
Introducció. Idees generals sobre la teoria de jocs

PID_00268970

Ignacio Sánchez-Cuenca

Temps mínim de dedicació recomanat: 2 hores



Ignacio Sánchez-Cuenca

La revisió d'aquest recurs d'aprenentatge UOC ha estat coordinada pel professor: Albert Batlle (2019)

Primera edició: setembre 2019
© Ignacio Sánchez-Cuenca
Tots els drets reservats
© d'aquesta edició, FUOC, 2019
Av. Tibidabo, 39-43, 08035 Barcelona
Realització editorial: FUOC

Cap part d'aquesta publicació, incloent-hi el disseny general i la coberta, no pot ser copiada, reproduïda, emmagatzemada o transmesa de cap manera ni per cap mitjà, tant si és elèctric com químic, mecànic, òptic, de gravació, de fotocòpia o per altres mètodes, sense l'autorització prèvia per escrit dels titulars dels drets.

Índex

Introducció.....	5
1. Breu història de la teoria de jocs.....	7
2. Teoria de jocs i relacions internacionals.....	9
3. Teoria de jocs i debat metodològic.....	12
Bibliografia.....	15

Introducció

A l'hora d'estudiar l'acció racional, se sol fer una distinció bàsica entre **situacions paramètriques** i **situacions estratègiques**:

1) En les **situacions paramètriques** l'agent coneix tots els paràmetres que afecten la seva decisió. Per exemple, quan un consumidor s'adreça al mercat a comprar un bé, els paràmetres rellevants són els preus dels béns i la restricció pressupostària del consumidor. En aquest cas, la decisió, a més de ser paramètrica, es duu a terme amb **certesa**: l'agent coneix els valors de tots els paràmetres (sap quant costen els diferents béns i de quants diners disposa per a gastar-s'hi).

En canvi, si un agent decideix comprar un bitllet de loteria, la seva decisió, fins i tot essent paramètrica, es pren amb **risc**, ja que l'agent no pot saber per endavant si el bitllet adquirit resultarà premiat o no. Això no obstant, l'agent pot conèixer la probabilitat de rebre el premi (si se sortegen 100.000 números, la probabilitat de guanyar serà 1/100.000).

Si l'agent ni tan sols coneix aquesta probabilitat, llavors diem que la seva decisió es duu a terme amb **incertesa**. Per exemple, si l'agent inverteix en borsa, no sap per endavant quina probabilitat té de guanyar o perdre. L'agent podrà tenir estimacions personals o subjectives de com evolucionarà la borsa, però no es tracta d'una probabilitat objectiva com la del cas de la loteria.

2) En les **situacions estratègiques**, els resultats de l'acció o elecció de l'agent no depenen solament de paràmetres. A més dels paràmetres, el resultat de l'acció depèn del que facin unes altres persones. Diem que **hi ha interacció estratègica** entre diferents agents quan l'acció de cada un depèn de les expectatives que tinguin sobre el que faran els altres. Suposem el cas més senzill, una situació estratègica entre dos agents, *A* i *B*. El que faci *A* depèn del que cregui que farà *B*, però al seu torn el que *B* faci depèn del cregui que farà *A*.

Pedra, paper, tisora

Pensem en un cas molt senzill, el joc de pedra, paper, tisora. *A* traurà tisora si pensa que *B* traurà paper, però *B* traurà pedra si pensa que *A* traurà tisora. Ara bé, si *A* sap que això és el que *B* pensa, en realitat hauria de treure paper, cas en què *B* hauria de treure tisora, i així successivament.

Com en l'exemple de «pedra, paper, tisora», podria semblar que l'encadenament de les expectatives dels agents que té lloc en una situació estratègica origina una mena de cercle viciós o una tornada a l'infinit que els impedeix prendre una decisió. Per sort, la teoria de jocs s'encarrega de demostrar que això no és així i que les situacions estratègiques tenen «solucions» racionals. En el cas del joc «pedra, paper, tisora», la teoria recomanaria a cada

jugador, per raons que s'exposen detalladament en el mòdul 3, que prenguis la decisió a l'atzar i triés amb la mateixa probabilitat cada una de les opcions possibles.

Mentre que la teoria de jocs s'ocupa de les situacions estratègiques, la teoria de la decisió (o teoria de la utilitat) estudia les situacions paramètriques.

Aquest repartiment del treball no implica, tanmateix, que es tracti de dues teories independents. Com veurem en el mòdul 2, la teoria de jocs es construeix sobre els fonaments que proporciona la teoria de la decisió.

Situacions estratègiques

Les situacions estratègiques es donen en multitud d'àmbits. Hi ha dependència estratègica quan dues empreses en un duopoli han de fixar el nivell de producció, quan dos jugadors d'escacs s'enfronten en una partida, quan un sindicat negocia amb una empresa, quan els partits polítics competeixen en les eleccions, quan els exèrcits lluiten en una batalla, quan diferents grups ètnics han d'organitzar la convivència en un mateix territori, quan en un parlament els representants estableixen coalicions per a formar un govern, quan una organització terrorista pressiona un estat, quan es fa una subhasta, etc.

1. Breu història de la teoria de jocs

Històricament, la teoria de jocs va començar analitzant jocs de cartes com el pòquer. El terme *joc* es va conservar fins i tot quan la teoria va abandonar l'estudi dels autèntics jocs i va passar a considerar situacions estratègiques en general. Un joc, en aquest sentit, és qualsevol situació estratègica. El primer treball important sobre aquesta matèria és el llibre *Theory of Games and Economic Behavior*, publicat el 1944 per Von Neumann (un gran físic i matemàtic) i Morgenstern (un economista austríac). En aquest llibre es proposava, entre altres coses, una nova teoria de la utilitat (les línies generals de la qual s'exposen en el mòdul 2) i una solució algorítmica per als jocs de suma zero, en què el que un guanya l'altre ho perd. Van demostrar que aquests jocs, tot i que poc freqüents en la realitat, tenen una solució senzilla i elegant des del punt de vista matemàtic.

Tanmateix, les principals aportacions es produeixen amb la publicació de diferents treballs sobre teoria de jocs a càrrec del matemàtic **John Nash** els anys de 1950 a 1953. Nash va proposar una noció general i simple d'equilibri – l'anomenat *equilibri de Nash*, que s'estudia en el mòdul 3–, entenen per *equilibri* una situació en què cap dels jugadors no té incentius per a canviar la seva elecció. Aquesta noció s'aplica de la mateixa manera als jocs de suma zero, en què la divergència d'interessos és total, i als jocs de suma diferent de zero, en què aquesta divergència és solament parcial. En un equilibri de Nash, els jugadors només van proveïts de la seva racionalitat (intenten de maximitzar la seva utilitat) i no poden arribar a acords entre que no se sostinguin sobre els mateixos interessos. Quan això succeeix, que no hi ha possibilitat d'establir acords amb un compliment garantit per una tercera part, es parla de jocs no cooperatius. En aquest curs només examinarem els jocs no cooperatius i deixarem de banda els cooperatius.

Les aplicacions i desenvolupaments de la teoria de jocs van trigar temps a fer-se notar. Fins ben avançats els anys seixanta no es fan nous avenços teòrics importants. **John Harsanyi** va proposar la seva teoria de jocs d'informació incompleta i **Richard Selten**, ja en els setanta, va oferir noves nocions d'equilibri més refinades, tenint en compte els problemes de credibilitat de les promeses i amenaces que poden intervenir en els jocs (vegeu el mòdul 4). A més, Harsanyi i Selten van col·laborar en un ambiciós projecte destinat a proposar una teoria general de l'equilibri vàlida per a qualsevol tipus de joc, que va culminar amb la publicació del llibre *A General Theory of Equilibrium Selection in Games* el 1988.

Aquests tres autors, Nash, Harsanyi i Selten, van rebre el premi Nobel d'Economia el 1994 per les seves decisives contribucions a la teoria de jocs. Durant els anys vuitanta es va avançar molt notablement en el que es coneixim-

Bibliografia

Von Neumann, J.; Morgenstern, O. (1944). *Theory of Games and Economic Behavior* (3a. ed. 1980). Princeton: Princeton University Press.

Bibliografia

Nash, J.F. (1996). *Essays on Game Theory*. Cheltenham: Edward Elgar.

Bibliografia

Harsanyi, J.; Selten, R. (1988). *A General Theory of Equilibrium Selection in Games*. Cambridge, Massachusetts: The MIT Press.

Kuhn, H.W. (ed.) (1997). *Classics in Game Theory*. Princeton: Princeton University Press.

Weintraub, E.R. (ed.) (1992). *Toward a History of Game Theory*. Durham: Duke University Press.

com els *refinaments de l'equilibri de Nash* i es van establir els primers models de negociació basats en la teoria de jocs no cooperatius (vegeu el mòdul "Jocs repetits"). Durant l'última dècada potser el més important ha estat el ràpid desenvolupament de models d'aprenentatge i racionalitat limitada, en què no és necessari fer supòsits tan exigents sobre la capacitat de càlcul i anticipació dels agents.

2. Teoria de jocs i relacions internacionals

La teoria de jocs va anar penetrant lentament però de manera inexorable en la teoria econòmica, i avui no és estrany que els manuals de microeconomia s'exposin en termes d'aquesta teoria. La teoria de jocs ha resultat extremadament útil en economia per a entendre tots els intercanvis entre agents en què hi ha informació asimètrica: una de les parts sap més que l'altra, té informació que els altres no tenen.

Tanmateix, la teoria de jocs ha tingut un impacte molt notable en altres disciplines. En biologia, gràcies al treball pioner de John Maynard Smith, es va demostrar que aquesta teoria es podia aplicar per a explicar un notable nombre de comportaments de les espècies animals. Sens dubte, no és que s'atribueixi racionalitat als animals, però el mecanisme de selecció natural exerceix un paper anàleg. Senzillament, els membres d'una espècie que desenvolupin certes estratègies produiran més descendència, de manera que els mecanismes evolutius acaben produint els mateixos equilibris que aquells als quals s'hauria arribat si els animals haguessin estat racionals.

En ciències socials, és en la ciència política on la teoria de jocs ha estat especialment important. En la mesura que la ciència política estudia moltes situacions estratègiques (negociacions entre estats, competició entre partits, relacions entre grups d'interès i governs, conflictes entre institucions, etc.), la teoria de jocs ha trobat un terreny especialment fèrtil. En aquest context, s'entén que la teoria de jocs proporciona l'instrumental analític més útil per a poder estendre l'enfocament d'elecció racional importat de l'economia a la ciència política. Així, es considera que la teoria de jocs és una part fonamental del que de vegades es coneix com a teoria de l'elecció racional. Dins de la teoria de l'elecció racional s'inclouen moltes altres coses, com els models espacials de la política o la teoria de l'acció col·lectiva, però en el fons es tracta de desenvolupaments especials de la teoria de jocs per a la ciència política.

Per bé o per mal, l'ús de models de la teoria de jocs s'ha tornat habitual en les principals revistes de ciència política. Cada vegada és més freqüent trobar models en les pàgines de l'*American Political Science Review*, *American Journal of Political Science*, *European Journal of Political Research*, *Journal of Theoretical Politics*, *Rationality & Society*, i fins i tot en revistes més «tradicionals» com *World Politics* i *International Organization*. El polítòleg ja no es pot permetre de desconèixer la teoria de jocs, ja que ha passat a ser una eina gairebé tan important com les tècniques estadístiques d'anàlisi de dades.

Des dels seus principis, la teoria de jocs va tenir les seves aplicacions en el camp de les relacions internacionals. Va ser precisament utilitzada per entendre el desenvolupament de la guerra freda entre els EUA i la Unió Soviètica i va in-

Bibliografia

Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

Bibliografia

Riker, W. (1992). «The entry of game theory in political science». A: Weintraub (ed.) (1992). *Toward a History of Game Theory* (pág. 207-224). Durham: Duke University Press.

fluir de manera important en l'estratègia militar de l'època: s'havia d'utilitzar l'arsenal nuclear? Sota quines circumstàncies? Amb quins objectius? Quina seria la resposta de l'altre jugador? Quin és l'avantatge de «copejar primer»? La idea intuïtiva i fonamental segons la qual els estats interactuen en un escenari anàrquic en el qual no existeix un «govern mundial» que limiti les seves estratègies i comportaments va poder desenvolupar-se a través de l'aparell deductiu de la teoria de jocs. Evidentment, més enllà de quines fossin les seves aplicacions concretes inicials, darrere aquest desenvolupament hi havia tres concepcions fonamentals: (1) els estats són els agents més importants de les relacions internacionals; (2) els estats actuen de manera interdependent en un context d'anarquia; i (3) els estats són actors racionals i amb el seu comportament busquen satisfer els seus propis interessos encara que això pugui arribar a lesionar els interessos d'altres estats (és a dir, els estats desitgen «maximitzar la seva utilitat»). Aquestes tres concepcions vinculen la teoria de jocs amb una concepció «realista» de les relacions internacionals. Així doncs, la teoria de jocs intenta explicar les relacions internacionals a través d'analitzar el comportament real dels estats. I la teoria de jocs relaciona aquest comportament observable dels estats amb les seves preferències: en la mesura en què se suposa que aquests actors polítics són racionals, s'entén que els estats sempre escolliran la millor alternativa possible, l'alternativa que maximitzi la seva funció d'utilitat. En resum, la teoria de jocs està fonamentalment interessada en l'optimització de les estratègies (i. e. seleccionar l'alternativa que proporciona més utilitat) o, en sentit contrari, està interessada en la ineficiència, a explicar com és possible que, encara que els actors siguin racionals, de vegades es produeixin resultats col·lectius que deixen els actors (estats) en una pitjor situació de la que podrien haver aconseguit.

En el camp de les relacions internacionals, l'aplicació de la teoria de jocs suposa la necessitat d'identificar amb precisió quins són els actors rellevants, quins són els seus objectius, quines són les seves possibles estratègies i quines són les recompenses associades a cada estratègia. És a dir, en primer lloc, cal determinar si realment els estats són l'actor principal de les relacions internacionals i si no hi ha un altre actor amb influència suficient en el resultat final (com, per exemple, organitzacions internacionals, grans empreses multinacionals o grups de pressió...), en aquest cas s'haurien d'incorporar a l'anàlisi. En segon lloc, també és necessari corroborar que l'estat sigui un actor unitari, és a dir, que no inclogui diferents actors amb preferències diferents o, dit en altres paraules, que sigui possible identificar de forma unívoca quin és l'ordre de preferències d'aquest estat entre les diferents alternatives. Finalment, hem de conèixer quants actors estan involucrats en el joc: certament molts exemples concrets de relacions internacionals tenen un caràcter bilateral (són jocs de dos jugadors) però altres exemples impliquen a més de dos jugadors.

En segon lloc, ha de ser possible identificar quins són els objectius de l'estat, és a dir, què és el que permet donar contingut a l'expressió «maximització de la utilitat». En l'àmbit de les relacions internacionals és normal entendre

que conceptes com «seguretat», «poder», «territori», «desenvolupament econòmic» siguin els paràmetres que utilitzen els estats a l'hora d'interpretar el valor relatiu de les possibles estratègies al seu abast.

Precisament, en tercer lloc, han de comprendre's quines són aquestes estratègies a l'abast dels estats. Una estratègia és un pla d'acció complet que cobreix totes les contingències possibles (fins i tot la probabilitat que es produeixi un fet fortuït). Molts dels models derivats de la teoria de jocs resumeixen les possibles estratègies dels estats a la idea binària de «cooperar» i «no cooperar». Òbviament, aquestes estratègies han de contemplar el fet que els altres estats també «juguen a les relacions internacionals» i que les seves decisions poden afectar les del primer (aquest és precisament el sentit de la paraula interdependència).

Finalment, ha de ser possible determinar els pagaments, les recompenses que cada actor rep al final del joc en funció de l'estratègia seguida. Aquestes recompenses (o payoffs) han de ser identificades abans de la construcció del joc. Moltes vegades solament és necessari identificar les recompenses de manera ordinal, és a dir solament és necessari associar els possibles resultats del joc a una posició relativa en l'ordre de preferències dels actors. En altres ocasions és possible establir una relació cardinal entre els pagaments amb la qual cosa el joc és molt més ric en termes de comparabilitat.

La teoria de jocs aplicada a les relacions internacionals tan sol pot oferir explicacions de caràcter limitat. La complexitat del món fa que els contextos estratègics, les preferències dels actors, les seves estratègies i les seves recompenses solament puguin interpretar-se correctament amb l'ajuda de disciplines com l'economia, la ciència política, la sociologia, el dret o la història. No obstant això, una vegada aquests elements estan clars, en emprar models de teoria de jocs, l'analista es compromet a ser transparent en els supòsits que fa. De la mateixa manera, el model garanteix que hi hagi una connexió lògica entre aquests supòsits i les hipòtesis últimes que es derivin de l'equilibri oposat. Es guanya, per tant, en rigor i claredat. Com a contrapartida, els models obliguen a deixar de costat un munt d'informació empírica detallada que pot ser extremadament interessant, però que, segons la teoria, no és necessària per entendre l'assumpte que s'estigui analitzant.

3. Teoria de jocs i debat metodològic

La teoria de l'elecció racional i la teoria de jocs han estat objecte d'una intensa discussió metodològica en la ciència política i també en la sociologia, sobre tot a partir de la publicació el 1994 del llibre de Donald Green i Ian Shapiro *Pathologies of Rational Choice*. S'ha acusat a aquesta teoria d'estar més preocupada per l'elegància formal dels models que per la seva rellevància empírica; igualment, s'ha objectat que, quan la teoria s'interessa per la realitat, sol ser immune als fracassos, ja que sempre es poden fer modificacions *ad hoc* dels models fins que aquests s'ajustin als fets. Segons el parer dels seus crítics, la teoria de l'elecció racional està llastrada pels seus plantejaments universalistes (aplicació sense restriccions del del supòsit de racionalitat) i pel seu excés d'ambició.

Evidentment, aquest no és el moment d'entrar en discussions metodològiques. Amb tot, sí és possible fer algunes observacions generals sobre els avantatges i inconvenients que planteja l'ús de teoria de jocs en la ciència política i en les relacions internacionals. La teoria de jocs parteix del supòsit comú a tota la teoria econòmica que els agents tracten de maximitzar la seva utilitat. El valor afegit de la teoria de jocs consisteix que especifica en què consisteix maximitzar la utilitat en situacions estratègiques. Concretament, estableix quines estratègies són racionals, ja que cada agent sap que tots els altres també tracten de maximitzar la seva utilitat. Així, es diu que una combinació d'estratègies constitueix un equilibri quan cap dels agents no pot augmentar unilateralment la seva utilitat canviant d'estratègia. La teoria de jocs calcula en cada joc què compta com a equilibri.

Com es tradueix això en la ciència política? Elaborant models formals de situacions polítiques estratègiques. Per a l'elaboració del model, és necessari prendre decisions sobre què és essencial i què és accidental o accessori en la descripció de la situació que es vol analitzar.

1) S'han d'identificar quins són els **actors** rellevants. Això no sol ser massa complicat, encara que de vegades s'hagin de fer simplificacions dràstiques.

2) Després, s'han d'especificar quines **preferències** tenen els actors. No sempre és fàcil fer-ho, ja que és necessari comptar amb alguna raó poderosa per a atribuir unes preferències i no unes altres. Si no es fa així, sempre rondarà la sospita que es van escollir unes preferències determinades per a aconseguir derivar un equilibri que coincideixi amb la realitat, cas en què és evident que el model no té cap valor explicatiu. Igualment, si és necessari, s'ha d'especificar també de quin tipus d'**informació** disposen els actors.

3) En tercer lloc, s'han d'aclarir quines **estratègies** o quines accions poden dur a terme els diferents actors. Normalment aquí també s'han de fer grans simplificacions.

Una vegada que s'han identificat els actors, les seves preferències (i informació) i les seves estratègies, es procedeix a «resoldre» el model, és a dir, a calcular quines combinacions d'estratègies poden ser equilibris quan els actors són racionals.

Els **equilibris** serveixen de base per a la derivació de conseqüències empíriques del model. Abans de tot, es pot comprovar si els resultats de la realitat coincideixen més o menys amb el model. Però el problema és que, si no coincideixen, sempre hi ha la possibilitat d'arreglar-ho modificant algun dels supòsits inicials sobre el nombre de jugadors, les seves preferències o la seva informació. Per això, per a garantir que el model es pugui posar a prova amb una mica més de rigor, el que convé més no és comparar una situació concreta amb l'equilibri predit pel model, sinó més aviat establir conclusions sobre com l'equilibri es modifica quan es modifiquen els valors de les variables independents que segons el model tenen pes explicatiu. D'aquesta manera, l'investigador podrà elegir diferents casos empírics, cada un amb diferents valors en les variables independents, i explicar les variacions trobades en la variable dependent a partir dels canvis en l'equilibri que es produeixen quan varien les variables independents.

En l'àmbit de les relacions internacionals els models de teoria de jocs són sobretot útils quan aconsegueixen proporcionar hipòtesis de treball a les quals no es podria haver arribat de no haver estat per l'exercici de modelització. Això no obstant, els models poden tenir altres usos que van més enllà de la seva aplicació empírica immediata. Freqüentment serveixen per a introduir claredat i precisió en qüestions que empíricament són extremadament complexes. Els models són importants perquè constitueixen un punt de referència bàsic a partir del qual es poden introduir noves variables (incertesa, diferents tipus de preferències, etc.) que contribueixin a aproximar-nos millor a la realitat.

No obstant això, no s'ha d'oblidar en cap cas que els models de teoria de jocs són únicament un instrument d'anàlisi per al polític. No hi ha cap obligació d'usar-los sempre, però tampoc no hi ha cap raó per a rebutjar-los per principi.

Bibliografia

Green, D.; Shapiro, I. (1994). *Pathologies of Rational Choice*. New Haven: Yale University Press.

Harsanyi, J.; Selten, R. (1988). *A General Theory of Equilibrium Selection in Games*. Cambridge, Mass: The MIT Press.

Kuhn, H.W. (ed.) (1997). *Classics in Game Theory*. Princeton: Princeton University Press.

Maynard Smith, J. (1982). *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

Nash, J.F. (1996). *Essays on Game Theory*. Cheltenham: Edward Elgar.

Riker, W. (1992). «The entry of game theory in political science». A: Weintraub (ed.). *Toward a History of Game Theory* (pàg. 207-224). Durham: Duke University Press.

Von Neumann, J.; Morgenstern, O. (1944). *Theory of games and economic behavior* (3a. ed. 1890). Princeton: Princeton University Press.

Weintraub, E.R. (ed.) (1992). *Toward a history of game theory*. Durham: Duke University Press.

