UNIVERSITÀ DI PISA

Decisioni in situazioni di complessità e conflitto

IL DILEMMA DEL PRIGIONIERO RIPETUTO

Tommaso Cavalieri [597707]

Laura Laid [486651]

Valeria Messina [516482]

Enrica Polidori [491265]

INDICE:

1. Introduzione *dilemma del prigioniero*

Il dilemma del prigioniero è un gioco ad informazione completa proposto negli anni cinquanta da Albert Tucker come problema di teoria dei giochi. Esso può essere descritto come segue. Due criminali vengono accusati di aver commesso un reato, gli investigatori arrestano entrambi e li chiudono in due celle diverse, impedendo loro di comunicare; essi hanno su di loro solamente prove per reati minori, punibili con 1 anno di detenzione, ma sospettano che abbiano commesso un crimine ben più grave, punibile con 5 anni di prigione. Ad ognuno dei criminali vengono date due scelte: **cooperare** oppure **non cooperare** con la polizia. Se uno dei due coopera e testimonia contro il suo partner, egli verrà liberato, condannando l'altro a 5 anni di prigione, ma se cooperano entrambi verranno puniti entrambi per il reato più grave, ma solamente a 3 anni vista la loro buona volontà. Risulta dunque evidente come ai due prigionieri converrebbe collaborare tra di loro e non cooperare con la polizia, essendo così condannati ad un anno ciascuno, ma entrambi sono tentati dal non collaborare tra di loro e confessare in quanto ciò potrebbe annullare la loro la pena. Risulta infatti che l'equilibrio di Nash, ovvero il profilo di strategie rispetto al quale nessun giocatore ha interesse ad essere l'unico a deviare, prevede che entrambi cooperino e vengano dunque condannati a 3 anni ciascuno.

1 \ 2	Cooperare	Non cooperare
Cooperare	3, 3	0, 5
Non cooperare	5, 0	1, 1

2. Giochi ripetuti finitamente ed infinitamente volte

Un gioco ripetuto **finitamente** volte è un gioco statico (*stage game*) che viene ripetuto un numero predefinito di volte, in ognuna delle quali ogni giocatore, prima di fare la sua scelta, osserva i risultati del periodo precedente. Se osserviamo un gioco di questo tipo, si può notare come, per induzione a ritroso (*backward induction*), sia T l'ultimo periodo nel quale si ripete il gioco, in T non ci sono incentivi a cooperare, poiché non ci sono perdite future di *payoff* di cui preoccuparsi. Così sarà anche in T-1, poiché già sapendo che non vi sarà cooperazione in T, non si ha alcun beneficio nel cooperare in T-1; procedendo così si può dedurre che non vi sarà cooperazione in alcun periodo. Nel caso in cui i due giocatori sappiano la durata fissata della ripetizione del gioco non vi è dunque alcuna differenza con la versione non ripetuta del medesimo gioco. Si aprono invece alcune possibilità differenti nel caso in cui la durata non sia certa, sia sconosciuta o il gioco duri infinitamente.

I giochi ripetuti un numero infinito di volte, sono chiamati anche "**supergiochi**", in essi ovviamente ogni giocatore ha vita finita e si assume che esso continui a giocare infinitamente. Ogni giocatore ha inoltre un **fattore di sconto δ** sulle future ricompense.

Tale fattore di sconto si può ottenere come $\delta = 1/(1+r)$, ipotizzando che r sia il tasso d'interesse offerto sul mercato (il guadagno che si avrebbe investendo questi soldi o depositandoli in banca). La ricompensa nel periodo corrente di ogni giocatore risulta dunque essere nient'altro che il **valore attuale delle ricompense future**. Assumendo che π_i sia il valore della ricompensa al tempo i, il valore attuale della ricompensa al tempo corrente (t=0) risulta essere pari a:

$$\pi = \pi_0 + \pi_1 * \delta + \pi_2 * \delta^2 + \pi_3 * \delta^3 + \cdots = \sum \pi_i * \delta^i$$

Se assumiamo che il valore di π_i sia costante nel tempo e dunque $\pi_i = \pi$ ad ogni i, risulta:

$$\pi + \pi * \delta + \pi * \delta^2 + \pi * \delta^3 + \cdots = \pi \frac{1}{1-\delta}$$

Si noti che il premio al periodo t=0 non dipende in alcun modo dal fattore di sconto. Essendo ragionevole assumere che $r>0$ e di conseguenza $0<\delta<1$, risulta che al crescere di i, δ^i diminuisce, ed essendo i premi futuri moltiplicati per questo fattore di sconto, anch'essi diminuiscono. Questo significa che i **premi più lontani nel futuro sono ritenuti sempre meno importanti dal giocatore che deve prendere delle decisioni oggi.**

Proviamo ora ad osservare un esempio concreto di questo tipo di gioco: **il dilemma del prigioniero ripetuto**. Nello specifico andiamo ad analizzare cosa succede nel caso in cui venga adottata da ogni giocatore la cosiddetta *trigger strategy*, ovvero la seguente:

1) Non confessare nel periodo 0;
2) continuare a non confessare fintanto che l'avversario non confessa;
3) se l'avversario tradisce e confessa, da quel momento in poi confessa in ogni fase del gioco fino alla fine.

Questa strategia, se venisse adottata da entrambe le parti, costituirebbe un accordo implicito a cooperare in ogni fase e punire il giocatore qualora non cooperasse, da qui il nome di strategia del "**dito sul grilletto" o strategia reattiva**, in quanto anche una sola deviazione innesca una 'punizione infinita'. **Dimostriamo ora che, se ciascun giocatore adotta questa strategia la coppia di strategie conseguente è un equilibrio di Nash** per il gioco ripetuto, purché il fattore di sconto δ assuma valori opportuni.

Stando alla strategia descritta sopra, in ogni fase del gioco i giocatori riceveranno il premio legato al non confessare, che indicheremo con ***π(NC, NC)***. Il valore attuale del premio sarà dunque $\pi(NC, NC)/(1-\delta)$. Se la strategia descritta è davvero un equilibrio di Nash i giocatori non dovrebbero avere alcun incentivo a deviare da essa, vogliamo dunque capire se tale incentivo esiste o meno. Supponiamo che il giocatore 1 stia considerando il fatto di confessare alla polizia e scegliere C nel periodo t. Da quel momento in poi l'altro giocatore adotterà la strategia di punirlo e scegliere sempre C, alla quale la risposta migliore risulterebbe proseguire con C all'infinito.

Per indicare le condizioni sotto le quali non è vantaggioso deviare, proviamo ad osservare le differenze tra il premio per il giocatore 1 nel caso in cui nessuno dei due giocatori confessi:

$$\pi = \sum_{i=0}^{\infty} \delta^i * \pi(NC, NC)$$

e nel caso in cui egli decide di deviare nel periodo t:

$$\pi = \sum_{i=0}^{t-1} \delta^i * \pi(NC, NC) + \delta^t * \pi(C, NC) + \sum_{i=t+1}^{\infty} \delta^i * \pi(C, C)$$

Si osservi che fino al periodo t queste due strategie danno lo stesso premio cumulato, in quanto prevedono le stesse azioni. Nel periodo t, tuttavia, i premi cambiano a causa della deviazione del giocatore 1. Esaminando i flussi dei premi in termini di valore attuale, si può affermare che la deviazione sarebbe vantaggiosa nel caso in cui risulti:

$$\delta^t * \pi(C, NC) + \sum_{i=t+1}^{\infty} \delta^i * \pi(C, C) > \sum_{i=t}^{\infty} \delta^i * \pi(NC, NC)$$

Sostituendo le sommatorie, tale equazione può essere riscritta nella seguente formula:

$$\delta^t * \pi(C, NC) + \delta^{t+1} * \frac{\pi(C, C)}{1-\delta} > \delta^t * \frac{\pi(NC, NC)}{1-\delta}$$

da cui si può ottenere, dopo qualche passaggio, che una deviazione è vantaggiosa se:

$$\delta < \frac{\pi(C, NC) - \pi(NC, NC)}{\pi(C, NC) - \pi(C, C)}$$

Si può dunque concludere che quanto più i giocatori scontano il futuro (ossia quanto minore risulta essere il fattore di sconto), tanto più è probabile che la cooperazione infinita non rappresenti una strategia di equilibrio nel gioco ripetuto infinitamente. **I giocatori infatti tenderanno a dare maggiore importanza al premio maggiore che riceveranno nel periodo in cui deviano t che del flusso infinito di premi minori che ne consegue**, siccome il valore attuale del primo è meno 'scontato' degli altri.

3. Definizione del gioco ripetuto

I giochi ripetuti sono una particolare categoria di giochi dinamici (ovvero quei giochi in cui si considera la possibilità che i giocatori intervengano, anche più volte, nel gioco secondo una sequenza prestabilita di mosse) che consistono nella ripetizione del medesimo gioco di base. Nella definizione di gioco ripetuto consideriamo di seguito il solo caso di giochi di base a mosse simultanee, poiché questo consente di semplificarne la presentazione, seppur con la perdita di qualche generalità.

Per comodità analitica, supponiamo che il gioco G(T) inizi al tempo 0 e prosegua fino al termine dell'orizzonte temporale predefinito, con una durata quindi di T+1 stadi (periodi).

Nel gioco ripetuto G(T) indichiamo con:

1. I è il numero dei giocatori (nel nostro caso I=2).
2. A è l'Insieme delle azioni che possono essere assunte dai giocatori a qualche punto del gioco: A = (A1, A2,..., AI) dove A_i è l'insieme delle azioni del giocatore i.
3. A_i^t l'insieme delle azioni del giocatore i nello stadio t del gioco; poiché tale insieme rimane invariato nei successivi stadi del gioco, si ha A_i^t = A_i, (t= 0,1,....,T); e con $A = A_i \times A_{-i}$ lo spazio delle azioni.
4. $a^\tau = \left(a_i^\tau, a_{-i}^\tau\right) \in A$ un profilo di azioni dei giocatori nel periodo τ = 0,1,..t-1
5. $h^t = \left\{\left(a^0, a^1, \dots, a^{t-1}\right)\right\}$ la storia del gioco costituita dalla successione dei profili di azione scelti dai giocatori nei successivi stadi del gioco; ovviamente h^t è un insieme vuoto
6. $s_i\left(h^t\right)$ la strategia del giocatore i nello stadio t, data la storia h^t formalmente, $s_i\left(h^t\right)$, è una regola d'azione che associa ad ogni possibile storia un'azione $a_i \in A_i$; nel nostro caso, Cooperare o Non Cooperare in funzione delle precedenti decisioni dei giocatori; una strategia pura s_i del giocatore *i* è quindi una successione di strategie pure $s_i\left(h^t\right)$ per tutti i possibili stadi del gioco; (s_i, s_{-i}) indica un profilo di strategie pure dei giocatori.
7. $U_i^T\left(s_i, s_{-i}\right) = \sum_{t=0}^{t} \beta^t u_i^t\left(a_i\left(h^t\right), a_{-i}\left(h^t\right)\right)$, il payoff del giocatore i definito come somma dei payoff $u_i\left(a_i\left(h^t\right), a_{-i}\left(h^t\right)\right)$ dei successivi stadi del gioco, attualizzati al presente al fattore di sconto per unità di tempo. Possiamo dare tre interpretazioni del fattore di sconto. La prima si fonda sulla considerazione, tradizionale nella teoria economica, che una somma di denaro disponibile nel futuro vale meno della corrispondente somma disponibile nel presente e vada quindi scontata al tasso di interesse di

mercato i: quindi $\beta = \frac{1}{1+i}$. La seconda, che si adatta ad una situazione in cui i payoff sono intesi come utilità, fa di β un fattore di sconto che esprime le preferenze intertemporali di un agente: quindi $\beta = \frac{1}{1+\theta}$. In questo caso possiamo considerare β come una misura del grado di impazienza del giocatore. La terza fa riferimento alla possibilità che il gioco possa non ripetersi nel futuro e traduce tale incertezza attraverso la probabilità *p* di continuazione del gioco nello stadio successivo. In entrambi i casi il fattore di sconto β è un numero strettamente compreso tra zero ed uno.

Possiamo a questo punto, riprendendo la Definizione precedente, dare una definizione formale di gioco ripetuto.

Definizione : Dato il gioco di base a mosse simultanee G, il medesimo gioco ripetuto T+1 volte, in cui gli esiti degli stadi precedenti del gioco (storia del gioco) sono noti prima che abbia inizio lo stadio successivo, è il gioco

$$G(T) = \left\{ I, A_i, h^t, s_i, U_i = \sum_t \beta^t u_i\left(a_i\left(h^t\right), a_{-i}\left(h^t\right)\right) \right\}_{i=1,t=0}^{I \quad T}$$

4. The Evolution of Cooperation, *Axelrod 1984*

4.1 Il concetto di cooperazione e strategia

Gran parte delle teorie economiche degli ultimi due secoli sono state costruite sull'idea che l'essere umano compia delle scelte razionali e finalizzate alla massimizzazione del proprio bene. Tuttavia, nel 1984 il politologo Axelrod si interroga sulla possibilità che le azioni umane possano creare azioni cooperative ed in particolare "sotto quali condizioni la cooperazione può emergere in un mondo di egoisti senza un'autorità centrale?".

Queste dinamiche sono state studiate dal politologo Axelrod nel suo ***The Evolution of Cooperation*** **(Axelrod, 1984)**. Proprio utilizzando il gioco del dilemma del prigioniero nella sua versione ripetuta, Axelrod ha cercato di studiare le dinamiche della cooperazione, cioè come possano nascere comportamenti cooperativi in situazioni in cui gli individui agiscono perseguendo i propri interessi. Axelrod ha osservato che se in un gioco con una sola occasione, oppure anche con un numero di occasioni finito e noto in anticipo, non c'è un incentivo alla cooperazione (abbiamo già visto prima che la scelta 'razionale' sarebbe quella di non cooperare), le cose cambiano nel caso che il gioco venga ripetuto un numero infinito di volte, oppure semplicemente che venga ripetuto più volte, senza però che i giocatori sappiano in anticipo quante. In questo caso possono emergere delle dinamiche che portano alla cooperazione. Obiettivo degli studi di Axelrod era cercare di capire "*sotto quali condizioni la cooperazione possa emergere in un mondo di egoisti in assenza di una autorità centrale*" guardando e confrontando i comportamenti adottati nella biologia. Egli infatti aveva osservato come l'esistenza di situazioni di cooperazione "spontanea", cioè senza la presenza di una autorità esterna che la imponesse, fosse una esperienza abbastanza comune.

4.2 La strategia TIT FOR TAT

Al fine di comprendere quali dinamiche si nascondessero dietro la nascita di azioni cooperative, Axelrod fece un esperimento invitando un certo numero di ricercatori a presentare strategie per il dilemma del prigioniero ripetuto. Queste strategie vennero confrontate fra di loro in una sorta di torneo, al fine del quale la strategia che risultava essere la vincitrice fu quella posta da Anatol Rapoport che sembrava essere di facile intuizione. Questa strategia che da Rapoport fu chiamata "Tit for Tat" (TFT) si potrebbe tradurre con le seguenti due regole:

1. Alla prima iterazione coopera;
2. Ad ognuna delle iterazioni successive copia la mossa fatta dal tuo avversario nell'interazione precedente.

La strategia TFT risulto' essere la migliori in diversi esperimenti che vennero condotti e la conclusione di Axelrod fu che in effetti comportamenti cooperativi possono emergere e mantenersi anche fra individui egoisti, purché si tratti di individui che vivano in un contesto sociale in cui ci siano più occasioni di incontro, cioè in situazioni in cui la probabilità che due individui si incontrino di nuovo sia sufficientemente grande. Nell'analizzare i risultati Axelrod cercò di individuare le caratteristiche comuni alle strategie che avevano prodotto nei suoi esperimenti i migliori risultati, che sintetizzò in alcune proprietà.

5. Simulazione del modello di Axelrod e proprietà

Giocatore B **Giocatore A**	**C** **Cooperazione**	**D** **Defection**
C **Cooperazione**	**R = 3** Premio per reciproca cooperazione	**S = 0** Payoff dello stupido
D **Defection**	**T = 5** Tentativo di defezione	**P = 1** Punizione per reciproca defezione

Il gioco descritto nella matrice sopra è definito da T > R > P > S e R > (S+T)/2 .

I valori numerici sono i payoff del Giocatore A.

Il modello è basato su una più realistica assunzione che il numero di interazioni non è fissato a priori, ma ci sono delle **probabilità *w* che**, dopo la corrente interazione, gli stessi 2 **individui si incontreranno ancora.**

w: probabilità che i due individui dopo la corrente iterazione si incontrino ancora.

include mobilità relativa, la durata della vita, la salute degli individui…

Per ogni valore di w, la strategia ALL D, la *"strategia della defezione incondizionata"* è la strategia **evolutivamente stabile.** Una strategia si dice stabile evolutivamente stabile se tutti usano quella strategia e nessuna strategia alternativa può invadere la popolazione.

ALL D "strategia della defezione incondizionata": strategia in cui il giocatore sceglie sempre di non cooperare con la polizia (D sta per "defeated", in inglese cambiare alleanza)

Ma questo non significa che non ci possano essere altre, altrettanto evolutivamente stabili. Infatti **quando w è sufficientemente grande, non esiste una sola strategia migliore nel comportamento degli altri nella popolazione.**

Per tale motivo oltre alla stabilità, dimostreremo che data una certa strategia, essa gode delle proprietà di robustezza e di vitalità iniziale. Prima di spiegare tali proprietà, bisogna evidenziare che il nostro modello, per motivi analitici, assume che le **scelte** tra gli individui avvengano **simultaneamente** (ovvero non sequenziali), e in un intervallo di tempo discreto, ovvero le scelte sono continue interazione nel tempo.

L'evoluzione della cooperazione può essere concettualizzata in **3 quesiti:**

- *Stabilità*: in quali condizioni una strategia di questo tipo potrà resistere all'invasione delle altre strategie?
- *Robustezza* : quale strategia in un ambiente variegato, cioè con la presenza di altre strategie, può resistere?
- *Vitalità iniziale*: come può una strategia cooperativa, ammesso che sia robusta e stabile, essersi creata all'interno di un ambiente prevalentemente non cooperativo?

Strategia TIT FOR TAT: strategia descritta dalle seguenti due regole:
1. *Alla prima iterazione l'individuo sceglie di cooperare con la polizia (C)*
2. *Ad ognune delle interazioni successive esso copia la mossa fatta dall'avversario nell'iterazione precedente (simultaneamente C o D)*

Stabilità

La stabilità evolutiva di una strategia riguarda le condizioni in cui una strategia potrà resistere all'invasione delle altre, una volta che si è insediata.

Una volta che TIT FOR TAT si è stabilita, può resistere all'invasione di qualsiasi possibile altra strategia, a condizione che gli individui che interagiscono abbiano una probabilità sufficientemente grande, w, di incontrarsi di nuovo. La prova è descritta nei paragrafi successivi.

Come primo passo nella prova notiamo che, poiché TIT FOR TAT **"ricorda" solo una mossa indietro;** un C dell'altro giocatore in qualsiasi round è sufficiente per ripristinare la situazione com'era all'inizio del gioco. Allo stesso modo, un D riporta la situazione come era al secondo turno dopo che un D era stato giocato nel primo. Poiché c'è una possibilità fissa, w, che l'interazione non finisca in una determinata mossa, una strategia non può essere ottima nel giocare con TIT FOR TAT a meno che non faccia la stessa cosa sia alla prima occorrenza di un dato stato sia ad ogni reset di quello stato.

Una regola si dice massimale se ha un payoff maggiore ovvero risulta la strategia migliore in un determinato momento per un determinato giocatore rispetto ad altre.

Dunque, se una regola è massimale e inizia con C, il secondo stato ha lo stesso stato del primo, una regola massimale continuerà con C e quindi coopererà sempre con TIT FOR TAT. Ma la stessa regola non farà meglio di quanto TIT FOR TAT fa con un altro TIT FOR TAT, e quindi non può invadere. Se, da una parte, una regola inizia con D, il primo D induce un cambiamento nello stato di TIT FOR TAT e ci sono due possibilità per il continuo che potrebbero essere massimali. Se D segue il "primo D", dopodichè questo diventerà massimale all'inizio implicando che è ovunque massimale da seguire D con D, facendo diventare la strategia equivalente a TUTTE LE D. Se C segue le D iniziali, il gioco si resetta come nella prima mossa, dunque deve essere massimale ripetere le sequenze DC indefinitamente. Questi punti mostrano che lo scopo di cercare un apparentemente infinito array di regole di comportamento per uno potenzialmente in grado di invadere TIT FOR TAT e più semplice di come si pensi.

Schema riassuntivo degli scenari della strategia TIT FOT TAT
<u>*Il gioco inizia con C*</u>*: resetta il gioco e ripristina la situazione allo stato iniziale* *=> segue DC simultaneamente*
<u>*Il gioco inizia con D:*</u> *cambio dello stato, due possibili scenari.* *A. D è seguito da D => strategia equivalente ad ALL D.* *B. D è seguito da C => si resetta lo stato e seguirà la strategia DC simultaneamente.*

Se né TUTTE LE D, né l'alternanza di D con C può invadere TIT FOR TAT, allora nessuna strategia può farlo. Per vedere se queste strategie possono invadere, notiamo che la probabilità che l'ennesima interazione occorra è w^{n-1}. Di conseguenza, l'espressione per il payoff totale è facilmente raggiungibile applicando i pesi 1, w, w^2...alla sequenza di payoff e sommando la serie risultante.

Quando TIT FOR TAT gioca un altro TIT FOR TAT (quando si resetta il gioco), ottiene un payoff di R ogni mossa per un totale di

R + wR + w^2R..., ovvero R / (1 - w).

ALL D che giocano con TIT FOR TAT, ottengono T alla prima mossa e P da allora in poi, così non può invadere TIT FOR TAT se

R / (1 - w) >= T + wP / (1 - w).

Allo stesso modo con **l'alternanza di D e C** TIT FOR TAT, ottiene un payoff di

T = wS + w^2T + w^3S …

T = (T + wS) / (1 - w^2)

Alternando D e C, cosicchè non può invadere TIT FOR TAT:

R / (1 - w) >= (T + wS) / (1 - w^2)

Con riferimento alla grandezza di w, troviamo che **nessuna di queste due strategie** (e dunque nessuna strategia in assoluto) **può invadere TIT FOR TAT,** se e solo se:

w >= (T - R) / (T - P) e w >= (T - R) / (R - S).

Questo dimostra che TIT FOR TAT è in evoluzione stabile se e solo se le interazioni tra gli individui hanno sufficientemente grandi probabilità di continuare.

Robustezza

Axelrod affronta il problema con un ingegnoso metodo di simulazione. Dopo avere invitato alcuni specialisti della teoria dei giochi a presentare una strategia per il gioco iterativo, codificata in un appropriato programma, organizza un torneo computerizzato, facendo interagire ripetutamente tra loro le quattordici strategie proposte più una strategia randomizzata.

Con sua grande sorpresa la più semplice fra tutte le strategie, vale a dire **Tit For Tat ("colpo su colpo")**, presentata da Anatol Rapoport, vince il torneo sbaragliando gli avversari. Alla prima mossa Tit For Tat (d'ora in poi, TFT) coopera, e in ogni mossa successiva sceglie fra C (cooperare) e D (defezionare) semplicemente riproducendo la scelta effettuata dall'avversario nella mossa immediatamente precedente. Questo significa che TFT è caratterizzata dal **massimo grado di 'reciprocità**' e, quindi, dalla massima disponibilità a cooperare con altre strategie che si dimostrino inclini alla cooperazione. Un risultato ancora più sorprendente viene ottenuto da Axelrod con un secondo torneo computerizzato al quale sono ammesse sessantadue strategie proposte da studiosi a conoscenza dei risultati del primo girone. Nonostante alcune di queste strategie siano ispirate dallo sforzo consapevole di perfezionare TFT, **la versione originale di TFT vince anche il secondo girone.** Il successo di TFT nel secondo girone, contro una così grande varietà di avversari, suggerisce l'ipotesi che TFT 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.

Il grado di robustezza di una strategia: misura quanto tempo tale strategia è in grado di invadere le decisioni e resistere nel tempo fino al tentativo di invasione da parte di altre strategie.

Per determinare il '**grado di robustezza**' di TFT e di strategie dello stesso genere, fondate su criteri di reciprocità e disponibilità a cooperare Axelrod effettua svariati esperimenti di simulazione basati sui modelli dinamici della teoria evoluzionistica dei giochi. Questi esperimenti consistono in una lunga serie di gare nel seguito delle quali le strategie di maggior successo sono ammesse a partecipare a un numero maggiore di incontri, mentre viene limitato il grado di partecipazione delle strategie meno fortunate. Ogni gara rappresenta una generazione di individui ciascuno dei quali impiega costantemente una determinata strategia nelle sue interazioni con altri membri della popolazione. Il successo di una strategia in ogni singola gara determinerà 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 si possono 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 "Defeziona sempre!"), diffondersi in questo mondo fino ad 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.

Vitalità iniziale

TIT FOR TAT come detto in precedenza non è la sola strategia che può essere evolutivamente stabile. In concreto ALL D è evoluzionatamente stabile indipendentemente da quale è la probabilità di interazione continua. Questo solleva il problema di come un trend evoluzionario di comportamento cooperativo potrebbe sempre iniziare al primo posto.

Questa domanda si pone in generale in ogni tentativo di spiegare i fenomeni cooperativi e non esiste ancora una teoria unica che spieghi come tutti abbiano inizio. Un punto di vista molto diffuso tra i teorici del gioco è che la cooperazione si realizza in contesti ripetuti perché è applicabile e reciprocamente vantaggiosa e perché le persone possono comunicare. Come chiariscono gli esempi biologici, **la comunicazione non è un prerequisito per la cooperazione**. Nella situazione della guerra di trincea, è improbabile che i combattenti possano comunicare direttamente, ma la strategia TIT-FOR- TAT si è affermata. Alla radice c’era la staticità della guerra che faceva sì che le stesse unità si trovassero per lunghi periodi di tempo le une di fronte alle altre.

La teoria delle **parentele genetiche** suggerisce una plausibile scappatoia dall'equilibrio di ALL D. La stretta correlazione degli interlocutori permette un vero altruismo-sacrificio da parte di un individuo a beneficio di un altro. Il vero altruismo può evolvere quando le condizioni di costo, beneficio e relazionalità producono guadagni netti agli individui correlati. Non defezionare nel Dilemma del Prigioniero con una sola mossa è un tipo di altruismo (l'individuo sta rinunciando ai proventi che avrebbe potuto prendere) e quindi può evolvere se i due interlocutori sono sufficientemente imparentati. In effetti, ricalcolare la matrice del payoff in modo tale che un individuo abbia un interesse parziale nel guadagno del partner (ovvero, il calcolo dei payoff in termini di idoneità inclusiva) può spesso eliminare le disuguaglianze $T > R$ e $P > S$, nel qual caso **la cooperazione diventa incondizionatamente favorita**. Così è possibile immaginare che i benefici della cooperazione in situazioni simili al dilemma del prigioniero possano cominciare a essere raccolti da gruppi di individui strettamente imparentati. Ovviamente, per quanto riguarda le coppie, un genitore e i suoi figli o una coppia di fratelli e sorelle sarebbero particolarmente promettenti, e in effetti sono noti molti esempi di cooperazione o di contenimento dell'egoismo in tali coppie. Una volta che i geni per la cooperazione esistono, la selezione promuoverà le strategie che basano il comportamento cooperativo sugli spunti dell’ ambiente. Fattori come la paternità promiscua e gli eventi a margini di gruppo non definiti porteranno sempre ad una relazione incerta tra i potenziali interlocutori. Il riconoscimento di eventuali migliori correlazioni di parentela e l'uso di questi spunti per determinare il comportamento cooperativo permetterà sempre di avanzare nella forma inclusiva. Quando è stata fatta una scelta cooperativa, uno spunto di relatività è semplicemente il fatto di ricambiare la collaborazione.Infine, **quando la probabilità che due individui si incontrino di nuovo è sufficientemente alta, la cooperazione basata sulla reciprocità può prosperare ed essere evolutivamente stabile in una popolazione senza alcuna relazione.**

Un caso di cooperazione che si adatta a questo scenario, almeno alle prime prove, è stato scoperto nei rapporti di deposizione delle uova in un branzino. I pesci, che sono ermafroditi, formano coppie e si può dire che si alternano nell'essere il partner ad alto investimento (deporre le uova) e il partner a basso investimento (fornire lo sperma per fecondare le uova).Si verificano fino a dieci riproduzioni in un giorno e vengono fornite solo poche uova ogni volta.Le coppie tendono a rompersi se i ruoli sessuali non sono divisi equamente.

Supponiamo che un piccolo gruppo di individui stia usando una strategia come TIT FOR TAT e che una certa proporzione, p, delle interazioni dei membri di questo cluster sia con altri membri del cluster. Allora il punteggio medio raggiunto dai membri del cluster nel giocare la strategia TIT FOR TAT è $p[RM(1 - w)] + (1 - p)[S + wP/(l - w)$. Quando p e w (probabilità che i due individui si incontrino in futuro) sono grandi abbastanza, un gruppo di individui TIT FOR TAT può poi diventare inizialmente vitale in un ambiente composto in prevalenza da TUTTI D. Il **clustering** è spesso associato alla parentela, e i due meccanismi possono rafforzarsi a vicenda nel promuovere l'iniziale fattibilità della cooperazione reciproca. Tuttavia, è possibile che il clustering sia efficace senza parentela (3). Abbiamo visto che TIT FOR TAT può insinuarsi in un cluster su una popolazione di TUTTI D, anche se TUTTI D è evolutivamente stabile. Ciò è possibile perché un gruppo di TIT FOR TAT offre a ciascun membro una probabilità non banale di incontrare un altro individuo che ricambierà la cooperazione. Se da un lato questo suggerisce un meccanismo per l'avvio della cooperazione, dall'altro solleva anche la questione se il contrario possa accadere una volta che una strategia come TIT FOR TAT si sia affermata. In realtà, c'è un'interessante asimmetria. Definiamo una buona strategia, come TIT FOR TAT, che non sarà mai la prima a defezionare. Ovviamente, quando due buone strategie interagiscono, entrambe ricevono R ad ogni mossa, che è il punteggio medio più alto che un individuo può ottenere quando interagisce con un altro individuo usando la stessa strategia. Pertanto, se una strategia è valida e stabile dal punto di vista evolutivo, non può essere invasa da un cluster. Ciò significa che quando w è abbastanza grande da rendere TIT FOR TAT una strategia evolutivamente stabile, può resistere all'intrusione di qualsiasi gruppo di qualsiasi altra strategia.

La storia cronologica che emerge da questa analisi è la seguente. **TUTTI D è lo stato primordiale ed è evolutivamente stabile**. Ciò significa che può resistere all'invasione di qualsiasi strategia che ha praticamente tutte le sue interazioni con TUTTI D. Ma la **cooperazione basata sulla reciprocità** può prendere piede attraverso **due meccanismi diversi**. In primo luogo, ci può essere una **parentela** tra le strategie degli altri soggetti, dando ai loro geni una certa partecipazione al successo dell'altro, alterando così la matrice di payoff efficace dell'interazione se vista dal punto di vista del gene piuttosto che dell'individuo. Un secondo meccanismo per superare la defezione totale **è che le strategie degli altri soggetti arrivino in un cluster** in modo che forniscano una non banale proporzione delle interazioni che ciascuno di essi ha, anche se sono così poche da fornire una proporzione trascurabile delle interazioni che gli individui di TUTTI D hanno. Quindi l'approccio del torneo dimostra che una volta che una varietà di strategie è presente, TIT FOR TAT è estremamente robusto. Si comporta bene in una vasta gamma di circostanze e sposta gradualmente tutte le altre strategie in una simulazione di una grande varietà di regole decisionali più o meno sofisticate. E se la probabilità che l'interazione tra due individui continui è abbastanza grande, allora TIT FOR TAT è di per sé evolutivamente stabile. Inoltre, la sua stabilità è particolarmente sicura perché può resistere all'intrusione di interi gruppi di strategie mutanti. Pertanto la cooperazione basata sulla reciprocità può iniziare in un mondo prevalentemente non cooperativo, può prosperare in un ambiente variegato e può difendersi una volta pienamente stabilita.

Un esempio di calcolo della vitalità strategica

Nel loro lavoro sull'evoluzione della cooperazione, Axelrod e Hamilton [1981] hanno sostenuto che Tit-For-Tat è efficace nel promuovere la cooperazione in parte perché è in grado di sopravvivere anche quando ci sono relativamente pochi individui che inizialmente utilizzano questa strategia. Più in generale, un policy maker interessato a promuovere una strategia di cooperazione che non sia già ampiamente diffusa deve accertare se una tale strategia possa mai prendere piede in una popolazione e, in caso affermativo, a quali condizioni. Possiamo quindi considerare la fattibilità iniziale come un vincolo per un policy maker riguardo a una selezione di strategie che possono essere promosse in modo sostenibile. Per quantificare l'idea di fattibilità strategica (in precedenza, questa nozione era in gran parte qualitativa), dobbiamo prima quantificare cosa significa per una strategia "sopravvivere". A tal fine, ci si riferisce alla teoria dei giochi evolutivi e, nello specifico, alle dinamiche dei replicatori [Friedman, 1991, Fudenberg e Tirole, 1991]. La dinamica dei replicatori inizia con una popolazione iniziale di agenti, con ogni agente che sceglie una singola strategia dall'intero gruppo S. Per ogni strategia s in questo gruppo possiamo di conseguenza quantificare la frazione di agenti, p_s , scegliendo quella strategia:

$$p_s = \textit{num di agenti che sceglie s/num di agenti totale}$$

Partendo da questa strategia iniziale di distribuzione p_s , la dinamica dei replicatori procede attraverso una sequenza di passi. In ogni round, gli agenti sono accoppiati casualmente per giocare un'istanza del gioco (il gioco iterato). Possiamo calcolare il payoff atteso per ogni strategia nella popolazione rispetto a tali accoppiamenti casuali; lasciamo che questo valore atteso per s sia $Eu(s)$. Prima di iniziare il turno successivo, la distribuzione di ogni strategia nella popolazione, p_s viene aggiornata a $p'_s \propto Eu(s)$ ed il processo si ripete. La dinamica dei replicatori procede su L iterazioni , e spesso (nell'esperienza pratica) converge verso un equilibrio simmetrico misto di strategia Nash del gioco, a patto che ogni strategia in S abbia una qualche rappresentazione nella popolazione agente iniziale (cioè, p_s > 0 per tutti s ∈ S). Supponiamo ora di selezionare una dinamica dei replicatori con una determinata distribuzioni di probabilità su una strategia S in cui la strategia obiettivo s è giocata con probabilità p_s . Lasciamo che le probabilità di altre strategie nella distribuzione del seme siano indicati da p_{-s}. Successivamente, la dinamica dei replicatori viene eseguita fino a quando non converge o raggiunge un limite L sul numero di turni (nella simulazione sono state utilizzate 200 round); lasciamo che $N(z, p_s, p_{-s})$ sia la strategia mista finale (proporzioni di popolazione) raggiunta dalla dinamica dei replicatori. Diciamo che la sopravvivenza della strategia s è alta se la sua probabilità sotto $N(z, p_s, p_{-s})$ è significativamente superiore a 0. Per quantificare quanto è possibile s quando è sottorappresentata nel gruppo di strategia iniziale, lasciamo $p_s = f\, min_{\, t \in S \setminus s} p_t$,con f∈[0, 1]; cioè, s è inizialmente giocata con una probabilità che è una frazione della più piccola probabilità con cui si gioca qualsiasi altra strategia. Se ad esempio viene usata f = ½ e una distribuzione uniforme su tutte le strategie diverse da s e vi sono 10 strategie (11 in totale) ogni strategia viene giocata con probabilità di circa 0.095 mentre la probabilità iniziale di s è di circa 0,048.