

Il tredicesimo appuntamento con

LA TEORIA DEI GIOCHI

Repetere? L'ultima volta dei giochi ripetuti

di **Fioravante Patrone**

12

FIORAVANTE PATRONE, curatore di questa rubrica sulla Teoria dei Giochi, è docente presso la Facoltà di Ingegneria dell'Università di Genova ed è autore di *Decisori (razionali) interagenti*, un'introduzione alla Teoria dei Giochi,



Dopo gli ultimi articoli *Repetita iuvant?* sul n. 72 e *Repetere, repetere, ..., repetere. Siste gradum!* sul n. 73 della *Lettera*, con questa puntata la smettiamo (per il momento...) di occuparci di giochi ripetuti! Non perché non vi sarebbe altro da dire, ma per non esagerare con la monotonia di un singolo tema. Mi limiterò, in questo ultimo intervento, a parlare del Dilemma del Prigioniero (DP), esaminando casi "intermedi" fra un gioco finitamente ed uno infinitamente ripetuto.

In realtà, come vedremo, farò intervenire anche altre "dimensioni" del problema.

Ricordo che il DP è questo gioco in forma strategica:

$I \backslash II$	L	R
T	3 3	1 4
B	4 1	2 2

Comincio con il considerare un gioco di durata *aleatoria* molto semplice: dopo ogni turno, si estrae una pallina da un'urna che ne contiene 99 bianche ed una nera. Se si estrae quella nera, il gioco termina. Se se ne estrae una bianca, la si rimette dentro e si gioca un altro turno.

È facile calcolare la probabilità che vi siano almeno n turni: occorre che per $n-1$ volte sia stata estratta una pallina bianca. La probabilità di questo evento è:

$$\left(\frac{99}{100}\right)^{n-1}$$

che tende a zero al tendere di n all'infinito. Sappiamo quindi che, con probabilità pari a 1, il gioco avrà una durata finita. I giocatori, però, non sanno se il turno che stanno giocando è l'ultimo, oppure no. Questa mancanza di certezza sul termine del gioco ha l'effetto di demolire la tentazione di "fare il furbo" all'ultimo turno.

È abbastanza facile provare che una coppia di *trigger strategy*, come quella descritta per il DP infinitamente ripetuto, è un equilibrio di Nash, sotto condizioni che costituiscono la ragionevole generalizzazione di quelle apparse nel caso del DP infinitamente ripetuto. Assumiamo che le preferenze intertemporali dei giocatori siano adeguatamente descritte mediante un fattore di sconto v (potremmo anche ammettere due fattori di sconto diversi per i due giocatori: non cambierebbe nulla di essenziale) e assumiamo che la probabilità di continuare il gioco sia pari a p , costante e *common knowledge* fra i giocatori (nell'esempio dell'urna con le palline, $p = 99/100$).

La *trigger strategy* che consideriamo è la seguente (descritta per il giocatore I , per l'altro è analoga): I gioca T , fintantoché non osserva II giocare R . Se questo avviene, dal turno successivo e per sempre I gioca B . Allora, se I e II adottano queste strategie, il *payoff* atteso per I sarà:

$$3 + pv3 + p^2v^23 + \dots + p^n v^n 3 + \dots$$

ovvero:

$$\sum_{n=0}^{\infty} (pv)^n 3 = 3 \frac{1}{1 - pv}.$$

Se invece I devia, ad esempio al primo turno, il suo *payoff* atteso è *minore o uguale* a:

$$4 + pv2 + p^2v^22 + \dots + p^n v^n 2 + \dots$$

ovvero:

$$4 + \sum_{n=1}^{\infty} (pv)^n 2 = 2 + \sum_{n=0}^{\infty} (pv)^n 2 = 2 + 2 \frac{1}{1 - pv}.$$

Risulta quindi:

$$3 \frac{1}{1 - pv} \geq 2 + 2 \frac{1}{1 - pv}$$

se e solo se è $pv \geq 1/2$.

Questa è esattamente la stessa condizione trovata per un gioco infinitamente ripetuto, semplicemente usando un fattore di sconto "corretto" in modo da tener conto della probabilità che il gioco possa terminare. Insomma, in questo caso abbiamo un tranquillissimo passaggio al limite: il risultato che otteniamo per $p \rightarrow 1$ è esattamente quello che si ha per il gioco infinitamente ripetuto.

Come detto la scorsa volta, ci sono molti modi per "passare al limite". Tra i tanti possibili, ne vorrei menzionare un altro che trovo molto interessante ed anche specifico della

teoria dei giochi (TdG). Si tratta di far andare all'infinito il "livello di intelligenza" del giocatore! Ricordo che una ipotesi standard della TdG classica è quella di avere giocatori infinitamente intelligenti. Ci si può chiedere quindi cosa avvenga se abbiamo una qualche limitazione sull'intelligenza dei giocatori e cosa succede man mano che questo vincolo diventi meno stringente.

Ovviamente una questione non banale è: come modellizziamo questo fatto? Alcune strade sono state percorse. Una di queste consiste nel modellizzare i giocatori come automi a stati finiti (Neyman, Rubinstein): in questo caso, il passaggio al limite si effettuerà semplicemente aumentando il numero di stati a disposizione dell'automa.

Un altro approccio, per molti versi più *naïf* e anche meno esplicito, che viene usato per descrivere la "razionalità limitata" dei giocatori, è quello di descriverli come ottimizzatori approssimati (vedi Simon). Questo porta, ad esempio, a considerare come "soluzione" di un gioco, anziché l'equilibrio di Nash, il cosiddetto ϵ -equilibrio. Dato un gioco in forma strategica (X, Y, f, g) , una coppia $(\bar{x}, \bar{y}) \in X \times Y$ è un ϵ -equilibrio quando:

$$f(\bar{x}, \bar{y}) \geq f(x, \bar{y}) - \epsilon \text{ per ogni } x \in X$$

$$g(\bar{x}, \bar{y}) \geq g(\bar{x}, y) - \epsilon \text{ per ogni } y \in Y.$$

Naturalmente è $\epsilon \geq 0$; in particolare, per $\epsilon = 0$ si ottiene la definizione di equilibrio di Nash.

Da questo punto di vista, un interessante risultato è stato ottenuto da Radner il quale dimostra che in un "gioco" di oligopolio ripetuto un numero finito di volte, se si usa l' ϵ -equilibrio come soluzione, si può avere "cooperazione" (cioè: *collusione*) tra i giocatori per un certo numero di turni. Questo numero dipende naturalmente dai *payoff*, dal valore di ϵ ed eventualmente dal fattore di sconto, se presente.

C'è ancora un altro risultato che ci tengo a menzionare, risultato dovuto a Kreps, Milgrom, Roberts e Wilson. L'allentamento dell'ipotesi di infinita intelligenza viene da loro declinato in altro modo ancora. Questo contributo recupera, seppur in modo molto stilizzato, un suggerimento che proviene dalla copiosa letteratura disponibile sugli esperimenti di gioco con il DP ripetuto. Si osserva correntemente come, anche quando il numero di turni di gioco sia esplicitamente fissato e pubblicamente noto ai giocatori sin dall'inizio, i giocatori cerchino spesso di trovare un qualche modo di coordinarsi in modo da giocare le strategie "cooperative" T ed L per lo meno un "congruo"

numero di turni. Una spiegazione, fra le tante, di questi dati sperimentali consiste nel ritenere che i giocatori non corrispondano al modello di decisori razionali e intelligenti postulato dalla TdG classica o che, per lo meno, ci sia qualche dubbio su questo (ad esempio, un giocatore potrebbe pensare che l'altro non sia "infinitamente" intelligente o potrebbe ritenere che l'altro non lo reputi "infinitamente" intelligente). Una versione semplificata del modello proposto dai quattro autori summenzionati presuppone che:

- il giocatore *I* è il solito decisore razionale ed intelligente;
- il giocatore *II* idem con probabilità p mentre, con probabilità $1-p$, è una macchinetta pre-programmata per giocare la strategia TFT (*Tit For Tat* ovvero "pan per focaccia").

La strategia TFT, che ha acquisito notorietà a seguito degli esperimenti di Axelrod cui farò cenno in seguito, è molto semplice: il giocatore (supponiamo sia *II*), gioca:

- *L* al primo turno;
- *L* se ha osservato *I* giocare *T* al turno precedente;
- *R* se ha osservato *I* giocare *B* al turno precedente.

Credo che sia ovvio perché la si chiami così, in particolare perché si dica "pan per focaccia".

Gli autori citati mostrano che questo gioco ha un equilibrio in cui vengono giocate per un certo numero di turni (dipendenti dal valore del parametro p e dal numero totale di turni n) le strategie *T* ed *L*. L'aspetto più interessante è che, qualora il giocatore *II* sia il classico decisore razionale ed intelligente, avrà convenienza a giocare *L* per qualche turno! Perché? Per sfruttare l'incertezza di *I*, inducendolo a credere di avere di fronte a sé la macchinetta TFT: esempio limpidissimo di cosa significhi "reputazione" (buona o cattiva non importa) e di come sia importante crearsene una. Se necessario, *mentendo con i propri atti*, anche se in questo caso si tratta di "bugie a fin di bene" visto che entrambi i giocatori si avvantaggiano di questa opinione sbagliata di *I* su chi sia davvero *II*.

Avendo menzionato TFT, non posso non citare Axelrod e i suoi due "tornei". Egli invitò ricercatori di varie discipline a mandargli una strategia da giocare in un torneo in cui ad ogni incontro due giocatori avrebbero giocato il DP, ripetuto 200 volte. Fu proprio la semplicissima TFT, inviata dallo psicologo Anatol Rapoport, a risultare la strategia vincitrice. Ancor più sorprendente fu la seconda "edizione" in cui Axelrod ricevette 62 risposte (la prima volta ne aveva ricevute 14). Vinse ancora TFT, ancora inviata da Rapoport.

Questi risultati mostrano come, in una situazione "reale" di interazioni ripetute, possa esservi spazio per strategie che consentano ai giocatori di ottenere un *payoff* significativamente migliore di quello che otterrebbero giocando ad ogni turno la strategia dominante. Si tratta, come penso sia ovvio, di una tematica molto intrigante che va al cuore della definizione stessa di cosa sia la razionalità in un contesto di interazione strategica. Non a caso, questo tema è stato affrontato in molteplici contesti e sotto ipotesi molto diverse: incompletezza informativa, possibilità limitata di osservare o di ricordare cosa sia avvenuto nei turni precedenti, intervento di fattori aleatori esogeni, comunicazione esplicita tra i giocatori (o implicita, usando le mosse come segnali), gioco giocato in una "popolazione" di agenti ecc.

Come vedete, se ne potrebbe parlare per anni... ma ho fatto una promessa. Pertanto chiudo, notando a mo' di chiosa finale, come la semplice idea di gioco ripetuto possa essere declinata in molti modi diversi. Questa ricchezza si traduce in modelli matematici anche molto sofisticati, che permettono di esprimere in termini formali idee che potrebbero anche dare l'impressione di non poter essere traducibili in "matematiche": oltre alla *reputazione*, cito i termini *punizione*, *stare all'erta*, *altruismo*, *imitazione* e molti altri. ■

BIBLIOGRAFIA

- R. Axelrod, "Effective Choice in the Prisoner's Dilemma" in *The Journal of Conflict Resolution*, 24, pp. 3-25, 1980.
- R. Axelrod, "More Effective Choice in the Prisoner's Dilemma" in *The Journal of Conflict Resolution*, 24, pp. 379-403, 1980.
- D. Kreps, P. Milgrom, J. Roberts e R. Wilson, "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma" in *J. of Econ. Theory*, 27, pp. 245-252, 1982.
- A. Neyman, "Bounded Complexity Justifies Cooperation in the Finitely Repeated Prisoners' Dilemma" in *Economics Letters*, 19, pp. 227-229, 1985.
- R. Radner, "Collusive behavior in non-cooperative epsilon equilibria of oligopolies with long but finite lives" in *J. of Econ. Theory*, 22, pp. 121-157, 1980.
- A. Rubinstein, "Finite Automata Play the Repeated Prisoner's Dilemma" in *J. of Econ. Theory*, 39, pp. 83-96, 1986.
- H.A. Simon, *Models of Bounded Rationality*, Vol. 1 e 2, MIT Press, Cambridge, MA, 1982.