10. Giochi a informazione incompleta con aspettative soggettive mutuamente incoerenti

di Pierpaolo Battigalli

1. Introduzione

Un *gioco con informazione incompleta* è una situazione di interazione strategica tra individui, o giocatori, in cui alcuni dei parametri che governano l'interazione, incluse le preferenze individuali rispetto ai possibili esiti, non sono noti a tutti. L'analisi dei giochi con informazione incompleta risulta centrale in molti settori della teoria economica e in particolare nella teoria dell'organizzazione. La teoria della contrattazione e fenomeni quali l'insorgenza di forme di cooperazione e di reputazione, sono sempre più spesso studiati con riferimento a giochi dinamici con informazione incompleta [si veda ad esempio Fudenberg e Tirole 1991, capp. 6-10].

L'informazione incompleta non va confusa con l'*informazione imperfetta*, cioè la situazione in cui qualche giocatore può trovarsi a scegliere un'azione senza aver perfettamente osservato tutte le mosse precedenti, siano esse mosse casuali o di altri individui. Mentre il grado di imperfezione dell'informazione può essere descritto in modo non ambiguo nell'ambito di un modello matematico (ad esempio un gioco in forma estensiva), le ipotesi circa la in/completezza dell'informazione fanno soprattutto riferimento all'interpretazione di un certo modello matematico e guidano il discorso teorico sulla adeguatezza dei vari concetti di soluzione [1].

Secondo la teoria dei giochi classica il concetto fondamentale di equilibrio non cooperativo, l'equilibrio di Nash, risulta significativo se si assume che vi sia informazione completa [si veda per esempio Harsanyi 1967-68]. Lo stesso vale per altri concetti di soluzione non cooperativa che prescindono dalla condizione di equilibrio, quali la razionalizzabilità e la dominanza iterata [Bernheim 1984, Pearce 1984]. Infatti affinché un individuo possa «calcolare» la soluzione è necessa-

[1] Hanno uno *status* analogo le ipotesi sulla possibilità di comunicazione e di assunzione di impegni vincolanti prima del gioco.

rio che conosca tutti gli aspetti del gioco, compresi quelli che riguardano gli avversari [2].

Il contributo fondamentale di Harsanyi [1967-69] ha permesso di ricondurre l'analisi dei giochi ad informazione incompleta nell'alveo della teoria tradizionale mediante una specie di «trucco». Sia ω lo «stato del mondo», cioè la lista dei parametri che caratterizzano una situazione di interazione strategica $G(\omega)$. L'incertezza sulla situazione strategica $G(\omega)$ è caratterizzata da una misura di probabilità ρ sull'insieme degli stati del mondo Ω. L'informazione di ogni giocatore i sul «vero» stato del mondo è descritta da una partizione $\mathscr{P}_i$ di Ω: se lo stato del mondo è ω, allora i sa soltanto che il vero stato si trova nell'insieme $P_i(\omega) \subseteq \Omega$, cioè nella cella della partizione $\mathscr{P}_i$ contenente ω. Si consideri il seguente schema: lo stato del mondo ω è scelto a caso secondo la distribuzione di probabilità ρ, quindi ogni giocatore i osserva $P_i(\omega)$ e infine viene giocato il gioco $G(\omega)$. Da un punto di vista puramente formale, ciò corrisponde ad un ben definito gioco $\Gamma = ((G(\omega))_{\omega\in\Omega}, \mathscr{P}_i, \rho)$ con *informazione imperfetta* su una mossa casuale iniziale, quella che seleziona lo stato del mondo. Supponiamo che valgano le seguenti ipotesi informali:

a) (*conoscenza delle partizioni informative*) i giocatori conoscono la struttura $(\Omega, (\mathscr{P}_i))$;

b) (*distribuzione a priori comune*) tutti credono che, per ogni individuo i in ogni stato del mondo ω, la probabilità di ω condizionata all'informazione di i sia data da $\rho(\omega)/\rho(P_i(\omega))$ [3].

Allora è appropriato analizzare la situazione d'interazione strategica con informazione incompleta *come se* si trattasse del gioco con informazione imperfetta Γ, applicando ad esso i concetti di soluzione usuali [4].

È sempre necessario fare delle ipotesi informali restrittive per po-

[2] La recente letteratura sull'apprendimento in situazioni d'interazione strategica ripetuta mostra che sotto certe condizioni (essenzialmente l'osservabilità delle scelte alla fine di ogni periodo) questi concetti di soluzione sono rilevanti per valutare il comportamento al limite del sistema, anche se l'informazione è incompleta (per un'introduzione a questa letteratura si può consultare Binmore [1991, cap. 9], Fudenberg e Kreps [1992] e la rassegna di Battigalli *et al.* [1992]).

[3] Si supponga per semplicità che $\mathscr{P}_i$ sia un insieme discreto e che per ogni ω, $\rho(P_i(\omega)) > 0$.

[4] L'imperfezione dell'informazione crea dei problemi nel formulare adeguate condizioni di credibilità, risultando priva di forza la condizione di perfezione nei sottogiochi. Si ricorre allora a raffinamenti del concetto di equilibrio di Nash (si veda Van Damme [1987] quali gli equilibri sequenziali [Kreps e Wilson 1982] o gli equilibri perfetti bayesiani [Fudenberg e Tirole 1991]).

per estendere il campo predicativo di una potente teoria formale oltre il suo dominio originale[5]. Quindi non è possibile affermare a priori che le ipotesi *a*) e *b*) sono troppo forti. Tali ipotesi possono essere valutate solo in relazione alle specifiche applicazioni e alla possibilità di elaborare approcci alternativi. In questo lavoro si propone di mantenere l'ipotesi *a*) e rimuovere l'ipotesi *b*) della distribuzione a priori comune (d'ora in poi DCP).

In molte delle situazioni economiche con informazione incompleta usualmente analizzate con la teoria di Harsanyi l'ipotesi DCP ha poco senso. Si tratta di situazioni in cui non esiste un processo casuale con caratteristiche note che seleziona ω, né esistono statistiche affidabili sulla popolazione dei potenziali giocatori da cui derivare la distribuzione ρ secondo una regola comunemente accettata, né simmetrie ovvie che potrebbero determinare le probabilità rilevanti secondo un principio di ragione insufficiente. Arricchire la nozione di stato del mondo, includendovi anche la descrizione delle aspettative soggettive dei giocatori [Mertens e Zamir 1985; Brandenburger e Dekel 1993], non sembra essere una soluzione adeguata del problema. Quando l'oggetto dell'incertezza è costituito da caratteristiche psicologiche dei giocatori sembra molto più ragionevole assumere che le aspettative siano genuinamente soggettive ed eterogenee, quindi non derivabili da una distribuzione a priori comune sullo spazio degli stati del mondo. Quindi non è necessario sottoscrivere un approccio soggettivistico estremo alla teoria delle probabilità per dubitare della ragionevolezza dell'ipotesi DCP[6].

È certamente vero che la distribuzione a priori comune ρ è un espediente analitico particolarmente utile per descrivere in modo implicito e compatto le aspettative soggettive dei giocatori, le aspettative sulle aspettative altrui . . . e così via. Si potrebbe quindi pensare che la rimozione dell'ipotesi DCP imponga di inserire esplicitamente nell'analisi gerarchie infinite di aspettative su aspettative, ma non è così. I recenti contributi alla teoria dei giochi con informazione completa legati alla nozione di razionalizzabilità sono particolarmente significativi da questo punto di vista.

In un gioco con informazione completa, l'incertezza di un giocatore riguarda soltanto il profilo di strategie che verrà selezionato dagli avversari. Ma ciò non rende la situazione radicalmente diversa da

[5] Si pensi ad esempio all'estensione al caso dell'incertezza del modello di equilibrio generale di Arrow e Debreu, che richiede l'esistenza di mercati completi.

[6] Si rimanda a Gul [1991] per una critica più approfondita del modello bayesiano di Harsanyi.

quella con informazione incompleta. Se le aspettative dei giocatori sono soggettive e non riconducibili ad una comune distribuzione di probabilità sullo spazio delle strategie pure, allora sembrerebbe che ogni giocatore debba formarsi delle aspettative sulle aspettative dei suoi avversari, aspettative sulle aspettative sulle aspettative . . . fino ad ottenere una gerarchia infinita di aspettative. In effetti è possibile analizzare le gerarchie infinite di aspettative senza cadere in regressi infiniti di tipo «vizioso» [Mertens e Zamir 1985, Brandenburger e Dekel 1993]. Tali gerarchie infinite di aspettative possono essere utilizzate per caratterizzare gli esiti compatibili con l'ipotesi che la razionalità sia conoscenza comune [Tan e Werlang 1992][7]. Tale caratterizzazione è certamente interessante da un punto di vista concettuale. Ma esiste una caratterizzazione equivalente e molto più semplice basata su una procedura iterativa. Si comincia eliminando le strategie che non sono risposta ottima ad alcuna distribuzione di probabilità soggettiva. Poiché tali strategie sono certamente irrazionali, ognuno è sicuro che gli avversari non le sceglieranno. Ad ogni passo successivo si eliminano le strategie che non sono risposte ottime ad alcuna distribuzione di probabilità soggettiva che assegna probabilità nulla alle strategie eliminate nei passi precedenti. Le strategie che non sono mai eliminate, dette strategie *razionalizzabili*, sono esattamente quelle la cui scelta è compatibile con la conoscenza comune della razionalità [Bernheim 1984, Pearce 1984].

Entrambe le caratterizzazioni della conoscenza comune della razionalità, sia quella che utilizza le gerarchie infinite di aspettative, sia quella basata sulla procedura iterata, possono essere estese al caso dell'informazione incompleta, basta tenere conto che esistono molte (infinite) aspettative ammissibili sullo stato del mondo. Nel paragrafo 2 si fornisce una estensione della procedura iterata, poiché questa risulta più adatta all'analisi dei giochi dinamici.

Il seguente «gioco di segnalazione» è particolarmente utile per chiarire la differenza tra l'approccio qui proposto e quello tradizionale. Si consideri un'impresa che deve costruire un impianto per la produzione semi-automatica di una certa merce e che deve assumere un lavoratore *specializzato* che faccia funzionare l'impianto. Alternativamente l'impresa può realizzare un investimento aggiuntivo per automatizzare completamente la produzione. È conoscenza comune che:

[7] Una proposizione è di conoscenza comune se è vera, tutti sanno che è vera, . . . (tutti sanno che)[k] tutti sanno che è vera, . . .

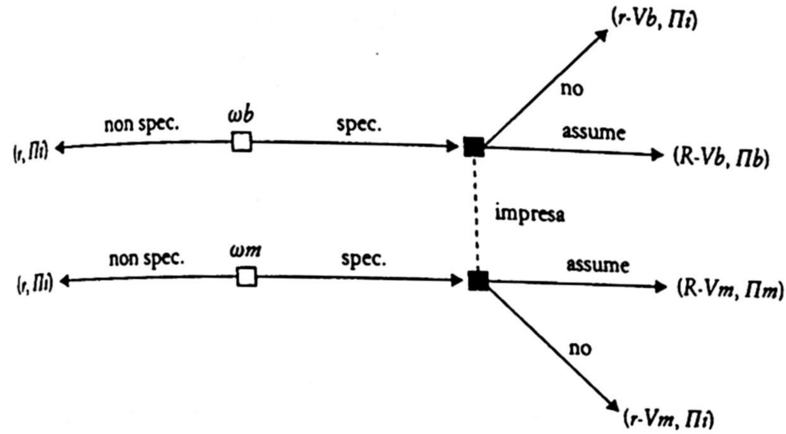

FIG. 10.1. Un gioco di segnalazione tra lavoratore e impresa.

a) ci sono due tipi di lavoratori (o stati del mondo): buoni (ω_b) e mediocri (ω_m);

b) ogni lavoratore può ottenere un lavoro non specializzato ad un salario r, mentre il salario per lavoratori specializzati è $R > r$;

c) i lavoratori buoni fanno molta meno fatica a specializzarsi dei lavoratori mediocri, in particolare valgono le disuguaglianze

$$R - V_b > r > R - V_m$$

dove V_t è l'equivalente monetario dello sforzo per specializzarsi di un lavoratore di tipo $\omega_t (t = b, m)$;

d) i lavoratori buoni sono quelli più produttivi; l'impresa può osservare il grado di specializzazione, ma non la produttività (cioè il tipo) dei lavoratori; il grado di specializzazione può essere interpretato come un segnale del tipo di lavoratore;

e) siano Π_b, Π_m, Π_i i profitti ottenuti dall'impresa rispettivamente assumendo un lavoratore buono, assumendo un lavoratore mediocre e facendo l'investimento addizionale invece di assumere, allora

$$\Pi_b > \Pi_i > \Pi_m.$$

La situazione è rappresentata nella figura 10.1.

La conclusione sembra abbastanza ovvia. Un lavoratore mediocre non si specializza perché nel migliore dei casi raggiunge una soddisfazione comunque inferiore a quella ottenuta con il lavoro non specializzato. Se l'impresa si trova di fronte ad un lavoratore specializzato, ne desume che deve trattarsi di un lavoratore buono e quindi lo assume.

Si noti che l'ipotesi DCP non gioca alcun ruolo in questo ragionamento. Invece secondo l'approccio tradizionale l'ipotesi DCP è anzitutto necessaria per trasformare la figura 10.1 in un gioco ben definito. Qualunque sia la distribuzione a priori ρ, la precedente soluzione corrisponde ad un equilibrio di Nash. Ma si può verificare che esiste anche un altro equilibrio, che soddisfa le usuali condizioni di razionalità sequenziale, in cui l'impresa non assume mai e quindi il lavoratore buono non trova conveniente specializzarsi. Per eliminare questo secondo equilibrio è necessario adottare «raffinamenti» potenti quali la «stabilità strategica» di Kohlberg e Mertens [1986].

Ma esiste un modo molto più parsimonioso di riprodurre formalmente il suddetto ragionamento, rimuovendo l'ipotesi DCP. Basta adottare l'artificio analitico secondo cui lo stato del mondo è scelto da una giocatrice fittizia del tutto indifferente all'esito, diciamo la Natura, e quindi applicare al gioco con tre persone così ottenuto la nozione di razionalizzabilità in forma estensiva di Pearce [1984]. L'indifferenza della Natura rappresenta l'ipotesi che qualsiasi aspettativa a priori sullo stato del mondo è ammissibile, perché qualunque scelta della Natura è razionalizzabile [Battigalli e Guaitoli 1988].

Questo esempio suggerisce due ordini di considerazioni:

i) l'analisi formale dei giochi con informazione incompleta è possibile anche se si rimuove l'ipotesi DCP e in alcuni casi può portare a risultati più forti rispetto a quelli ottenibili con alcuni dei concetti di equilibrio della teoria tradizionale;

ii) alcuni risultati della teoria dei giochi con informazione incompleta sono indipendenti dalle restrizioni imposte dall'analisi di equilibrio ed è interessante e metodologicamente corretto evidenziare tali risultati e le ipotesi fondamentali da cui discendono [cfr. Cho 1994, Watson 1993].

Il resto di questo lavoro è così organizzato. Nel paragrafo 2 è introdotto il modello generale di riferimento. Nell'ambito di tale modello viene definita una nozione di soluzione iterata chiamata ($\mathcal{F}$, Δ)-razionalizzabilità, dove $\mathcal{F}$ e Δ sono dei parametri che rappresentano restrizioni intersoggettive sulle aspettative ammissibili. Nel paragrafo 3 è analizzato a scopo illustrativo un particolare modello di gioco ripetuto, generalizzando dei risultati di Fudenberg e Levine [1989] e Watson [1993] sulla reputazione di un giocatore «di lungo perioche affronta una successione di oppositori «di breve periodo». Il paragrafo 4 offre alcune considerazioni conclusive.

La razionalizzabilità in giochi ad informazione incompleta con aspettative soggettive mutuamente incoerenti

In questa sezione viene formulata una definizione piuttosto generale di gioco con informazione incompleta senza distribuzione a priori comune e viene proposto un concetto di soluzione affine alla razionalizzabilità per giochi in forma estensiva [Pearce 1984, Battigalli 1993b].

2.1. Formulazione generale del modello

Seguendo Harsanyi [1967-68], considero un insieme di stati del mondo dato dal prodotto cartesiano $\Theta := \left(\prod_{i=1}^{n} \Theta_i \right)$ («$:=$» significa «uguale per definizione»), dove $n \leq \infty$, e Θ_i è l'insieme dei tipi possibili del giocatore i. Ogni vettore $\theta = (\theta_1, \theta_2, \ldots) \in \Theta$ corrisponde ad un gioco $G(\theta)$ con n (possibilmente un'infinità numerabile di) persone in forma estesa con *memoria perfetta*. I giochi statici costituiscono un caso particolare. Ogni giocatore i conosce il proprio tipo θ_i e ignora i tipi degli altri giocatori. Dunque θ_i rappresenta l'informazione privata di i sul gioco $G(\theta)$. Si assume per semplicità che soltanto le funzioni di payoff del gioco $G(\theta)$ dipendono da θ. Dunque le possibili sequenze di mosse e l'informazione acquisibile durante il gioco sulle mosse già avvenute sono indipendenti da θ e tali sono anche gli insiemi S_i delle strategie dei giocatori. Affinché sia significativo il concetto di soluzione che verrà proposto in seguito, bisogna supporre che la struttura informativa e l'applicazione $\theta \rightarrow G(\theta)$ siano conoscenza comune.

Sono qui opportune alcune considerazioni sul concetto di strategia. Nel caso di un gioco con informazione incompleta può esserci ambiguità sul significato del termine «strategia». Infatti una strategia è un piano contingente che prescrivere l'azione da eseguire secondo l'informazione a disposizione al momento dell'esecuzione. Se si immagina che esista uno stadio precedente al gioco, in cui non è ancora stato selezionato lo stato del mondo θ e in cui i giocatori contemplano i futuri sviluppi da «dietro un velo d'ignoranza», allora è corretto includere l'informazione privata poi acquisita sullo stato del mondo nel dominio delle strategie. Ma nel nostro contesto è più appropriata l'interpretazione secondo cui «il mondo comincia» in un particolare stato θ e i giocatori «nascono» con la loro informazione privata. Di conseguenza, nell'immaginare le possibili contingenze su cui basare le scelte, i giocatori considereranno soltanto le informazioni acquisibili sulle mosse precedenti, dando per scontata la loro informazione privata. Si deve inoltre tenere presente che S_i potrebbe rappresentare uno spazio di strategie opportunamente ristretto, ad esempio le strategie attuabili da automi finiti o le strategie con memoria (uniformemente) limitata.

Ad ogni gioco $G(\theta)$ corrisponde una forma strategica

$$(S_1, S_2, \ldots; U_1(\cdot; \theta), U_2(\cdot; \theta), \ldots)$$

dove ogni $U_i(\cdot; \theta)$ è una applicazione a valori reali con dominio $S: = \prod_{i=1}^{n} S_i$, che rappresenta la funzione di utilità Von Neumann-Morgenstern del giocatore i.

Ad ogni giocatore $i = 1, 2, \ldots$ corrisponde un insieme di stati d'informazione H_i («insiemi d'informazione» nella terminologia usuale dei giochi in forma estensiva), che costituisce il dominio delle sue strategie. Risulta utile includere in ogni H_i anche la «informazione nulla» $\emptyset$, cioè lo stato di informazione in cui non si è ancora osservato alcunché sulle mosse precedenti (risulterà chiaro dal contesto quando il simbolo $\emptyset$ è usato in questo senso e quando ha il significato usuale di insieme vuoto). Se in $G(\theta)$ il giocatore i attua delle scelte solo dopo avere osservato qualcosa, basta supporre che nello stato d'informazione $\emptyset$ egli abbia come unica scelta disponibile quella di «attendere». Per fissare le idee può essere opportuno considerare il caso particolare in cui $G(\theta)$ è un gioco multi-stadio con azioni osservate e quindi ogni $h \in H_i$ è una storia (una successione finita di azioni) osservata da i prima di prendere una decisione.

Ad ogni $h \in H_i$ corrisponde un «insieme d'informazione in forma strategica» $S(h) \subseteq S$, cioè l'insieme dei profili strategici la cui attuazione induce h. $S_i(h)$ e $S_{-i}(h)$ denotano le proiezioni $S(h)$ su S_i e $S_{-i}: = \Pi_{j \neq i} S_j$ rispettivamente. Per definizione $S(\emptyset): = S$. Si osservi che, in virtù dell'ipotesi di memoria perfetta, $S(h) = S_i(h) \times S_{-i}(h)$. Infine consideriamo la corrispondenza inversa a $S_i(h)$, indicando con $H_i(s_i): = \{h \in H_i | s_i \in S_i(h)\}$ l'insieme degli stati d'informazione di i la cui occorrenza non è preclusa dall'attuazione della strategia s_i. Tutte queste definizioni sono riportate nella seguente tabella riassuntiva.

Notazione	Terminologia
$1, 2, \dots, n \leq \infty$	giocatore i
$\theta_i \in \Theta_i$	il tipo del giocatore i
$\theta: = (\prod_{i=1}^{n} \Theta_i)$	lo stato del mondo
$\theta_{-i} \in (\prod_{j \neq i} \Theta_j)$	la componente dello stato θ non nota ad i
$G(\theta)$	il vero gioco nello stato θ
$s_i \in S_i$	strategia di i
$s \in S: = \prod_{i=1}^{n} S_i)$	profilo di strategie
$s_{-i} \in S_{-i}: = \prod_{j \neq i} S_j$	profilo di strategie degli avversari di i
$U_i(\cdot; \theta): (\prod_{i=1}^{n} S_i) \to \mathbb{R}$	funzione di utilità in forma strategica
$h \in H_i$	stato di informazione di i
$S(h) = S_i(h) \times S_{-i}(h)$	profili di strategie che inducono h
$H_i(s_i): = \{h \in H_i \mid s_i \in S_i(h)\}$	stati d'inform. di i non preclusi da s_i

Gli insieme $S(h)$ possono essere utilizzati per dare una rappresentazione in forma strategica delle aspettative condizionate dei giocatori quindi per formulare una definizione di razionalità sequenziale basata sulla funzione di utilità $U_i(\cdot;\ \theta)$.

È necessario introdurre delle ipotesi di tipo tecnico per garantire che i concetti formulati nel successivo paragrafo siano ben definiti. Il lettore non incline alla matematica può soprassedere, tenendo presente che tutti i concetti risultano ben definiti nel caso in cui i giochi $G(\theta)$ sono finiti.

Ipotesi 1. Per ogni $i = 1, 2, \dots$, Θ_i e S_i sono spazi metrici compatti, la funzione U_i: $S \times \Theta \to \mathbb{R}$ è continua su $S \times \Theta$ e per ogni $h \in H_i$, $S(h)$ è chiuso (si intende che ogni spazio prodotto è dotato della topologia prodotto corrispondente alle topologie su S_i, Θ_i, $i = 1, 2, \dots$).

2.2. Sistemi di probabilità condizionate e $(\mathcal{F}, \Delta)$-razionalizzabilità

Si assume che le aspettative dei giocatori siano rappresentate da sistemi di probabilità condizionate (si veda Rényi [1955] per una formulazione generale e Myerson [1986] per il caso finito). Ciò significa che i giocatori aggiornano le loro aspettative soggettive utilizzando la regola di Bayes ogniqualvolta sia possibile.

Per ogni spazio metrico M, $\mathscr{B}(M)$ indica la sigma-algebra di Borel (la minima sigma-algebra contenente tutti gli aperti di M) e $\Delta(M)$ indica l'insieme delle misure di probabilità con dominio $\mathscr{B}(M)$.

DEFINIZIONE 1. Sia M uno spazio metrico compatto. Un *sistema di probabilità condizionate* su M è una applicazione $\mu(\cdot|\cdot)$: $\mathscr{B}(M) \times \mathscr{B}(M) \rightarrow \mathbb{R}$ che soddisfa i seguenti assiomi:

[1.*i*] per ogni $A, B \in \mathscr{B}(M)$, $B \subseteq A$ implica $\mu(A|B) = 1$,

[1.*ii*] per ogni $A \in \mathscr{B}(M) \setminus \{\emptyset\}$, $\mu(\cdot|A) \in \Delta(M)$,

[1.*iii*] per ogni $A, B, C \in \mathscr{B}(M)$, $\mu(A \cap B | C) = $
$= \mu(A | B \cap C)\mu(B|C)$.

L'insieme dei sistemi di probabilità condizionate su M è indicato da $\Delta^*(M)$.

Osservazione. Il simbolo $\emptyset$ è qui usato con il significato usuale di insieme vuoto. Gli assiomi [1.*ii*] (normalizzazione) e [1.*iii*] (regola di aggiornamento) implicano che $\mu(A|B) = \mu(A \cap B|B)$ e $\mu(A|\emptyset) = 1$ per ogni $A, B \in \mathscr{B}(M)$. L'ultima uguaglianza è soltanto un'utile convenzione.

Se M è finito, si ottiene la nozione di sistema di probabilità condizionate introdotta nella teoria dei giochi da Myerson [1986].

Spesso è possibile derivare risultati forti sulle aspettative e sul comportamento dei giocatori da ipotesi di indipendenza stocastica che possono risultare naturali o almeno interessanti nel particolare contesto in esame. Le tipiche ipotesi d'indipendenza stocastica considerate nella teoria dei giochi riguardano le relazioni tra strategie e/o tipi di giocatori diversi. La seguente definizione fornisce un'utile formalizzazione della nozione intuitiva d'indipendenza stocastica con riferimento ai sistemi di probabilità condizionate.

DEFINIZIONE 2. Si consideri un prodotto cartesiano di spazi metrici compatti $M = \prod_{i=1}^{m} M_i$ $(m \leq \infty)$ e un sistema di probabilità condizionali $\mu \in \Delta^*(M)$. Sia $\mathscr{I} = \{I_1, I_2, \ldots\}$ una partizione dell'insieme indice $\{1, 2, \ldots, m\}$. Per ogni $I_j \in \mathscr{I}$ si definisca l'applicazione $\mu_j(\cdot|\cdot)$: $\mathscr{B}(\Pi_{i \neq I_j} M_i) \times \mathscr{B}(\Pi_{i \notin I_j} M_i)$, $\rightarrow \mathbb{R}$ mediante le seguenti uguaglianze:

[2.i] per ogni $A, B \in \mathscr{B}(\Pi_{i \in I_j} M_i)$, $\mu_{I_j}(A|B) = \mu(A \times (\Pi_{k \notin I_j} M_k)|B \times (\Pi_{k \notin I_j} M_k))$. μ_{I_j} è la *marginalizzazione* di μ su I_j. Il sistema di probabilità condizionate μ è $\mathscr{F}$*-indipendente* se

[2.ii] per ogni $A_j, B_j \in \mathscr{B}(\Pi_{i \in I_j} M_i)$, $C_{-j} \in \mathscr{B}(\Pi_{k \notin I_j} M_k)$,

$$\mu(A_j \times C_{-j}|B_j \times C_{-j}) = \mu_j(A_j|B_j).$$

L'insieme dei sistemi di probabilità condizionata $\mathscr{F}$-indipendenti su è indicato da $\mathscr{F}\Delta^*(M)$[8].

Osservazione. Ogni marginalizzazione di μ è un ben definito sistema probabilità condizionali.

In quanto segue si definisce l'insieme delle strategie razionalizzabiper ogni tipo θ_i e risulta perciò opportuno applicare la definizione 2 ponendo $M_j = (S_j \times \Theta_j)$ e $M = \prod_{j=1}^{n} (S_j \times \Theta_j)$ oppure $M = \Pi_{j \neq i}(S_j \times \Theta_j)$, dove il primo insieme prodotto è lo spazio d'incertezdal punto di vista di un osservatore esterno e il secondo insieme prodotto è lo spazio d'incertezza dal punto di vista di i.

Si assume che ogni possibile tipo θ_i sia dotato di un sistema di probabilità condizionate su $\Pi_{j \neq i}(S_j \times \Theta_j)$ che soddisfa certe restrizioe che scelga una strategia s_i che massimizza la sua utilità attesa condizionata ad ogni stato informativo h non precluso dall'attuazione s_i.

DEFINIZIONE 3. Una strategia $\hat{s}_i$ è razionale per il tipo $\hat{\theta}_i$ rispetto a $\hat{\Delta}_{-i} \subseteq \Delta^*(\Pi_{j \neq i}(S_j \times \Theta_j))$, scritto $(\hat{s}_i, \hat{\theta}_i) \in r_i(\hat{\Delta}_{-i})$, se esiste $\mu \in \hat{\Delta}_{-i}$ tale che

[3] per ogni $h \in H_i(\hat{s}_i)$, per ogni $s_i \in S_i(h)$,
$\int_{S_{-i} \times \Theta_{-i}} [U_i(\hat{s}_i, s_{-i}; \hat{\theta}_i, \theta_{-i}) - U_i(s_i, s_{-i}; \hat{\theta}_i, \theta_{-i})]\mu(ds_{-i}, d\theta_{-i}|S_{-i}(h) \times \Theta_{-i}) \geq 0$

Il prossimo passo consiste nel definire le strategie razionalizzabili per ogni tipo mediante una procedura iterativa. Essenzialmente si tratta

[8] Si veda Battigalli e Veronesi [1992] per una caratterizzazione di [2.ii] in termini di teoria delle decisioni e per riferimenti bibliografici su questa nozione d'indipendenza stocastica.

di un'estensione della definizione originale di Pearce [1984], formulata per giochi finiti in forma estensiva con informazione completa [9]. Questa procedura risolutiva formalizza un principio di «induzione in avanti» (*forward induction*): non solo ogni giocatore ritiene a priori (cioè nello stato informativo $\emptyset$) che i suoi avversari siano razionali; oltre a ciò, tutte le osservazioni sulle mosse avversarie vengono interpretate in base all'ipotesi che gli avversari siano razionali, ogniqualvolta ciò sia non contraddittorio. La forza del principio di induzione in avanti è ben illustrata dall'esempio introduttivo (figura 10.1). Anche se l'impresa fosse completamente sorpresa dalla scelta del lavoratore di istruirsi, essa assumerebbe che il lavoratore sia del tipo «buono», cioè l'unico tipo che può ritenere razionale istruirsi.

Mentre l'ipotesi che ogni giocatore si aspetti a priori un comportamento razionale da parte degli avversari costituisce un elemento essenziale di ogni teoria della razionalità strategica, il principio di induzione in avanti è molto più controverso. Qui si mantiene tale principio perché viene spesso invocato nell'analisi dei giochi con informazione incompleta, ma risulterà chiaro come rimuoverlo dall'analisi [10].

Il concetto di soluzione qui proposto considera come dati primitivi dell'analisi i seguenti elementi:

a) una partizione $\mathscr{F}$ dell'insieme dei giocatori, essendo inteso che è conoscenza comune che se due giocatori appartengono a celle diverse della partizione allora i loro tipi e le loro scelte strategiche sono mutuamente indipendenti in senso stocastico;

b) un insieme ristretto $\Delta \subseteq \mathscr{F}\Delta^*(\sum_{j=1}^{n} (S_j \times \Theta_j))$ di sistemi ammissibili di probabilità condizionali, che rappresenta uno standard comune, essendo inteso che è conoscenza comune che le aspettative di ogni giocatore i siano rappresentate dalla marginalizzazione μ_{-i} di un qualche sistema di probabilità condizionate $\mu \in \Delta$; in altri termini è conoscenza comune che per ogni i e ogni θ_i il corrispondente μ_{-i} appartiene alla proiezione di Δ su $\Delta^*(\Pi_{j\neq i}(S_j \times \Theta_j))$, cioè all'insieme

$$\begin{aligned}\Delta_{-i} := \{&\mu_{-i} \in \Delta^*(\Pi_{j\neq i}(S_j \times \Theta_j)) \mid \exists \mu \in \Delta,\ \forall A, B \in \\ &\mathscr{B}(\Pi_{j\neq i}(S_j \times \Theta_j)),\\ &\mu_{-i}(A|B) = \mu(A \times (S_i \times \Theta_i) \mid B \times (S_i \times \Theta_i))\}\end{aligned}$$

[9] Si rimedia inoltre ad un errore concettuale commesso da Pearce nel formalizzare l'ipotesi d'indipendenza [cfr. Battigalli 1993b].

[10] Battigalli [1993b] fornisce una discussione approfondita del principio di induzione in avanti unitamente ad una caratterizzazione delle aspettative razionalizzabili.

Osservazione. Sia $\mathscr{F}_{-i}$ la partizione ottenuta da $\mathscr{F}$ eliminando i, allora $\Delta_{-i} \subseteq \mathscr{F}_{-i}\Delta^*(\Pi_{j\neq i}(S_j \times \Theta_j))$.

Nella seguente definizione si evidenzia la dipendenza della soluzione da $\mathscr{F}$, anche se non sarebbe strettamente necessario dato che l'insieme Δ riassume tutte le restrizioni sulle aspettative, comprese le condizioni d'indipendenza stocastica.

DEFINIZIONE 4. Siamo dati $\mathscr{F}$ e Δ e si definiscano induttivamente gli insiemi $R_i(k, \mathscr{F}, \Delta)$ ($k = 0, 1, 2, \ldots, i = 1, 2, \ldots$) come segue:

[4.0] per ogni $i = 1, 2, \ldots, R_i\,(0; \mathscr{F}, \Delta) = S_i \times \Theta_i$

[4.k] per ogni $i = 1, 2, \ldots, k = 1, 2, \ldots$

a) $H_i(k-1; \mathscr{F}, \Delta) = \{h \in H_i \mid \exists(s, \theta) \in (\prod_{j=1}^{n} R_j(k-1, \mathscr{F}, \Delta)),$

$s \in S(h)\}$,

b) $\Delta_{-i}\,(k-1; \mathscr{F}, \Delta) = \{\mu \in \Delta_{-i} \mid$

$\forall h \in H_i(k-1, \mathscr{F}, \Delta),$

$\mu(cl(R_{-i}(k-1, \mathscr{F}, \Delta))) \mid S_{-i}(h) \times \Theta_{-i}) = 1\}$

(dove $cl(X)$ indica la chiusura di un insieme X),

c) $R_i(k; \mathscr{F}, \Delta) = R_i(k-1; \mathscr{F}, \Delta) \cap r_i(\Delta_{-i}(k-1; \mathscr{F}, \Delta))$.

Una strategia $\hat{s}_i$ è $(k, \mathscr{F}, \Delta)$-*razionalizzabile* per il tipo $\hat{\theta}_i$ se $(\hat{s}_i, \hat{\theta}_i) \in R_i(k, \mathscr{F}, \Delta)$, $\hat{s}_i$ è $(\mathscr{F}, \Delta)$-*razionalizzabile* per il tipo $\hat{\theta}_i$ se $(\hat{s}_i, \hat{\theta}_i) \in R_i(\infty, \mathscr{F}, \Delta) := \bigcap_{k=1}^{\infty} R_i(k, \mathscr{F}, \Delta)$.

L'equazione [4.*k*.*a*] definisce gli stati d'informazione del giocatore i che possono essere indotti da profili strategici $(k-1)$-razionalizzabili; la [4.*k*.*b*] dice che il giocatore i crede con probabilità uno che gli avversari attuino strategie $(k-1)$-razionalizzabili, ogni volta che tale credenza non contraddice il suo stato d'informazione; la [4.*k*.*c*] dice che una strategia è k-razionalizzabile se è $(k-1)$-razionalizzabile e se è sequenzialmente razionale per qualche θ_i rispetto a $\Delta_{-i}(k-1; \mathscr{F}, \Delta)$.

TEOREMA 1. Per ogni partizione $\mathscr{F}$, se $(S \times \Theta)$ è finito e $\Delta = \mathscr{F}\Delta^*(\prod_{j=0}^{n}(S_j \times \Theta_j))$, allora esiste un intero K^* tale che per ogni $k \geq K^*$ e ogni $i = 1, 2, \ldots, R_i(k; \mathscr{F}, \Delta) = R_i(\infty, \mathscr{F}, \Delta) \neq \emptyset$.

La dimostrazione di questo risultato può essere facilmente desunta dalle dimostrazioni contenute in Battigalli [1993b].

3. Un esempio: la reputazione di un giocatore «di lungo periodo» che affronta una successione di oppositori

Per esemplificare i concetti generali e astratti sopra esposti consideriamo un modello introdotto da Fudenberg e Levine [1989] [si veda anche Fudenberg e Tirole 1991, cap. 9] e successivamente generalizzato da Watson [1993]. In entrambi i casi l'analisi riguarda un gioco ad informazione incompleta in senso classico, ma Fudenberg e Levine considerano gli equilibri Nash-Bayesiani, mentre Watson considera una versione debole della razionalizzabilità. Qui si estende il campo predicativo dell'analisi di Watson eliminando l'ipotesi DCP e inoltre si evidenzia il ruolo di un'ipotesi di indipendenza stocastica di cui Watson si avvale implicitamente.

Sia (A_1, A_2, u_1, u_2) un gioco finito a due persone con mosse simultanee, dove u_i: $A_1 \times A_2 \rightarrow \mathbb{R}$ è la funzione di payoff del giocatore i. Consideriamo una situazione di interazione infinitamente ripetuta con azioni osservate, in cui il giocatore 1 affronta in successione delle «copie» $2_1, 2_2, 2_3 \ldots$ del giocatore 2, dove 2_n indica il giocatore che affronta 1 allo stadio n-esimo del gioco. Scopo dell'analisi è individuare delle condizioni generali sotto le quali il giocatore 1 si aspetta di ottenere nel lungo periodo un payoff da «leader» nel senso di Stackelberg costruendosi una opportuna reputazione.

Supponiamo, solo per semplicità, che per ogni azione a_1 esista un'unica risposta ottima del giocatore 2: $br_2(a_1)$: $=$ $\mathrm{argmax}_{a_2 \in A_2}[u_2(a_1, a_2)]$. Il *payoff di Stackelberg* del gioco costituente (A_1, A_2, u_1, u_2), indicato con u_1^*, è definito come segue:

[5] $$u_1^* := \max_{a_1 \in A_1} \{u_1[(a_1, br_2(a_1))]\}$$

Fissiamo una *azione di Stackelberg*, cioè una delle azioni che raggiungono il massimo in [5] e indichiamola con a_1^*. L'ipotesi fondamentale del modello è che i giocatori 2_n hanno informazione incompleta sui payoff di 1 e attribuiscono probabilità strettamente positiva un tipo θ^* per il quale scegliere sempre a_1^* è l'unico piano d'azione razionale. Nei termini del paragrafo 2, vale quanto segue. Posto $H^0 := \{\emptyset\}$ e $H^n := (A_1 \times A_2)^n$, $n = 1, 2, \ldots$, si ha $H_1 = \bigcup_{n \geq 0} H^n$, $H_{2_n} = H^{n-1}$, $S_1 = \{s_1 : H_1 \rightarrow A_1\}$, $S_{2_n} = \{s_2 : H_{2_n} \rightarrow A_2\}$. L'insieme degli stati del mondo coincide con l'insieme dei tipi del giocatore 1 (cioè $\Theta = \Theta_1$) e contiene almeno il tipo «normale» $\hat{\theta}$, corrispondente alla funzione di payoff u_1 e un tipo θ^* che ha interesse a scegliere sempre a_1^*[11]. Più precisamente, dato un profilo di strategie $s = (s_1, s_{2_1}, s_{2_2}, \ldots)$, indichiamo con $a^n(s) = (a_1^n(s), a_2^n(s))$ la coppia di azioni indotta allo stadio n-esimo; allora $U_{2_n}(s, \theta) = u_2(a^n(s))$ per ogni θ, $U_1(s, \hat{\theta}) = (1 - \delta) \sum_{\tau=1}^{\infty} \delta^{\tau-1} u_1(a^{\tau}(s))$, $U_1(s, \theta^*) = (1-\delta) \sum_{\tau=1}^{\infty} \delta^{\tau-1} v_1^*(a^{\tau}(s))$, dove $\delta \in (0, 1)$ è il tasso di sconto e $v_1^* : A_1 \times A_2 \rightarrow \mathbb{R}$ è una funzione di payoff tale che

$$[6] \qquad \min_{a_2 \in A_2} [v_1^*(a_1^*, a_2)] > \max_{a \in (A_1 \setminus \{a_1^*\}) \times A_2} [v_1^*(a)].$$

È banale verificare che, se vale la [6], per ogni sistema di probabilità condizionali del giocatore 1 le strategie sequenzialmente ottime per il tipo θ^* nel senso della [3] sono date dalla strategia s_1^* che prescrive sempre l'azione di Stackelberg e dalle strategie s_1 equivalenti nelle realizzazioni a s_1^*. Si tratta quindi delle strategie s_1 tali che per ogni $s_{-1} = (s_{2_1}, s_{2_2}, \ldots) \in S_{-1} := S_{2_1} \times S_{2_2} \times \ldots$ e per ogni n, $(a_1^n(s_1, s_{-1}), (a_2^n(s_1, s_{-1})) = (a_1^*, a_2(s_1^*, s_{-1}))$. Indichiamo con $[s_1^*]$ l'insieme delle strategie equivalenti nelle realizzazioni a s_1^* e con H^* l'insieme delle storie indotte da tali strategie, cioè quelle in cui 1 sceglie sempre l'azione di Stackelberg.

Rimangono da determinare le restrizioni sulle aspettative condizionate dei giocatori, cioè le proprietà dell'insieme Δ introdotto nel paragrafo 2. Sia $\mu(\theta^*) := \mu(S_1 \times \{\theta^*\} \times S_{-1} | S_1 \times \Theta \times S_{-1})$ la probabilità a priori di θ^*. Innanzitutto ipotizziamo che esista un $\varepsilon > 0$ tale che per ogni $\mu \in \Delta$, $\mu(\theta^*) \geq \varepsilon > 0$. Inoltre assumiamo che $\Delta \subseteq \mathcal{F}\Delta^* ((S_1 \times \Theta_1) \times S_{-1})$ con $\mathcal{F} = \{\{1\}, \{2_1, \ldots, 2_n, \ldots\}\}$, $\Theta_1 = \Theta$. In altri termini, è comune conoscenza che tutti i giocatori ritengono (s_1, θ) e s_{-1} stocasticamente indipendenti nel senso della [2]. Come dimostreremo in seguito, ciò implica che le scelte dei giocatori 2_n non influenzano le aspettative condizionate riguardanti il *tipo*

[11] Alternativamente si può supporre che θ^* sia un tipo che può scegliere solo a^*_1.

del giocatore 1 [12]. Questa ipotesi d'indipendenza stocastica non è esplicita nell'analisi di Watson [1993], ma si può dimostrare che senza di essa le conclusioni di Watson non sono valide.

L'ultima ipotesi riguarda la «dispersione» delle aspettative dei diversi giocatori 2_n condizionatamente a storie in cui il giocatore 1 sceglie sempre l'azione di Stackelberg. Sostanzialmente si richiede che tali aspettative condizionate diventino sufficientemente simili con il passare del tempo. Per formalizzare tale ipotesi definiamo una semimetrica [13] sull'insieme dei sistemi condizionali [cfr. Watson 1993]. Si indichi con $S_i(h^n, a_i) := \{s_i \in S_i(h^n) \,|\, s_i(h^n) = a_i\}$ l'insieme delle strategie di i che non precludono il realizzarsi della storia h^n e scelgono $a_i^{n+1} = a_i$, se si realizza h^n. Dato $\mu \in \Delta^* \; ((S_1 \times \Theta) \times S_{-1})$, sia $\mu(a_1^* \,|\, h^n) := \mu \; (S_1 \; (h^n, a_1^*) \times \Theta \times S_{-1} \,|\, S_1(h^n) \times \Theta \times S_{-1} \; (h^n))$ la probabilità che 1 scelga l'azione di Stackelberg, data h^n. Poniamo

$$[7] \qquad d(\mu', \mu'') := \mathrm{Sup}_{h^n \in H^*} \{ \,|\, \mu'(a^*_1 \,|\, h^n) - \mu''(a^*_1 \,|\, h^n) \,|\, \}$$

È facilmente verificabile che [7] definisce una semimetrica su $\Delta^*((s_1 \times \Theta_1) \times S_{-1})$. Assumiamo che Δ sia *compatto* rispetto a $d(\cdot,\cdot)$ [14].

Sotto queste ipotesi si può affermare che, se il tipo «normale» $\hat{\theta}$ del giocatore 1 si aspetta che gli oppositori giochino strategie $(2, \mathcal{F}, \Delta)$-razionalizzabili e sconta poco il futuro, allora il suo massimo payoff atteso di lungo periodo è limitato inferiormente da un valore vicino al payoff di Stackelberg u^*_1.

Per comodità ripetiamo qui la definizione di razionalizzabilità del paragrafo precedente adattandola (con una modifica non sostanziale) al presente modello. Si ricordi che, per un dato insieme di sistemi condizionali $\hat{\Delta}_{-i}$, $r_i \; (\hat{\Delta}_{-i})$ è l'insieme delle coppie (s_i, θ_i) che sono sequenzialmente razionali rispetto ad un qualche $\mu_{-i} \in \hat{\Delta}_{-i}$; $R_i(k, \mathcal{F}, \Delta)$ è l'insieme delle coppie (s_i, θ_i) che sono k-razionalizzabili date le restrizioni sui sistemi condizionali ammissibili. Nel presente contesto

[12] Si tratta dell'ipotesi «do not signal what you do not know» di Fudenberg e Tirole [1991] [cfr. Battigalli 1993a].

[13] Una semimetrica su uno spazio M è una funzione simmetrica d: $M \times M \rightarrow \mathbb{R}_+$ che soddisfa la disuguaglianza triangolare. A differenza di una metrica, una semimetrica non soddisfa necessariamente la proprietà $(\forall (x, y) \in M \times M, \; d(x, y) = 0 \Rightarrow x = y)$. Dato uno spazio semimetrico (M, d), si ottiene uno spazio metrico identificando con x tutti i punti y per i quali $d(x, y) = 0$. Tutti gli usuali concetti topologici sono formulabili rispetto a tale spazio metrico.

[14] Dall'analisi di Watson [1993] si evince facilmente come generalizzare questa ipotesi.

$(k;\mathcal{T}, \Delta) \subseteq S_1 \times \Theta_1 = S_1 \times \Theta$, $R_{2_n}(k;\mathcal{F}, \Delta) \subseteq S_{2_n}$. La definizione ricorsiva è data da:

$$H(k-1;\mathcal{F}, \Delta) = \{h \in H \mid \exists (s_1, \theta, s_{-1}) \in R_1(k-1, \mathcal{F}, \Delta)) \times$$
$$\times R_{-1}(k-1;\mathcal{F}, \Delta), (s_1, s_{-1}) \in S(h)\},$$
$$\Delta(k-1;\mathcal{F}, \Delta) = \{\mu \in \Delta \mid \forall h \in H(k-1;\mathcal{F}, \Delta), \mu(cl(R(k-1, \mathcal{F},$$
$$\Delta)) \mid S_1(h) \times \Theta_1 \times S_{-1}) = 1\},$$
$$R_i(k; \mathcal{F}, \Delta) = R_i(k-1;\mathcal{F}, \Delta) \cap r_i(\Delta_{-i}(k-1;\mathcal{F}, \Delta)).$$

TEOREMA 2. Sotto le ipotesi sopra elencate con riguardo al gioco con informazione incompleta e all'insieme Δ, esiste un intero positivo $K(\Delta)$ tale che, per ogni $\mu_1 \in \Delta_{-1}(2; \mathcal{F}, \Delta)$

$$\operatorname{argsup}_{s_1 \in S_1} \{\textstyle\int S_{-1} U_1(s_1, s_{-1}, \hat{\theta})\, \mu_{-1}(ds_{-1})\} \geq$$
$$\geq \delta^{K(\Delta)} u_1^* + (1 - \delta^{K(\Delta)}) \min_{a \in A_1 \times A_2} u_1(a)$$

I passaggi fondamentali della dimostrazione di questo teorema seguono le linee di Fudenberg e Levine [1989] e Watson [1993]. Tuttavia la dimostrazione viene qui riportata integralmente sia per completezza, sia per illustrare il ruolo delle restrizioni sui sistemi di probabilità condizionali (non utilizzati nei suddetti lavori). Tale ruolo è evidenziato soprattutto dal seguente risultato di inferenza statistica [15].

Si indichi con $\#X$ la numerosità di un insieme finito X. Per ogni storia $h^n = (a^1, a^2, \ldots, a^n)$ sia $h^t = (a^1, a^2, \ldots, a^t)$ $(t \leq n)$.

LEMMA 1. Sia $\mu \in \Delta^*((S_1 \times \Theta_1) \times S_{-1})$ tale che

i) $\mu \in \mathcal{F}\, \Delta^*((S_1 \times \Theta_1) \times S_{-1})$, con $\mathcal{F} = \{\{1\}, \{2_1, 2_2, \ldots\}\}$,

ii) $\mu(\theta^*) \geq \varepsilon > 0$,

iii) $\mu([s_1^*] \mid \theta^*) := \mu([s_1^*] \times \{\theta^*\} \times S_{-1} \mid S_1 \times \{\theta^*\} \times S_{-1}) = 1.$

Allora per ogni $\zeta \in (0, 1)$ e ogni $h^n \in H^*$

$$\#\{t < n \mid \mu(a_1^* \mid h^t) < \zeta\} < \frac{\log \varepsilon}{\log \zeta}$$

[15] Il corrispondente risultato in Fudenberg e Levine [1989] considera soltanto le strategie miste di equilibrio e gli eventi che hanno probabilità positiva secondo tali strategie e la distribuzione a priori sui tipi.

Dimostrazione. Si fissi una storia $h^n = ((a_1^*, a_2^1), \ldots, (a_1^*, a^n{}_2)) \in \mathrm{H}^*$.

Si osservi che, poiché $h^n \in H^*$, $[s_1^*] \subseteq S_1 (h^t, a_1^*) \subseteq S_1(h^t)$, $t = 0, 1, \ldots n$.

Dimostriamo anzitutto che

[8] $\forall t \in \{1, \ldots, n\}, \mu(a_1^* \mid h^{t-1}) \geq \varepsilon > 0.$

Infatti, applicando le ipotesi *i*) (vedi [2]), *ii*) e *iii*) si ottiene

$\mu(a_1^* \mid h^{t-1})\!: = \mu(S_1 (h^{t-1}, a_1^*) \times \Theta \times S_{-1} (h^{t-1}) \mid S_1 (h^{t-1}) \times \Theta \times S_{-1} (h^{t-1})) =$

$= \mu_1 (S_1 (h^{t-1}, a_1^*) \times \Theta \mid S_1 (h^{t-1}) \times \Theta) \geq \mu_1 ([s_1^*] \times \Theta \mid S_1 (h^{t-1}) \times \Theta \geq$

$\geq \mu_1 ([s_1^*] \times \Theta \mid S_1 \times \Theta) = \mu_1 ([s_1^*] \times \{\theta^*\} \mid S_1 \times \{\theta^*\}) \mu_1 (S_1 \times \{\theta^*\} \mid S_1 \times \Theta) \geq 1 \cdot \varepsilon > 0.$

Grazie all'ipotesi d'indipendenza *i*) si ottiene

[9] $\forall t \in \{1, \ldots, n\}, \mu(\theta^* \mid h^t) = \mu_1 (S_1 (h^{t-1}, a_1^*) \times \{\theta^*\} \mid S_1 (h^{t-1}, a_1^*) \times \Theta).$

Infatti:

$\mu(\theta^* \mid h^t)\!: = \mu(S_1 (h^{t-1}, a_1^*) \times \{\theta^*\} \times S_{-1} (h^{t-1}, a_2^t) \mid S_1 (h^{t-1}, a_1^*) \times \Theta \times S_{-1} (h^{t-1}, a_2^t)) =$

$= \mu_1 (S_1 (h^{t-1}, a_1^*) \times \{\theta^*\} \mid S_1 (h^{t-1}, a_1^*) \times \Theta.$

Applicando [8], [9], *i*), *iii*) e la regola di Bayes si ottiene

[10] $\forall t \in \{1, \ldots, n), \mu(\theta^* | h^t) = \mu(\theta^* | h^{t-1}) / \mu(a^*{}_1 | h^{t-1}).$

Infatti:

$\mu(\theta^* | h^t) = \mu_1(S_1(h^{t-1}, a_1^*) \times \{\theta^*\} | S_1(h^{t-1}, a_1^*) \times \Theta) =$

$= \mu_1(S_1(h^{t-1}, a_1^*) \times \{\theta^*\} | S_1(h^{t-1}) \times \Theta) / \mu_1(S_1(h^{t-1}, a^*{}_1) \times$

$\times \Theta | S_1(h^{t-1}) \times \Theta) =$

$$= \frac{\mu_1(S_1(h^{t-1}, a_1^*) \times \{\theta^*\} \mid S_1(h^{t-1}) \times \{\theta^*\})\mu_1(S_1(h^{t-1}) \times \{\theta^*\} \mid S_1(h^{t-1}) \times \Theta)}{\mu_1(S_1(h^{t-1}, a_1^*) \times \Theta \mid S_1(h^{t-1}) \times \Theta)} =$$

$$= \frac{\mu_1(S_1(h^{t-1}, a_1^*) \times \{\theta^*\} \mid S_1(h^{t-1}) \times \{\theta^*\})\mu(\theta^* \mid h^{t-1})}{\mu(a_1^* \mid h^{t-1})};$$

inoltre

$$\mu_1(S_1(h^{t-1}, a_1^*) \times \{\theta^*\} \mid S_1(h^{t-1}) \times \{\theta^*\}) \geq$$
$$\geq \mu_1([s_1^*] \times \{\theta^*\} \mid S_1(h^{t-1}) \times \{\theta^*\}) \geq \mu_1([s_1^*] \times \{\theta^*\} \mid S_1 \times \{\theta^*\}) = \mu([s_1^*] \mid \theta^*) = 1.$$

La [10] porge

$$[11] \quad \mu(\theta^* \mid h^n) = \mu(\theta^*)(\prod_{t=0}^{n-1} 1/\mu(a_1^* \mid h^t)) \geq \varepsilon(\prod_{t=0}^{n-1} 1/\mu(a_1^* \mid h^t)).$$

Dalla [11], posto $k = \# \{t<n \mid \mu(a_1^* \mid h^t) < \zeta\}$, si ottiene

$$1 \geq \mu(\theta^* \mid h^n) \geq \varepsilon(\prod_{t=0}^{n-1} 1/\mu(a_1^* \mid h^t)) > \varepsilon/\zeta^k.$$

Passando ai logaritmi si ottiene la tesi.

L'intuizione sottostante all'equazione [10], e quindi al lemma 1, è relativamente semplice. La probabilità condizionata dei tipi e delle strategie (non delle azioni) del giocatore 1 dipende solo dalle azioni osservate di 1 e non da quelle degli oppositori. Si assume che il tipo θ^* abbia probabilità a priori positiva e che scelga solo a_1^*. Dopo aver osservato a_1^* per $t-1$ periodi la probabilità di $a_1^t = a_1^*$ è positiva, perché non si può scartare l'ipotesi che il tipo vero sia θ^*. Se tale probabilità è bassa, ciò significa che si attribuisce una probabilità elevata ad un insieme di tipi che scelgono a_1^* nei primi $t-1$ priodi e un'azione diversa nel t-esimo periodo. Se si osserva a_1^* anche nel t-esimo periodo, la probabilità condizionata dei suddetti tipi si annulla e quindi la probabilità condizionata di θ^* aumenta in modo rilevante.

Dimostrazione del teorema 2 [cfr. Watson 1993]. Poiché $br_2(a_1^*)$ è l'unica risposta ottima contro a_1^*, esiste un valore critico $\rho \in (0, 1)$ tale che

[12] $\mu(a_1^* \mid h^n) > \rho \Rightarrow$
$$br_2(a_1^*) = \text{argmax}_{a_2 \in A_2} \{\Sigma_{a_1} \mu(a_1 \mid h^n)\, u_2\,(a_1, a_2)\}.$$

Sia $B_r\,(\hat{\mu})\text{:} = \{\mu \mid d(\mu, \hat{\mu}) < r\}$ la sfera aperta di raggio r centrata in $\hat{\mu}$ e si fissi ρ come nella [12]. Poiché $\Delta(1; \mathcal{F}, \Delta)$ è contenuto in un compatto, esistono $K(\rho)$ sistemi condizionali $\mu^1, \ldots, \mu^{K(\rho)} \in \Delta\,(1; \mathcal{F}, \Delta)$ tali che

[13] $$\Delta(1; \mathcal{F}, \Delta) \subseteq [B_{(1-\rho)/2}\,(\mu^1) \cup \ldots \cup B_{(1-\rho)/2}\,(\mu^{K(\rho)})]$$

Per la definizione di $d(\cdot, \cdot)$

[14] $$\forall h^n \in H^*, \forall k \in \{1, \ldots, K(\rho)\}, \forall \mu \in B_{(1-\rho)/2}\,(\mu^k),$$
$$\mu\,(a_1^* \mid h^n) < \rho \Rightarrow \mu^k\,(a_1^* \mid h^n) < (\rho + 1)/2$$

Si osservi che, per ogni $\mu \in \Delta(1; \mathcal{F}, \Delta)$ valgono le ipotesi del lemma 1. Quindi

[15] $$\forall h^n \in H^*, \forall k \in \{1, \ldots, K(\rho)\}, \ \# \{t < n \mid \mu^k(a_1^* \mid h^t) < (\rho + 1)/2\} < \frac{\log \varepsilon}{\log [(\rho + 1)/2]}$$

[14] e [15] implicano che

$$\# \{t < n \mid \exists \mu \in B_{(\rho-1)/2}\,(\mu^k), \mu(a_1^* \mid h^t) < \rho\} < < \frac{\log \varepsilon}{\log [(\rho + 1)/2]},$$

da cui si ottiene

[16] $$\forall h^n \in H^*, \ \# \{t < n \mid \exists \mu \in \Delta(1; \mathcal{J}, \Delta), \mu(a_1^* \mid h^t) < \rho\} < K(\rho) \log \varepsilon / \log [(\rho + 1)/2]\text{:} = K(\Delta),$$

[15] e [16] implicano

[17] $$\# \{t < n \mid \exists \mu \in \Delta)\,(2; \mathcal{J}, \Delta), \mu(br_2\,(a_1^*) \mid h^t) < 1\} < K(\Delta),$$

Si supponga che le aspettative del giocatore 1 siano date da $\mu \in \Delta\,(2; \mathcal{F}, \Delta)$ e venga attuata la strategia s_1^*. Grazie alla [17], la cosa peggiore che 1 si può aspettare è che per i primi $K(\Delta)$ periodi (cioè quelli scontati) venga scelta l'azione a_2 che risulta maggiormente dan-e che solo in seguito venga scelta $br_2(a_1^*)$. Da ciò segue la tesi teorema 2.

Si osservi che i risultati di questo paragrafo dipendono solo dalle restrizioni sulle aspettative condizionate e dall'ipotesi che *a priori* ci si aspetta un comportamento razionale da parte degli avversari. Il principio di «induzione in avanti» non ha qui alcuna rilevanza.

Conclusioni

In questo saggio ho proposto un nuovo approccio ai giochi con informazione incompleta, nel quale – contrariamente ad Harsanyi [1967-68] – non si assume che le aspettative dei giocatori sullo stato mondo, condizionate alla propria informazione privata, siano deducibili da una distribuzione a priori comune. Le aspettative condizionate dei diversi giocatori possono quindi risultare mutuamente incorenti (*inconsistent*) nel senso di Harsanyi.

In questo contesto risulta naturale rinunciare ad un'analisi basata sul concetto di equilibrio, richiamandosi invece alla nozione di razionalizzabilità. Avendo individuato quali elementi della situazione strategica sono di conoscenza comune, si caratterizzano le implicazioni della conoscenza comune di tali elementi e della razionalità dei giocatori. Il paragrafo 2 propone una formalizzazione generale ed astratta questo concetto di soluzione.

Tra gli elementi di comune conoscenza possono apparire delle restrizioni sulle aspettative condizionate dei giocatori, che sono un dato del modello. Tra queste restrizioni possono giocare un ruolo rilevante delle condizioni d'indipendenza stocastica tra le strategie e/o i tipi di giocatori diversi.

Il concetto di soluzione proposto nel paragrafo 2 incorpora un principio di «induzione in avanti», secondo il quale si cerca di dare una spiegazione razionale delle scelte osservate degli avversari ogniqualvolta ciò sia possibile. Un gioco di segnalazione presentato nell'Introduzione illustra la forza di tale principio.

Nel paragrafo 3 si analizza un gioco ripetuto con informazione incompleta per illustrare il ruolo delle restrizioni sulle aspettative condizionali, e in particolare dell'ipotesi d'indipendenza stocastica. Grazie a tali restrizioni e assumendo che esistano almeno due livelli di mutua conoscenza della razionalità, si ottengono interessanti conclusioni sugli «effetti di reputazione», che in precedenza erano state derivate nell'ambito della tradizionale analisi di equilibrio. L'analisi del paragrafo 3 prescinde dal principio di induzione in avanti.

Ho cercato di argomentare che un approccio ai giochi con informazione incompleta alternativo a quello tradizionale è auspicabile e può produrre risultati interessanti. Naturalmente soltanto l'applicazione di tale approccio ad una serie di modelli specifici potrà dare indicazioni sulla sua fecondità.

Riferimenti bibliografici

Battigalli, P. (1990), *Incomplete Information Games with Private Priors*, Istituto di Economia Politica, Università Bocconi, dattiloscritto.

— (1993a), *Strategic Independence and Perfect Bayesian Equilibria*, Rapporto Interno n. 93.011, Dipartimento di Economia e Produzione, Politecnico di Milano.

— (1993b), *Strategic Rationality Orderings and the Best Rationalization Principle*, Rapporto Interno n. 93.014, Dipartimento di Economia e Produzione, Politecnico di Milano.

Battigalli, P., Gilli, M., Molinari, M. C. (1992), *Learning and Convergence to Equilibrium in Repeated Strategic Interactions: an Introductory Survey*, in «Ricerche Economiche», 46, pp. 335-377.

Battigalli, P., Guaitoli, D. (1988), *Conjectural Equilibria and Rationalizability in a Macroeconomic Game with Incomplete Information*, Quaderni di ricerca n. 1988-6, Istituto di Economia Politica, Università Bocconi.

Battigalli, P., Veronesi, P. (1992), *A Note on Stochastic Independence without Savage Null Events*, Istituto di Economia Politica, Università Bocconi, dattiloscritto.

Bernheim, D. (1984), *Rationalizable Strategic Behavior*, in «Econometrica», 52, pp. 1007-1028.

Binmore, K. (1991), *Fun and Games: A Text on Game Theory*, Lexington, Mass, D.C. Heath.

Brandenburger, A., Dekel, E. (1993), *Hierarchies of Beliefs and Common Knowledge*, in «Journal of Economic Theory», 59, pp. 189-198.

Cho, I-K. (1994), *Stationarity, Rationalizability and Bargaining*, in «Review of Economic Studies», 61, pp. 357-374.

Damme, E. van (1987), *Stability and Perfection of Nash Equilibria*, Berlin, Springer Verlag.

Fudenberg, D., Kreps, D. (1992), *Lectures on Learning and Equilibrium in Strategic Form Games*, CORE Lecture Series, CORE Foundation, Louvain-La-Neuve.

Fudenberg, D., Levine, D. (1989), *Reputation and Equilibrium Selection in Games with a Patient Player*, in «Econometrica», 57, pp. 759-778.

Fudenberg, D., Tirole, J. (1991), *Perfect Bayesian Equilibria*, in «Journal of Economic Theory», 53, pp. 236-260.

(1991), *Game Theory*, Cambridge, Mass., The MIT Press.
F. (1991), *On the Bayesian View in Game Theory and Economics*, Research Paper 1991, Graduate School of Business, Stanford University.
Harsanyi, J. (1967-68), *Games with Incomplete Information Played by Bayesian Players. Parts I, II, III*, in «Management Sciences», 14, pp. 3, 5, 7.
Harstad, R. M., Phlips, L. (1993), *Future Market Contracting when You don't Know Who the Optimist Are*, Working Paper Eco n. 93-3, Istituto Universitario Europeo, Firenze.
Kohlberg, E., Mertens, J. F. (1986), *On the Notion of Stability of Equilibria*, in «Econometrica», 54, pp. 1003-1037.
Kreps, D., Wilson, R. (1982), *Sequential Equilibria*, in «Econometrica», 50, pp. 863-894.
Mertens, F., Zamir, S. (1985), *Formulation of Bayesian Analysis for Games with Incomplete Information*, in «International Journal of Game Theory», 14 pp. 1-29.
Myerson, R. (1986), *Multistage Games with Communication*, in «Econometrica», 54, pp. 323-358.
Pearce, D. (1984), *Razionalizable Strategic Behaviour and the Problem of Perfection*, in «Econometrica», 52, pp. 1029-1050.
Rényi, A. (1955), *On A New Axiomatic Theory of Probability*, in *Acta Mathematica Academiae Scientiarum Hungaricae*, 6, pp. 285-335.
– (1956), *On Conditional Probability Spaces Generated by a Dimensionally Ordered Set of Measures*, in *Theory of Probability and its Applications*, 1, pp. 61-71.
Tan, T., Werlang, S. (1988), *The Bayesian Foundation of Solution Concepts of Games*, in «Journal of Economic Theory», 45, pp. 370-391.
Watson, J. (1993), *A «Reputation» Refinement without Equilibrium*, in «Econometrica», 61, pp. 199-205.