Jocs amb informació incompleta

Una de professors guillats

Hi ha un professor (jugador 1) i els estudiants (jugador 2). El professor decideix si avaluar mitjançant un examen difícil (estratègia *d*) o fàcil (estratègia *f*). Els estudiants decideixen, ignorant el tipus d'examen, si estudiar (estratègia *e*) o no estudiar (estratègia *n*). A més, els estudiants ignoren si el professor que tenen és el típic professor o està guillat. En aquest segon cas, al professor tant li és tot i actua de forma que s'inverteixen les preferències dels estudiants sobre els resultats del joc. Per exemple, el resultat més preferit per als estudiants és que l'examen sigui fàcil i ells no hagin d'estudiar. Quan el professor està pertorbat, posa un examen fàcil i els estudiants no estudien, però després suspèn a tothom amb un zero amb independència del que hagin respost, fent que aquest sigui el pitjor resultat per als estudiants.

		2: *e*	2: *n*
1	*d*	3 1	1 0
1	*f*	2 2	0 3

professor "normal"

		2: *e*	2: *n*
1	*d*	0 2	0 3
1	*f*	0 1	0 0

professor "trastornat"

Fig. 1

La diferència entre aquest tipus de joc i els considerats fins ara és que la informació d'algun dels jugadors és <u>incompleta</u>, això és, algun jugador ignora algun dels elements que descriuen el joc[1]. En general, aquesta ignorància pot expressar-se en termes d'ignorància sobre els pagaments d'algun dels altres jugadors. Això fa que es pugui entendre un <u>joc d'informació incompleta com un joc on algun jugador té incertesa sobre els pagaments d'algun altre jugador</u>.

Al joc anterior, els estudiants no saben si estan jugant el joc de l'esquerra de la Fig. 1 o el joc de la dreta. John C. Harsanyi (1967-68), premi Nobel d'Economia el 1994[2], va proposar una manera de <u>transformar un joc d'informació incompleta en un d'informació imperfecta</u>. La idea és que <u>un jugador *i* que ignora els pagaments d'un altre jugador *j* pot ser tractat com un jugador que té incertesa sobre el tipus de jugador *j* al qual s'enfronta</u>. Aleshores n'hi ha prou amb considerar que la naturalesa tria, d'acord amb una certa distribució de probabilitat, el tipus de jugador *j* amb el qual *i* juga. Atès que *i* ignora la decisió de la naturalesa, el joc queda transformat en un d'informació imperfecta. Per exemple, al cas de la Fig. 1, els estudiants poden assumir que la naturalesa tria el professor "normal" amb probabilitat p i tria el professor "trastornat" amb probabilitat $1 - p$. El joc resultant (un joc amb una distribució de probabilitat sobre els tipus que pot assumir algun jugador) s'anomena <u>joc baiesià</u>.

---

[1] No s'ha de confondre "informació imperfecta" amb "informació incompleta". La informació imperfecta es refereix a la ignorància d'algun jugador sobre les <u>decisions</u> d'altres jugadors. La informació incompleta es refereix a la ignorància d'algun jugador sobre les <u>característiques</u> d'algun altre jugador (per exemple, els seus pagaments).

[2] http://nobelprize.org/nobel_prizes/economics/laureates/1994/harsanyi-autobio.html.

Definició de joc baiesià

Un joc (simultani) baiesià està format pels següents elements.

• Un conjunt finit i no buit N d'n jugadors.

• Per a cada jugador $i \in N$, un conjunt finit i no buit S_i d'accions del jugador i.

• Per a cada jugador $i \in N$, un conjunt finit i no buit T_i de tipus del jugador i (i sap el seu tipus).

• Una distribució de probabilitat $\mu \in \Delta(T)$ sobre el conjunt $T = T_1 \times T_2 \times \ldots \times T_n = \prod_{i \in N} T_i$ de totes les combinacions de tipus dels n jugadors, consistent amb el fet que cada i sap el seu tipus.

• Per a cada jugador $i \in N$, una funció de pagaments $u_i(s, t)$ que especifica el pagament del jugador i quan es juga el vector $s \in \prod_{i \in N} S_i$ i cada jugador j és del tipus t_j que estableix $t \in T$.

La novetat que aporta un joc baiesià és fer dependre els pagaments d'un jugador, no només de les decisions dels jugadors, sinó dels tipus dels jugadors. Per exemple, a la Fig. 1, el professor pot ser de dos tipus: t_{11} = normal i t_{12} = trastornat; en canvi, els estudiants són només d'un tipus, t_2. Per tant, el conjunt de combinacions de tipus és $T = \{(t_{11}, t_2), (t_{12}, t_2)\}$. La descripció ha de donar la distribució de probabilitat μ, que s'assumeix coneguda per tots els jugadors. En aquest cas, la probabilitat és indeterminada: $\mu(t_{11}, t_2) = p$ i, per tant, $\mu(t_{12}, t_2) = 1 - p$.

La distribució marginal de probabilitat d'un tipus és la suma de les probabilitats de tots els t que contenen aquell tipus. Per exemple, la probabilitat del tipus t_{11} és $\mu(t_{11}, t_2) = p$. En canvi, la probabilitat del tipus t_2 és $\mu(t_{11}, t_2) + \mu(t_{12}, t_2) = 1$.

La distribució de probabilitat $\mu(t_{-i}|t_i)$ de la combinació tipus t_{-i} condicionada pel tipus t_i (o simplement la probabilitat de t_{-i} donat t_i) és la probabilitat obtinguda aplicant la fórmula de Bayes: $\mu(t_{-i}|t_i)$ és la probabilitat $\mu(t_i, t_{-i})$ dividida per la suma, per a tot t_{-i}', de les probabilitats $\mu(t_i, t_{-i}')$. A la distribució de l'exercici 1 de la llista d'exercicis, $\mu(4|1) = \mu(1, 4) / (\mu(1, 4) + \mu(1, 5)) = \frac{1}{10} / (\frac{1}{10} + \frac{4}{10}) = \frac{1}{5}$. La idea és que la probabilitat condicionada de t_{-i} donat t_i captura la informació que rep el jugador i sobre els tipus dels altres jugadors quan sap que el seu propi tipus és t_i. En el cas anterior, si el jugador 1 sap que el seu tipus és 1, la probabilitat total de les combinacions de tipus que contenen el tipus 1 és $\frac{1}{10}$ [la probabilitat que la combinació sigui (1, 4)] més $\frac{4}{10}$ [la probabilitat que la combinació sigui (1, 5)]. La proporció que (1, 4) representa d'aquesta probabilitat total $\frac{5}{10}$ és $\frac{1}{10} / \frac{5}{10} = \frac{1}{5}$, ja que (quan el tipus 1 succeeix) la probabilitat de ser del tipus 4 és quatre vegades inferior a la probabilitat de ser del tipus 5 ($\frac{1}{10}$ contra $\frac{4}{10}$). Això fa que la probabilitat condicionada $\mu(4|1)$, del 20%, sigui quatre vegades inferior a la probabilitat condicionada $\mu(5|1)$, que és del 80%.

Estratègies en un joc baiesià

Una estratègia d'un jugador i en un joc baiesià consisteix a assignar una distribució de probabilitat sobre S_i a cada tipus del jugador i. Per tant, una estratègia del jugador i és una funció $\sigma_i : T_i \to \Delta(S_i)$. Un vector d'estratègies d'un joc baiesià consisteix a especificar una estratègia per a cada jugador. L'extensió de la funció de pagaments de cada jugador en cas que les estratègies siguin mixtes es fa de la manera habitual: sumant pagaments ponderats per les probabilitats.

Una estratègia d'un jugador en un joc baiesià és una funció que assigna una estratègia (possiblement mixta) a cada tipus del jugador. Al joc de la Fig. 1, una estratègia del jugador 1 especificaria què fa el professor si és del tipus normal i què fa si és del tipus pertorbat. Per exemple, l'estratègia σ_1 tal que $\sigma_1(t_{11}) = d$ i tal que $\sigma_1(t_{12})$ assigna probabilitat ½ a d és l'estratègia on el professor normal proposa un examen difícil als estudiants i on el professor pertorbat proposa un examen difícil amb probabilitat ½.

Quins serien els pagaments del vector (σ_1, σ_2), on σ_1 és l'estratègia definida anteriorment i σ_2 és l'estratègia on e es juga amb probabilitat ⅓? En el cas del jugador 2, els seus pagaments serien la suma del pagament esperat quan el professor és del tipus normal, $p\cdot[\frac{1}{3}\cdot 1 + \frac{2}{3}\cdot 0]$, i del pagament esperat quan el professor és del tipus trastornat, $(1 - p)\cdot[\frac{1}{3}\cdot(\frac{1}{2}\cdot 2 + \frac{1}{2}\cdot 1) + \frac{2}{3}\cdot(\frac{1}{2}\cdot 3 + \frac{1}{2}\cdot 0)]$.

Definició d'equilibri baiesià

Un <u>equilibri baiesià</u> (també anomenat equilibri de Nash-Bayes) d'un joc baiesià és un vector $\sigma = (\sigma_i)_{i\in N} \in \prod_{i\in N} \Delta(S_i)$ d'estratègies mixtes que satisfà, per a tot jugador $i \in N$, tot tipus $t_i \in T_i$ del jugador i i tota estratègia pura $s_i \in S_i$ del jugador i,

$$\sum_{t_{-i}\in T_{-i}} \mu(t_{-i}|t_i)\cdot u_i([\sigma_i(t_i), \sigma_{-i}(t_{-i})], (t_i, t_{-i})) \geq \sum_{t_{-i}\in T_{-i}} \mu(t_{-i}|t_i)\cdot u_i([s_i, \sigma_{-i}(t_{-i})], (t_i, t_{-i})) \qquad (1)$$

on $\sigma_{-i}(t_{-i})$ són les estratègies assignades a cada tipus dels jugadors diferent de l'i.

La definició anterior no diu més que un equilibri baiesià és un equilibri de Nash del joc baiesià. Que el vector σ sigui un equilibri baiesià vol dir que <u>cap tipus de cap jugador no té una estratègia (pura) que li dóna més pagament esperat que el pagament que obté amb el vector σ</u>.

Exemple de càlcul d'equilibris baiesians

Calculem els equilibris baiesians del joc de la Fig. 1. Cal procedir jugador a jugador, i tipus a tipus de cada jugador. Comencem pel jugador 1 i el seu primer tipus: t_{11} = "normal". Atès que els jugadors sempre saben quin és el seu propi tipus, quan 1 és de tipus t_{11} sap que està jugant el joc del costat esquerre a la Fig. 1. En aquest joc, d és fortament dominant, de manera que, a tot equilibri baiesià σ, $\sigma_1(t_{11}) = d$: el tipus t_{11} del jugador 1 sempre tria d amb probabilitat 1.

Considerem ara el tipus t_{12} (= trastornat) del jugador 1. En aquest cas, quan 1 és del tipus t_{12}, 1 sap que juga el joc del costat dret a la Fig. 1, on tota estratègia li maximitza el pagament. Això no vol dir que tota estratègia que jugui el tipus t_{12} sigui un equilibri baiesià, perquè cal tenir en compte les restriccions que imposa el jugador 2.

El jugador 2 només té un tipus, de manera que no pot determinar en quina de les dues matrius de la Fig. 1 està jugant. Aquí és on es pot il·lustrar més clarament (1). El problema del jugador 2 és que la mateixa estratègia s'aplica tant al tipus t_{11} com al tipus t_{12} del jugador 1, atès que 2 no pot saber contra quin tipus del jugador 1 està jugant. Per tant, per a què l'estratègia triada sigui part d'un equilibri baiesià cal que sigui "suficientment bona" quan s'enfronta al que fan els dos tipus. La condició (1) expressa aquesta idea considerant la suma de pagaments: la suma dels pagaments (ponderats per la probabilitat de cada tipus del jugador 1) que 2 espera obtenir contra cada tipus del jugador 1 ha de ser la màxima possible.

Determinem el pagament esperat de 2 si 2 tria e. Quan e es juga contra el tipus t_{11} (ens trobem a la matriu esquerra de la Fig. 1), el pagament esperat és 1, perquè ja s'ha determinat que el tipus t_{11} del jugador 1 tria d en tot equilibri baiesià. Quan e es juga contra el tipus t_{12} (ens trobem a la matriu dreta de la Fig. 1), el pagament esperat és $2d + 1\cdot(1 - d)$, on d és la probabilitat amb què el tipus t_{12} tria d. Atès que la probabilitat d'enfrontar-se amb el tipus t_{11} és p i la probabilitat d'enfrontar-se al tipus t_{12} és $1 - p$ el pagament esperat de 2 quan tria e és

$$p\cdot 1 + (1 - p)\cdot(2d + 1\cdot(1 - d)),$$

això és,

$$p + (1 - p)\cdot(1 + d). \tag{2}$$

Determinem el pagament esperat de 2 si 2 tria n. Quan n es juga contra el tipus t_{11}, el pagament esperat de 2 és 0, ja que el tipus t_{11} del jugador 1 tria d a tot equilibri baiesià. Quan n es juga contra el tipus t_{12}, el pagament esperat de 2 és $3d + 0\cdot(1 - d) = 3d$. Com abans, ponderant cada pagament per la probabilitat del tipus corresponent i sumant, resulta que el pagament esperat de 2 quan tria n és

$$p\cdot 0 + (1 - p)\cdot 3d =$$

això és,

$$(1 - p)\cdot 3d. \tag{3}$$

Ara cal considerar les tres possibilitats: que jugar e sigui part d'un equilibri baiesià; que ho sigui jugar n; o que ho sigui una mixta entre e i n.

• <u>Cas 1: equilibris baiesians on 2 tria e</u> (per tant, on es tria n amb probabilitat zero). Per a què triar e sigui millor que triar n cal que (2) sigui més gran que (3). Així, per a què $p + (1 - p)\cdot(1 + d) \geq (1 - p)\cdot 3d$ cal que $d \leq \frac{1}{2(1-p)}\cdot$ Per consegüent, tot vector σ tal que $\sigma_1(t_{11}) = d$, $\sigma_1(t_{12})$ assigna probabilitat a d no superior a $\frac{1}{2(1-p)}$ i $\sigma_2 = e$ és un equilibri baiesià.

• <u>Cas 2: equilibris baiesians on 2 tria n</u>. Ara cal que $d \geq \frac{1}{2(1-p)}$ i tot vector τ tal que $\tau_1(t_{11}) = d$, $\tau_1(t_{12})$ assigna probabilitat a d no inferior a $\frac{1}{2(1-p)}$ i $\tau_2 = n$ és un equilibri baiesià.

• <u>Cas 3: equilibris baiesians on 2 tria e i n amb probabilitat positiva</u>. En aquest cas, és necessari que (2) sigui igual a (3), de manera que $d = \frac{1}{2(1-p)}$. D'aquí que tot vector d'estratègies υ tal que $\upsilon_1(t_{11}) = d$, $\upsilon_1(t_{12})$ assigna probabilitat $\frac{1}{2(1-p)}$ a d i υ_2 assigna qualsevol probabilitat (entre 0 i 1) a e és un equilibri baiesià.

Possibles inconvenients de tenir més informació

En jocs amb un únic jugador, el jugador no pot estar pitjor si disposa de més informació: com a últim recurs, el jugador pot decidir ignorar la informació addicional. En jocs amb més d'un jugador la cosa és diferent: tenir més informació pot ser perjudicial, ja que és possible que <u>si un</u>

jugador rep nova informació i la resta de jugadors ho saben aleshores el pagament del jugador que es torna més informat pot reduir-se. El joc de la Fig. 2 permet d'il·lustrar aquest fet.

1 \ 2	*c*	*d*	*e*
a	2 2	1 0	1 3
b	4 4	0 0	0 6

p = ½

1 \ 2	*c*	*d*	*e*
a	2 2	1 3	1 0
b	4 4	0 6	0 0

$1 - p$ = ½

Fig. 2

El joc de la Fig. 2 representa el joc baiesià on tots dos jugadors ignoren si estan jugant a la matriu de l'esquerra o a la matriu de la dreta. Tots dos atribueixen la mateixa probabilitat d'½ d'estar jugant a una de les matrius. Comprovem d'entrada que en tot equilibri baiesià el jugador 2 tria *c*. Si 1 tria *a*, aleshores el pagament esperat per a 2 quan tria *c* és 2, en tant que *d* i *e* proporcionen un pagament d'1'5 (= 0·½ + 3·½). Si 1 tria *b*, el pagament esperat per a 2 quan tria *c* és 4, en tant que *d* i *e* proporcionen un pagament de 3 (= 0·½ + 6·½). Donat *c*, la millor resposta d'1 és *b*. Conclusió: [*b*, *c*] és l'únic equilibri baiesià del joc de la Fig. 2. Els dos vectors de pagaments (4, 4) d'aquest equilibri s'indiquen mitjançant un rectangle a la Fig. 2.

Ara suposem que només el jugador 2 és informat de la matriu on s'està jugant. En la matriu esquerra, *e* és una estratègia fortament dominant; a la matriu dreta, *d* és una estratègia fortament dominant. Donat que 2 només tria *e* o *d*, la millor resposta d'1 (tot i ignorar en quina matriu es troba) és sempre *a*. Conclusió: [*a*, (*e*, *d*)] és l'únic equilibri baiesià del joc de la Fig. 2 quan el jugador 2 és informat de la matriu en què juga. Els dos vectors de pagaments (1, 3) d'aquest equilibri s'indiquen mitjançant un cercle a la Fig. 2. Clarament, el jugador 2 es troba pitjor en aquest cas que quan no estava informat.

Els possibles inconvenients de perdre una mínima informació

La Fig. 3 permet d'il·lustrar el fet que una petita variació en la informació dels jugadors pot alterar substancialment el conjunt d'equilibris baiesians. El joc de la part esquerra de la Fig. 3 té dos tipus d'equilibris: l'equilibri [*b*, *d*] i el conjunt d'equilibris [*a*, *c*] tal que $a \geq$ ½ i $c = 1$. El segon equilibri dóna pagaments més grans que el primer a tots dos jugadors, malgrat que *a* és una estratègia feblement dominada.

1 \ 2	*c*	*d*
a	3 3	0 2
b	3 0	1 1

1 \ 2	*c*	*d*
a	1 1	0 2
b	2 0	3 3

Fig. 3

Modifiquem el joc inicial de la forma següent: per a $\varepsilon > 0$ tan petit com es vulgui, es juga la matriu esquerra a la Fig. 3 amb probabilitat $1 - \varepsilon$ i es juga la matriu dreta amb probabilitat ε. El jugador 2 sap a quina matriu s'està jugant (observa l'elecció de la naturalesa entre una matriu o l'altra) però el jugador 1 no, i atribueix la probabilitat $1 - \varepsilon$ al fet de jugar la matriu esquerra. Al nou joc, quan es juga la matriu dreta, *d* és una estratègia fortament dominant per al jugador 2. En vista d'això, el pagament esperat del jugador 1 quan tria *a* és

$$(1 - \varepsilon)[3c + 0\cdot(1 - c)] + \varepsilon\cdot 0$$

que equival a

$$3(1 - \varepsilon)c,$$

on *c* és la probabilitat amb què el jugador 2 tria *c* a la matriu esquerra (ja s'ha indicat que 2 tria *d* a la matriu dreta). De manera anàloga, el pagament esperat del jugador 1 quan tria *b* és

$$(1 - \varepsilon)[3c + 1\cdot(1 - c)] + \varepsilon\cdot 3$$

que equival a

$$3(1 - \varepsilon)c + (1 - \varepsilon)(1 - c) + 3\varepsilon.$$

Atès que $0 < \varepsilon < 1$ i $0 < c < 1$, és evident que el pagament de *b* és superior al d'*a*. Així doncs, en tot equilibri baiesià del nou joc, el jugador 1 tria *b*. Donat *b*, el millor per al jugador 1 a la matriu esquerra és *d*. El vector de pagaments assolit és $(1 - \varepsilon)(1, 1) + \varepsilon(3, 3) = (1 + 2\varepsilon, 1 + 2\varepsilon)$.

Recapitulem. En el joc inicial, el vector de pagaments (3, 3) s'assolia a un equilibri. Després d'una petita pertorbació del joc inicial que introdueix una mínima incertesa en el jugador 1, l'únic vector de pagaments d'un equilibri dóna menys a tots els jugadors.

Iteració del coneixement: una de suspesos i aprovats

Un cop corregit l'examen final, un professor de Microeconomia Superior reuneix en el seu despatx a tres estudiants (1, 2 i 3) que han fet l'examen. A continuació, el professor diu, davant de tothom, que lliurarà a cada estudiant un full on es diu si els altres dos estudiants han aprovat o suspès. Dit això, lliura a cadascú d'ells un full que diu "Els altres dos estudiants que t'acompanyen han suspès". Els estudiants no poden llegir els fulls lliurats als altres i el professor no permet que es comuniquin entre sí. Un cop cada estudiant ha llegit el seu full, el professor pregunta davant tothom si, sobre la base de la informació que han rebut des del moment d'entrar al despatx, algun dels presents pot determinar amb certesa absoluta si ha aprovat o suspès. Ningú no respon, indicant que ningú no té la completa certesa de si ha aprovat o suspès.

Arribats a aquest punt, el professor diu davant tothom: "Almenys un de vosaltres ha suspès. Us donaré tres oportunitats. A cada oportunitat, aprovaré a qui digui que té la certesa absoluta que estava suspès". El professor dóna la primera oportunitat i ningú no diu res. Dóna la segona, i ningú no diu res. Dóna la tercera i aprova a tothom (perquè tothom diu que pot demostrar que estava suspès). Com és possible això si, des del primer cop que el professor ha preguntat si algú sabia la seva nota, només s'ha anunciat un fet que cada estudiant ja sabia (que algú estava suspès)?

Per a explicar el raonament que van seguir els estudiants, considerem el següent model. Hi ha 8 estats del món, representats a continuació (on S = suspès i A = aprovat).

Estat del món	Qualificació de l'estudiant 1	de l'estudiant 2	de l'estudiant 3
1	A	A	A
2	A	A	S
3	A	S	A
4	A	S	S
5	S	A	A
6	S	A	S
7	S	S	A
8	S	S	S

Cada estat del món es correspon amb la qualificació (aprovat o suspès) de cada estudiant. Per exemple, l'estat del món 6 és aquell on suspenen els estudiants 1 i 3 i on aprova l'estudiant 2. Com al cas dels jocs seqüencials (i al joc del correu electrònic), podem <u>representar la informació d'un agent mitjançant una partició dels possibles estats on es pot trobar l'agent</u>. Als jocs seqüencials, els possibles estats d'un jugador s'associaven amb els seus nodes de decisió i la partició del seu conjunt de nodes donava lloc als conjunts d'informació. S'interpretava que tot el que quedava dins un element de la partició (tots els nodes que pertanyen a un conjunt d'informació) era indistingible per al jugador. En el cas present, tota partició del conjunt d'estats del món representa la informació de què disposa (el que sap) l'agent que té la partició.

Abans d'entrar al despatx del professor, és raonable suposar que la partició que té cada estudiant és la partició trivial {{1, 2, … , 8}} on tot els estats es troben dins el mateix element. Això s'interpreta en el sentit que cap estudiant no sap ni la seva qualificació ni la de ningú altre: no té cap informació. A l'altre extrem, la partició {{1}, {2}, {3}, … , {7}, {8}} és la partició que expressa màxima informació: a cada estat se sap que s'està en aquell estat.

La recepció d'informació permet de refinar les particions. Per exemple, quan el professor diu que lliurarà a cada estudiant un full on es diu si els altres dos estudiants han aprovat o suspès, la partició d'1 queda refinada tal com s'indica a la Fig. 4, perquè ara 1 sap si 2 i 3 aproven o suspenen. Els estats 1 i 5 representen l'esdeveniment on els estudiants 2 i 3 aproven. Per això, si el full diu que 2 i 3 aproven, 1 sabrà que l'estat del món és l'1 (on ell aprova) o el 5 (on suspèn). De manera similar, 2 i 6 representen l'esdeveniment on l'estudiant 2 aprova i el 3 suspèn. Quan aquest és el cas, el full del professor donaria aquesta informació, permetent a 1 distingir el conjunt {2, 6} de la resta de subconjunts.

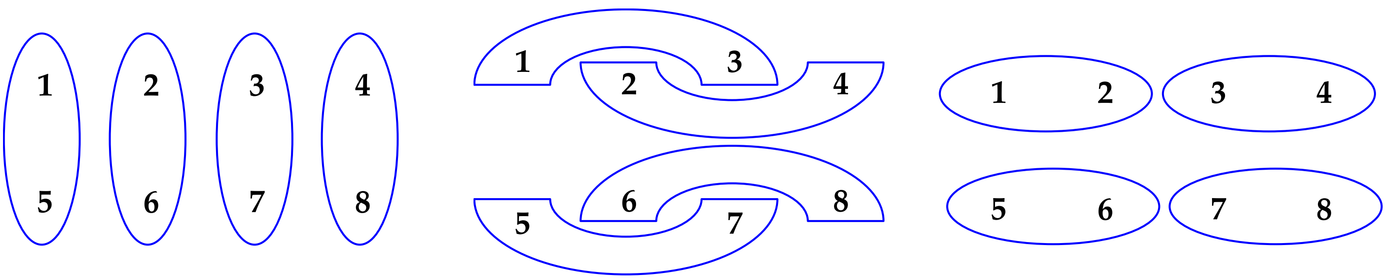

Fig. 4 Fig. 5 Fig. 6

La Fig. 5 mostra com es refina la partició de l'estudiant 2 quan el professor fa l'anunci del que contenen els fulls que lliurarà. En aquest cas, 2 pot distingir el conjunt {1, 3} de la resta, perquè {1, 3} representa l'esdeveniment que 1 i 3 aproven, però no pot distingir l'estat 1 del 3, atès que 2 no sap si ha aprovat o suspès. La Fig. 6 mostra la partició de l'estudiant 3. Ell, per exemple, no pot distingir l'estat 1 del 2, perquè aquests són els estats on 1 i 2 aproven; ni el 3 del 4, perquè són els estats on 1 aprova i 2 suspèn; ni el 5 del 6, perquè són els estats on 1 suspèn i 2 aprova; ni el 7 del 8, perquè són els estats on 1 i 2 suspenen.

Quan el professor anuncia que algun dels estudiants està suspès, diu una cosa que tothom sabia però que no tothom sabia que tothom ho sabia. Dir que algú està suspès implica que l'estat del món <u>no</u> és l'1, que és l'únic estat on ningú no està suspès. Això ja ho sabia cada estudiant (que l'estat del món no és l'1) quan va llegir el seu full: 1 va descobrir que l'estat del món era 4 o 8; 2, que era 6 o 8; i 3, que era 7 o 8. Però el que ningú <u>no sabia</u> era que tothom sabia que l'estat del món no és l'1, perquè ningú no sabia les notes escrites als fulls dels altres. Així que l'anunci del professor transmet informació als estudiants sobre el que els estudiants saben: ara tothom sap que l'estat del món és diferent d'1 (perquè l'anunci del professor és públic davant dels tres estudiants), tothom sap que tothom ho sap, tothom sap que tothom sap que tothom ho sap[3]…

Ara és el torn de les tres oportunitats. El professor canta "Primera oportunitat" i ningú no diu res, perquè cap estudiant no sap si està suspès o no. L'estudiant 1 sap que l'estat del món és 4 o 8, però no sap quin (si és 4, està aprovat; si és 8, suspès). El 2 sap que és 6 o 8, però no quin. I el 3 sap que és 7 o 8, però no quin.

Paradoxalment, aquest silenci dels estudiants durant la primera oportunitat és <u>informatiu</u>. En primer lloc, l'anunci inicial del professor dient el tipus d'informació (no la informació concreta) que contenen els fulls fa que cada estudiant no només estigui al corrent de quina és la seva partició, sinó de <u>quina és la partició dels altres</u>. Així, per exemple, l'1 sabrà que el 2 sap la nota de l'1 i del 3 i, en conseqüència, 1 sap que la partició del 2 és la de la Fig. 5.

<u>Quan 1 calla durant la primera oportunitat, està revelant a 2 i 3 que l'estat del món en què es troben no pot ser el 5</u>. Perquè si l'estat fos el 5, 1 sabria que 2 i 3 estan aprovats (és el que hi hauria escrit al full que li ha passat el professor) i com 1 sabria que almenys hi ha un estudiant

[3] Es diu que un esdeveniment (o un fet) és <u>coneixement comú</u> entre un grup de persones si tothom al grup sap que l'esdeveniment té lloc, si tothom sap que tothom sap que l'esdeveniment té lloc, si tothom sap que tothom sap que tothom sap que l'esdeveniment té lloc, *ad infinitum*. Recuperant la conversa telefònica de la nota a peu de pàgina anterior, *i* i *j* no tenen manera d'assolir coneixement comú sobre el fet de trobar-se al lloc i hora proposades perquè cal un nombre infinit de confirmacions. Quan *j* confirma per primera vegada la proposta d'*i*, *j* sap que han quedat però fins que *i* no rebi la confirmació de *j*, *i* no sap que *j* sap que han quedat. Quan *i* rep la confirmació de *j*, ara *i* sap *j* ho sap, però, fins que *j* no rebi la confirmació d'*i* que *i* sap que *j* ho sap, *j* no sabrà que *i* sap que *j* ho sap. Quan *j* sàpiga que *i* sap que *j* ho sap encara haurà d'enviar una confirmació a *i* per a què, quan *i* la rebi, *i* sàpiga que *j* sap que *i* sap que *j* ho sap… En canvi, si *i* i *j* estiguessin l'un davant l'altre i, en aquestes condicions, *i* fes la proposta i *j* l'acceptés, es podria entendre que hi ha coneixement comú, perquè la transmissió de la informació es fa públicament davant de tots dos, de manera que quan *i* fa la proposta *veu* que *j* la rep i *j veu* que *i* ho veu, i *i* veu que *j* veu que *i* ho veu… A l'exemple del professor, quan aquest anuncia públicament que algú ha suspès aleshores esdevé coneixement <u>comú</u> entre els estudiants que l'estat del món no és l'1 (un fet que abans de l'anunci era <u>merament</u> conegut pels estudiants).

suspès (perquè el professor ho ha anunciat), podria concloure que l'estudiant suspès és ell. Això li permetria demostrar que estava suspès i ho hauria dit. Atès que no ha dit res, l'estat del món no és el 5. Aquesta conclusió la saben tots els estudiants (i saben que la saben, i saben que saben que la saben, i saben que…).

Però hi ha més. <u>Quan 2 calla durant la primera oportunitat, està revelant a 1 i 3 que l'estat del món en què es troben no pot ser el 3</u>. Com abans, si l'estat fos el 3, 2 sabria que 1 i 3 estan aprovats i com 1 sabria que almenys hi ha un estudiant suspès, arribaria a la conclusió que està suspès. Així que tothom sap que l'estat del món no és el 3 (i tothom sap que tothom ho sap i tothom sap que tothom sap que tothom ho sap…).

I encara hi ha més. <u>Quan 3 calla durant la primera oportunitat, està revelant a 1 i 2 que l'estat del món en què es troben no pot ser el 2</u>. Si tal fos el cas, 3 sabria que 1 i 2 estan aprovats. I sabent que algú ha suspès, sabria que el suspès és ell. Per tant, tothom sap (…) que l'estat del món no és el 2.

Transcorreguda la primera oportunitat en silenci, el professor canta "Segona oportunitat" i ningú no diu res. De fet, per l'anunci del professor que hi havia algú suspès, és coneixement comú entre els estudiants que l'estat no és l'1. Els silencis de la primera oportunitat fan que sigui coneixement comú entre ells que l'estat tampoc no és el 5, el 3 o el 2. A l'estudiant 1 això no li resol el dubte de si es troben a l'estat 4 o al 8; ni a l'estudiant 2 li resol el dubte de si es troben al 6 o al 8; ni al 3 si es troben al 7 o al 8.

Però, com a la primera ronda, els silencis parlen. Quan 1 calla, de fet diu que l'estat no pot ser ni el 6 ni el 7. La raó és que el 6 és l'únic estat dels que es consideren possibles on l'1 seria informat que 2 ha aprovat i 3 ha suspès. Com que en l'estat 6 l'estudiant 1 suspèn, sabria que ha suspès. El mateix s'aplica al 7, que és l'únic estat dels que es consideren possibles on l'1 seria informat que 3 ha aprovat i 2 ha suspès. Per aquest motiu, 1 podria distingir l'estat 7 de la resta d'estats considerats com a possibles i, com a resultat, sabria que està suspès.

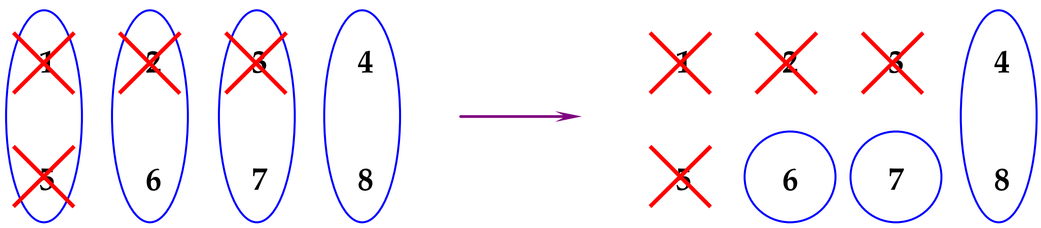

Fig. 7

La Fig. 7 explica aquesta conclusió en termes de model de particions. L'anunci del professor que algú ha suspès elimina l'estat 1 com a possible estat. Els silencis durant la primera ronda han descartat els estats 2, 3 i 5. Per tant, és com si l'estructura informativa de l'estudiant 1 estigués representada per la part dreta de la Fig. 7, on l'estat 6 és inconfusible i on l'estat 7 també ho és. Atès que l'estudiant 1 no ha dit res, l'explicació és que l'estat on els estudiants es troben no és el 6 ni el 7.

El silenci de l'estudiant 2 porta a tothom a la conclusió que l'estat no és el 4, perquè dels estats considerats possibles després de la primera ronda (4, 6, 7 i 8), l'estat 4 és l'únic on l'estudiant 2 seria informat que 1 ha aprovat i 3 ha suspès. Per consegüent, 2 podria distingir l'estat 4 i, si estiguessin a l'estat 4, hauria dit a la segona ronda que està suspès. Com no ho ha dit, tothom arriba a la conclusió que l'estat no és el 4.

Si suposem que la decisió de callar o parlar dels estudiants és simultània, aleshores quan s'inicia la tercera oportunitat de parlar tots tres saben que l'únic estat possible és el 8 i tothom manifesta simultàniament al professor que pot demostrar que estava suspès.

L'exemple posa de manifest la importància de la iteració del coneixement (coneixement sobre el coneixement dels altres) a l'hora de prendre decisions estratègiques. En aquest cas, l'anunci del professor que hi ha algú que ha suspès implica augmentar el que els estudiants saben sobre el que els propis estudiants saben i aquest nou coneixement és el desencadenant d'un procés que porta a descobrir un fet que, abans de l'anunci, era un fet desconegut: la qualificació pròpia.

Bibliografia

- Eichberger, Jürgen (1993): *Game Theory for Economists*. Academic Press: San Diego (capítol 5).
- Fudenberg, Drew i Tirole, Jean (1991): *Game Theory*. The MIT Press, Cambridge, Massachusetts, p. 209–216 i 554–562.
- Harsanyi, John C. (1967–68): "Games with incomplete information played by 'Bayesian' players", *Management Science* 14, 159–182, 320–334, 486–502.
- Myerson, Roger B. (1991): *Game Theory. Analysis of Conflict*. Harvard University Press: Cambridge, Massachusetts, p. 127–131.
- Osborne, Martin J. (2004): *An Introduction to Game Theory*. Oxford University Press: Nova York (capítol 9).
- Ratliff, Jim (1997): Graduate-Level Course in Game Theory (capítol 6). http://www.virtualperfection.com/gametheory/.
- Vega-Redondo, Fernando (2003): *Economics and the Theory of Games*. Cambridge University Press: Cambridge, UK (capítol 6).