato da interessi dichiaratamente più descrittivi ed esplicativi che prescrittivi.

Resta aperto, naturalmente, il problema di sapere “cosa possiamo fare” per creare le condizioni più adatte all’affermarsi di comportamenti giusti, altruistici e così via. Questo problema è solo un aspetto del più vasto insieme di questioni relativo al rapporto tra le dimensioni descrittiva e quella normativa e “operativa” della teoria politica. A questo proposito, sarà sufficiente osservare che, nell’ambito dell’approccio “ecologico” di Hayek, Skyrms e dei teorici dei giochi, un’approfondita conoscenza teorica dei meccanismi sociali induce, per lo meno, alla cautela nel “disegnare” l’estensione della sfera d’intervento del potere politico in una “società di uomini liberi”. D’altra parte, tale cautela non implica affatto l’idea che sia in ogni caso indesiderabile modificare i sistemi di norme in uso in una determinata società. Per esempio, Axelrod (1984, capitolo 7) offre alcuni suggerimenti ai “riformatori” desiderosi di “modificare i termini stessi dell’interazione in modo da stimolare la cooperazione” (*ibidem*, p. 26; cfr. anche Binmore 2005, cap. 12). D’altra parte, la “prospettiva del riformatore” (*ivi*) è ben lontana dalla visione utopica dell’ingegnere sociale intenzionato a rimodellare la società da cima a fondo. Per questo motivo, contrariamente a quello che qualcuno potrebbe pensare, la prospettiva del riformatore è compatibile anche con la concezione evolucionistica di von Hayek, purché si fondi sulla consapevolezza che ogni modifica effettuata su una parte di un ordine spontaneo complesso può avere, oltre agli attesi e benefici “effetti locali”, anche imprevedibili e indesiderabili effetti su molte altre parti di quell’ordine.

Riteniamo che tale consapevolezza possa giovare anche della comprensione dei meccanismi di interazione e di evoluzione dell’ordine sociale esplorati dall’approccio giochistico. Le brevi osservazioni fatte finora dovrebbero suggerire almeno la possibilità che le teorie dei giochi rappresentino un utile strumento concettuale non solo per lo scienziato sociale, ma anche per chi intenda contribuire alla fondazione del liberalismo politico nell’ambito della tradizione di ricerca evolucionistica che ha in Hayek uno tra i suoi più illustri esponenti.



ELENCO DELLE FIGURE

- Figura 2.1. Teoria della scelta razionale, teoria delle decisioni e teoria dei giochi
- Figura 2.2. Teorie dei giochi
- Figura 3.1 Un gioco contro natura
- Figura 3.2. La morra cinese.
- Figura 3.3: Gioco con totale coincidenza di interessi
- Figura 3.4 Il gioco della guida
- Figura 3.5: Gioco con totale conflitto di interessi
- Figura 3.6 La battaglia del Mare di Bismarck (1)
- Figura 3.7: Gioco con parziale coincidenza di interessi
- Figura 3.8 La battaglia dei sessi
- Figura 3.9 La battaglia del Mare di Bismarck (2)
- Figura 3.10. Gioco con un unico equilibrio di Nash
- Figura 3.11: Il dilemma del prigioniero
- Figura 4.1. Il dilemma dei beni pubblici
- Figura 4.2 Posizioni teoriche sui beni pubblici
- Figura 4.3. Due strutture di incentivi per il problema dei beni pubblici
- Figura 4.4. Il gioco del cumulo di neve (1)
- Figura 4.5. Il gioco dei beni pubblici in forma generale
- Figura 4.6. Gioco del cumulo di neve (2)
- Figura 4.7. Il gioco della caccia al cervo
- Figura 5.1. Riga avverte Colonna che sceglierà R2
- Figura 5.2. Est minaccia Nord di rispondere con b ad a
- Figura 5.3. Riga minaccia Colonna di rispondere con R2 a C1
- Figura 5.4. Est promette a Nord di rispondere con b ad A
- Figura 5.5. Riga e Colonna promettono di scegliere, rispettivamente, R1 e C1.
- Figura 5.6. Est minaccia Nord di rispondere con B ad a e gli promette di rispondere con b ad A .
- Figura 5.7. Riga trasforma il primo gioco (a) nel secondo (b) impegnandosi su R2 e nel terzo (c) impegnandosi su R1.
- Figura 5.8. Il supergioco determinato dal gioco in figura 5.7



BIBLIOGRAFIA

- Alexander, J. M. (2000). "Evolutionary Explanations of Distributive Justice". *Philosophy of Science*, 67, pp. 490-516.
- Alexander, J. M. (2007). *The Structural Evolution of Morality*. Cambridge University Press, Cambridge.
- Antiseri D. e Pellicani L. (1992). *L'individualismo metodologico. Una polemica sul mestiere dello scienziato sociale*. Franco Angeli, Milano.
- Arielli, E. e Scotto, G. (1998). *I conflitti. Introduzione a una teoria generale*. Bruno Mondadori, Milano.
- Aron, R. (1963). *Le grand Débat. Initiation à la stratégie atomique*. Calmann-Lévy, Parigi. Trad. it.: *Il grande dibattito. Introduzione alla strategia atomica*. Il Mulino, Bologna, 1965.
- Arnhart, L. (2005). *Darwinian Conservatism*, Imprint Academic, Exeter (UK).
- Axelrod, R. (1984). *The Evolution of Cooperation*. Basic Books, New York. Trad. it.: *Giochi di reciprocità. L'insorgenza della cooperazione*. Feltrinelli, Milano, 1985.
- Axelrod, R. (1997). *The Complexity of Cooperation*. Princeton University Press, Princeton.
- Axelrod, R. (2000). "On Six Advances in Cooperation Theory". *Analyse & Kritik*, 22, pp. 130-151. URL: http://www.analyse-und-kritik.net/2000-1/AK_Axelrod_2000.pdf.
- Axelrod, R. e Hamilton, W. D. (1981). "The Evolution of Cooperation". *Science*, 211, 1390-1396.
- Aumann, R., e Hart, S. (2008). *Razionalità, cooperazione, conflitto. Intervista sulla teoria dei giochi*. Morcelliana, Brescia.
- Barkow, J., Cosmides, L. e Tooby, J. (1992). *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. Oxford University Press, New York.
- Baron, J. (1988). *Thinking and Deciding*. Cambridge University Press, Cambridge. Quarta edizione: 2007.
- Barry, B. (1965). *Political Argument*. Routledge & Kegan Paul, Londra.
- Beaufre, A. (1963). *Introduction à la stratégie*. Librairie Armand Colin, Parigi. Trad. it.: *Introduzione alla strategia*, Il Mulino, Bologna, 1966.
- Bellhouse, D. (2007). "The problem of Waldegrave". *Journal Électronique d'Histoire des Probabilités et de la Statistique*, 3.2. URL: <http://eudml.org/doc/117579>.
- Beito, D. T., Gordon, P. e Tabarrok, A. (2002), a cura di. *The Voluntary City: Choice, Community, and Civil Society*. The Independent Institute, Oakland

- (California). Trad. it. parziale: *La città volontaria*. Rubbettino/Leonardo Facco, Soveria Mannelli (Catanzaro)/Treviglio (Bergamo), 2010.
- Bicchieri, C. (1993). *Rationality and Coordination*. Cambridge University Press, Cambridge. Trad. it.: *Razionalità e azione collettiva*, Milano, Feltrinelli, 1998.
- Bicchieri, C. (2006). *The Grammar of Society. The Nature and Dynamics of Social Norms*. Cambridge University Press, Cambridge.
- Bicchieri, C. e Muldoon, R. (2011). "Social Norms". *The Stanford Encyclopedia of Philosophy*, a cura di E. N. Zalta. URL: <http://plato.stanford.edu/archives/spr2011/entries/social-norms/>.
- Binmore, K. (1990). *Essays on the foundations of game theory*. Basil Blackwell, Oxford.
- Binmore, K. (1994). *Game Theory and the Social Contract. Playing fair*. The MIT Press, Cambridge.
- Binmore, K. (1998). *Game Theory and the Social Contract. Just playing*. The MIT Press, Cambridge.
- Binmore K. (2005), *Natural Justice*, Oxford University Press, Oxford.
- Binmore, K. (2007). *Game Theory: A Very Short Introduction*. OUP Oxford. Trad. it.: *Teoria dei giochi*. Codice, Torino, 2008.
- Boettke, P. (2005). "Anarchism as a progressive research program in political economy". In *Anarchy, State and Public Choice*, a cura di E. P. Stringham. Edward Elgar Publishing, Cheltenham (UK), pp. 178-190.
- Boettke, P. J., C. J. Coyne, e P. T. Leeson (2008). "The Continuing Relevance of F. A. Hayek's Political Economy". In: *Explorations in Austrian Economics* a cura di R. Koppl. Vol. 11 di *Advances in Austrian Economics*. Emerald, Bingley (UK), pp. 79-98.
- Boudon, R. (1979). *La logique du social*. Hachette, Parigi. Trad. it.: *La logica del sociale*, Mondadori, Milano, 1980.
- Bouillon, H. e Kliemt, H. (2007), a cura di. *Ordered anarchy: Jasay and his Surroundings*. Ashgate Publishing, Farnham (UK), 2007.
- Boyd, R. e Richerson, P. (1985). *Culture and the Evolutionary Process*, University of Chicago Press, Chicago.
- Bowlby, J. (1969). *Attachment*. Basic Books, New York. Trad. it.: *Attaccamento e perdita*, Bollati Boringhieri, Torino, 1999.
- Bowles, S. (2004). *Microeconomics: Behavior, Institutions, and Evolution*. Princeton University Press, Princeton. Trad. it.: *Microeconomia*, Dipartimento di Economia Politica dell'Università di Siena, <http://www.econ-pol.unisi.it/bowles-microeconomia/>.
- Bowles, S. e Gintis, H. (2011). *A Cooperative Species: Human Reciprocity and Its Evolution*. Princeton University Press, Princeton.
- Braithwaite, R. B. (1955). *Theory of games as a tool for the moral philosopher*. Cambridge University Press, Cambridge
- Buchanan, J. (1975). *The Limits of Liberty: Between Anarchy and Leviathan*. University of Chicago Press, Chicago. Trad. it.: *I limiti della libertà*, Rusconi, Milano, 1998.
- Buchanan, J. (2002). "Game theory, mathematics, and economics". *Journal of Economic Methodology*, 8 (1), pp. 27-32.

- Buchanan, J. (2007). "Introducing Tony de Jasay". In Bouillon e Kliemt (2007), pp. 3-4.
- Caldwell, B. (2002). "Hayek and Cultural Evolution". In *Fact and Fiction in Economics*, a cura di U. Mäki, Cambridge University Press, Cambridge, pp. 285-303.
- Camerer, C. (2003). *Behavioral Game Theory*. Princeton University Press, Princeton.
- Carnap, R. (1962). *The Logical Foundations of Probability* (seconda edizione). The University of Chicago Press, Chicago.
- Cavalli-Sforza, L. L. e Feldman, M. (1981). *Cultural Transmission and Evolution: A Quantitative Approach*. Princeton University Press, Princeton.
- Cevolani, G. (2008). "Giochi, dilemmi sociali e scelte collettive". In de Jasay (1991), pp. 13-56.
- Cevolani, G. (2011). "Hayek in the lab", *Logic & Philosophy of Science*, 9.1, pp. 429-436.
- Cevolani, G. e Festa, R. (2011). "Giochi di anarchia. Beni pubblici, teoria dei giochi e anarco-liberalismo". *Nuova Civiltà delle Macchine*, 1-2, pp. 163-180.
- Cevolani, G. e R. Festa (2012a). "Giochi di altruismo. L'approccio evoluzionistico alla cooperazione umana". In: Ridley, M. *Le origini della virtù*. IBL Libri, Torino, pp. 7-38.
- Cevolani, G. e R. Festa (2012b). "L'ingranaggio della cooperazione. Teorie dei giochi, cooperazione spontanea e produzione di beni pubblici". In: *Libertates*, a cura di C. Lottieri e D. Velo Dalbrenta. IBL Libri, Torino. In corso di pubblicazione.
- Clavien, C. e Chapuisat, M. (2012). "Altruism across Disciplines: One Word, Multiple Meanings". In corso di pubblicazione in *Biology and Philosophy*. DOI: 10.1007/s10539-012-9317-3.
- Corbellini, G. (2009). "Politica darwiniana vs darwinismo politicizzato". In: Rubin (2002), pp. 7-38.
- Cubeddu, R. (1997). *Atlante del liberalismo*, Ideazione, Roma.
- D'Arms, J., Batterman, R. e Górný, K. (1998). "Game Theoretic Explanations and Evolution of Justice", in *Philosophy of Science*, 65, pp. 76-102.
- Darwin, C. (1871). *The Descent of Man, and Selection in Relation to Sex*. John Murray, Londra. Trad. it.: *L'origine dell'uomo*, Studio Tesi, Pordenone, 1991.
- Davis, M. D. (1983). *Game theory: a nontechnical introduction*. Dover Publications, New York.
- Dawkins, R. (1976). *The Selfish Gene*. Oxford University Press, Londra. Trad. it.: *Il gene egoista*, Mondadori, Milano, 1994.
- de Bruin, B. (2005). "Game Theory in Philosophy". *Topoi* 24 (2):197-208.
- de Bruin, B. (2010). *Explaining Games: The Epistemic Programme in Game Theory*. Springer, New York.
- de Finetti, B. (1937). "La prévision: ses lois logiques, ses sources subjectives". *Annales de l'Institut Henri Poincaré*, 7, 1-68. Trad. it. in *La logica dell'incerto*, a cura di M. Mondadori, Il Saggiatore, Milano, 1989, pp. 71-147.

- de Jasay, A. (1985). *The State*. Basil Blackwell, Oxford. Nuova edizione: Liberty Fund, Indianapolis, 1998. Trad. it. in corso di pubblicazione per IBL Libri, Torino.
- de Jasay, A. (1989). *Social Contract, Free Ride: A Study of the Public Goods Problem*. Clarendon Press, Oxford. Nuova edizione: Liberty Fund, Indianapolis, 2008.
- de Jasay, A. (1991). *Choice, Contract and Consent: A Restatement of Liberalism*. Institute of Economic Affairs, Londra. Trad. it.: *Scelta, contratto e consenso*, Rubbettino e Leonardo Facco, Soveria Mannelli e Treviglio, 2008.
- de Jasay, A. (1992). "Indivisibility, Probability and Game Theory". *Humane Studies Review*, 7.1. Replica a Sugden (1992).
- de Jasay, A. (1997). *Against Politics*. Routledge, Londra.
- de Jasay, Anthony (1998). "Public goods theory". In P. Boettke, *The Elgar Companion to Austrian Economics*, Edward Elgar Publishing, Cheltenham (UK), pp. 276-284.
- de Jasay, A. (2002). *Justice and its Surroundings*. Liberty Fund, Indianapolis.
- de Jasay, A. (2010). *Political economy, concisely*. Liberty Fund, Indianapolis.
- Devlin, K., (2008). "The Unfinished Game. Pascal, Fermat and the Seventeenth-Century Letter that Made the World Modern". Basic Books, Londra.
- Dimand, M. A. e R. W. Dimand. *A History of Game Theory. From the Beginnings to 1945*, vol. I, Routledge, Londra e New York, 1996.
- Dimand, R. W. (2014). *A History of Game Theory*, vol. II, Routledge, Londra e New York (in corso di pubblicazione).
- Di Nuoscio, E. (2006), *Il mestiere dello scienziato sociale*, Liguori, Napoli.
- Dionisio, F. e Gordo, I. (2006). "The Tragedy of the Commons, the Public Goods Dilemma, and the meaning of Rivalry and Excludability in Evolutionary Biology". *Evolutionary Ecology Research* 8, pp. 321-332.
- Dugatkin, L. A. (2006). *The Altruism Equation*. Princeton University Press, Princeton.
- Dugatkin, L. A. (2011). *The Prince of Evolution: Peter Kropotkin's Adventures in Science and Politics*. CreateSpace, Charleston (SC).
- Elster, J. (1989a). *The Cement of Society*, Cambridge: Cambridge University Press. Trad. it.: *Il cemento della società. Uno studio sull'ordine sociale*, Il Mulino, Bologna, 1995.
- Elster, J. (1989b). *Nuts and Bolts for the Social Sciences*, Cambridge: Cambridge University Press. Trad. it.: *Come si studia la società*, Il Mulino, Bologna, 1993.
- Fehr, E. e Fischbacher, U. (2003). "The Nature of Human Altruism". *Nature*, 425 (6960), pp. 785-791.
- Fehr, E. e Gintis, H. (2007). "Human nature and social cooperation", *Annual Review of Sociology*, 33, pp. 1-22.
- Festa, R. (1999). "Modelli evolutivisti del contratto sociale". *Etica & Politica / Ethics & Politics*, 1999, 1, URL: <http://www.units.it/etica/1999>.
- Festa, R. (2001). "Come evolvono le norme sociali: la prospettiva della teoria dei giochi". *Biblioteca della libertà*, 158, pp. 75-98.
- Festa, R. (2003). "Minacce e promesse. Logica e metodologia della deterrenza". In: *La guerra. Una riflessione interdisciplinare*, a cura di G. Manganaro Favaretto. Edizioni Università di Trieste, Trieste, pp. 247-298.

- Festa, R. (2004). "Teoria dei giochi e strategie della deterrenza", *Logic and Philosophy of Science*, 2. URL: http://www.univ.trieste.it/episteme/L&PS_Vol-2No1/festa_L&PS_Vol2No1.pdf
- Festa, R. (2006). "Giochi di società. La complessità sociale nelle teorie ABM e nelle teorie dei giochi". *élite*, 3, pp. 15-30.
- Festa, R. (2007). "Teoria dei giochi ed evoluzione delle norme morali". *Etica & Politica*, IX (2), pp. 148-181.
- Festa, R. (2008). "Teoria dei giochi, metodo delle scienze sociali e filosofia della politica". Prefazione a de Jasay (1991), pp. 5-12.
- Festa, R. (2012). "Giocare con Hayek. Teoria dei giochi e concezione hayekiana dell'evoluzione culturale". In: R. de Mucci (a cura di), *An Austrian in Italy. Festschrift in Honour of Professor Dario Antiseri*, Rubbettino, Soveria Mannelli (Catanzaro), 2012.
- Festa, R. e Crupi, V. (2009). "Dai giochi d'azzardo ai test diagnostici: la teoria della decisione nella medicina clinica". *Logic and Philosophy of Science*, 7, pp. 69-94.
- Foss, N. (2000). "Austrian Economics and Game Theory: A Stocktaking and an Evaluation". *Review of Austrian Economics*, 13, pp. 41-58.
- Freedman, L. (1986). "Le prime due generazioni di strateghi nucleari". In: *Guerra e strategia nell'età contemporanea*, a cura di N. Lablanca, Marietti, Genova, 1992, pp. 283-324.
- Friedman, Daniel (2008), *Morals and markets*, Palgrave Macmillan. Trad. it.: *Morale e mercato*, IBL Libri, Torino, 2012.
- Friedman, David (1989). *The Machinery of Freedom*. Open Court Publishing Company, Chicago. Trad. it.: *L'ingranaggio della libertà*, Macerata, Liberilibri, 1997.
- Gauthier, D. (1969). *The Logic of 'Leviathan': the Moral and Political Theory of Thomas Hobbes*. Clarendon Press, Oxford.
- Gauthier, D. (1984). "Deterrence, Maximization, and Rationality", *Ethics* 94, 474-495. Ristampato in D. Gauthier, *Moral Dealing. Contract, Ethics, and Reason*, Cornell University Press, Ithaca e Londra, 1990.
- Gauthier, D. (1986). *Morals by Agreement*. Clarendon Press, Oxford.
- Gibbons, R. (1992). *A Primer in Game Theory*. Simon & Schuster, Englewood Cliffs (NJ). Trad. it.: *Teoria dei giochi*, Il Mulino, Bologna, 1994.
- Gintis, H. (2000). *Game Theory Evolving*. Princeton University Press.
- Gintis, H. (2007). "The evolution of private property". *Journal of Economic Behavior and Organization*, 64, pp. 1-16.
- Gintis, H. (2009). *The Bounds of Reason*. Princeton University Press, Princeton.
- Gintis, H., Bowles, S., Boyd, R. e Fehr, E. (2005), a cura di. *Moral Sentiments and Material Interests: The Foundations of Cooperation In Economic Life*. MIT Press, Cambridge.
- Gould, S. J. (1991). *Bully for Brontosaurus*, W. W. Norton, New York.
- Govindan, S. e Wilson, R. B. (2008). "Nash equilibrium, refinements of", in *The New Palgrave Dictionary of Economics*, seconda edizione a cura di S. N. Durlauf e L. E. Blume. Palgrave Macmillan.
- Guala, F. (2005). *The Methodology of Experimental Economics*. Cambridge University Press, Cambridge.

- Guala, F. (2006). *La filosofia dell'economia*, Il Mulino, Bologna.
- Guala, F. (2012). "The Normativity of Lewis Conventions". *Synthese*, in corso di pubblicazione (DOI: 10.1007/s11229-012-0131-x).
- Grüne-Yanoff, T., e Weirich, P. (2010). "The Philosophy and Epistemology of Simulation: A Review". *Simulation & Gaming*, 41.1, pp. 20-50.
- Grüne-Yanoff, T. e Lehtinen, A. (2012). "Philosophy of Game Theory". In: *Handbook of the Philosophy of Economics*, a cura di U. Mäki, Elsevier.
- Hacking, I. (1975). *The Emergence of Probability*. Cambridge University Press, Cambridge. Trad. it.: *L'emergenza della probabilità*. Il Saggiatore, Milano, 1987.
- Hardin, G. (1968). "The Tragedy of the Commons". *Science*, 162, pp. 124-142.
- Hardin, R. (2007). *David Hume: Moral and Political Theorist*. Oxford University Press, Oxford.
- Hargreaves Heap, S., Hollis, M., Lyons, B., Sugden, R. e Weale, A. (1992). *The Theory of Rational Choice. A Critical Guide*, Blackwell Publishers, Cambridge. Trad. it.: *La teoria della scelta razionale. Una guida critica*. Laterza, Roma-Bari, 1996.
- Hargreaves Heap, S. P., e Varoufakis, Y. (2004). *Game Theory: a Critical Text*. Routledge, Londra. Prima edizione: 1995.
- Harms W. e Skyrms, B. (2007). "Evolution of Moral Norms". In: *Oxford Handbook on the Philosophy of Biology*, a cura di M. Ruse. Oxford University Press, Oxford.
- Harsanyi, J. (1955). "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility". *Journal of Political Economy* 63, pp. 309-321. Trad. it.: "Benessere cardinale, etica individualistica e confronti interpersonali di utilità", in *L'utilitarismo*, Il Saggiatore, Milano, 1994, pp. 137-155.
- Harsanyi, J. (1968). "Individualistic and Functionalistic Explanations in Light of Game Theory: The Example of Social Status". In: *Problems in the Philosophy of Science*, a cura di I. Lakatos e A. Musgrave. North-Holland, Amsterdam, 1968, pp. 305-322.
- Hauert, C., Holmes, M. e Doebeli, M. (2006). "Evolutionary Games and Population Dynamics: Maintenance of Cooperation in Public Goods Games". *Proceeding of the Royal Society B*, 273, pp. 2565-2571.
- Hausman, D. M. (2012). "Philosophy of Economics". In: *The Stanford Encyclopedia of Philosophy*, a cura di E. N. Zalta. URL: <http://plato.stanford.edu/entries/economics/>
- Hayek, F. A. (1942-1943). "Scientism and the Study of Society". *Economica*, 35, pp. 267-291 e 37, pp. 34-63. Trad. it.: "Lo scientismo e lo studio della società", in *Conoscenza, mercato, pianificazione*. Bologna, Il Mulino, 1988, pp. 97-210.
- Hayek, F. A. (1952). *The Counter-revolution of Science: Studies on the Abuse of Reason*. Londra, Collier-Macmillan. Trad. it.: *L'abuso della ragione*. Firenze, Vallecchi, 1967
- Hayek, F. A. (1967). *Studies in Philosophy, Politics and Economics*. Londra, Routledge and Kegan Paul. Trad. it.: *Saggi di filosofia, politica ed economia*. Rubettino, Soveria Mannelli, 1998.

- Hayek, F. A. von (1973-1979). *Law, Legislation, and Liberty* (3 vol.). University of Chicago Press, Chicago. Trad. it.: *Legge, legislazione e libertà*. Il Saggiatore, Milano, 1994.
- Haywood, O. G. (1954). "Military Decision and Game Theory". *Journal of the Operations Research Society of America*, 2.4, pp. 365-385.
- Hobbes, T. (1651). *Leviathan*, Andrew Crooke, Londra. Trad. it.: *Leviatano*. Rizzoli, Milano, 2011.
- Hofstadter, D. R. (1985). *Metamagical Themas: Questing for the Essence of Mind and Pattern*. Basic Books, New York.
- Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press, Ann Arbor, Michigan. Seconda edizione: 1992.
- Holland, J. H. (1992). "Genetic Algorithms". *Scientific American*, 267, pp. 66-72. Trad. it.: "Algoritmi genetici", *Le Scienze*, 289, 1992, pp. 50-57.
- Holcombe, R. (2009). "The Behavioral Foundations of Austrian Economics". *The Review of Austrian Economics* 22.4, pp. 301-313.
- Hoppe, H.-H. (1989a). "Fallacies of the public goods theory and the production of security". *Journal of Libertarian Studies*, 9.1, pp. 27-46.
- Hoppe, H.-H. (1989b). *A Theory of Socialism and Capitalism*. Kluwer Academic Publishers, Dordrecht.
- Hume, D. (1739). *A Treatise of Human Nature*, John Noon, Londra. Trad. it.: *Trattato sulla natura umana*, Laterza, Bari, 1982.
- Kahn, H. (1962/1984). *Thinking about the Unthinkable*, Horizon Press, New York. Seconda ediz. aggiornata: *Thinking about the Unthinkable in the 1980s*, Simon & Schuster, New York, 1984.
- Kahn, H. (1965). *On Escalation. Methaphors and Scenarios*, Frederick A. Praeger Publishers, New York. Trad. it.: *La filosofia della guerra atomica. Esempi e schemi*, Edizione del Borghese, Milano, 1966.
- Kahneman, D. (2011). *Thinking Fast and Slow*, Allen Lane. Trad. it.: *Pensieri lenti e veloci*, Mondadori, Milano, 2012.
- Kahneman, D., McFadden, D. e Smith, V. L (2005). *Critica della ragione economica*. Il Saggiatore, Milano.
- Kelly Y. J. (2009). "Mises, Morgenstern, Hoeselitz, and Nash: The Austrian Connection To Early Game Theory". *The Quarterly Journal of Austrian Economics* 12, pp. 37-42.
- Kennedy, D. e Norman, C. (2005). "What Don't We Know?" *Science*, 309 (5731), p. 75.
- Kliemt, H. e Lahno, B. (1992). "Social contract, free ride". *Constitutional Political Economy*, 3.2, pp. 267-271. Recensione di de Jasay (1989).
- Kohlberg, E. (1990). "Refinement of Nash Equilibrium: The Main Ideas". In: *Game Theory and Applications*, a cura di T. Ichiishi, A. Neyman e Y. Tauman. Academic Press, San Diego.
- Kollock, P. (1998). "Social Dilemmas: The Anatomy of Cooperation". *Annual Review of Sociology*, 24, pp. 183-214.
- Koopl, R. (2004), a cura di. *Evolutionary Psychology and Economic Theory*. Advances in Austrian Economics, vol. 7. Elsevier, Oxford.

- Kreps, D. M. (1990). *Game theory and economic modelling*. Oxford University Press. Trad. it.: *Teoria dei giochi e modelli economici*. Il Mulino, Bologna, 1992.
- Kropotkin, P. (1902). *Mutual Aid: A Factor of Evolution*. Heinemann, Londra. Trad. it.: *Il mutuo appoggio: un fattore dell'evoluzione*, Edizioni Anarchismo, Trieste, 2008.
- Kuipers, T. A. F. (2000). *From instrumentalism to constructive realism*. Kluwer, Dordrecht.
- Infantino, L. (1995). *L'ordine senza piano*. Armando, Roma, terza edizione: 2008.
- Jordan, J. (1994), a cura di. *Gambling on God: Essays on Pascal's Wager*. Rowman & Littlefield, Lanham (MD).
- Leinfellner, W. e Köhler, E. (1998), a cura di. *Game Theory, Experience, Rationality. Foundations of Social Sciences, Economics and Ethics*, Dordrecht, Kluwer
- Lemennicier, B. (2006). "Fallacies in the theories of the emergence of the state". *Journal of Libertarian Studies*, 20 (3), pp. 3-28.
- Lewens, T. (2008). "Cultural Evolution". In *The Stanford Encyclopedia of Philosophy*, a cura di E. N. Zalta. URL: <http://plato.stanford.edu/archives/fall2008/entries/evolution-cultural/>.
- Lewis, D. (1969). *Convention. A Philosophical Study*. Harvard University Press, Harvard. Trad. it.: *Convenzione. Uno studio filosofico*, Bompiani, Milano, 1974.
- Lloyd, E. (2012). "Units and Levels of Selection". In *The Stanford Encyclopedia of Philosophy*, a cura di E. N. Zalta, URL: <http://plato.stanford.edu/archives/spr2012/entries/selection-units/>.
- Luce, R. D. e Raiffa, H. (1957). *Games and Decisions. Introduction and Critical Survey*. John Wiley and Sons, New York. Nuova edizione: Dover, New York, 1989.
- Lumsden, C. e Wilson, E. O. (1981). *Genes, Mind and Culture: The Coevolutionary Process*. Harvard University Press, Cambridge.
- Luttwak, E. N. (1987). *Strategy: the Logic of War and Peace*. Harvard University Press, Harvard. Trad. it.: *Strategia*, Rizzoli, Milano, 1989.
- Marciano, A. (2009). "Why Hayek is a Darwinian (after all)? Hayek and Darwin on Social Evolution". *Journal of Economic Behavior & Organization*, 71.1, pp. 52-61.
- Maynard Smith, J. (1982). *Evolution and the Theory of Games*, Cambridge University Press, Cambridge.
- Maynard Smith, J. e Price, G. R. (1973). "The Logic of Animal Conflict". *Nature*, 246, pp. 15-18.
- McAdams, R. H. e Nadler, J. (2005). "Testing the Focal Point theory of Legal Compliance: The Effect of Third-Party Expression in an Experimental Hawk/Dove Game". *Journal of Empirical Legal Studies*, 2.1, pp. 87-123.
- McAllister, J. W. (1996). "Beauty and Revolution in Science". Cornell University Press, Ithaca, New York. Trad. it.: *Bellezza e rivoluzione nella scienza*, Milano, McGraw-Hill, 1998.
- McClennen, E. (1990). *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge University Press, New York.

- Menger C. (1983). *Untersuchungen über die Methode der Socialwissenschaften und der Politischen Oekonomie insbesondere*, Lipsia. Trad. it.: *Sul metodo delle scienze sociali*. Macerata, Liberilibri, 1996.
- Mises, L., von (1949). *Human Action*. William Hodge, Londra. Trad. it.: *L'azione umana*. Torino, UTET, 1959.
- Mises, L., von (1962). *The Ultimate Foundation of Economic Science. An Essay on Method*. Princeton, D. Van Nostrand. Terza edizione: Indianapolis, Liberty Fund, 2006.
- Morgenstern, O. (1935). "Vollkommene voraussicht und wirtschaftliches gleichgewicht". *Zeitschrift für Nationalökonomie* 6, pp. 337-57. Trad. inglese: "Perfect foresight and economic equilibrium". In *Selected Economic Writings of Oskar Morgenstern*, a cura di A. Schotter. New York University Press, New York, 1976, pp. 169-83. Trad. it: "Previsione perfetta ed equilibrio economico", in Morgenstern (1969).
- Morgenstern, O. (1937). *The Limits of Economics*. Londra: William Hodge.
- Morgenstern, O. (1969). *Teoria dei giochi*. Boringhieri, Torino.
- Morrow, J. (1994). *Game theory for political scientists*. Princeton University Press.
- Motterlini, M. e Piattelli Palmarini, M. (2005). "Introduzione". In: Kahneman, McFadden e Smith (2005), pp. 9-24.
- Motterlini, M. e Guala, F. (2011). *Mente mercati decisioni. Introduzione all'economia cognitiva e sperimentale*. Bocconi, Milano.
- Narveson, J. (1998). *The Libertarian Idea*. Temple University Press, Philadelphia. Nuova edizione: Broadview Press, Peterborough (Ontario), 2001.
- Neumann, J. von (1928). "Zur Theorie der Gesellschaftspiele". *Mathematische Annalen*, 100, pp. 295-320. Trad. inglese: "On the theory of games of strategy". In: *Contributions to the Theory of Games*, a cura di A.W. Tucker e R.D. Luce, Princeton University Press, Princeton 1959, pp. 13-42.
- Neumann, J. von e Morgenstern, O. (1944). *Theory of Games and Economic Behaviour*. Princeton University Press, Princeton.
- Niiniluoto, I. (1987). *Truthlikeness*. Reidel, Dordrecht.
- Nozick, R. (1974). *Anarchy, State, Utopia*. Basic Books, Londra. Trad. it.: *Anarchia, stato, utopia*, Il Saggiatore, Milano, 2000.
- Nozick, P. (1997). *Socratic Puzzles*. Harvard, Harvard UP. Trad. it.: *Puzzles socratici*. Milano, Raffaello Cortina, 1999
- Okasha, S. (2009). "Biological Altruism". In *The Stanford Encyclopedia of Philosophy*, a cura di E. N. Zalta, URL: <http://plato.stanford.edu/archives/win2009/entries/altruism-biological/>.
- Olson, M. (1965). *The Logic of Collective Action: Public Goods and the Theory of Groups*, Harvard University Press. Seconda edizione: 1971.
- Oprea, R. e B. Powell (2010). "Why Austrians Should Quit Worrying and Learn to Love the Lab". In: *What is so Austrian about Austrian Economics?* a cura di R. Koppl e S. Horwitz. Vol. 14 di *Advances in Austrian Economics*. Emerald, Bingley (UK), pp. 145-163.
- Osborne, M. J. (2003). *An introduction to game theory*. Oxford University Press.
- Osborne, M. e Rubinstein, A. (1994). *A course in game theory*. MIT Press. URL: <http://books.osborne.economics.utoronto.ca/>

- Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press.
- Ostrom, E., Burger, J., Field, C. B., Norgaard, R. B., e Policansky, D. (1999). "Revisiting the Commons: Local Lessons, Global Challenges". *Science*, 284 (5412), pp. 278-282.
- Parfit, D. (1984). *Reason and Persons*. Oxford University Press, Oxford. Trad. it.: *Ragioni e persone*, Il Saggiatore, Milano, 1989.
- Paternotte, C. e Grose, J. (2012). "Social Norms and Game Theory: Harmony or Discord?". *The British Journal for the Philosophy of Science*. In corso di pubblicazione (DOI: 10.1093/bjps/axs024).
- Powell, B. (2009). "Public choice and the economic analysis of anarchy: a survey". *Public Choice* 140, pp. 503-538.
- Powell, B. and Wilson, B. (2008). "An experimental investigation of Hobbesian jungles". *Journal of Economic Behavior and Organization*, 66 (3-4), pp. 669-686.
- Powell, B. e Stringham, E. (2009). "Public choice and the economic analysis of anarchy: a survey". *Public Choice*, 140.3, pp. 503-538.
- Ramsey, F.P. (1931). "Truth and probability". In: *Foundations of Mathematics and Other Essays*. Routledge & Kegan, Londra, pp. 156-198. Anche in: *Philosophical Papers*. Cambridge University Press, Cambridge 1990, pp. 52-94.
- Rasmussen, E. (2006). *Games and Information: An Introduction to Game Theory*. Wiley. Trad. it. della prima edizione: *Teoria dei giochi e informazione*. Hoepli, Milano, 1993.
- Rawls, J. (1971). *A Theory of Justice*. Cambridge: Harvard University Press. Trad. it.: *Una teoria della giustizia*, Feltrinelli, Milano, 1982.
- Rescher N. (1998). *Complexity. A Philosophical Overview*. Transaction Publishers, New Brunswick e Londra.
- Richerson, P. e Boyd, R. (2005). *Not by Genes Alone: How Culture Transformed Human Evolution*. University of Chicago Press, Chicago.
- Riolo, R. L. (1992). "Survival of the Fittest". *Scientific American*, 267, pp. 114-116. Trad. it.: "La sopravvivenza dei bit più adatti". *Le Scienze*, 289, 1992, pp. 86-88.
- Ross, D. (2012). "Game Theory". In: *The Stanford Encyclopedia of Philosophy*, a cura di E. N. Zalta. URL = <<http://plato.stanford.edu/archives/win2012/entries/game-theory/>>.
- Rothbard M. (1973). *For a New Liberty*. Collier Books, New York. Seconda edizione: 1978. Trad. it.: *Per una nuova libertà*, Liberilibri, Macerata, 1996 (seconda edizione: 2004).
- Rothbard, M. (1981). "The Myth of Neutral Taxation". *The Cato Journal*, 1.2, pp. 519-564. Ora in: *The Logic of Action Two*, Edward Elgar Publishing, Cheltenham (UK), 1997, pp. 56-108.
- Rousseau, J.-J. (1755), *Discours sur l'origine et les fondaments de l'inégalité parmi le hommes*. Trad. it.: *Scritti politici*, vol. I, Laterza, Bari, 2005.
- Rubin, P. H. (2002). *Darwinian politics: the evolutionary origin of freedom*. Rutgers University Press, Piscataway (NJ). Trad. it.: *La politica secondo Darwin. L'origine evolutiva della libertà*. IBL Libri, Torino 2009.

- Savage, L.J. (1954). *The Foundations of Statistics*. Wiley, New York.
- Schelling, T.C. (1960). *The Strategy of Conflict*. Harvard University Press, Cambridge. Seconda edizione: 1980. Trad. it.: *La strategia del conflitto*, Bruno Mondadori, Milano, 2006.
- Schelling, T. C. (1966). *Arms and influence*. Yale University Press, New Haven, 1966. Trad. it.: *La diplomazia della violenza*, Il Mulino, Bologna, 1968.
- Schelling, T. C. (1978). *Micromotives and Macrobehavior*. New York, W. W. Norton and Co., 1978. Seconda edizione aggiornata: 2006. Trad. it.: *Micromotivazioni della vita quotidiana*, Milano, Bompiani, 2008.
- Schelling, T. C. (1984). *Choice and Consequence. Perspectives of an Errant Economist*. Harvard University Press, Cambridge, Mass., 1984
- Schelling, T. C. e Halperin, M. H. (1961). *Strategy and Arms Control. Controlled Response and Strategic Warfare*. Twentieth Century Fund, New York.
- Schwalbe, U., e Walker, P. (2001). Zermelo and the Early History of Game Theory. *Games and Economic Behavior*, 34(1), 123-137.
- Seagren, C. (2011). "Examining Social Processes With Agent-based Models". *The Review of Austrian Economics*, 24.1, pp. 1-17.
- Selten, R. (1975). "A reexamination of the perfectness concept for equilibrium points in extensive games". *International Journal of Game Theory*, 4, pp. 25-55.
- Selten, R. (1998). "Game Theory, Experience, Rationality". In Leinfellner e Köhler (1998), pp. 9-34.
- Skyrms, B. (1994). "Darwin Meets 'The Logic of Decision': Correlation in Evolutionary Game Theory". *Philosophy of Science*, 61, pp. 503-528.
- Skyrms, B. (1996). *The Evolution of Social Contract*. Cambridge University Press, Cambridge.
- Skyrms, B. (2004). *The Stag Hunt and the Evolution of Social Structure*. Cambridge University Press, Cambridge.
- Simon, H. (1982-1997). *Models of Bounded Rationality*. MIT Press, Cambridge.
- Simon, H. (1983). *Reason in Human Affairs*. Stanford University Press, Stanford, 1983. Trad. it.: *La ragione nelle vicende umane*, Bologna, Il Mulino, 1984.
- Singer, P. (2000). *A Darwinian Left: Politics, Evolution and Cooperation*. Yale University Press, New Haven (CT).
- Smith, A. (1776). *An Inquiry into the Nature and Causes of the Wealth of Nations*. Strahan e Cadell, Londra. Trad. it.: *La ricchezza delle nazioni*. Newton Compton, Roma, 2005.
- Smith, V. (1999). "Reflections on *Human Action* 50 years later", *Cato Journal* 19, pp. 195-214. Trad. it.: "Riflessioni su *L'azione umana*, cinquant'anni dopo", IBL Occasional Paper, 61, 2008. URL: http://brunoleonimedia.servingfreedom.net/OP/61_Smith.pdf
- Smith, V. (2005). "Hayek and Experimental Economics". *The Review of Austrian Economics* 18.2, pp. 135-144.
- Smith, V. (2008). *Rationality in Economics*. Cambridge University Press, New York. Trad. it.: *La razionalità in economia*, IBL Libri, Torino, 2010.
- Sober, E. e Wilson, D. S. (1999). *Unto Others: The Evolution and Psychology of Unselfish Behavior*. Harvard University Press, Oxford.

- Sokal, A. e Bricmont J. (1999). *Impostures Intellectuelles*. LGF - Livre de Poche, Parigi. Trad.it.: *Imposture intellettuale*, Garzanti, Milano, 1999.
- Sugden, R. (1986). *The Economics of Rights, Co-operation and Welfare*. Basil Blackwell, Oxford. Seconda edizione: Palgrave Macmillan, Basingstoke (UK), 2004.
- Sugden, R. (1989). "Spontaneous Order". *The Journal of Economic Perspectives*, 3.4, pp. 85-97.
- Sugden, R. (1992). "Suckers, free riders and public goods". *Humane Studies Review*, 7. URL: <http://osf1.gmu.edu/~ihs/w91review.html>. Recensione di de Jasay (1989).
- Sugden, R. (2001). "The evolutionary turn in game theory". *Journal of Economic Methodology* 8.1, pp. 113-130.
- Sugden, R. e Zamarron, I. E. (2006). "Finding the key: The riddle of focal points". *Journal of Economic Psychology*, 27.5, pp. 609-621.
- Stringham, E. (2007), a cura di. *Anarchy and the Law: the Political Economy of Choice*, Transaction Publishers, Piscataway (NJ).
- Taylor, M. J. (1976). *Anarchy and Cooperation*. Wiley, New York.
- Taylor, M. J. (1987). *The Possibility of Cooperation*. Cambridge University Press, Cambridge.
- Taylor, M. J. e Ward, H. (1982). "Chickens, whales and lumpy goods: alternative models of public goods provision". *Political Studies*, 30.3, pp. 350-370.
- Tooby, J. e Cosmides, L. (1992). "The Psychological Foundations of Culture", in Barkow, Cosmides e Tooby (1992), pp. 19-136.
- Trivers, R. L. (1971). "The Evolution of Reciprocal Altruism". *The Quarterly Review of Biology*, 46(1), 35-57.
- Tullock, G. (1972). *Explorations in the Theory of Anarchy*. Center for the Study of Public Choice, Virginia Polytechnic Institute and State University.
- Ullmann-Margalit, E. (1977). *The Emergence of Norms*. The Clarendon Press, Oxford.
- Ullmann-Margalit, E. (1978). "Invisible-hand explanations". *Synthese* 39.2, pp. 263-291.
- Vanderschraaf, P. (1999). "Game Theory, Evolution and Justice". *Philosophy and Public Affairs* 28.4, pp. 325-328.
- Verbeek, B. e Morris, C. (2010). "Game theory and ethics". In: *The Stanford Encyclopedia of Philosophy*, a cura di E. N. Zalta. <http://plato.stanford.edu/archives/sum2010/entries/game-ethics/>
- Wade, M. J., Bijma, P., Ellen, E. D. e Muir, W. (2010). "Group Selection and Social Evolution in Domesticated Animals". *Evolutionary Applications*, 3(5-6), pp. 453-465.
- Walker, P. (2012). "A Chronology of Game Theory". URL: http://www.econ.canterbury.ac.nz/personal_pages/paul_walker/gt/hist.htm
- Weibull, J. (1997). *Evolutionary Game Theory*. MIT Press, Cambridge (MA).
- White, L. H. (2008). "The Research Program of Austrian Economics", in *Explorations in Austrian Economics*, a cura di R. Koppl. Emerald, pp. 11-24.
- Wilkinson, G. S. (1984). "Reciprocal Food Sharing in the Vampire Bat". *Nature*, 308 (5955), pp. 181-184.

- Wilson, D. S. e Sober, E. (1994). "Reintroducing Group Selection to the Human Behavioral Sciences". *Behavioral and Brain Sciences* 17 (4), pp. 585-654.
- Wilson, D. S. e Wilson, E. O. (2007). "Rethinking the Theoretical Foundation of Sociobiology". *The Quarterly Review of Biology*, 82 (4), pp. 327-348.
- Wilson, D. S. e Wilson, E. O. (2008). "Evolution 'for the Good of the Group'". *American Scientist*, 96 (5), 380-389. Trad. it.: "Evolgere per il bene del gruppo", *Le Scienze*, 490, 2009, pp. 90-99.
- Wilson, E. O. (1975). *Sociobiology. The New Synthesis*. Harvard University Press, Cambridge. Trad. it.: *Sociobiologia. La nuova sintesi*, Bologna, Zanichelli, 1979.
- Wilson, E. O. (2005). "Kin Selection as the Key to Altruism: its Rise and Fall". *Social Research* 72 (1), pp. 159-166.
- Wohlstetter, A. (1959). "The Delicate Balance of Terror". *Foreign Affairs*, 37, n. 2
- Young, H. P. (2008). "Social norms". In *The New Palgrave Dictionary of Economics* (seconda edizione) a cura di S. N. Durlauf e L. E. Blume. Palgrave Macmillan.
- Zywicki, T. J. (2000). "Was Hayek Right about Group Selection After All? Review Essay of *Unto Others. The Evolution and Psychology Of Unselfish Behaviour*". *Review of Austrian Economics*, 13, pp. 81-95.



ETEROTOPIE

Collana diretta da *Pierre Dalla Vigna* e *Salvo Vaccaro*

1. Nerozzi Bellman Patrizia (a cura di), *Internet e le muse. La rivoluzione digitale nella cultura umanistica*
2. Vaccaro Salvo (a cura di), *Il secolo deleuziano*
3. Berni Stefano, *Soggetti al potere. Per una genealogia del pensiero di Michel Foucault*
4. Carbone Paola (a cura di), *Congenialità e traduzione*
5. Marzocca Ottavio, *Transizioni senza meta. Oltremarxismo e antieconomia*
6. Carbone Paola (a cura di), *Le comunità virtuali*
7. Fadini Ubaldo, *Principio metamorfosi. Verso un'antropologia dell'artificiale*
8. Mello Patrizia (a cura di), *Spazi della patologia, patologia degli spazi*
9. Petrilli Susan, Ponzio Augusto, *Fuori campo. I segni del corpo tra rappresentazione ed eccedenza*
10. Carmagnola Fulvio, *La specie poetica. Teorie della mente e intelligenza sociale*
11. Deleuze Gilles, *La passione dell'immaginazione. L'idea della genesi nell'estetica di Kant*
12. De Michele Girolamo, *Tiri Mancini. Walter Benjamin e la critica italiana*
13. Riccio Franco, Vaccaro Salvo (a cura di), *Nietzsche in lingua minore*
14. Carbone Paola, *Patchwork Theory. Dalla letteratura postmoderna all'ipertesto*
15. Ferri Paolo, *La rivoluzione digitale. Comunità, individuo e testo nell'era di Internet*
16. Foucault Michel, *Spazi altri. I luoghi delle eterotopie*
17. Bataille Georges, *La condizione del peccato*
18. Carbone Paola (a cura di), *eLiterature in ePublishing*
19. Dal Bo Federico, *Società e discorso. L'etica della comunicazione in Karl Otto Apel e Jacques Derrida*
20. Deleuze Gilles, *Istinti e istituzioni*
21. Paquot Thierry, *L'utopia ovvero un ideale equivoco*
22. Pirrone Marco Antonio, *Approdi e scogli. Le migrazioni internazionali nel Mediterraneo*
23. Ponzio Augusto, *Individuo umano, linguaggio e globalizzazione nella filosofia di Adam Schaff*
24. Simone Anna, *Divenire sans papiers. Sociologia dei dissensi metropolitani*
25. Vaccaro Salvo (a cura di), *La censura infinita. Informazione in guerra, guerra all'informazione*
26. Artaud Antonin, *CsO. Il corpo senz'Organi*
27. Moulian Tomás, *Una rivoluzione capitalista. Il Cile, primo laboratorio mondiale del neoliberalismo*
28. Thea Paolo, *Il vero cioè il falso. Invenzione, riconoscimento e rivelazione nell'arte*
29. Amato Pierandrea (a cura di), *La biopolitica. Il potere e la costituzione della soggettività*
30. Bertuccioli Manolo, *Carlos Castaneda e i navigatori dell'infinito*
31. Bonaiuti Gianluca, Simoncini Alessandro (a cura di), *La catastrofe e il parassita. Scenari della transizione globale*

32. Buchbinder David, *Sii uomo! Studio sulle identità maschili*
33. Cozzo Andrea, *Conflittualità nonviolenta. Filosofia e pratiche di lotta comunicativa*
34. Deleuze Gilles, *Fuori dai cardini del tempo, Lezioni su Kant*
35. Galluzzi Francesco, *Roba di cui sono fatti i sogni. Arte e scrittura nella modernità*
36. Leghissa Giovanni, *Il gioco dell'identità. Differenza, alterità, rappresentazione*
37. Maistrini Maria, *Il figurale in J.-F. Lyotard*
38. Montanari Moreno, *Il Tao di Nietzsche*
39. Vaccaro Salvo, *Globalizzazione e diritti umani. Filosofia e politica della modernità*
40. Bazzanella Emiliano, *Il ritornello. La questione del senso in Deleuze-Guattari*
41. Fabbri Lorenzo, *L'addomesticamento di Derrida. Pragmatismo/ Decostruzione*
42. Marcenò Serena, *Le tecnologie politiche dell'acqua. Governance e conflitti in Palestina*
43. Piana Gabriele, *Conoscenza e riconoscimento del corpo*
44. Prebisch Raul, *La crisi dello sviluppo argentino. Dalla frustrazione alla crescita vigorosa*
45. Scopelliti Paolo, *Psicanalisi surrealista. L'influenza del surrealismo su Hesnard, Lacan, Deleuze e Guattari*
46. Vaccaro Salvo, *Biopolitica e disciplina. Michel Foucault e l'esperienza del GIP (Group d'Information sur les prisons)*
47. Vercelloni Luca, *Viaggio intorno al gusto. L'odissea della sensibilità occidentale dalla società di corte all'edonismo di massa*
48. Caronia Antonio, Livraghi Enrico, Pezzano Simona, *L'arte nell'era della produttività digitale*
49. Dino Alessandra (a cura di), *La violenza tollerata. Mafia, poteri, disobbedienza*
50. Rodda Fabio, *Cioran, l'antiprofeta. Fisionomia di un fallimento*
51. Scolari Raffaele, *Paesaggi senza spettatori. Territori e luoghi del presente*
52. Pastore Luigi, Limnatis G. Nectarios (a cura di), *Prospettive del postmoderno Vol.1. Profili epistemici*
53. Poidimani Nicoletta, *Oltre le monoculture del genere*
54. Pastore Luigi, Limnatis G. Nectarios (a cura di), *Prospettive del postmoderno Vol.2. Profili epistemici*
55. Bellini Paolo, *Cyberfilosofia del potere. Immaginari, ideologie e conflitti della civiltà*
56. Bazzanella Emiliano, *Etica del tardocapitalismo*
57. Cuttita Paolo, *Segnali di confine. Il controllo dell'immigrazione nel mondo-frontiera*
58. De Conciliis Eleonora (a cura di), *Dopo Foucault. Genealogie del postmoderno*
59. Di Benedetto Giovanni, *Il naufragio e la notte. La questione migrante tra accoglienza, indifferenza ed ostilità*
60. Pagliani Piero, *Naxalbari-India. L'insurrezione nella futura "terza potenza mondiale"*
61. Vaccaro Giovanbattista, *Per la critica della società della merce*
62. Vinale Adriano (a cura di), *Biopolitica e democrazia*
63. Demichelis Lelio, Leghissa Giovanni (a cura di), *Biopolitiche del lavoro*
64. Corradi Luca, Perocco Fabio (a cura di), *Sociologia e globalizzazione*
65. Bellini Paolo (a cura di), *La rete e il labirinto. Tecnologia, identità e simbolica politica*

66. Dalla Vigna Pierre, *A partire da Merleau-Ponty. L'evoluzione delle concezioni estetiche merleau-pontyane nella filosofia francese e negli stili dell'età contemporanea*
67. Riccioni Ilaria (a cura di), *Comunicazione, cultura, territorio. Contributi della sociologia contemporanea*,
68. Pasquino Monica, Plastina Sandra (a cura di), *Fare e disfare. Otto saggi a partire da Judith Butler*
69. Bertoldo Roberto, *Anarchismo senza anarchia. Idee per una democrazia anarchica*
70. Del Bono Serena, *Foucault, pensare l'infinito. Dall'età della rappresentazione all'età del simulacro*
71. Dino Alessandro e Licia A. Callari (a cura di), *Coscienza e potere. Narrazioni attraverso il mito*
72. Farci Manolo, Pezzano Simona (a cura di), *Blue lit stage. Realtà e rappresentazione mediatica della tortura*
73. La Grassa Gianfranco, *Tutto torna ma diverso. Capitalismo o capitalismi?*
74. Dalla Vigna Pierre, *La Pattumiera della storia. Beni culturali e società dello spettacolo*
75. Palumbo Antonino, Vaccaro Salvo (a cura di), *Governance e democrazia. Tecniche del potere e legittimità dei processi di globalizzazione*
76. Vaccaro Giovanbattista (a cura di), *Al di là dell'economico. Per una critica filosofica dell'economia*
77. Meattini Valerio, Pastore Luigi (a cura di), *Identità, individuo, soggetto tra moderno e postmoderno*
78. Dino Alessandra (a cura di), *Criminalità dei potenti e metodo mafioso*
79. Scolari Raffaele, *Filosofi e del mastodontico. Figure contemporanee del sublime della grande dimensione*
80. Trasatti Filippo, *Leggere Deleuze attraverso Millepiani*
81. Manicardi Enrico, *Liberi dalla civiltà. Spunti per una critica radicale ai fondamenti della civilizzazione: dominio, cultura, paura, economia, tecnologia*
82. Vaccaro Gianbattista, *Antropologia e utopia. Saggio su Herbert Marcuse*
83. Trasatti Filippo, Filippi Massimo (a cura di), *Nell'albergo di Adamo. Gli animali, la questione animale e la filosofia*
84. Franck Giorgio, *Il feticcio e la rovina. Società dello spettacolo e destino dell'arte*
85. Marzocca Ottavio (a cura di), *Governare líambiente? La crisi ecologica tra poteri, saperi e conflitti*
86. Grossmann Henryk, *Il crollo del capitalismo. La legge dell'accumulazione e del crollo del sistema capitalista*
87. Pullia Francesco, *Dimenticare Cartesio. Ecosofia per la compresenza*
88. Bazzanella Emiliano, *Religio I. Senso e fede nel tardocapitalismo*
89. Foucault Michel, *La società disciplinare*
90. Palano Damiano, *Volti della paura. Figure del disordine all'alba dell'era biopolitica*
91. Simone Anna, *I corpi del reato. Sessualità e sicurezza nelle società del rischio*
92. De Gaspari Mario, *Malacità. La finanza immobiliare contro la società civile*
93. Ruta Carlo, *Guerre solo ingiuste. La legittimazione dei conflitti e l'America dall'Vietnam all'Afghanistan*
94. Frazzetto Giuseppe, *Molte vite in multiversi. Nuovi media e arte quotidiana*

95. Bazzanella Emiliano, *Religio II. La religione del soggetto*
96. Brindisi Gianvito, de Conciliis Eleonora (a cura di), *Lavoro, merce, desiderio*
97. Casaccia Alessandro, *I paradossi della società competitiva*
98. Castanò Ermanno, *Ecologia e potere. Un saggio su Murray Bookchin*
99. d'Errico Stefano, *Il socialismo libertario ed umanista oggi fra politica ed anti-politica*
100. Tursi Antonio, *Politica 2.0. Blog, Facebook, YouTube, WikiLeaks: ripensare la sfera pubblica*
101. Lombardi Chiara, *Mondi nuovi a teatro. L'immagine del mondo sulle scene europee di Cinquecento e Seicento: spazi, economia, società*
102. Petrillo Antonello (a cura di), *Società civile in Iraq. Retoriche sullo "scontro di civiltà" nella terra tra i due fiumi*
103. Paolo Bellini, *Mitopie tecnopolitiche. Stato, nazione, impero e globalizzazione*
104. Palumbo Antonino, Segreto Viviana (a cura di), *Globalizzazione e governance delle società multiculturali*
105. Bertoldo Roberto, *Nullismo e letteratura. Al di là del nichilismo e del postmoderno debole. Saggio sulla scientificità dell'opera letteraria*
106. Ruggero D'Alessandro, *La comunità possibile. La democrazia consiliare in Rosa Luxemburg e Hannah Arendt*
107. Tessari Alessandro (a cura di), *Sindrome giapponese. La catastrofe nucleare da Chernobyl a Fukushima*
108. Bonazzi Matteo, Carmagnola Fulvio, *Il fantasma della libertà. Inconscio e politica al tempo di Berlusconi*, 2011
109. Mario De Gaspari, *La Bolla immobiliare. Le conseguenze economiche delle politiche urbane speculative*, 2011
110. Bruni Sara Elena Anna, Colavero Paolo, Nettuno Antonio (a cura di), *L'animale di gruppo. Etologia e psicanalisi di gruppo. Riflessioni gruppal da un seminario urbinato*, 2011
111. Segreto Viviana, *«Il padre di tutte le cose» Appunti per una pedagogia del conflitto*, 2011
112. Alessandra Dino (a cura di), *Poteri criminali e crisi della democrazia*, 2011
113. Serena Marcenò, *Biopolitica e sovranità. Concetti e pratiche di governo alle soglie della modernità*
114. Cosimo Degli Atti, *Soggetto e verità. Michel Foucault e l'Etica della cura di sé*
115. Pascal Boniface, *Verso la quarta guerra mondiale*
116. Guido Dalla Casa, *L'ecologia profonda. Lineamenti per una nuova visione del mondo*
117. *Il clown. Il meglio di Wikileaks sull'anomalia italiana*, introduzione di Marco Marsili
118. Carlo Grassi, *Sociologia della cultura tra critica e clinica. Bataille, Barthes, Lyotard*
119. Friedrich Georg Jünger, Ernst Jünger, *Guerra e guerrieri. Discorso*
120. Emma Palese, *Benvenuti a Gattaca. Corpo liquido, pedicopolitica, genetocrazia*
121. Anna Simone (a cura di), *Sessismo democratico. L'uso strumentale delle donne nel neo liberismo*
122. Matthew Calarco, *Zoografie. La questione dell'animale da Heidegger a Derrida*
123. Luigi Vergallo, *Economia reale ed economia sommersa nel riminese in prospettiva storica*

124. Salvo Vaccaro (a cura di), *L'onda araba. I documento delle rivolte*
125. Valeria Nuzzo, *L'immagine per il paesaggio e l'architettura. Percorsi didattici per la scuola*
126. Félix Guattari, *Una tomba per Edipo. Introduzione di Gilles Deleuze*
127. Raffaele Federici, *Sociologie del segreto*
128. Luca Taddio, *Global revolution. Da Occupy Wall Street a una nuova democrazia*
129. Enrique Dussel, *Indignados*
130. James Tobin, *Tobin Tax*
131. Jean-François Lyotard, *Istruzioni pagane*
132. Delfo Cecchi, *Cibo, corpo, narrazione. Sondaggi estetici*
133. Mario Giorgetti Fumel, Federico Chicchi (a cura di), *Il tempo della precarietà*
Sofferenza soggettiva e disagio della postmodernità
134. Spartaco Pupo, **Robert Nisbet e il conservatorismo sociale**
135. Giuseppina Tumminelli, **Strategie di ri-produzione. Aziende agricole e strutture familiari nella Sicilia centro-occidentale**
136. Iris Gavazzi, **Il vampiresco. Percorsi nel brutto**
137. Ferruccio Capelli, **Indignarsi è giusto**
138. Enrico Manicardi, *L'ultima era. Comparsa, decorso, effetti di quella patologia sociale ed ecologica chiamata civiltà*
139. Manuele Bellini, *Corpo e rivoluzione. Sulla filosofia di Luciano Parinetto*
140. Giovan Battista Vaccaro, *Le idee degli anni Sessanta*
141. Milena Meo, *Il corpo politico. Biopotere, generazione e produzione di soggettività femminili*
142. Massimiliano Vaghi, *L'idea dell'India nell'Europa moderna (secoli XVII-XX)*
143. Gianluca Cuozzo, *Mr. Steve Jobs. Sognatore di computer*
144. Paolo Cuttitta, *Lo spettacolo del confine. Lampedusa tra produzione e messa in scena della frontiera*
145. Emiliano Bazzanella, *Religio III. Logica e follia*
146. Emma Palese, *La filosofia politica di Zygmunt Bauman. Individuo, società, potere, etica, religione nella liquidità del nostro tempo*
147. Emma Palese, *Mostri, draghi e vampiri. Dal meraviglioso totalizzante alla naturalizzazione delle differenze*
148. Matteo Bonazzi, *Lacan e le politiche dell'inconscio. Clinica dell'immaginario contemporaneo*
149. Eleonora de Conciliis, *Il potere della comparazione. Un gioco sociologico*
150. *L'apartheid in Palestina. Il rapporto Human Rights Watch sui territori arabi occupati da Israele*
151. Fulvio Carmagnola, *Clinamen. Lo spazio estetico nell'immaginario contemporaneo*
152. Francesco Pullia, *Al punto di arrivo comune. Per una critica della filosofia del mattatoio*
153. Maurizio Soldini, *Hume e la bioetica*
154. Gianluca Cuozzo, *Gioco d'azzardo. La società dello spreco e i suoi miti*
155. Andrea Gilardoni, *Distruzioni. Potere & Dominio I*
156. Andrea Gilardoni, *(Dis)obbedienza. Meccanismi, strategie, argomenti. Potere & Dominio II*
157. Nicoletta Vallorani, *Millennium London, Of Other Spaces and the Metropolis*
158. Giuseppe Armocida, Gaetana S. Rigo (a cura di), *Dove mi ammalavo. La geografia medica nel pensiero scientifico del XIX secolo*

159. Salvo Torre, *Dominio, natura, democrazia. Comunità umane e comunità ecologiche*
160. Tindaro Bellinvia, *Xenofobia, sicurezze, resistenza. L'ordine pubblico in una città "rossa" (il caso Pisa)*
161. Amalia Rossi, Lorenzo D'Angelo (a cura di), *Antropologia, risorse naturali e conflitti ambientali*
162. Augusto Illuminati, *Teologia dei quattro elementi, Manifesto per un politeismo politico*
163. Giovanni Leghissa, *Neoliberalismo, Un'introduzione critica*
164. Anna Sica, Alison Wilson, *The Murray Edwards Duse Collection*
165. Stefano Cardini (a cura di), *Piazza Fontana. 43 anni dopo. Le verità di cui abbiamo bisogno*
166. Isacco Turina, *Chiesa e biopolitica. Il discorso cattolico su famiglia, sessualità e vita umana da Pio XI a Benedetto XVI*
167. Felice Papparo, *Perdere tempo*
168. Ugo Maria Olivieri, *Il dono della servitù. Étienne de La Boétie tra Machiavelli e Montaigne*
169. Giovanna D'Amia, *Milano e Parigi. Sguardi incrociati.*
170. Vittorio Morfino (a cura di) *Machiavelli: tempo e conflitto*
171. Andrea Gilardoni, *Potere potenziale*
172. Laura Sanò, *Donne e violenza*
173. Marilena Parlati, *Oltre il moderno. Orrori e tesori del lungo Ottocento inglese*
174. Damiano Palano, *La democrazia e il nemico*
175. Andrea Rabbito, *Il moderno e la crepa*
176. Pierre Dalla Vigna, *Estetica e ideologia*
177. Paola Gandolfi, *Rivolte in atto*
178. Chiara Simonigh (a cura di) *Pensare la complessità. Per un umanesimo planetario*
179. Carmelo Buscema, *L'epocalisse finanziaria. Rivelazioni (e rivoluzione) nel mondo digitalizzato*
180. Lidia Lo schiavo, *Governance Globale, Governamentalità, Democrazia*
181. Alessandra Vicentini, *Anglomanie settecentesche*
182. Francesco Saverio Festa, *Un'altra "teologia politica"?*
183. Daniela Calabrò, *L'ora meridiana. Il pensiero inoperoso di Jean-Luc Nancy tra ontologia, estetica e politica*
184. Mimmo Pesare, *Comunicare Lacan. Attualità del pensiero lacaniano per le scienze sociali*
185. Riccardo Ciavolella, *Antropologia politica e contemporaneità. Un'indagine critica sul potere presente*
186. Carlo Calcagno, *Impotenza. Storia di un'ossessione*
187. Marta Sironi, *Ridere dell'arte. L'arte moderna nella grafica satirica europea tra Otto e Novecento*
188. Gianpaolo Di Costanzo, *Assi mediani. Per una topografia sociale della provincia di Napoli*
189. Terrence Des Pres, *Il sopravvivate. Anatomia della vita nei campi di morte*, a cura di Adelmina Albini e Stefanie Golisch
190. Francesca Nicoli, *Giù le mani dalla modernità*
191. Leonardo Vittorio Arena, *La durata infinita del non suono*
192. Anselm Jappe, *Contro il denaro*

193. Giovanni Comboni, Marco Frusca, Andrea Tornago (a cura di), *L'abitare e lo scambio. Limiti, confini, passaggi*,
194. Gianluca Cuozzo, *Regno senza grazia. Oikos e natura nell'era della tecnica*
195. Elisa Virgili, *Ermafroditi*
196. Flavia Conte (a cura di), *Conversazioni sul postmoderno. Letture critiche del nostro tempo*
197. Alessandra MR D'Agostino, *Sesso mutante. I transgender si raccontano*
198. Gianfranco La Grassa, *L'altra strada. Per uscire dall'impasse teorica*
199. Paolo Mottana (a cura di), *Spacco tutto! Violenza e educazione*
200. Licia Michelangeli e Vittorio Ugo Vicari (a cura di), *Mode società e cultura nella Sicilia del secolo d'oro*
201. Roberto Bertoldo, *Istinto e logica della mente. Una prospettiva oltre la fenomenologia*
202. Giuseppe Raciti, *Ho visto Jünger nel Caucaso. Jonathan Littell, Max Aue e Ernst Jünger*
203. Furio Semerari (a cura di), *Etica ed estetica del volto*
204. Leonardo Grimoldi, *Storia e utopia. Saggio sul pensiero di Ignazio Silone*
205. Laura Bazzicalupo, *Dispositivi e soggettivazione*
206. Oscar Ricci, *Celebrità 2.0. Sociologia delle star nell'epoca dei new media*
207. Rosanna Castorina, Gabriele Roccheggiani, *Paradossi della fragilità. Critica della normalizzazione sociale, tra neuroscienze e filosofia politica*
208. Antonio Tursi, *Non solo cyber. Frammenti di un discorso mediologico*

