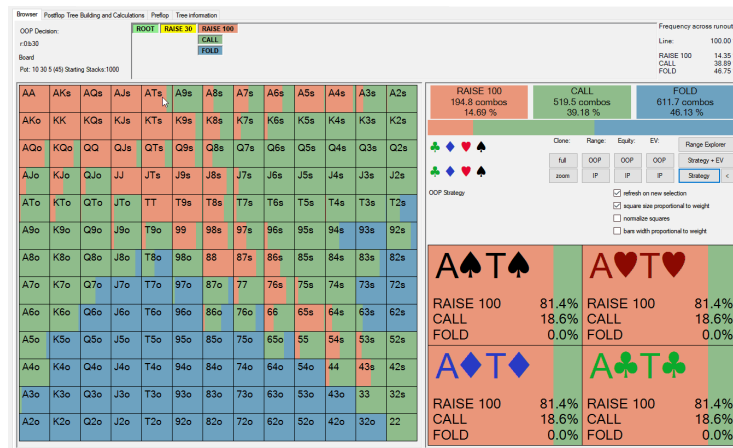


Teoria de jocs:

El rol de la informació

en els equilibris de Nash



Treball de Final de Grau
 Juny de l'any 2025
 Autor: Rodrigo Salinas Pérez
 Tutor: Roberto Rubio

Resum

La teoria de jocs és una eina potent per a l'anàlisi de la presa de decisions estratègiques en una gran varietat de disciplines, des de l'economia fins a la intel·ligència artificial. Aquest treball explora els conceptes fonamentals d'aquesta teoria, amb especial èmfasi en el paper que juguen les asimetries d'informació en la determinació dels equilibris. Comencem amb les definicions de nocions bàsiques com els jocs en forma normal, les estratègies pures i mixtes, i l'equilibri de Nash. S'exposen i demostren resultats centrals com l'existència d'equilibris de Nash, tant en estratègies pures com en mixtes. A continuació, ens endinsem en els jocs extensius per analitzar processos de decisió seqüencials, distingint entre informació perfecta i imperfecta. Presentem conceptes com l'equilibri en *subjocs* i l'equilibri seqüencial, destacant la seva rellevància en entorns dinàmics on les decisions es prenen en diferents etapes. L'eix central del treball és l'estudi dels jocs amb informació incompleta, mitjançant l'enfocament de Harsanyi. Seguint els treballs fonamentals d'aquest autor [5, 6, 7], mostrem com aquests jocs es poden modelitzar com a jocs bayesians. Al llarg del treball, es combinen el rigor formal amb exemples concrets i models que il·lustren els resultats. La teoria es fonamenta principalment en les obres clàssiques de [11] i [10], entre d'altres. Finalment, s'examina com l'estructura de la informació i les creences condicionen les estratègies òptimes, mostrant com la informació incompleta modifica profundament els resultats estratègics.

Abstract

Game theory is a powerful tool for analyzing strategic decision-making in a wide range of disciplines, from economics to artificial intelligence. This thesis explores the fundamental concepts of game theory, with a particular emphasis on the role of informational asymmetries in determining equilibria. We begin with an introduction, where basic notions such as normal-form representations, pure and mixed strategies, and Nash equilibrium are defined. Core results such as the existence of Nash equilibria—both in pure and mixed strategies—are stated and proved. We then move on to extensive-form games to analyze sequential decision-making processes, distinguishing between perfect and imperfect information. Concepts such as subgame perfect equilibrium and sequential equilibrium are presented, highlighting their importance in dynamic environments where decisions are made in stages. The central focus of the thesis is the study of games with incomplete information, following Harsanyi's approach. Building on the foundational work of this author [5, 6, 7], we show how such games can be modeled as Bayesian games. Throughout the thesis, formal rigor is combined with concrete examples and models that illustrate the theoretical results. The framework is primarily based on the classical works of [11] and [10], among others. Finally, we examine how the structure of information and the players' beliefs influence optimal strategies, showing how incomplete information significantly alters strategic outcomes.

Índex

1	Introducció	1
2	Teoria de Jocs	3
2.1	Introducció a la Teoria de jocs	3
2.2	Estratègies en jocs	4
2.3	Equilibris de Nash	5
2.4	Minimax	9
2.5	Jocs extensius d'informació perfecta	12
2.6	Jocs extensius d'informació imperfecta	17
2.7	Jocs amb informació incompleta	20
2.7.1	Teoria i Model de Harsanyi	21
2.7.2	Equivalència entre Jocs	22
2.7.3	Jocs Bayesians	23
3	Un cas d'estudi, Kuhn pòquer	29
3.1	Anàlisi teòric	29
3.2	Aproximació computacional, CFR	32
3.2.1	Esquematització teòrica del CFR	32
3.2.2	Kuhn Poker Modificat	33
4	Conclusions	36
A	Appendix A: Codi CFR	37

1 Introducció

La teoria de jocs és una branca de les matemàtiques que s'ocupa de l'estudi formal de les situacions d'interacció estratègica entre diversos agents racionals, coneguts com a jugadors. Des de la seva formulació clàssica al segle XX, aquesta disciplina ha esdevingut una eina fonamental per comprendre la presa de decisions en escenaris tan diversos com la negociació política, la competència econòmica, la dinàmica evolutiva i, més recentment, la intel·ligència artificial.

Aquest treball es centra en l'estudi sistemàtic dels jocs no cooperatius, amb especial atenció als jocs amb informació incompleta. S'hi duu a terme un recorregut rigorós que comença amb els jocs en forma normal i avança cap als jocs extensius, tractant progressivament els casos d'informació perfecta, imperfecta i, finalment, incompleta. L'estudi no es limita a les definicions formals, sinó que posa un èmfasi especial en la conceptualització dels equilibris de Nash, els equilibris perfectes en *subjocs*, els equilibris seqüencials i els equilibris bayesians, tot analitzant-ne les implicacions i les relacions mútues.

Un dels punts clau del treball és la incorporació de l'enfocament bayesià desenvolupat per John Harsanyi, que permet representar formalment els jocs amb informació incompleta com a jocs amb tipus, introduint així un marc general per estudiar l'estratègia òptima quan els jugadors tenen coneixement asimètric. Aquesta formalització dona lloc a una rica varietat de comportaments estratègics que es poden analitzar amb gran precisió.

Aquesta fonamentació teòrica es complementa amb l'anàlisi detallada de diversos jocs específics, entre els quals destaca el Kuhn pòquer, una versió simplificada del pòquer però molt il·lustrativa dels jocs d'estratègia amb informació incompleta. A partir d'aquest joc, i com a aplicació pràctica del marc teòric desenvolupat, es presenta breument una implementació de l'algoritme Counterfactual Regret Minimization (CFR), que permet aproximar equilibris en jocs extensius de manera iterativa. Tot i que aquesta part computacional no constitueix el nucli del treball, serveix per exemplificar com les tècniques matemàtiques es poden traslladar a aplicacions concretes en la resolució d'estratègies òptimes.

En conjunt, aquest treball pretén oferir una visió completa i coherent del desenvolupament dels conceptes fonamentals de la teoria de jocs en contextos d'informació incompleta, així com de la seva aplicabilitat en situacions representatives. La combinació d'una anàlisi teòrica rigorosa amb exemples pràctics vol proporcionar una comprensió profunda dels mecanismes que governen la presa de decisions en entorns estratègics amb incertesa.

Contribució

Les meves contribucions més destacades en aquest treball inclouen l'estudi detallat de casos representatius, la sistematització dels conceptes clau i la presentació clara dels resultats per facilitar-ne la comprensió. La meva aportació pot veure's reflectida en els següents aspectes clau:

- **Rigor formal combinat amb claredat expositiva:** He fet un esforç per combinar el formalisme matemàtic amb explicacions clares i exemples accessibles, facilitant així la comprensió d'un tema tècnic tant per a públic especialitzat com per a estudiants iniciats en la disciplina.
- **Integració del model de Harsanyi:** He sintetitzat i explicat amb detall la transformació de jocs amb informació incompleta en jocs bayesians, fent especial èmfasi en la seva rellevància pràctica i el seu impacte en la teoria moderna dels jocs. Aquesta secció aporta una visió estructurada i coherent d'una temàtica complexa que sovint es presenta fragmentada.
- **Aplicació pràctica amb l'algoritme CFR:** He inclòs una aproximació computacional al joc de Kuhn pòquer mitjançant l'algoritme *Counterfactual Regret Minimization* (CFR), que tot i no ser el nucli del treball, suposa una connexió amb el món de la intel·ligència artificial i les aplicacions reals d'aquesta teoria.
- **Autonomia en el desenvolupament del contingut:** Tot el contingut ha estat elaborat de forma autònoma, seleccionant, contrastant i estructurant referències clàssiques i recents per oferir una visió global, però també crítica, de la teoria de jocs i les seves extensions.
- **Exemples originals i estudi propi del Kuhn pòquer:** He aportat exemples propis per il·lustrar conceptes complexos i he estudiat una modificació del joc de Kuhn pòquer que no havia estat estudiada, per mostrar com canvis estructurals afecten els equilibris i el comportament estratègic dels jugadors.

Amb aquest treball espero haver sintetitzat i divulgat d'una manera clara i concisa diferents aspectes complexos de l'àmbit de la teoria de jocs i haver-los desenvolupat d'una manera rigorosa i entenedora per al lector.

2 Teoria de Jocs

La teoria de jocs és una branca de les matemàtiques que estudia les interaccions estratègiques entre agents racionals, coneguts com a jugadors. Aquesta disciplina proporciona eines per analitzar com els individus prenen decisions en situacions d'interdependència mútua, on els resultats de les seves eleccions depenen no només de les seves pròpies accions, sinó també de les dels altres. A través de models matemàtics, la teoria de jocs permet abordar situacions tan diverses com conflictes polítics, estratègies empresarials o comportaments animals, entre altres.

En aquest apartat, basant-nos en referències clàssiques com [11], [10] o [5], introduïrem els conceptes bàsics d'aquesta disciplina, analitzarem quin rol té la informació que tenen els jugadors per a la presa de les seves decisions i en quins casos podrem tenir decisions òptimes o el que més endavant anomenarem equilibris.

2.1 Introducció a la Teoria de jocs

Comencem amb les primeres definicions bàsiques d'aquesta disciplina per, més endavant, poder endinsar-nos en models cada cop més complexes.

Definició 2.1 (Joc estratègic). Un joc estratègic en forma normal és una tupla (N, A, u) formada per:

- Un conjunt finit no buit, N , el qual anomenarem conjunt de jugadors,
- $A = \times_{i \in N} A_i$ el producte cartesià de conjunts no buits A_i indexats per N , als quals anomenem conjunt d'accions possibles (o estratègies) per al jugador i . Al conjunt A l'anomenarem perfil estratègic. Encara que no sempre és així, en aquest treball assumirem, per simplicitat, que A és finit.
- $u = \times_{i \in N} u_i$ on $u_i : A \rightarrow \mathbb{R}$ és una funció anomenada funció de pagament del jugador i , la qual assigna al perfil estratègic $a \in A$ el pagament que rep cada jugador i si es juga a .

Tenint en compte aquesta definició, veiem que per al jugador i una jugada $a \in A_i$ serà millor que una altra $b \in A_i$ quan $u_i(a) \geq u_i(b)$. Per indicar aquesta prioritat entre jugades, farem servir la notació $a \succeq_i b$. De manera anàloga, podem definir un joc estratègic en forma normal com $(N, A, (\succeq_i))$.

Definició 2.2 (Joc de suma zero). Un **joc de suma zero** és un joc estratègic en què la suma dels guanys i les pèrdues dels jugadors és sempre igual a zero. És a dir: Per a tota $a \in A$, $\sum_{i \in N} u_i(a) = 0$. En altres paraules, no hi ha creació de valor ni en guanys ni en pèrdues.

Matemàticament, en un joc de suma zero de 2 jugadors es compleix que:

$$u_1(a_1, a_2) = -u_2(a_1, a_2)$$

on u_1 i u_2 són els pagaments dels jugadors 1 i 2 respectivament, i a_1 i a_2 són les estratègies escollides per cada jugador.

Definició 2.3 (Matriu de pagaments). En un joc estratègic de dos jugadors, la *matriu de pagament* és una representació matricial dels guanys associats a cada combinació d'estratègies pures.

Sigui $A = \{a_1, a_2, \dots, a_m\}$ el conjunt d'estratègies del jugador 1 i $B = \{b_1, b_2, \dots, b_n\}$ el conjunt d'estratègies del jugador 2. Aleshores, la matriu de pagament per al jugador 1 és una matriu $M \in \mathbb{R}^{m \times n}$ tal que:

$$M_{ij} = u_1(a_i, b_j).$$

On $u_1(a_i, b_j)$ és el guany del jugador 1 quan juga a_i i el jugador 2 juga b_j . De manera anàloga, el jugador 2 pot tenir una matriu de pagament $M' \in \mathbb{R}^{m \times n}$ on:

$$M'_{ij} = u_2(a_i, b_j).$$

En un joc de suma zero, es compleix que $M' = -M$.

2.2 Estratègies en jocs

En molts jocs, els jugadors poden optar per prendre la seva decisió de manera determinista o bé prendre la decisió de manera probabilística. Aquesta distinció entre el que anomenarem estratègies pures i estratègies mixtes és clau per entendre com els jugadors, prenen les seves decisions. Ens referirem a estratègies pures com a aquelles que són deterministes, és a dir, amb la informació que el jugador té sobre el joc, sempre prendrà la mateixa decisió. En canvi, ens referirem a estratègies mixtes quan, en una mateixa situació, un jugador determinarà la seva estratègia de manera probabilística. Dins la teoria de jocs, per referir-se a la distribució de probabilitat sobre les possibles accions d'un jugador, s'utilitza la paraula loteria.

Definició 2.4 (Loteria). Sigui A_i el conjunt d'estratègies pures disponibles per al jugador i , una loteria (o *estratègia mixta*) és una funció:

$$\alpha_i : A_i \rightarrow [0, 1] \quad \text{tal que} \quad \sum_{a \in A_i} \alpha_i(a) = 1.$$

On $\alpha_i(a)$ representa la probabilitat que el jugador i jugui l'acció $a \in A_i$.

Definició 2.5 (Extensió mixta). L' **extensió mixta** del joc estratègic (N, A, u) és el joc estratègic $(N, (\Delta(A_i)_{i \in N}), (U_i)_{i \in N})$, on:

- $\Delta(A_i)$ és el conjunt de loteries sobre A_i i
- $U_i : \times_{j \in N} \Delta(A_j) \rightarrow \mathbb{R}$ assigna a cada $\alpha \in \times_{j \in N} \Delta(A_j)$ el valor esperat segons u_i de la loteria sobre A que és induïda per α .

De manera que

$$U_i(\alpha) = \sum_{a \in A} \left(\prod_{j \in N} \alpha_j(a_j) \right) u_i(a)$$

si A és finit.

Aquest enfocament ens permet considerar situacions en què els jugadors no es limiten a triar una única estratègia de manera determinista, sinó que assignen probabilitats a diferents accions possibles.

Remarca 2.6. Veiem clar que les estratègies pures són un subconjunt de les estratègies mixtes on la distribució de probabilitat pren valors en $\{0, 1\}$.

Definició 2.7 (Estratègia estrictament dominada). En un joc amb n jugadors, es diu que una estratègia a_i està estrictament dominada per una estratègia a_i^* si a_i^* proporciona sempre una utilitat estrictament superior a a_i , independentment de les estratègies dels altres jugadors. Així doncs, si

$$u_i(a_1, \dots, a_i^*, \dots, a_n) > u_i(a_1, \dots, a_i, \dots, a_n)$$

per a qualsevol estratègia $a_i \neq a_i^*$ i per a qualsevol combinació d'estratègies dels oponents $(a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n)$. Anomenem a_i^* l'estratègia estrictament dominant.

2.3 Equilibris de Nash

En aquest apartat arribem a un dels conceptes clau dins la teoria de jocs, els equilibris de Nash. Aquests ens permetran poder estudiar els diferents jocs, de tal manera que puguem obtenir una estratègia “òptima”.

Definició 2.8 (Equilibri de Nash). Donat un joc estratègic $G = (N, A, u)$. Sigui $(a_{-i}, \hat{a}_i) = (a_1, a_2, \dots, a_{i-1}, \hat{a}_i, a_{i+1}, \dots, a_N)$. Un equilibri de Nash es defineix com un perfil estratègic $a^* \in A$ amb la propietat que, per a tot i dins N i per a tot $\hat{a}_i \in A_i$ es satisfà

$$u_i(a^*) \geq u_i(a_{-i}^*, \hat{a}_i).$$

És a dir, trobem un equilibri de Nash al perfil estratègic a^* quan, cap jugador j vol prendre una altra estratègia $\hat{a}_j$ perquè el seu guany serà menor. Posem un parell d'exemples senzills:

Exemple 2.9. Considerem el joc següent:

	A	B
A	(1, 1)	(0, 0)
B	(0, 0)	(1, 1)

En aquest joc, dos jugadors trien entre les accions A i B simultàniament. Si tots dos trien el mateix moviment, reben un pagament d'1. Si trien accions diferents, reben 0.

- (A, A) és un equilibri de Nash a^* .
- (B, B) també és un equilibri de Nash a^{**} .

Cap jugador i té una estratègia $\hat{a}_i$ tal que desviant-se del perfil estratègic a^* o a^{**} , pot millorar el seu pagament, per tant, són equilibris de Nash.

Exemple 2.10 (Dilema dels presoners). *Dos sospitosos d'un crim són tancats en cel·les separades. Si tots dos confessen, cadascun complirà una condemna de tres anys de presó. Si només un d'ells confessa, aquest quedarà en llibertat, mentre que el seu company rebrà una pena de quatre anys. En canvi, si cap dels dos confessa, seran condemnats per un delictes menor i passaran un any a la presó. Si representem aquest dilema amb una matriu de pagaments, veurem que és un joc en què la millor opció per als dos jugadors seria no confessar (si sumem les dos funcions pagament, és la que té el valor màxim). Tot i això, cadascun d'ells té un incentiu per traïr l'altre, ja que, independentment del que faci el seu company, sempre li convé més confessar. Això fa que l'únic equilibri de Nash sigui la situació en què tots dos confessen.*

	Confessar	No confessar
Confessar	$(-3, -3)$	$(0, -4)$
No confessar	$(-4, 0)$	$(-1, -1)$

Definició 2.11 (Equilibri de Nash en estratègies mixtes). Un perfil estratègic, d'estratègies mixtes $\alpha^* = (\alpha_1^*, \dots, \alpha_n^*)$ és un equilibri de Nash en estratègies mixtes si, per a tot jugador i , es compleix que

$$U_i(\alpha_i^*, \alpha_{-i}^*) \geq U_i(\alpha_i, \alpha_{-i}^*) \quad \forall \alpha_i \in \Delta(A_i).$$

Això significa que cap jugador té incentius per a desviar-se unilateralment de la seva estratègia mixta α_i^* , ja que fer-ho no millora la seva utilitat esperada.

El concepte de “resposta òptima” que porta l'equilibri de Nash, ens porta directament a la funció de millor resposta.

Definició 2.12 (Funció de millor resposta). Definim $B_i : A_{-i} \rightarrow \mathcal{P}(A_i)$ la funció de millor resposta per al jugador i com una funció que va de les estratègies de tots els jugadors menys del jugador i a les parts de A_i . Aquesta funció assigna a $a_{-i} \in A_{-i}$ el conjunt de les millors accions per al jugador i donat que els altres jugadors segueixen $a_{-i} \in A_{-i}$, és a dir:

$$B_i(a_{-i}) = \{a_i \in A_i : (a_{-i}, a_i) \succeq_i (a_{-i}, a') \text{ per a tot } a' \in A_i\}.$$

Posem un exemple clàssic per entendre amb profunditat el concepte d'equilibri de Nash en estratègies mixtes.

Exemple 2.13 (Pedra-Paper-Tisora). *Un exemple clàssic d'un joc que no té equilibri de Nash en estratègies pures és el joc de **Pedra-Paper-Tisora**. En aquest joc, dos jugadors trien simultàniament entre tres opcions: pedra (P), paper (L) o tisora (T). El resultat es determina de la següent manera:*

- La pedra guanya a la tisora.
- La tisora guanya al paper.
- El paper guanya a la pedra.
- Si tots dos jugadors trien la mateixa opció, hi ha un empat.

La matriu de pagaments per al jugador 1 és la següent:

	P	L	T
P	(0, 0)	(-1, 1)	(1, -1)
L	(1, -1)	(0, 0)	(-1, 1)
T	(-1, 1)	(1, -1)	(0, 0)

Per determinar si hi ha un equilibri de Nash en estratègies pures, analitzem les possibles estratègies:

- Si el jugador 1 tria pedra (P), el jugador 2 preferirà triar paper (L) per guanyar.
- Si el jugador 1 tria paper (L), el jugador 2 preferirà triar tisora (T).
- Si el jugador 1 tria tisora (T), el jugador 2 preferirà triar pedra (P).
- Aquest raonament és simètric per al jugador 2.

Atès que sempre existeix un incentiu per desviar-se cap a una estratègia que millori el resultat, no existeix un equilibri de Nash en estratègies pures. Ara bé, sí existeix un **equilibri de Nash en estratègies mixtes**:

Suposem que cada jugador tria les seves estratègies de manera mixta. Anomenem p_1 , p_2 , i p_3 les probabilitats de triar Pedra, Paper i Tisora, respectivament, pel jugador 1, i q_1 , q_2 , i q_3 les probabilitats del jugador 2. Així, tenim:

$$p_1 + p_2 + p_3 = 1 \quad (\text{per al jugador 1})$$

$$q_1 + q_2 + q_3 = 1 \quad (\text{per al jugador 2})$$

Per trobar les probabilitats que garanteixen l'equilibri de Nash, calculem els pagaments esperats per a cada possible estratègia. El valor esperat per al jugador 1 quan tria Pedra és:

$$V(\text{Pedra}) = 0 \cdot q_1 + (-1) \cdot q_2 + 1 \cdot q_3 = -q_2 + q_3$$

El valor esperat per al jugador 1 quan tria Paper és:

$$V(\text{Paper}) = 1 \cdot q_1 + 0 \cdot q_2 + (-1) \cdot q_3 = q_1 - q_3$$

El valor esperat per al jugador 1 quan tria Tisora és:

$$V(\text{Tisora}) = (-1) \cdot q_1 + 1 \cdot q_2 + 0 \cdot q_3 = -q_1 + q_2$$

Veiem que el pagament esperat del jugador 1 depèn totalment de l'estratègia del jugador 2, com és un joc de suma zero, si el jugador 2 minimitza el guany del jugador 1, està maximitzant el seu. Sabent això, com el guany del primer jugador és, com a mínim, el mínim entre aquests valors atesos, el jugador 2 vol minimitzar z tal que :

$$z \leq -q_2 + q_3$$

$$z \leq q_1 - q_3$$

$$z \leq -q_1 + q_2$$

$$q_1 + q_2 + q_3 = 1 \quad \text{amb} \quad q_1, q_2, q_3 \geq 0$$

Amb el mètode Símplex es conclou ràpidament que $q_1 = q_2 = q_3 = \frac{1}{3}$. De manera anàloga es pot calcular p_1, p_2 i p_3 .

Aquesta estratègia fa que cap jugador pugui millorar el seu guany esperat canviant unilateralment la seva estratègia, ja que qualsevol desviació faria que l'oponent pugui explotar la nova distribució de probabilitat.

Conclusió: El joc Pedra-Paper-Tisores no té un equilibri de Nash en estratègies pures, però sí en mixtes on cada jugador tria cada acció amb probabilitat $1/3$. Aquesta estratègia assegura que cap jugador tingui un avantatge sobre l'altre a llarg termini.

Així doncs, veiem quina condició necessitem per a que hi existeixi l'equilibri de Nash a un joc estratègic. Recordem primer un enunciat topològic:

Teorema 2.14 (Teorema del punt fix de Kakutani). *Sigui X un subconjunt compacte i convex de $\mathbb{R}^n$ i sigui $f : X \rightarrow \mathcal{P}(X)$ tal que:*

- *Per a tot $x \in X$, el conjunt $f(x)$ és no buit i convex.*
- *El graf de f és tancat (és a dir, per a totes les successions $\{x_n\}$ i $\{y_n\}$ tals que $y_n \in f(\{x_n\})$ per a tot n , si $x_n \rightarrow x$ i $y_n \rightarrow y$, llavors tenim $y \in f(x)$).*

Llavors existeix $x^ \in X$ tal que $x^* \in f(x^*)$.*

Amb aquest enunciat ja podem provar un dels resultats fonamentals en teoria de jocs (veure, per exemple, [11, Thm. 20.3]).

Teorema 2.15 (Existència de l'equilibri de Nash). *El joc estratègic $(N, (A_i), (\succeq_i))$ té un equilibri de Nash si per a tot $i \in N$:*

- *El conjunt A_i d'accions del jugador i és un subconjunt compacte, convex i no buit d'un espai euclidià.*

I la relació de preferència $\succeq_i$ és :

- *contínua,*
- *quasi-còncava sobre A_i .*

Recordem que, una funció f es diu que és quasi-concava si es compleix:

$$f(\lambda x + (1 - \lambda)y) \geq \min\{f(x), f(y)\}$$

És a dir, el valor de la funció en qualsevol combinació convexa de dos punts no és menor que el mínim dels valors en aquests punts. La relació de preferència serà quasi-còncava si u ho és.

Demostració. Definim $B : A \rightarrow A$ com $B(a) = \times_{i \in N} (B_i(a_{-i}))$ on B_i és la funció de millor resposta Definició 2.12. Per a cada $i \in N$, el conjunt $B_i(a_{-i})$ és no buit perquè $\succeq_i$ és contínua i A_i és compacte, i és convex perquè $\succeq_i$ és quasi-còncava sobre A_i . A més, B té un graf tancat ja que cada $\succeq_i$ és contínua. Per tant, pel Teorema 2.14, B té un punt fix. És trivial observar que qualsevol punt fix de la funció B correspon a un equilibri de Nash. $\square$

Un cop demostrat aquest resultat, és natural preguntar-se per què un joc pot admetre un equilibri en estratègies mixtes però no en estratègies pures. La resposta queda clara un cop provat el següent teorema trobat a [11]:

Teorema 2.16 (Equilibris de Nash en estratègies mixtes). *Tot joc estratègic finit, té un equilibri de Nash en estratègies mixtes.*

Demostració. Sigui $G = (N, A, u)$ un joc estratègic, i per a cada jugador i , sigui m_i el nombre d'elements del conjunt A_i . Llavors, podem identificar el conjunt $\Delta(A_i)$ d'estratègies mixtes del jugador i amb el conjunt de vectors

$$(p_1, p_2, \dots, p_{m_i})$$

per als quals $p_k \geq 0$ per a tot k i

$$\sum_{k=1}^{m_i} p_k = 1,$$

on p_k representa la probabilitat amb què el jugador i utilitza la seva k -èsima estratègia pura. És fàcil veure que aquest conjunt és no buit, convex i compacte. Atès que el pagament esperat és lineal en les probabilitats, la funció de pagament de cada jugador en l'extensió mixta de G és contínua i quasi-còncava en la seva pròpia estratègia. Així, l'extensió mixta de G satisfà tots els requisits de Teorema 2.15. $\square$

2.4 Minimax

Ara farem una breu introducció a les estratègies *Minimax*, atesa la seva gran importància dins la teoria de jocs. Tot i això no ens seran de gaire ajuda per entendre quin rol exerceix la informació en els equilibris de Nash. Sense els resultats fonamentals de Von Neumann sobre el *Minimax*, no hi hauria teoria de jocs com la coneixem ara. Una pregunta fonamental dins la teoria de jocs és: com podem prendre la millor decisió davant d'un adversari racional? En jocs de

suma zero, on el guany d'un jugador representa exactament la pèrdua de l'altre, aquest problema es torna especialment rellevant. El **teorema *Minimax***, proposat per John von Neumann el 1928, proporciona una resposta elegant a aquesta qüestió. Aquest resultat fonamental estableix que, en qualsevol joc estratègic finit de suma zero amb dues persones, existeix un valor òptim que cada jugador pot garantir mitjançant una estratègia adequada. Concretament, cada jugador pot minimitzar la seva màxima pèrdua, assegurant-se el millor resultat possible contra un adversari òptim. En aquest cas utilitzarem les referències de [11] i [2]. Per simplificar assumirem un joc amb només $N = 2$, per tant sota aquestes condicions, en un joc de suma zero, els beneficis d'un són les pèrdues de l'altre i viceversa. (Les matrius de pagament del joc són iguals en valor absolut, però amb el signe canviat).

Definició 2.17 (Estratègies *minimax* i *maximin*). L'estratègia *maximin* és aquella que pretén maximitzar el benefici mínim que pot obtenir independentment de l'estratègia que prengui el rival. Recíprocament, l'estratègia *minimax* és aquella que tracta de minimitzar el màxim benefici que pot assolir l'oponent independentment de l'estratègia que prengui.

És a dir, si suposem que, en un joc de suma zero, el jugador 1 pren l'estratègia *maximin* i el jugador 2 l'estratègia *minimax* i, sigui B el benefici del jugador 1 (el benefici de 2 serà $-B$ per tractar-se d'un joc de suma zero) tenim una fita superior i una inferior per a aquest B :

$$\max_{a_1 \in A_1} \min_{a_2 \in A_2} U(a_1, a_2) \leq B \leq \min_{a_2 \in A_2} \max_{a_1 \in A_1} U(a_1, a_2).$$

Definició 2.18 (Punt de sella). Sigui G un joc de suma zero per 2 jugadors i M la seva matriu de pagaments. Anomenarem punt de sella a l'element a_{ij} si és el més gran de la seva columna i el més petit de la seva fila.

Remarca 2.19. Observem, per tant, que si una estratègia (a_1^*, a_2^*) condueix a un punt de sella, cap jugador té incentius a desviar-se d'aquesta estratègia. És el màxim de la columna, per tant, si el jugador 2 no canvia l'estratègia a 1 no li convé desviar-se i el més petit de la fila (el més gran amb el signe canviat), per tant si 1 no canvia l'estratègia, tampoc al jugador 2 li convé fer-ho. Així doncs, és trivial que punt de sella $\implies$ equilibri de Nash.

Remarca 2.20. Al definir-se a la matriu de pagaments les diferents files com les diferents estratègies del jugador 1 i les columnes com les estratègies del jugador 2. És també trivial que, si

$$\max_{a_1 \in A_1} \min_{a_2 \in A_2} U(a_1, a_2) = \min_{a_2 \in A_2} \max_{a_1 \in A_1} U(a_1, a_2).$$

Aleshores existeix com a mínim un punt de sella.

Exemple 2.21 (Estratègies minimax i maximin). *Considerem un joc de suma zero entre dos jugadors, on el jugador 1 tria una de les seves estratègies A_1, A_2 i el jugador 2 tria entre B_1, B_2, B_3 . La següent taula mostra la matriu de*

pagaments M del jugador 1 (com és un joc de suma zero, la matriu del jugador 2 serà $-M$):

	B_1	B_2	B_3
A_1	3	-1	2
A_2	1	0	-2

El jugador 1 vol maximitzar el seu guany mínim trobem $\max_{a_1 \in A_1} \min_{a_2 \in A_2} U(a_1, a_2)$. Així doncs, calcularem primer el mínim de cada fila:

$$\min(A_1) = \min(3, -1, 2) = -1.$$

$$\min(A_2) = \min(1, 0, -2) = -2.$$

El jugador 1 escull l'estratègia que maximitza aquest mínim:

$$\max(-1, -2) = -1.$$

Per tant, l'estratègia maximin per al jugador 1 és jugar A_1 . Sabem que el pagament del jugador 1 està acotat inferiorment per -1 . El jugador 2 vol minimitzar el màxim guany del jugador 1. Per seguir aquesta estratègia calculem els màxims de cada columna:

$$\max(B_1) = \max(3, 1) = 3.$$

$$\max(B_2) = \max(-1, 0) = 0.$$

$$\max(B_3) = \max(2, -2) = 2.$$

El jugador 2 escull l'estratègia que minimitza aquest màxim:

$$\min(3, 0, 2) = 0.$$

Per tant, l'estratègia minimax per 2 és jugar B_2 . Amb aquest resultat, també obtenim fita superior per al benefici del jugador 1, que és 0. Si el valor maximin coincideix amb el valor minimax, tenim un punt de sella. En aquest cas:

$$-1 \neq 0.$$

Per tant sigui F el benefici del jugador 1, sabem que $F \in (-1, 0)$, efectivament si els dos jugadors juguen les seves estratègies maximin i minimax, prenen les accions (A_1, B_2) que li donarà un pagament de -1 al jugador 1, que és dins aquest interval.

Es poden aplicar les mateixes estratègies *maximin* i *minimax* en estratègies mixtes on l'estratègia *maximin* intenta trobar x_0 tal que, sigui X el conjunt d'estratègies mixtes per al jugador 1, Y el conjunt d'estratègies mixtes per al jugador 2, i M la matriu de pagament del joc:

$$v_1 = \min_{y \in Y} x_0 M y^T = \max_{x \in X} \min_{y \in Y} x M y^T.$$

Aquest valor, v_1 , s'anomena valor del joc per al jugador 1, i de manera corresponent, el valor del joc per al jugador 2 és:

$$v_2 = \max_{x \in X} xMy_0^T = \min_{y \in Y} \max_{x \in X} xMy^T.$$

Von Neumann va arribar a la següent conclusió:

Teorema 2.22 (Minimax, Von Neumann). *Si G un joc estratègic de suma zero de 2 jugadors aleshores: $v_1 = v_2$.*

Per raons de breuetat, no inclourem la demostració d'aquest teorema atès que no és essencial per al tema central d'aquest treball. La demostració es pot consultar a [12].

2.5 Jocs extensius d'informació perfecta

Abans d'introduir els jocs d'informació incompleta, cal presentar els jocs extensius, començant per la informació completa perfecta, fins a arribar a la imperfecta, per això acudirem a [11].

Fins ara només hem considerat jocs on l'estratègia de cada jugador és fixa des de l'inici del joc, en canvi, hi ha moltes situacions, on les decisions es prenen de manera seqüencial i l'acció prèvia d'un jugador afecta l'estratègia de l'altre. Per això, a la noció de joc estratègic cal afegir diferents nocions, com la de "torn", i sorgeix la definició de joc extensiu, que definirem primerament en el cas d'informació perfecta:

Definició 2.23 (Joc extensiu d'informació perfecta). *Un joc extensiu d'informació perfecta és una tupla (N, H, P, u) on els components d'aquest joc són:*

- Un conjunt de jugadors $N = \{1, 2, \dots, n\}$.
- Un conjunt H de seqüències (finites o infinites) que compleixen les següents propietats:
 - La seqüència buida $\emptyset$ és membre de H .
 - Si $(a^k)_{k=1, \dots, K} \in H$ (on K pot ser finit o infinit) i $L < K$, llavors $(a^k)_{k=1, \dots, L} \in H$.

Cada element de H és una *història*, on cada component d'una història és una acció presa per un jugador. Una història $(a^k)_{k=1, \dots, K}$ és *terminal* si no hi ha cap moviment a^{K+1} tal que $(a^k)_{k=1, \dots, K+1} \in H$ o si no hi ha cap moviment rellevant que es pugui fer posteriorment. El conjunt d'històries terminals es denota per Z .

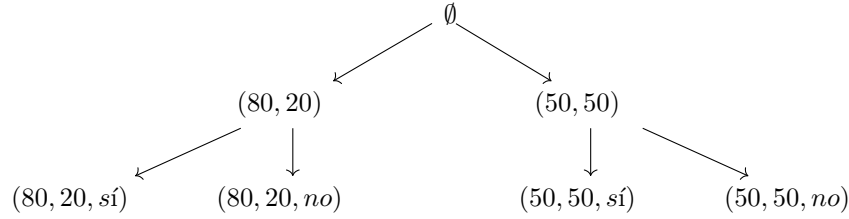
- Una funció $P : H \setminus Z \rightarrow N$ que assigna a cada història no terminal (cada element de $H \setminus Z$) un jugador de N . Aquesta funció P s'anomena la funció del jugador i descriu quin jugador pren una decisió a cada història. $(P(h))$ és el jugador que ha de fer un moviment després d'haver arribat a h

- La funció $u = \times_{i \in N} u_i$ on $u_i : Z \rightarrow \mathbb{R}$ assigna a cada història terminal un pagament per a cada jugador.

Aquest joc es caracteritza per una informació perfecta en què tots els jugadors coneixen totalment la història actual en qualsevol punt del joc.

Aquests tipus de jocs es representen com a arbres decisius com al següent exemple.

Exemple 2.24 (Joc extensiu amb informació perfecta). *Suposem un joc de 2 jugadors on s'han de repartir 100 fitxes de la següent manera, el jugador 1 pot escollir entre quedar-se 80 i donar-li al altre jugador 20 o repartir 50 – 50. I el jugador 2, un cop se li fa l'oferta, ha de dir sí o no, si accepta la oferta es compleix i si la rebutja, els dos es queden sense fitxes. En aquest cas, el joc es representa com:*



En aquest cas:

- $N = \{1, 2\}$
- $H = \{\emptyset, (80, 20), (50, 50), (80, 20, \text{sí}), (80, 20, \text{no}), (50, 50, \text{sí}), (50, 50, \text{no})\}$.
- $P(\emptyset) = 1$ i sigui h qualsevol altre element de $H \setminus Z$ aleshores $P(h) = 2$.
- $u((80, 20, \text{sí})) = (80, 20)$, $u((80, 20, \text{no})) = (0, 0)$, $u((50, 50, \text{sí})) = (50, 50)$, $u((50, 50, \text{no})) = (0, 0)$

Abans d'introduir el concepte d'estratègia en aquest tipus de jocs, cal definir el conjunt d'accions disponibles en cada moment de decisió.

Definició 2.25 (Accions disponibles). Donada una història no terminal $h \in H \setminus Z$, definim el conjunt d'accions disponibles per al jugador $P(h)$ immediatament després d'arribar a h com:

$$A(h) = \{a : (h, a) \in H\}.$$

Així doncs, $A(h)$ conté totes les accions a que poden seguir la seqüència h i que, concatenades amb h , donen lloc a una nova història vàlida dins del conjunt H .

Definició 2.26 (Estratègia pura en joc extensiu amb informació perfecta). Una **estratègia pura del jugador i** en un joc extensiu amb informació perfecta és una funció s_i que assigna una acció $a \in A(h)$ a cada història no terminal

$h \in H \setminus Z$ tal que $P(h) = i$. Un **perfil estratègic** $s = (s_1, \dots, s_n)$ conté una estratègia per a cada jugador del conjunt N , i s'utilitza la notació s_{-i} per referir-se al perfil estratègic excepte la del jugador i .

La interpretació d'una estratègia en un joc extensiu és “tenir una resposta preparada per a cada estratègia que té disponible el rival”. Per exemple, a l'exemple d'abans el jugador 2 té una estratègia (a, b) si pren la decisió a quan el jugador 1 decideix fer $(80, 20)$ i pren la decisió b si el jugador 1 decideix fer $(50, 50)$. Utilitzarem la notació $s_i(h)$ per denominar l'estratègia que ha de prendre el jugador i , seguint els preceptes de s_i , un cop ja ha succeït fins aleshores la història h .

Definició 2.27 (Resultat). Donat un joc extensiu (N, H, P, u) , i un perfil estratègic $s = (s_i)_{i \in N}$, definim el **resultat** $R(s)$ de s com la història terminal (possiblement infinita) que resulta quan cada jugador $i \in N$ segueix els preceptes de la seva estratègia s_i .

Per simplicitat, assumirem que les històries terminals són finites. Formalment, $R(s) = (a_1, a_2, \dots, a_k) \in Z$, on $Z \subseteq H$ és el conjunt d'històries terminals i tal que per a cada $0 \leq k < K$, es compleix:

$$s_{P(a_1, \dots, a_k)}(a_1, \dots, a_k) = a_{k+1}.$$

És a dir, a cada història no terminal $(a_1, \dots, a_k)$, el jugador que ha de jugar, $P(a_1, \dots, a_k)$, escull l'acció a_{k+1} d'acord amb la seva estratègia s_i .

Amb aquestes definicions podem definir finalment l'equilibri de Nash en un joc extensiu amb informació perfecta.

Definició 2.28 (Equilibri de Nash en joc extensiu amb informació perfecta).

Un equilibri de Nash d'un joc extensiu amb informació perfecta

(N, H, P, u) és un perfil estratègic s^* tal que per a cada jugador $i \in N$ tenim

$$R(s_{-i}^*, s_i^*) \succeq_i R(s_{-i}^*, s_i) \quad \text{per a tota estratègia } s_i \text{ del jugador } i.$$

Encara que l'equilibri de Nash és un concepte essencial dins l'àmbit de la teoria de jocs, als jocs extensius no s'aplica massa bé ja que pot ser inconsistent amb la natura seqüencial d'aquests i incoherent amb la realitat i preferències del jugador, posem un exemple:

Exemple 2.29. Sigui un joc extensiu G de 2 jugadors on el jugador 1 pot escollir entre les accions (A, B) i el jugador 2 entre les accions (L, R) quan el jugador 1 escollirà A , s'acaba el joc directament però quan el jugador 1 escollirà B i tenim els següents pagaments:

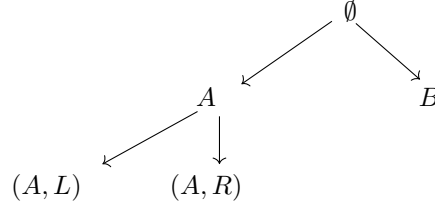
$$1. u(A, L) = (0, 0),$$

$$2. u(A, R) = (2, 1),$$

3. $u(B, L) = (1, 2)$ i

4. $u(B, R) = (1, 2)$.

Representat pel següent arbre decisiu:



Amb la definició d'equilibri de Nash a jocs extensius, tenim en aquest joc 2 equilibris de Nash, (A, R) i (B, L) . El perfil estratègic (B, L) és un equilibri de Nash perquè, donat que el jugador 2 tria L després de la història A , és òptim per al jugador 1 triar B al començament del joc (si tria A en lloc de B , aleshores, donada l'elecció del jugador 2, obté 0 en lloc de 1), i donada l'elecció del jugador 1 de B , és òptim per al jugador 2 triar L (ja que la seva elecció no fa cap diferència en el resultat). Podem veure que el fet que un jugador pugui reavaluar el seu joc a cada història no terminal, fa que aquest equilibri manqui de sentit. Si es produís la història A , sembla que el jugador 2 triaria R en lloc de L , ja que obté una major recompensa fent-ho. L'equilibri (B, L) es manté per l'amenaça del jugador 2 de triar L si el jugador 1 tria A . Aquesta amenaça no és creïble ja que el jugador 2 no té cap manera de comprometre's a aquesta elecció. Així doncs, el jugador 1 pot estar segur que si tria A , el jugador 2 triarà R ; atès que ell prefereix el resultat (A, R) a l'equilibri de Nash (B, L) , té un incentiu per desviar-se de l'equilibri i triar A .

Amb aquest exemple podem veure que l'equilibri de Nash no sembla ideal per a jocs extensius, per això arribarem a un altre equilibri ideal per a aquest tipus de jocs.

Definició 2.30 (Subjoc). El subjoc, d'un joc extensiu d'informació perfecta $G = (N, H, P, u)$ que segueix la història h , és el joc extensiu d'informació perfecta $G(h) = (N, H|_h, P|_h, u|_h)$. On $H|_h$ és el conjunt de seqüències h' d'accions per a les quals $(h, h') \in H$, P_h està definit per $P|_h(h') = P(h, h')$ per a cada $h \in H|_h$, i $u|_h(h') = u(h, h')$.

La noció d'equilibri que volem definir requereix que cada jugador prengui la seva estratègia òptima tenint en compte l'estratègia del seu rival a **cada** història del joc. Denominarem $s_i|_h$ l'estratègia induïda per $G(h)$ i $R|_h$ la funció resultat del joc $G(h)$.

Definició 2.31 (Equilibri perfecte en subjocs). Un equilibri perfecte en subjocs d'un joc extensiu amb informació perfecta $G = (N, H, P, u)$ és un perfil estratègic s^* tal que per a cada jugador $i \in N$ i cada història no terminal $h \in H \setminus Z$ per a la qual $P(h) = i$, tenim que

$$u_i|_h(R|_h(s_{-i}^*|_h, s_i^*|_h)) \geq u_i|_h(R|_h(s_{-i}^*|_h, s_i|_h))$$

per a cada estratègia s_i del jugador i en el subjoc $G(h)$.

És a dir, un equilibri perfecte en *subjocs* és un perfil estratègic s^* en G tal que per a tota història no terminal h , el perfil estratègic $s^*|_h$ és equilibri de Nash en $G(h)$. Per veure quan podem assegurar aquesta mena d'equilibris, encara ens queda un enunciat.

Lema 2.32 (Propietat de la desviació única). *Sigui $G = (N, H, P, u)$ un joc extensiu amb informació perfecta i finit. El perfil estratègic s^* és un equilibri perfecte en subjocs de G si i només si per a cada jugador $i \in N$ i cada història no terminal $h \in H$ per a la qual $P(h) = i$, tenim que*

$$u_i|_h(R|_h(s_{-i}^*|_h, s_i^*|_h)) \geq u_i|_h(R|_h(s_{-i}^*|_h, s_i|_h))$$

per a cada estratègia s_i del jugador i en el subjoc $G(h)$ que només es diferencia en l'acció que indica després de la història inicial del joc $G(h)$.

Demostració. Si s^* és un equilibri perfecte en *subjocs* de G , aleshores compleix aquesta condició. Suposem que s^* no és un equilibri perfecte en *subjocs*; suposem que el jugador i pot desviar-se de manera rentable en el subjoc $G(h')$. Aleshores existeix una estratègia de desviació rentable s_i del jugador i en $G(h')$ per a la qual $s_i(h) \neq (s_i^*|_{h'})(h)$ per a un nombre de històries h no més gran que la longitud de $G(h')$, ja que G és finit, aquest nombre també ho és. Així, entre totes les estratègies de desviació rentables possibles de jugador i en $G(h')$, tria una estratègia s_i per a la qual la quantitat d'històries h tals que $s_i(h) \neq (s_i^*|_{h'})(h)$ sigui la més petita possible. Ara, h^* és la història més llarga de $G(h')$ (la més avançada dins l'arbre) per a la qual $s_i(h) \neq (s_i^*|_{h'})(h)$. La història inicial de $G(h^*)$ és l'única història en $G(h^*)$ en què l'estratègia prescrita per s_i difereix de la prescrita per $s_i^*|_{h'}$. A més, $s_i^*|_h$ és una desviació rentable en $G(h^*)$, ja que d'altra manera hi hauria una desviació rentable en $G(h')$, que diferiria de $s_i^*|_h$ després de menys històries que s_i . Així, $s_i|_{h^*}$ és una desviació rentable en $G(h^*)$ que es diferencia de $s_i^*|_{h^*}$ en l'estratègia que indica després de la història inicial del joc $G(h^*)$. $\square$

Ara ja estem preparats per veure el següent:

Teorema 2.33 (Existència equilibri perfecte en *subjocs*). *Tot joc extensiu amb informació perfecta, que sigui finit (té un nombre d'històries finites), té un equilibri perfecte en subjocs.*

Demostració. Anomenarem $\ell(G)$ la longitud de la història més llarga en G . Sigui $G = (N, H, P, u)$ un joc extensiu finit amb informació perfecta. Construïm un equilibri perfecte en *subjocs* de G per inducció sobre $\ell(G(h))$; al mateix temps, definim una funció T que associa una història terminal a cada història $h \in H$ i mostrem que aquesta història és un resultat d'equilibri perfecte en *subjocs* del subjoc $G(h)$. Si $\ell(G(h)) = 0$ (és a dir, h és una història terminal de G), definim $T(h) = h$. Suposem que $T(h)$ està definit per a tot $h \in H$ amb $\ell(G(h)) \leq k$ per a algun $k \geq 0$. Sigui h^* una història per la qual $\ell(G(h^*)) = k + 1$ i $P(h^*) = i$.

Com que $\ell(G(h^*)) = k + 1$, tenim que $\ell(G(h^*, b)) \leq k$ per a tot $b \in A(h^*)$. Definim $a_i(h^*)$ per ser un maximitzador del pagament del jugador i de $T(h^*, b)$ sobre $b \in A(h^*)$, i definim $T(h^*) = T(h^*, s_i(h^*))$. Per inducció, hem definit un perfil estratègic s en G ; per *Lemma 2.32*, aquest perfil estratègic és un equilibri perfecte en *subjocs* de G . $\square$

Aquest model seqüencial també té les seves limitacions ja que ens requereix que tots els jugadors sàpiguen en tot moment la història del joc. Si afegim incertesa sobre aquest fet, arribem a la definició de joc extensiu amb informació imperfecta.

2.6 Jocs extensius d'informació imperfecta

Tot i que els jocs extensius amb informació perfecta proporcionen un marc ric per analitzar la racionalitat seqüencial, la realitat sovint presenta escenaris amb incertesa sobre l'estat del joc. Això ens porta al concepte dels jocs amb informació imperfecta, que són aquells on, si bé els jugadors saben a la perfecció les condicions del joc, regles, pagament, estratègies i número de jugadors, desconeixen la història, és a dir, tenen informació parcial sobre les accions preses prèviament.

Definició 2.34 (Joc extensiu amb informació imperfecta). Un joc extensiu amb informació imperfecta és una tupla $(N, H, P, f_c, (\mathcal{I}_i)_{i \in N}, u)$ on (N, H, P) estan definits exactament iguals que als jocs extensius d'informació completa i:

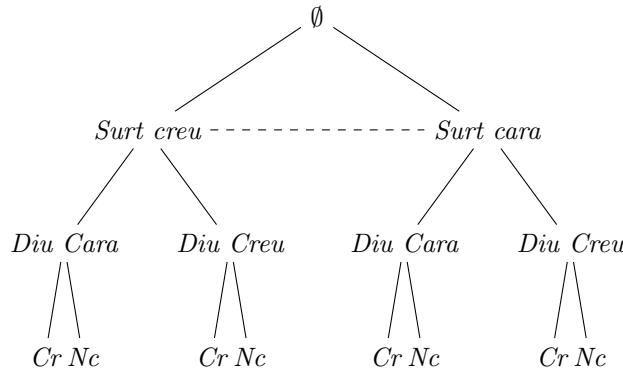
- Una funció f_c que associa a cada història h per a la qual $P(h) = c$ una mesura de probabilitat $f_c(\cdot|h)$ sobre $A(h)$, on cada una d'aquestes mesures de probabilitat és independent de totes les altres. I $f_c(a|h)$ és la probabilitat que a ocorri després de la història h .
- Per a cada jugador $i \in N$, una partició $\mathcal{I}_i$ de $\{h \in H : P(h) = i\}$. Denotarem $\mathcal{I}_i$ com la **partició d'informació** del jugador i ; quan escollim un conjunt de la partició $\mathcal{I}_i$ per al jugador i el denotem per I_i^j . Aquest conjunt l'anomenarem **conjunt d'informació** j del jugador i . Tindrem la propietat que $A(h) = A(h')$ sempre que h i h' pertanyin al mateix conjunt d'informació.
- Per a cada jugador $i \in N$, una funció de pagament $u_i : Z \rightarrow \mathbb{R}$ definida sobre Z , tal que les preferències del jugador i sobre loteries sobre Z poden ser representades mitjançant el valor esperat de u_i .

La partició d'informació $\mathcal{I}_i$ és un element clau en aquesta definició ja que representa el fet que aquests jocs són d'informació imperfecta. Interpretarem totes les històries dins I_i^j com indistingibles per al jugador i , per tant ell sap que és dins alguna història de I_i^j però no ben bé dins quina.

Exemple 2.35. El Jugador 1 llença una moneda en secret. El resultat pot ser Cara (C) o Creu (X). Després, decideix entre dir la veritat o mentir sobre el resultat. El Jugador 2 escolta el que diu el Jugador 1, però no veu la moneda. Llavors decideix si Creure (Cr) o No creure (Nc) el que se li diu. Suposem els pagaments següents:

- Si Jugador 1 diu la veritat i Jugador 2 creu: $u_1 = 2, u_2 = 2$
- Si Jugador 1 menteix i Jugador 2 creu: $u_1 = 3, u_2 = -1$
- Si Jugador 2 no creu i era mentida: $u_1 = 0, u_2 = 1$
- Si Jugador 2 no creu i era veritat: $u_1 = 1, u_2 = 0$

En aquest cas, el conjunt d'informació del jugador 2 si el jugador 1 diu que ha sortit cara és: $(I_1)_2 = \{(surt cara, diu cara), (surt creu, diu cara)\}$. Veiem que en aquest cas I_2 és el conjunt de tots aquests conjunts d'informació com $(I_1)_2$ o $(I_2)_2$ que és $(I_2)_2 = \{(surt cara, diu creu), (surt creu, diu creu)\}$. Aquest tipus de jocs, es representen amb arbres de decisió, i els conjunts d'informació queden units per línies discontinues, de la següent manera:



Per simplicitat de l'arbre, només he escrit l'acció actual i no la història sencera, encara que hauria de ser així, per exemple abaix de tot, en comptes de només Nc, hauria de posar, per exemple, (Surt creu, diu cara, Nc).

En teoria de jocs extensius, els conceptes d'equilibri de Nash i d'equilibri perfecte en subjocs han estat fonamentals per caracteritzar el comportament racional dels jugadors. Tanmateix, quan la informació disponible per als jugadors és imperfecta (és a dir, quan un jugador no pot distingir amb certesa entre diferents històries del joc), aquests conceptes presenten limitacions:

- D'igual manera que abans, l'equilibri de Nash en aquests jocs, presenta limitacions per no estar contemplant la seqüencialitat del mateix.

- L'equilibri perfecte en *subjocs* no és aplicable perquè en jocs amb informació imperfecta els *subjocs* tal com es defineixen (comencen en un únic historial conegut) poden no existir. Per tant, no podem descompondre el joc en *subjocs* independents per aplicar el equilibri perfecte en *subjocs* com havíem fet amb els jocs extensius d'informació perfecta.

En conseqüència, per analitzar correctament jocs amb informació imperfecta, necessitem refinaments més subtils, com l'equilibri seqüencial, estudiat en profunditat per Kreps i Wilson a [8]. Per simplicitat només en parlarem sobre jocs de record perfecte, és a dir, aquells on els jugadors no obliden jugades prèvies, que abans sí coneixien. Per dificultat i extensió de la prova d'existència d'aquest tipus d'equilibris, en aquest cas només la farem d'una manera esque-màtica. La prova completa es pot veure tant a [8] com a [3]. Encara necessitem definir conceptes per introduir la prova.

Definició 2.36 (Estratègia comportamental). Sigui G un joc extensiu finit d'informació imperfecta. Una *estratègia comportamental* per al jugador i és una aplicació:

$$\sigma_j : \mathcal{I}_i \rightarrow \Delta(A(\mathcal{I}_i))$$

que assigna a cada I_i^j una loteria de $A(I_i^j)$ on $A(I_i^j)$ són les accions disponibles per al jugador i quan és al conjunt d'informació I_i^j .

Anomenarem un perfil estratègic comportamentals a σ . A un perfil que assigna a cada conjunt d'informació del jugador i una estratègia comportamental.

Definició 2.37 (Sistema de creences). Un *sistema de creences* μ assigna a cada conjunt d'informació $I_i^j \in \mathcal{I}_i$ una distribució de probabilitat $\mu(\cdot | I_i^j)$ sobre les històries consistents amb I_i^j .

Definició 2.38 (Equilibri seqüencial). Un parell (σ, μ) format per:

- un perfil estratègic comportamentals σ ,
- un sistema de creences μ ,

és un *equilibri seqüencial* si:

1. (Racionalitat seqüencial) Per a cada jugador i i per cada conjunt d'informació $I_i^j \in \mathcal{I}_i$, l'estratègia σ_j maximitza l'esperança de guany donat μ .
2. (Consistència) μ es deriva de σ per la regla de Bayes quan sigui aplicable, i de manera coherent si no, és a dir:

$$\mu(h | I_i^j) = \frac{\mathbb{P}_\sigma(h)}{\sum_{h' \in I_i^j} \mathbb{P}_\sigma(h')} \quad (1)$$

La demostració de l'existència d'equilibris seqüencials (veure, per exemple, [8, Proposition 1], o a [11, Proposition 249.1]) en jocs extensius d'informació imperfecta finits de record perfecte consta de tres passos principals. Com he mencionat abans, només farem una esquematització de la prova per la complexitat i extensió que té:

Teorema 2.39 (Existència d'equilibri seqüencial). *Tot joc extensiu d'informació imperfecta finit de record perfecte té un equilibri seqüencial.*

Esquema de la demostració. Anirem per passos:

Pas 1: Aproximació per estratègies totalment mesclades

Definim una seqüència $\{\sigma^k\}$ d'estratègies comportamentals *totalment mesclades*, és a dir:

- En cada σ^k , totes les accions en qualsevol conjunt d'informació es juguen amb probabilitat estrictament positiva.

Això garanteix que, sota σ^k , totes les històries tenen probabilitat positiva.

Pas 2: Definició de creences via la regla de Bayes

Per cada σ^k , definim el sistema de creences μ^k segons la regla de Bayes(1) on $\mathbb{P}_{\sigma^k}(h)$ és la probabilitat d'arribar a h sota σ^k .

Pas 3: Límit i compacitat

Donada la compacitat dels espais de les estratègies i de les creences:

- Existeix una subseqüència $(\sigma^{k_j}, \mu^{k_j})$ que convergeix a un parell (σ, μ) .

Pas 4: Propietats del parell límit

El parell (σ, μ) satisfà:

- (Consistència) μ és consistent amb σ per construcció.
- (Racionalitat seqüencial) Les estratègies σ_j maximitzen la utilitat esperada a cada conjunt d'informació, respectant μ .

La consistència es desprèn del fet que μ és el límit de creences construïdes via Bayes. La racionalitat seqüencial es deriva per continuïtat de les funcions de pagament i del fet que les desviacions eren negligentment millorables en cada σ^k . Per tant el parell (σ, μ) és un equilibri seqüencial del joc extensiu finit considerat. $\square$

2.7 Jocs amb informació incompleta

Fins ara hem estat analitzant diferents tipus de joc. En els jocs estratègics veiem com els jugadors tenen accés a la mateixa informació i sobre la base d'unes regles i taules de pagament ben clares i definides, podien prendre decisions òptimes amb estratègies mixtes. Posteriorment hem afegit la seqüencialitat, la qual dona lloc als jocs extensius, en què hem analitzat tant el cas d'informació perfecta com d'imperfecta. En aquests tres casos, tenim en comú que els jugadors tenen informació sobre l'entorn bàsic del joc (nombre de jugadors,

accions disponibles, guanys en funció de les accions realitzades). En molts entorns reals, els jugadors poden tenir informació privada que els altres no coneixen. Això dona lloc als anomenats **jocs d'informació incompleta** (quan informació rellevant sobre el joc deixa de ser coneixement comú), en els quals cada jugador ha de formar expectatives sobre les característiques dels altres. L'any 1968, John Harsanyi va desenvolupar la teoria dels jocs amb informació incompleta. En [5] va introduir el model bàsic que permet analitzar situacions on els jugadors tenen coneixement asimètric sobre les característiques del joc. Posteriorment, els equilibris bayesians van ser definits en [6], i la formalització de la distribució de probabilitat es va establir en [7]. El que va aconseguir Harsanyi, és dotar de context als jocs d'informació incompleta que, fins al moment no s'havien sabut estudiar, va saber modelitzar-los en jocs Bayesians com analitzarem més endavant. En un joc amb informació incompleta, almenys un dels jugadors no coneix alguna característica rellevant del joc, en concret els casos més representatius són:

- El nombre de jugadors.
- Les accions disponibles per a un jugador.
- Les funcions de pagament.

Això genera incertesa i obliga a cada jugador a formar creences sobre els possibles estats dels altres participants.

2.7.1 Teoria i Model de Harsanyi

Harsanyi va observar que tot joc d'informació incompleta, on pot faltar-hi informació de qualsevol tipus rellevant al joc es pot reduir al cas en què ens falta informació de la funció de pagament. Per exemple, si manca informació sobre les estratègies o accions disponibles per als jugadors, aleshores podem agafar una representació equivalent del joc, per a cada jugador, existeix un “conjunt universal d'accions” que inclou totes les seves accions concebibles, i aquests conjunts són coneguts per tots els jugadors. Suposem que en el joc original l'acció a del jugador j no és disponible. Com que el jugador j coneix el seu conjunt d'accions, pot distingir quan a és disponible i quan no ho és. Per tant, n'hi ha prou amb definir la funció de pagaments en el nou joc de manera que l'estratègia a pugui ser sempre estrictament dominada quan a no apareix en el joc original. Aquest anàlisi el podem fer perquè des d'un punt de vista de Teoria de Jocs és equivalent un joc on l'estratègia a no existeix a un on l'estratègia a sigui físicament possible però mai escollida.

Suposem ara que hi ha incertesa sobre el conjunt de jugadors. El raonament és anàleg al del cas anterior. Podem considerar un joc amb un “conjunt universal de jugadors” i afegir una estratègia per a cada jugador indicant “estic jugant” o “no estic jugant”. N'hi ha prou amb definir la funció de pagaments de tal manera que els jugadors tinguin els incentius adequats en els diferents estats

possibles. En el cas de “no estic jugant”, l’estratègia “no estic jugant” és estrictament dominant en aquells estats en què el jugador no estava present en el joc original. Finalment, les incerteses sobre la funció de resultats o sobre les preferències dels jugadors es poden modelar conjuntament definint les utilitats dels jugadors directament en l’espai de les estratègies. Així, tota la informació incompleta es pot reduir a un desconeixement de les funcions de pagaments.

L’aproximació teòrica de Harsanyi és fer que cada jugador tracti el joc amb enfocament bayesià, és a dir, cada jugador assignarà una distribució de probabilitat conjunta P_i *subjectiva* a les variables que no conegui i intentarà maximitzar l’esperança de la seva funció de pagament d’acord amb aquesta funció de probabilitat. L’aspecte més rellevant del seu treball és la introducció del concepte Θ d’espai de tipus de jugador, format per aquelles variables desconegudes que afecten a la funció de pagament, Harsanyi introdueix un jugador a aquests jocs, la *Natura*, i el que fa aquesta *Natura* és fixar un $\theta \in \Theta$ abans de començar el joc. Cada jugador i només observa una informació parcial d’aquest estat, representada per θ_i . Cada jugador està informat del seu tipus però no dels tipus de la resta de jugadors.

2.7.2 Equivalència entre Jocs

Que dos jocs siguin equivalents des d’un punt de vista de la teoria de jocs vol dir que es produeixen els mateixos resultats estratègics, incloent-hi els mateixos equilibris de Nash. Atès que hem vist que tots els tipus d’informació incompleta es poden reduir a la falta d’informació sobre la funció de pagament, estudiem un joc d’informació incompleta G des del punt de vista d’un jugador j . En conseqüència, en aquest apartat parlarem sempre en funció de les seves creences.

La notació per la funció de pagament del jugador i és $u_i(a_1, \dots, a_n)$ on a_k és l’estratègia del jugador k , però com assumim que no tots els jugadors tenen tota la informació o coneixen totes les funcions de pagament, durant aquesta secció utilitzarem la notació:

$$u_i(a_1, \dots, a_n) = V^*(a_1, a_2, \dots, a_n, x_{0i}, x_{1i}, \dots, x_{ni}).$$

On V^* és una funció que sí és coneguda per tothom i x_{ki} són les dades rellevants per a la funció u_i que segons j són desconegudes per a alguns jugadors, però conegudes per k i li direm x_0 a la informació que ningú té. Per comoditat, definirem $x = (x_1, x_2, \dots, x_n)$ on

$x_k = (x_{k1}, x_{k2}, \dots, x_{kn})$, d’aquesta manera tenim que aquest vector x_k representa el que j creu que k sap sobre cada u_i , excepte el coneixement comú. En general $x^i = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n)$ seran desconeguts per i . En aquest anàlisi veiem que apareixen (i apareixeran) diverses variables que contribueixen a la visió del joc dels diferents jugadors. Cada jugador i li assignarà a les variables desconegudes per ell una distribució de probabilitat P_i . Li direm b_k a les variables desconegudes pel jugador i de P_k , la distribució de probabilitat *subjectiva* del jugador k a les variables desconegudes per ell. Definirem d’igual

manera que x , el vector b i d'igual manera que x^i , el vector b^i . Per tant adquireix sentit la notació $P_i(x_0, x^i, b^i)$ ja que P_i és la distribució de probabilitat sobre les variables desconegudes per ell, que són x_0 (desconegudes per a tothom), x^i (algú les coneix, però i no) i b^i (algú les coneix, però i no). En general tenim que els jugadors que no siguin i , no coneixeran la distribució subjectiva P_i , però utilitzarem la notació $P_i(x_0, x^i, b^i) = R_i(x_0, x^i, b^i | b_i)$ on R_i és una distribució coneguda per a tots els jugadors. Totes les probabilitats *subjectives* P_i , en realitat seran les condicionades de R_i . Ara introduïm el vector atribut: $c_i = (x_i, b_i)$ i d'igual manera obtindrem c i c^i . Està clar que el vector c_i conté tota la informació que té el jugador i . Per tant la informació incompleta ve del fet que, el valor real d'aquest vector c_i serà només conegut per al jugador i però desconegut per la resta de jugadors. Per aquest motiu direm posteriorment que en un joc d'informació incompleta, els jugadors pertanyen a diferents possibles "tipus" de jugadors, que corresponen als diferents possibles valors del vector c_i . Ara definim $V_i(a_1, \dots, a_n, x, b)$ com el valor atès del pagament del jugador i segons la seva pròpia funció de probabilitat subjectiva (per tant diferent de V^* definit abans). Podem escriure en general $V_i(a_1, \dots, a_n, c)$ així com $P_i(c^i) = R_i(c^i | c_i)$ i amb aquesta construcció podem definir:

Definició 2.40 (Joc d'informació incompleta). La forma estàndard d'un joc d'informació incompleta G de N jugadors és la tupla: $G = \{A, c, V, R\}$.

Harsanyi, per modelitzar aquests jocs en uns que ens seran més fàcils d'estudiar, utilitza l'enfocament bayesià, donant lloc al que anomenem Jocs Bayesians.

2.7.3 Jocs Bayesians

Harsanyi assegura que d'aquesta manera, amb la introducció de la *Natura* i la seva intervenció amb l'assignació de tipus, podem modelitzar un joc d'informació incompleta, en un d'informació completa però imperfecta. Completa perquè tots els jugadors coneixen l'estructura del joc i les distribucions de probabilitats, però imperfecta perquè desconeixen els tipus de la resta. No s'ha de confondre aquest tipus d'informació imperfecta amb els jocs extensius d'informació imperfecta, ja que en ells té un rol essencial la seqüencialitat i en els altres, no necessàriament. Gràcies a Harsanyi, podem modelitzar tots els jocs d'informació incompleta com a Jocs Bayesians, de la següent manera.

Definició 2.41 (Joc bayesià). Un joc bayesià de N jugadors és una tupla $G = (\Theta, A, P, u)$ on

- $\Theta = \theta_1 \times \theta_2 \times \dots \times \theta_N$ on cada θ_i és un conjunt, anomenat de possibles tipus de jugador i , aquest tipus l'entenem com a informació rellevant en possessió d'un jugador, que inclou dades sobre certs valors, creences o probabilitats subjectives referents a valors desconeguts.
- $A = A_1 \times A_2 \times \dots \times A_N$ on cada A_i és un conjunt (conjunt de estratègies o accions).

- P és una distribució de probabilitat conjunta sobre els tipus de jugador, de manera que $p(\theta) > 0$ per a tot $\theta \in \Theta$.
- $u : \Theta \times A \longrightarrow \mathbb{R}$ és la funció que anomenem, de pagaments.

L'equivalència que veurem entre aquests joc és la següent:

Definició 2.42 (Bayes-equivalent). Siguin G un joc d'informació incompleta i G^* un joc bayesià. Es diu que G i G^* són **Bayes-equivalents** si, per a cada jugador j es compleixen les condicions següents:

1. Els dos jocs han de tenir els mateixos **conjunts d'estratègies** $A_1, \dots, A_n$ i s'ha de complir que $c_i = \theta_i$ per a tot $i \in N$.
2. S'ha de complir l'equivalència $(u_1, \dots, u_n) = (V_1, \dots, V_n)$.
3. Les distribucions de probabilitat definides a dalt han de satisfer

$$R_i(c^i \mid c_i) = P(\theta^i \mid \theta_i),$$

Aquesta equivalència garanteix que els dos jocs generen els mateixos incentius estratègics per als jugadors, tot i que poden diferir en com es representa la informació disponible per a cada jugador. És a dir, són equivalents des del punt de vista de teoria de jocs.

En general, per entendre la diferència, ho veiem de la següent manera:

- Joc d'informació incompleta és un terme informal i conceptual, es refereix a la situació de falta de coneixement.
- Joc bayesià és la representació formal del joc d'informació incompleta.

La gran aportació de Harsanyi va ser mostrar que no cal desenvolupar una teoria completament distinta per a jocs d'aquest tipus: és suficient amb ampliar el marc de jocs d'informació completa i incorporar aquests tipus i creences. Així, qualsevol joc d'informació incompleta es pot representar com a un bayesià. Analitzem un exemple senzill per entendre què vol dir cada concepte dins la definició de joc bayesià.

Exemple 2.43 (Subhastes de primer preu). *Una subhasta de primer preu és un mecanisme d'ofertes on múltiples participants presenten ofertes per un bé, i l'oferta més alta guanya. El guanyador paga exactament l'import de la seva oferta. Aquí tenim que:*

- Cada jugador té un valor privat v_i per a l'objecte subhastat, que no és conegut per a la resta de participants, aquest valor és el que abans hem anomenat tipus de jugador, per tant en aquest cas $\theta_i = v_i$
- L'estratègia de cada jugador A_i és la quantitat que ha ofert pel bé subhastat.
- P és la distribució de probabilitat del valor de l'objecte.

- *El pagament és 0 si la nostra subhasta no és la màxima entre tots els jugadors i quan la nostra subhasta sí és la màxima, el pagament és la diferència entre el nostre tipus (el valor que li assignem a l'objecte) i la nostra estratègia (quant hem pagat per l'objecte).*

Definició 2.44 (Estratègia pura en jocs Bayesians). Una estratègia pura per a cada jugador en un joc bayesià G és una aplicació $a_i : \theta_i \longrightarrow A_i$. I escriurem el conjunt de totes les estratègies pures com A^{θ_i} .

Per exemple a l'exemple anterior, una estratègia seria apostar la meitat del meu tipus (el valor privat que jo li dono al objecte). La noció d'equilibri de Nash bayesià es defineix de manera anàloga com s'ha fet a les anteriors definicions d'equilibri de Nash, com per exemple a *Definició 2.8* o a *Definició 2.28*. Amb la diferència essencial que, tal i com hem definit un joc bayesià, en aquest cas estem maximitzant el valor atès del pagament per cada jugador. Ara que ja sabem modelitzar els jocs d'informació incompleta, podem veure dins [6] un dels teoremes més importants de la teoria de jocs, que ens permet estudiar de manera eficient els jocs d'informació incompleta, trobant equilibris de Nash dins aquests.

Ara tornarem amb la notació dels jocs d'informació incompleta per poder acabar de veure l'equivalència entre el model de Harsanyi i els jocs d'informació incompleta. Li direm estratègia normalitzada a $a_k^*(c_k)$, que pren un element del vector atribut i determina una estratègia (una estratègia pura per a jocs d'informació incompleta), podem escriure les funcions de pagament d'abans com

$$g_i = V_i(a_1^*(c_1), \dots, a_n^*(c_n), c_1, \dots, c_n). \quad (2)$$

Si agafem esperances a ambdues bandes, respecte al vector aleatori c amb distribució $R^*(c)$, definim:

$$\mathbb{E}(g_i) = W_i(a_1^*, \dots, a_n^*) = \int_C V_i(a_1^*, \dots, a_n^*, c) d_c R^*(c). \quad (3)$$

Tenim en compte que en aquest tipus de jocs, els jugadors tractaran el seu pagament atès com el seu pagament efectiu, aquesta forma, però, té el problema que està presa amb la probabilitat no-condicionada, però a l'estratègia real, el jugador sabrà el seu tipus (o atribut), per tant utilitzarà la probabilitat condicionada:

$$\mathbb{E}(g_i | c_i) = Z_i(a_1^*, \dots, a_n^* | c_i) = \int_{c_i} V_i(s_1^*, \dots, s_n^*, c_i, c^i) d_{c_i} R^*(c^i | c_i). \quad (4)$$

Veiem amb aquests termes una definició de punt d'equilibri en un joc d'informació incompleta:

Definició 2.45 (Gairebé uniformement una millor resposta). Sigui a_i^* una estratègia normalitzada tal que maximitza el pagament atès condicionat

$$\mathbb{E}(g_i | c_i) = Z_i(a_1^*, \dots, a_n^* | c_i)$$

si la resta de estratègies normalitzades $a_1^*, \dots, a_{i-1}^*, a_{i+1}^*, \dots, a_n^*$ són fixes i suposem que aquesta estratègia normalitzada compleix aquesta propietat per a tots els possibles valors c_i del vector atribut amb la possible excepció d'un petit conjunt C_i^* de valors de c_i , que té una massa de probabilitat total zero. És a dir, suposem que l'esdeveniment $E = \{c_i \in C_i^*\}$ rep una probabilitat zero segons la distribució de probabilitat marginal

$$R^*(c_i) = \int_{C_i^*} d_{(c_i)} R^*(c_i, c^i). \quad (5)$$

La qual es deriva de la distribució de probabilitat bàsica $R^*(c_1, \dots, c_i, \dots, c_n) = R^*(c_i, c^i)$. Llavors (si maximitza el pagament condicionat) direm que a_i^* és *gairebé uniformement* una millor resposta a

$a_1^*, \dots, a_{i-1}^*, a_{i+1}^*, \dots, a_n^*$.

Definició 2.46 (Punt d'equilibri). Sigui G un joc d'informació incompleta, suposem que en una n -tupla de les estratègies normalitzades donada $a^* = (a_1^*, \dots, a_n^*)$, cada component a_i^* ($i = 1, \dots, n$) és gairebé uniformement una millor resposta als altres $(n - 1)$ components $a_1^*, \dots, a_{i-1}^*, a_{i+1}^*, \dots, a_n^*$. En aquest cas, direm que a^* és un *punt d'equilibri* en el joc G .

Teorema 2.47 (Equivalència d'equilibris de Nash). *Sigui G un joc d'informació incompleta i G^* un joc bayesià, Bayes-equivalent a G . Aleshores tenim que a^* , un perfil estratègic, és un punt d'equilibri a G si i només si a^* és equilibri de Nash bayesià a G^* .*

Demostració. Per les equacions d'abans, podem obtenir:

$$W_i(a_1^*, \dots, a_i^*, \dots, a_n^*) = \int_{C_i^*} Z_i(a_1^*, \dots, a_i^*, \dots, a_n^* | c_i) d_{(c_i)} R^*(c_i). \quad (6)$$

On $R^*(c_i)$ és la distribució de probabilitat marginal definida a la definició de gairebé uniformement una millor resposta. Per provar la part de suficiència del teorema, suposem que a^* no és un punt d'equilibri en G . Això significa que com a mínim un dels components de a^* , per exemple a_i^* , no és gairebé uniformement una millor resposta als altres components de a^* . Per tant, ha d'existir algun conjunt C_i^* de valors possibles de c_i en els quals la funció $Z_i(\cdot | c_i)$ podria augmentar si es reemplaçaren les estratègies ordinàries $a_i = a_i^*(c_i)$ seleccionades per l'estratègia normalitzada a_i^* per alguna estratègia ordinària alternativa $a_i^{**} = a_i^{**}(c_i)$. És a dir, per a tot $c_i \in C_i^*$, cal que es compleixi:

$$Z_i[a_i^{**}(c_i), (a^*)^i] > Z_i[a_i^*(c_i), (a^*)^i]. \quad (7)$$

A més, la distribució de probabilitat $R^*(c_i)$ ha d'assignar una massa de probabilitat total no nul·la a aquest conjunt C_i^* . Ara, sigui a_i^{**} una estratègia normalitzada del jugador i , que selecciona les estratègies $a_i^{***}(c_i)$ per a tots els punts c_i en el conjunt C_i^* , però coincidint amb a_i^* a tot arreu més. És a dir,

$$\begin{aligned} a_i^{***}(c_i) &= a_i^{**}(c_i) \quad \text{per a tots } c_i \in C_i^* \\ a_i^{***}(c_i) &= a_i^*(c_i) \quad \text{per a tots } c_i \notin C_i^* \end{aligned} \quad (8)$$

Llavors, hem de tenir

$$W_i[a_i^{***}, (a^*)^i] > W_i[a_i^*, (a^*)^i]. \quad (9)$$

En conseqüència, en el joc G^* , la n -tupla a^* no és un equilibri de Nash bayesià. Així doncs, si a^* no és un punt d'equilibri G , llavors no pot ser un equilibri de Nash bayesià en G^* , en contra de la nostra hipòtesi. Per demostrar la part de necessitat del teorema, suposem que a^* no és un equilibri de Nash bayesià en el joc G^* . Això significa que almenys un component de a^* , diguem a_i^* , pot ser reemplaçat per alguna estratègia normalitzada alternativa a_i^{***} de manera que s'incrementi el valor numèric de la funció W_i . És a dir, hi ha d'haver alguna estratègia normalitzada a_i^{***} que compleixi (8). Però, tenint en compte l'equació (7), implica que, per algun conjunt C_i^* de valors possibles de c_i , hem de tenir

$$Z_i[a_i^{***}(c_i), (a^*)^i] > Z_i[a_i^*(c_i), (a^*)^i]. \quad (10)$$

A més, aquest conjunt C_i^* ha de tenir una massa de probabilitat total no nul·la segons la distribució de probabilitat $R^*(c_i)$. En conseqüència, a^* no pot ser un punt d'equilibri en el joc G . Per tant, si a^* no és un equilibri de Nash bayesià en G^* , llavors tampoc pot ser un punt d'equilibri en el joc G , cosa que contradueix la nostra hipòtesi. $\square$

Exemple 2.48 (Trobarequilibri de Nash bayesià). *Procedirem amb un exemple similar al que es pot trobar a [7]. On modificarem els números concrets i adaptarem els càlculs corresponents. Suposem un joc de suma zero:*

- $N = 2$
- $\Theta = \theta_1 \times \theta_2$, $\theta_1 = \{a_1, a_2\}$ i $\theta_2 = \{b_1, b_2\}$.
- $A = A_1 \times A_2$, $A_1 = \{y_1, y_2\}$ i $A_2 = \{z_1, z_2\}$
- La distribució de probabilitat està regida per la següent taula:

	b_1	b_2
a_1	0.2	0.3
a_2	0.4	0.1

- Els pagaments efectius (diferent a u) segueixen les següents matrius, en funció dels tipus (θ_1, θ_2) de cada jugador
En cas que estiguem a (a^1, b^1) :

	z^1	z^2
y^1	2	3
y^2	-2	8

En cas de (a^1, b^2) :

	z^1	z^2
y^1	-1	-12
y^2	0	3

En cas de (a^2, b^1) :

	z^1	z^2
y^1	7	-1
y^2	10	-3

En cas de (a^2, b^2) :

	z^1	z^2
y^1	4	10
y^2	1	8

Per simplicitat en la construcció de l'exemple he agafat matrius on existeixin punts de sella.

Com tots els espais de tipus de jugadors Θ és finit, i cadascú sap el seu tipus, es poden escriure les matrius de probabilitat condicionada, és a dir el jugador 1, sabent el seu tipus, estima el tipus θ_2 del jugador 2, segons la següent probabilitat:

$$\begin{aligned}\mathbb{P}(\text{Jugador 2} = b_1 | \text{Jugador 1} = a_1) &= 0.4 \\ \mathbb{P}(\text{Jugador 2} = b_2 | \text{Jugador 1} = a_1) &= 0.6 \\ \mathbb{P}(\text{Jugador 2} = b_1 | \text{Jugador 1} = a_2) &= 0.8 \\ \mathbb{P}(\text{Jugador 2} = b_2 | \text{Jugador 1} = a_2) &= 0.2\end{aligned}$$

Corresponentment per al jugador 2 tenim:

$$\begin{aligned}\mathbb{P}(\text{Jugador 1} = a_1 | \text{Jugador 2} = b_1) &= 0.33 \\ \mathbb{P}(\text{Jugador 1} = a_2 | \text{Jugador 2} = b_1) &= 0.67 \\ \mathbb{P}(\text{Jugador 1} = a_1 | \text{Jugador 2} = b_2) &= 0.75 \\ \mathbb{P}(\text{Jugador 1} = a_2 | \text{Jugador 2} = b_2) &= 0.25\end{aligned}$$

Per simplicitat treballarem en estratègies pures, on analitzarem el valor esperat de les estratègies $y^{nt} = (y^n, y^t)$ on y^n ($n = 1, 2$) és l'estratègia pura que seguirà el jugador 1 quan sigui de tipus a_1 i y^t ($t = 1, 2$) serà l'estratègia pura que prengui quan sigui de tipus a_2 , de igual manera definirem, per al jugador 2, $z^{uv} = (z^u, z^v)$.

Amb uns càlculs bastant senzills, obtenim la matriu d'esperances següent:

	z^{11}	z^{12}	z^{21}	z^{22}
y^{11}	3.3	0.6	0.3	-2.4
y^{12}	4.2	1.6	-0.8	-3.4
y^{21}	2.8	4.3	1.6	3.1
y^{22}	3.7	5.3	0.5	2.1

Veiem que a (y^{21}, z^{21}) tenim un punt de sella tal i com l'hem definit a la secció de minímax, per tant tenim un equilibri de Nash bayesià. És a dir, sempre que el jugador 1 sigui de tipus a_1 li convindrà jugar y^2 i sempre que sigui a_2 li convindrà jugar y^1 . Equivalentment amb el jugador 2. Seguint aquesta estratègia, el jugador 1 té un guany esperat de 1.6.

3 Un cas d'estudi, Kuhn pòquer

3.1 Anàlisi teòric

El *Kuhn pòquer* és un joc simplificat de pòquer desenvolupat per Harold Kuhn el 1950 [9]. És un exemple clàssic dins la teoria de jocs de informació incompleta i s'utilitza sovint per il·lustrar conceptes d'estratègies òptimes i equilibri de Nash.

Definició del joc

El *Kuhn pòquer* es juga amb dos jugadors i una baralla reduïda de només tres cartes diferents: per exemple, un 1, un 2 i un 3. Cada jugador rep una única carta de manera aleatòria i secreta.

Fases del Joc

El joc segueix els passos següents:

1. Cada jugador, abans d'obtenir les cartes, posa una fitxa al centre.
2. Cadascun dels dos jugadors rep una carta de manera aleatòria i no poden veure la carta de l'oponent.
3. Hi ha una única ronda d'apostes:
 - El primer jugador pot optar per *passar* o *apostar*.
 - Si el primer jugador passa, el segon jugador pot també passar o apostar.
 - Si hi ha una aposta, l'altre jugador pot igualar-la (i es comparen les cartes) o retirar-se (i perdre la ronda immediatament).
4. El jugador amb la carta més alta guanya el pot si s'arriba a l'abatiment.

El *Kuhn pòquer* és un joc d' **informació incompleta**, ja que els jugadors no coneixen la carta de l'altre. Això implica que han de prendre decisions basades en probabilitats i en el comportament de l'oponent. Un aspecte clau del joc és la possibilitat de fer un catxa, ja que un jugador amb una carta baixa pot intentar enganyar l'adversari apostant agressivament.

Trobar l'equilibri de Nash bayesià

Escriurem les estratègies del jugador 1 com (x_1, x_2, x_3) on cada x_i pot prendre els valors $\{pp, pb, bb\}$, on x_i representa l'acció que es fa si es té la carta i , el p representa l'estratègia passiva i b l'agressiva, és a dir, pp representa, no apostar, i si ens aposta abandonar, pb , no apostar, però igualar l'aposta i bb significa apostar. Ho farem de manera similar amb el jugador 2 amb (y_1, y_2, y_3) . Per al jugador 2, pp significa que, si el jugador 1 no aposta, ell tampoc ho farà i si aposta abandonarà, pb que si el primer no aposta, ell tampoc apostarà, però si aposta pagarà, bp que si el jugador 1 aposta, ell abandonarà, però que si passa,

ell apostarà i bb que si l'altre aposta, ell igualarà i si passa, apostarà. D'aquesta manera ens queda la següent matriu de pagaments:

$x_i \backslash y_j$	pp	pb	bp	bb
pp	± 1	± 1	-1	-1
pb	± 1	± 1	± 2	± 2
bb	1	± 2	1	± 2

on els signes $+$ es prenen si $i > j$ i el $-$ altrament. Abans de continuar, cal mencionar que, encara que la natura d'aquest joc és seqüencial, la manera amb que l'estem modelitzant fa que el puguem entendre com a un joc estratègic, encara que d'informació incompleta, i és per aquest motiu que l'equilibri de Nash sí que aplica i no hem d'intentar trobar el seqüencial. D'aquesta manera veiem que el jugador 1 té 27 estratègies pures i el jugador 2 en té 64, per reduir aquesta quantitat, veurem que hi ha algunes estratègies pures que són estrictament dominants *Definició 2.7*, per fer-ho assumirem que l'altre jugador també és un jugador racional i juga amb l'estratègia òptima:

1. Pagar una aposta prèvia quan es té la carta 1. Si tenim la carta 1, vol dir que el nostre rival té o bé la carta 2 o bé la carta 3. Per tant, com al pagar la seva aposta, sempre perdrem, $-2 = u(\text{pagar}) < u(\text{no pagar}) = -1$.
2. No pagar una aposta prèvia quan es té la carta 3. Similarment a abans, com sempre guanyarem quan s'ensenyin les cartes, l'estratègia té benefici 2 mentre que no pagar té benefici -1 .
3. No apostar amb la carta 3, quan el rival passa. Ja que, si abandona tindrem benefici 1 i si paga tindrem benefici 2, mentre que si passem sempre tindrem benefici 1.
4. El jugador 1 sempre ha de passar amb la carta 2, ja que, com el jugador 2 és racional mai ens pagarà amb una carta pitjor (1) però sempre ens pagarà amb una carta millor (3).

Aquests conceptes de quan aplicar estratègies pures, ens donen conceptes elementals que s'apliquen fins i tot al pòquer convencional, com per exemple no apostar amb una mà de valor mig, apostar per valor o abandonar mans dèbils. Hi ha altres conceptes, com els d'apostar de catxa o induir acció (fer que ens aposti de catxa amb una carta pitjor que la nostra) que s'introdueixen en aquesta simplificació només amb estratègies mixtes.

El concepte d'apostar de catxa es pot veure de la següent manera: quan el jugador 1 té una carta 2 i passa (estratègia òptima, vist abans), si el jugador 2 sempre passés amb un 1 i sempre apostés amb un 3, per al primer jugador seria molt fàcil jugar, ja que abandonaria sempre quan el segon apostés, ja que sabia que, al revelar-se les cartes, sempre perdria. Però si el primer jugador, quan té la carta 2 sempre abandona quan se li aposta, aleshores al segon jugador li convé apostar molt més la carta 1 per fer fora del joc a una carta millor que la seva i guanyar el pot.

Així es veu, que ens convé fer un catxa de tant en tant, per guanyar el pot a cartes millors que la nostra, d'igual manera a vegades al primer jugador amb la millor carta li interessa passar, per a que el segon jugador li faci un catxa i pagar. Per tant, per obtenir l'estratègia òptima en estratègies mixtes d'aquest pòquer simplificat, ens caldrà introduir 5 paràmetres que ens diguin amb quina probabilitat haurem de fer certes accions, en concret:

1. p_1 de fer $x_1 = bb$,
2. p_2 de fer $x_2 = pb$,
3. p_3 de fer $x_3 = bb$,
4. q_1 de fer $y_1 = pb$ i
5. q_2 de fer $y_2 = bb$.

Si tenim en compte les 6 diferents maneres de repartir les cartes entre els 2 jugadors, que aquests són racionals i per tant, en les situacions indicades sempre jugaran les estratègies pures dominants i quan no hi ha estratègia dominant, seguiran la distribució que assignin aquests paràmetres. Ja que sinó, la persona que juga una estratègia que no és dominant, no pot assolir el seu equilibri de Nash i l'altra persona podria explotar la seva manera "subòptima" de jugar. Així doncs podem calcular el pagament de cada jugador. Per exemple si el jugador 1 té la carta 1 i el jugador 2 té la carta 2, el primer jugador perdrà 1 fitxa amb probabilitat p_1q_2 , guanyarà 2 fitxes amb probabilitat $p_1(1 - q_2)$. Calculem aquest valor amb totes les possibilitats i, sigui a^* l'estratègia mixta on s'apliquen totes aquestes probabilitats i també les estratègies pures òptimes mencionades abans, aleshores el valor atès d' a^* és:

$$u_1(a^*) = \frac{1}{6}[p_1(1 - 3q_2) + p_2(3q_1 - 1) + p_3(q_2 - q_1) - q_1]. \quad (11)$$

Observem que quan $q_1 = q_2 = \frac{1}{3}$, aleshores tenim que $1 - 3q_2 = 3q_1 - 1 = q_2 - q_1 = 0$ i l'estratègia (p_1, p_2, p_3) no afecta al pagament del primer jugador, essent per tant, l'equilibri de Nash per al jugador 2.

Per a trobar l'equilibri de Nash per al primer jugador, haurem de reescriure la seva funció de pagament de la següent manera:

$$u_1(a^*) = \frac{1}{6}[q_1(3p_2 - p_3 - 1) + q_2(p_3 - 3p_1) + (p_1 - p_2)]. \quad (12)$$

Aquesta equació té la propietat que, si $3p_2 - p_3 - 1 = p_3 - 3p_1 = 0$ aleshores, l'estratègia del segon jugador (q_1, q_2) no influeix sobre el pagament del primer, per tant serà l'equilibri de Nash del jugador 1. Això es compleix quan $p_1 = \frac{p_3}{3}$ i $p_2 = \frac{p_3+1}{3}$. Per tant encara que l'estratègia del jugador 2 no pot variar (només té 1 equilibri de Nash), l'estratègia òptima del jugador 1 depèn d'un paràmetre. Per tant té infinits equilibris de Nash. Si apliquem aquests equilibris a la funció de pagament del jugador 1, veiem que aquesta té un valor esperat de $-\frac{1}{18}$ i per tant el jugador 2 té un valor atès del joc de $\frac{1}{18}$ quan es juga en equilibri.

3.2 Aproximació computacional, CFR

L'estudi del Kuhn pòquer ofereix una oportunitat ideal per explorar estratègies òptimes en jocs d'informació incompleta. Tot i la simplicitat d'aquest joc, la seva estructura encapsula molts dels elements presents en variants més complexes del pòquer i ens serveix com a pas previ a l'anàlisi general del joc. És en aquest context que introduïm el mètode “Counterfactual Regret Minimization” (CFR) introduït a [14], un algoritme iteratiu que ha demostrat ser extraordinàriament efectiu per aproximar equilibris de Nash en jocs extensius amb informació imperfecta i incompleta. Tot i que els equilibris de Nash poden tenir certes limitacions teòriques en aquest tipus de jocs —com ara la manca de credibilitat de certes estratègies fora de l'equilibri o la seva sensibilitat a desviacions fora del camí principal del joc—, CFR ha esdevingut una eina central en la resolució pràctica de jocs com el pòquer degut a la seva capacitat per generar estratègies difícils d'explotar. A diferència d'enfocaments centrats en la resolució directa de tot el joc, CFR aprofita la descomposició del joc en informacions parcials (conjunts d'informació) i minimitza iterativament el “regret” contrafactual, que definirem a continuació. Això permet, no només una millor eficiència computacional, sinó també la generació de estratègies que, tot i ser aproximades, s'acosten a l'equilibri de Nash amb una robustesa notable.

3.2.1 Esquematització teòrica del CFR

Encara que no és el motiu principal del treball, farem una introducció teòrica a aquest algoritme per la seva importància actual dins el context de jocs d'informació incompleta i sobretot el pòquer, ja que durant els últims anys, el desenvolupament d'aquest algoritme ha donat a lloc a una evolució massiva del joc i de l'anomenat pels jugadors GTO (Game Theoretical Optimization). En aquesta part del treball utilitzarem la referència original on es va introduir per primera vegada el CFR, [14].

Definició 3.1 (ϵ -Equilibri de Nash). Dins un joc extensiu de 2 jugadors. Un **ϵ -equilibri de Nash** és un perfil estratègic s on

$$u_1(s) + \epsilon \geq \max_{s'_1 \in S_1} u_1(s'_1, s_2) \quad u_2(s) + \epsilon \geq \max_{s'_2 \in S_2} u_2(s_1, s'_2). \quad (13)$$

Amb aquesta aproximació de l'equilibri de Nash, haurem d'intentar trobar un mètode que minimitzi aquesta ϵ . Amb CFR ho podem fer d'una manera molt elegant introduint el concepte de “regret”.

Definició 3.2 (“Regret” mitjà). Suposem que es juga repetidament un joc extensiu amb informació incompleta. Sigui s_i^t l'estratègia del jugador i a la ronda t , el “regret” mitjà és:

$$R_i^T = \frac{1}{T} \max_{s_i^* \in S_i} \sum_{t=1}^T (u_i(s_i^*, s_{-i}^t) - u_i(s_i^t)). \quad (14)$$

Sigui $\pi^s(h)$ la probabilitat que ocorri una història h si els jugadors trien accions d'acord amb s . Podem descompondre

$$\pi^s = \prod_{i \in N \cup \{Natura\}} \pi_i^s(h). \quad (15)$$

en la contribució de cada jugador a aquesta probabilitat. Així, $\pi_i^s(h)$ és la probabilitat que si el jugador i juga segons s per a totes les històries h' que són “avantpassades” de h amb $P(h') = i$, que el jugador i prengui l'acció corresponent en h . Definim $\pi_{-i}^s(h)$ com el producte de totes les contribucions dels jugadors (inclosa la Natura, com deia Harsanyi) excepte el jugador i . Per $I_i^j \subseteq \mathcal{I}$, definim $\pi^s(I_i^j) = \sum_{h \in I_i^j} \pi^s(h)$ com la probabilitat d'arribar a un determinat conjunt d'informació donat s , amb $\pi_i^s(I_i^j)$ i $\pi_{-i}^s(I_i^j)$ definits de manera similar. El valor total per al jugador i d'un perfil estratègic és aleshores el pagament esperat del node terminal resultant,

$$u_i(s) = \sum_{h \in Z} u_i(h) \pi^s(h). \quad (16)$$

Definició 3.3 (Estratègia mitjana). Definim $\bar{s}_i^t$ com l'estratègia mitjana del jugador i des del temps 1 fins al T . Per a cada conjunt d'informació $I_i^j \in \mathcal{I}_i$, i per a cada acció $a \in A(I_i^j)$, definim:

$$\bar{s}_i^t(I_i^j)(a) = \frac{\sum_{t=1}^T \pi_i^{s^t}(I_i^j) s^t(I_i^j)(a)}{\sum_{t=1}^T \pi_i^{s^t}(I_i^j)}. \quad (17)$$

Amb aquestes definicions, tenim prou per veure que minimitzant aquest “regret”, arribarem a una solució que aproxima molt bé l'equilibri de Nash.

Teorema 3.4. *En un joc de suma zero de dos jugadors, si després de T iteracions el “regret” mitjà de cada jugador és com a màxim ϵ , llavors $\bar{s}^T$ és un 2ϵ -equilibri de Nash.*

Aquest és un teorema molt conegut dins d'aquest àmbit (per exemple, a [13, Theorem 1]). Amb aquests conceptes, Zinkevich, va ser capaç, a [14], de trobar una cota superior del “regret”, que s'anava a 0 quan $T \rightarrow \infty$. La qual cosa succeeix si es segueix el següent algoritme:

Definim $R_i^{T,+}(I_i^j, a) = \max(R_i^T(I_i^j, a), 0)$, llavors l'estratègia per al temps $T + 1$ és:

$$s_i^{T+1}(I_i^j)(a) = \begin{cases} \frac{R_i^{T,+}(I_i^j, a)}{\sum_{a \in A(I_i^j)} R_i^{T,+}(I_i^j, a)} & \text{si } \sum_{a \in A(I_i^j)} R_i^{T,+}(I_i^j, a) > 0 \\ \frac{1}{|A(I_i^j)|} & \text{altrament.} \end{cases} \quad (18)$$

3.2.2 Kuhn Poker Modificat

Amb la base teòrica sobre teoria de jocs i CFR, voldríem trobar una solució computacional a un problema que no té solució analítica per complexitat de

l'arbre decisiu. La variant introduïda afegeix una nova acció legal a certs nodes del joc: una segona pujada quan un jugador ja ha pujat una vegada. Això incrementa el nombre de seqüències i nodes terminals. Per poder resoldre'l volem implementar el CFR dins aquest joc modificat. Gràcies a [1] i sobre tot a les instruccions i codi base que proporciona Jonathan Gardner a [4] vam poder fer una modificació del codi. En [4] es plantegen els diferents conjunts d'informació de la següent manera: " $Ka_1...a_n$ " on K és la carta que s'assigna al principi del joc $K \in \{1, 2, 3\}$ i a_i és l'acció que s'ha pres al torn i , $a_i \in \{p, b, r\}$. On p representa passar o abandonar, b representa apostar quan no hi ha una aposta anterior o pagar una aposta d'aquest tipus i r representa la nova acció, tornar a pujar un cop ja s'ha apostat o pagar una pujada. Per tant el conjunt d'informació $1pb$ vol dir que tinc la carta 1 i al primer torn he passat i ell al seu torn m'ha apostat o per exemple, $3p$ vol dir que sóc el segon jugador, tinc la carta 3 i el jugador 1 ha passat al seu torn. Aquesta interpretació té l'avantatge que és molt fàcil accedir al conjunt d'informació previ del nostre rival, ja que la carta 1 es descarta (la tenim nosaltres) i per tant ell només pot estar a $2p$ o a $3p$. D'aquesta manera, amb les estratègies, és molt fàcil trobar les probabilitats condicionades de que ell estigui a un o altre conjunt d'informació. Amb aquest model, quan aplicàrem l'algoritme CFR i vam obtenir els següents resultats (no són únics):

Table 1: Taula per al jugador 1

InfoSet	Strat:B	Strat:P	Strat:R
3	0.65	0.35	–
2	0	1	–
1	0.22	0.78	–
3br	–	0.02	0.98
2br	–	0.98	0.02
1br	–	0.99	0.01
3pb	0.01	0	0.99
2pb	0.55	0.45	0
1pb	0	1	0

Table 2: Taula per al jugador 2

InfoSet	Strat:B	Strat:P	Strat:R
3p	0.98	0.02	–
2p	0	1	–
1p	0.33	0.67	–
3b	0.03	0	0.97
2b	0.33	0.67	0
1b	0	1	0
3pbr	–	0.01	0.98
2pbr	–	0.98	0.02
1pbr	–	0.99	0.01

Aquesta extensió mostra que el mètode CFR és prou flexible per adaptar-se a modificacions del joc. Permetre una pujada addicional enriqueix l'anàlisi estratègica i mostra resultats consistents amb la teoria de jocs de suma zero. Com per exemple, no pagar mai amb la carta 1 o fer la nova acció r amb la carta 3 sempre que pot. Aquestes són estratègies dominants. Tanmateix, sorprèn que en aquesta variant del joc les estratègies obtingudes no arriben al valor de *catxa* amb r . Encara que inicialitzant els valors del programa en diferents valors, hi ha una lleugera diferència en els resultats que es proporcionen, mai apareix un valor de catxa amb r . Això es pot explicar per diversos motius:

- Pot ser que en un joc tan limitat, l'incentiu per fer un catxa de re-pujada no sigui prou.
- Un catxa amb re-pujada amb la carta 1 pot formar part del joc “explotatiu”, per a arribar a un equilibri on hi hagi catxa a la re-pujada, el jugador 2 hauria de pagar aquesta seqüència un percentatge de les vegades i aquest fet, segurament augmenta el “regret” perquè quan perd contra la carta 3, perd més.

Això obre la porta a una investigació més profunda sobre com els equilibris canvien en funció de les accions disponibles i la profunditat de l'arbre de decisions.

4 Conclusions

Aquest treball ha tingut com a objectiu principal l'estudi profund de la teoria de jocs amb especial èmfasi en el paper fonamental que hi exerceix la informació —tant la seva disponibilitat com la seva estructura— en la determinació dels equilibris estratègics. Al llarg del treball, hem dut a terme una anàlisi rigorosa dels diferents tipus de jocs: des dels jocs en forma normal fins als jocs extensius amb informació perfecta, imperfecta i finalment incompleta, culminant amb l'aplicació d'aquest marc conceptual a un cas d'estudi concret mitjançant una aproximació computacional.

En primer lloc, hem presentat de manera formal la noció d'equilibri de Nash, tant en estratègies pures com en mixtes, destacant-ne l'existència sota determinades condicions topològiques. Aquesta base teòrica ens ha permès entendre com els jugadors, actuant de manera racional i tenint en compte les decisions dels altres, poden arribar a situacions estables on cap d'ells té incentius per a desviar-se unilateralment de la seva estratègia. A continuació, hem analitzat els jocs extensius, primerament amb informació perfecta, on l'equilibri perfecte en *subjocs* ofereix un refinament essencial respecte a l'equilibri de Nash tradicional. Aquest anàlisi s'ha ampliat amb els jocs extensius d'informació imperfecta, on els conceptes d'informació incompleta i de creences cobren una rellevància fonamental. En aquests escenaris, el concepte d'equilibri seqüencial, introduït per Kreps i Wilson, s'erigeix com l'eina adequada per capturar l'estratègia òptima dels jugadors.

El nucli central del treball ha estat la teoria de jocs amb informació incompleta, modelitzada mitjançant l'enfocament bayesià proposat per John Harsanyi. Aquest model transforma jocs amb asimetries d'informació en jocs bayesians amb tipus privats, permetent un anàlisi coherent dins del marc de la teoria de jocs estàndard. Hem pogut demostrar que tota la incertesa pot ser encapsulada en els tipus dels jugadors i que, sota aquest enfocament, es poden definir equilibris bayesians que generalitzen la noció d'equilibri de Nash.

Un dels principals valors afegits d'aquest treball ha estat l'aplicació pràctica d'aquests conceptes mitjançant l'estudi del joc de Kuhn pòquer. L'estudi d'aquest s'ha vist enriquit amb una aproximació computacional mitjançant l'algoritme de minimització de “regret” contrafactual (CFR), que ha permès simular l'evolució estratègica cap a un equilibri en contextos de repetició i aprenentatge iteratiu.

En resum, aquest treball ha posat de manifest la importància de la informació i de les creences en la teoria de jocs, destacant com la seva presència (o absència) condiona la naturalesa dels equilibris i el comportament racional dels agents. A més, ha demostrat que l'articulació entre el rigor formal i la implementació computacional permet no només analitzar millor els jocs teòrics, sinó també traslladar aquests coneixements a entorns pràctics i aplicats.

A Appendix A: Codi CFR

Aquesta és la modificació del codi proporcionat per [4]. Cal mencionar que són molt similars ja que el gruix de l'algoritme estava proporcionat al codi i només vaig haver de modificar algunes definicions per afegir la pujada.

```
1 from __future__ import annotations
2 from tabulate import tabulate
3 import sys
4
5 infoSets: dict[str, InfoSetData] = {}
6 sortedInfoSets = []
7
8 RANKS = ["K", "Q", "J"]
9 ACTIONS = ["b", "p", "r"]
10 TERMINAL_ACTION_STR_MAP = {
11     "pp",
12     "bb",
13     "bp",
14     "pbb",
15     "pbp",
16     "brr",
17     "brp",
18     "pbrp",
19     "pbr",
20 }
21 INFO_SET_ACTION_STRS = {
22     "",
23     "p",
24     "b",
25     "pb",
26     "br",
27     "pbr",
28 }
29 def isIllegalAction(history: str, action: str) -> bool:
30     new_hist = history + action
31     if history == "" and action == "r":
32         return True
33     if 'pr' in new_hist:
34         return True
35     if 'rb' in new_hist:
36         return True
37     return False
38
39 def getDecidingPlayerForInfoSetStr(infoSetStr: str):
40     return (len(infoSetStr)-1)%2
41
42 class InfoSetData:
43     def __init__(self):
44         self.actions: dict[str, InfoSetActionData] = {
45             "b": InfoSetActionData(initStratVal=1 / len(ACTIONS)),
46             "p": InfoSetActionData(initStratVal=1 / len(ACTIONS)),
47             "r": InfoSetActionData(initStratVal=1 / len(ACTIONS)),
48         }
49         self.beliefs: dict[str, float] = {}
50         self.expectedUtil: float = None
51         self.likelihood: float = None
```

```

52
53 @staticmethod
54 def printInfoSetDataTable(infoSets: dict[str, InfoSetData]):
55     from tabulate import tabulate
56
57     rows = []
58
59     for infoSetStr in sortedInfoSets:
60         infoSet = infoSets[infoSetStr]
61
62         strategy = [f'{infoSet.actions[a].strategy:.2f}' for a
63 in ACTIONS]
64         utility = [f'{infoSet.actions[a].util:.2f}' if infoSet.
65 actions[a].util is not None else " " for a in ACTIONS]
66
67         opponent_pockets = getPossibleOpponentPockets(
68 infoSetStr[0])
69         beliefs = [f'{infoSet.beliefs.get(p, 0.0):.2f}' for p
70 in opponent_pockets]
71
72         row = [
73             infoSetStr,
74             *strategy,
75             *utility,
76             f'{infoSet.expectedUtil:.2f}' if infoSet.
77 expectedUtil is not None else " ",
78             f'{infoSet.likelihood:.4f}' if infoSet.likelihood
79 is not None else " ",
80             *beliefs
81         ]
82         rows.append(row)
83
84     headers = [
85         "InfoSet",
86         "Strat:B", "Strat:P", "Strat:R",
87         "Util:B", "Util:P", "Util:R",
88         "ExpUtil", "Likelihood",
89         "Belief:1", "Belief:2"
90     ]
91
92     print(tabulate(rows, headers=headers, tablefmt="fancy_grid"
93 ))
94
95 def getStrategyTableData(self):
96     return [f'{self.actions[action].strategy:.2f}' for action
97 in ACTIONS]
98
99 def getUtilityTableData(self):
100     return [f'{self.actions[action].util:.2f}' for action in
101 ACTIONS]
102
103 def getGainTableData(self):
104     return [f'{self.actions[action].cumulativeGain:.2f}' for
105 action in ACTIONS]
106
107 def getBeliefTableData(self):
108     return [f'{self.beliefs[oppPocket]:.2f}' for oppPocket in
109 self.beliefs.keys()]

```

```

98
99 class InfoSetActionData:
100     def __init__(self, initStratVal):
101         self.strategy = initStratVal
102         self.util = None
103         self.cumulativeGain = initStratVal
104 def getPossibleOpponentPockets(pocket):
105     return [rank for rank in RANKS if rank != pocket]
106
107 def getAncestralInfoSetStrs(infoSetStr) -> list[InfoSetData]:
108     if len(infoSetStr) == 1:
109         raise ValueError(f'no ancestors of infoSet={infoSetStr}')
110     possibleOpponentPockets = getPossibleOpponentPockets(infoSetStr
111 [0])
112     return [oppPocket + infoSetStr[1:-1] for oppPocket in
113 possibleOpponentPockets]
114
115 def getDescendantInfoSetStrs(infoSetStr, action):
116     history = infoSetStr[1:]
117
118     if isIllegalAction(history, action):
119         return []
120
121     oppPockets = getPossibleOpponentPockets(infoSetStr[0])
122     actionStr = history + action
123     return [oppPocket + actionStr for oppPocket in oppPockets]
124
125 def playerOnePocketIsHigher(pocket1,pocket2):
126     if pocket1=='K':
127         return True
128     if pocket1=='J':
129         return False
130     if pocket2=='K':
131         return False
132     if pocket2=='J':
133         return True
134     raise ValueError('this should not occur bc one player must have a
135 K or J')
136
137 def calcUtilityAtTerminalNode(pocket1,pocket2,actionStr: str):
138     if actionStr=='pp':
139         if playerOnePocketIsHigher(pocket1,pocket2):
140             return 1,-1
141         else:
142             return -1,1
143     elif actionStr=='pbp':
144         return -1,1
145     elif actionStr=='bp':
146         return 1,-1
147     elif actionStr=='bb' or actionStr=='pbb':
148         if playerOnePocketIsHigher(pocket1,pocket2):
149             return 2,-2
150         else:
151             return -2,2
152     elif actionStr=='brp':
153         return -2,2
154     elif actionStr=='pbrp':

```

```

152     return 2,-2
153 elif actionStr== 'brr' or actionStr== 'pbr':
154     if playerOnePocketIsHigher(pocket1, pocket2):
155         return 3,-3
156     else:
157         return -3,3
158 else:
159     raise ValueError(f'unexpected actionStr={actionStr}')
160
161
162 def initInfoSets():
163     for actionsStrs in sorted(INFO_SET_ACTION_STRS, key=lambda x:
164         len(x)):
165         for rank in RANKS:
166             infoSetStr = rank + actionsStrs
167             infoSets[infoSetStr] = InfoSetData()
168             sortedInfoSets.append(infoSetStr)
169
170 def updateBeliefs():
171     for infoSetStr in sortedInfoSets:
172         infoSet = infoSets[infoSetStr]
173         if len(infoSetStr) == 1:
174             possibleOpponentPockets = getPossibleOpponentPockets(
175                 infoSetStr[0])
176             for oppPocket in possibleOpponentPockets:
177                 infoSet.beliefs[oppPocket] = 1 / len(
178                     possibleOpponentPockets)
179         else:
180             ancestralInfoSetStrs = getAncestralInfoSetStrs(
181                 infoSetStr)
182             lastAction = infoSetStr[-1]
183             tot = 0
184             for oppInfoSetStr in ancestralInfoSetStrs:
185                 oppInfoSet=infoSets[oppInfoSetStr]
186                 tot += oppInfoSet.actions[lastAction].strategy
187             for oppInfoSetStr in ancestralInfoSetStrs:
188                 oppInfoSet=infoSets[oppInfoSetStr]
189                 oppPocket = oppInfoSetStr[0]
190                 infoSet.beliefs[oppPocket]=oppInfoSet.actions[
191                     lastAction].strategy / tot
192     return
193
194 def updateUtilitiesForInfoSetStr(infoSetStr):
195     playerId = getDecidingPlayerForInfoSetStr(infoSetStr)
196     infoSet = infoSets[infoSetStr]
197     beliefs = infoSet.beliefs
198     for action in ACTIONS:
199         history=infoSetStr[1:]
200         actionStr=history+action
201         if isIllegalAction(history, action):
202             infoSet.actions[action].util = -100
203             continue
204         descendentInfoSetStrs = getDescendantInfoSetStrs(infoSetStr
205             ,action)
206         utilFromInfoSets,utilFromTerminalNodes=0,0
207         for descendentInfoSetStr in descendentInfoSetStrs:
208             probOfThisInfoSet = beliefs[descendentInfoSetStr[0]]

```

```

203
204         pockets=[infoSetStr[0],descendentInfoSetStr[0]]
205         if playerId==1:
206             pockets=list(reversed(pockets))
207         if actionStr in TERMINAL_ACTION_STR_MAP:
208             utils=calcUtilityAtTerminalNode(*pockets,actionStr)
209             utilFromTerminalNodes+=probOfThisInfoSet*utils[
playerIdx]
210         else:
211             descendentInfoSet = infoSets[descendentInfoSetStr]
212             for oppAction in ACTIONS:
213                 oppHistory = actionStr
214                 destinationInfoSetStr = infoSetStr+action+
oppAction
215                 destinationActionStr = destinationInfoSetStr
[1:]
216                 if isIllegalAction(oppHistory, oppAction):
217                     infoSet.actions[action].util = -100
218                     continue
219                 probOfOppAction = descendentInfoSet.actions[
oppAction].strategy
220                 if destinationActionStr in
TERMINAL_ACTION_STR_MAP:
221                     utils=calcUtilityAtTerminalNode(*pockets,
destinationActionStr)
222                     utilFromTerminalNodes+=probOfThisInfoSet*
probOfOppAction*utils[playerIdx]
223                 else:
224                     destinationInfoSet = infoSets[
destinationInfoSetStr]
225                     utilFromInfoSets+=probOfThisInfoSet*
probOfOppAction*destinationInfoSet.expectedUtil
226
227                     infoSet.actions[action].util=utilFromInfoSets+
utilFromTerminalNodes
228                     infoSet.expectedUtil = 0
229                     for action in ACTIONS:
230                         actionData = infoSet.actions[action]
231                         infoSet.expectedUtil+=actionData.strategy*actionData.util
232
233 def calcInfoSetLikelihoods():
234     for infoSetStr in sortedInfoSets:
235         infoSet=infoSets[infoSetStr]
236         infoSet.likelihood=0
237         possibleOppPockets=getPossibleOpponentPockets(infoSetStr[0])
238         if len(infoSetStr)==1:
239             infoSet.likelihood=1/len(RANKS)
240         elif len(infoSetStr)==2:
241             for oppPocket in possibleOppPockets:
242                 oppInfoSet = infoSets[oppPocket + infoSetStr[1:-1]]
243                 infoSet.likelihood+=oppInfoSet.actions[infoSetStr[-1]].
strategy/(len(RANKS)*len(possibleOppPockets))
244         else:
245             for oppPocket in possibleOppPockets:
246                 oppInfoSet = infoSets[oppPocket + infoSetStr[1:-1]]
247                 infoSetTwoLevelsAgo = infoSets[infoSetStr[:-2]] # grab the
closest ancestral infoSet with the same owner as the infoSet

```

```

248     for which we seek to calculate likelihood
        ancestorLikelihood = infoSetTwoLevelsAgo.likelihood/len(
            possibleOppPockets) # note, each oppInfoSet is essentially
            slicing up the infoSetTwoLevelsAgo because they're each
            assuming a specific oppPocket. Therefore, we must divide by len
            (possibleOppPockets)
249         infoSet.likelihood+=ancestorLikelihood*oppInfoSet.actions[
            infoSetStr[-1]].strategy
250
251 def calcGains():
252     totAddedGain=0.0
253     for infoSetStr in sortedInfoSets:
254         infoSet = infoSets[infoSetStr]
255         for action in ACTIONS:
256             utilForActionPureStrat = infoSet.actions[action].util
257             gain = max(0,utilForActionPureStrat-infoSet.expectedUtil)
258             totAddedGain+=gain
259             infoSet.actions[action].cumulativeGain+=gain * infoSet.
                likelihood
260     return totAddedGain
261 def updateStrategy():
262     for infoSetStr in sortedInfoSets:
263         infoSet = infoSets[infoSetStr]
264         gains = [infoSet.actions[action].cumulativeGain for action in
            ACTIONS]
265         totGains = sum(gains)
266         for action in ACTIONS:
267             gain = infoSet.actions[action].cumulativeGain
268             infoSet.actions[action].strategy = gain/totGains
269
270 def setInitialStrategiesToSpecificValues():
271     infoSets['K'].actions['b'].strategy=1/3
272     infoSets['K'].actions['p'].strategy=1/3
273     infoSets['K'].actions['r'].strategy=1/3
274
275     infoSets['Q'].actions['b'].strategy=1/3
276     infoSets['Q'].actions['p'].strategy=1/3
277     infoSets['Q'].actions['r'].strategy=1/3
278
279     infoSets['J'].actions['b'].strategy=1/3
280     infoSets['J'].actions['p'].strategy=1/3
281     infoSets['J'].actions['r'].strategy=1/3
282
283     infoSets['Kpb'].actions['b'].strategy=1/3
284     infoSets['Kpb'].actions['p'].strategy=1/3
285     infoSets['Kpb'].actions['r'].strategy=1/3
286
287     infoSets['Qpb'].actions['b'].strategy=1/3
288     infoSets['Qpb'].actions['p'].strategy=1/3
289     infoSets['Qpb'].actions['r'].strategy=1/3
290
291     infoSets['Jpb'].actions['b'].strategy=1/3
292     infoSets['Jpb'].actions['p'].strategy=1/3
293     infoSets['Jpb'].actions['r'].strategy=1/3
294
295     infoSets['Kbr'].actions['b'].strategy=1/3
296     infoSets['Kbr'].actions['p'].strategy=1/3

```

```

297 infoSets['Kbr'].actions['r'].strategy=1/3
298
299 infoSets['Qbr'].actions['b'].strategy=1/3
300 infoSets['Qbr'].actions['p'].strategy=1/3
301 infoSets['Qbr'].actions['r'].strategy=1/3
302
303 infoSets['Jbr'].actions['b'].strategy=1/3
304 infoSets['Jbr'].actions['p'].strategy=1/3
305 infoSets['Jbr'].actions['r'].strategy=1/3
306
307 # player2
308 infoSets['Kb'].actions['b'].strategy=1/3
309 infoSets['Kb'].actions['p'].strategy=1/3
310 infoSets['Kb'].actions['r'].strategy=1/3
311 infoSets['Kp'].actions['b'].strategy=1/3
312 infoSets['Kp'].actions['p'].strategy=1/3
313 infoSets['Kp'].actions['r'].strategy=1/3
314
315 infoSets['Qb'].actions['b'].strategy=1/3
316 infoSets['Qb'].actions['p'].strategy=1/3
317 infoSets['Qb'].actions['r'].strategy=1/3
318 infoSets['Qp'].actions['b'].strategy=1/3
319 infoSets['Qp'].actions['p'].strategy=1/3
320 infoSets['Qp'].actions['r'].strategy=1/3
321
322 infoSets['Jb'].actions['b'].strategy=1/3
323 infoSets['Jb'].actions['p'].strategy=1/3
324 infoSets['Jb'].actions['r'].strategy=1/3
325 infoSets['Jp'].actions['b'].strategy=1/3
326 infoSets['Jp'].actions['p'].strategy=1/3
327 infoSets['Jp'].actions['r'].strategy=1/3
328
329 if __name__ == "__main__":
330     initInfoSets()
331
332     numIterations=300000
333     totGains = []
334
335     numGainsToPlot=100
336     gainGrpSize = numIterations//numGainsToPlot
337     if gainGrpSize==0:
338         gainGrpSize=1
339
340     for i in range(numIterations):
341         updateBeliefs()
342
343         for infoSetStr in reversed(sortedInfoSets):
344             updateUtilitiesForInfoSetStr(infoSetStr)
345
346         calcInfoSetLikelihoods()
347         totGain = calcGains()
348         if i%gainGrpSize==0:
349             totGains.append(totGain)
350             print(f'TOT_GAIN {totGain: .3f}')
351         updateStrategy()
352
353 InfoSetData.printInfoSetDataTable(infoSets)

```

References

- [1] Professor Bryce. Counterfactual regret minimization (agt 26). https://www.youtube.com/watch?v=ygDt_AumPr0, 2023.
- [2] John L. Casti. *Five Golden Rules: Great Theories of 20th-century Mathematics and Why They Matter*. Wiley & Sons, 1996.
- [3] Shyam S. Chakrabarti and Ievgen Topolyan. An extensive form-based proof of the existence of sequential equilibrium. *Economics Theory Bulletin*, 4: 355–365, 2016. doi: 10.1007/s40505-016-0098-8.
- [4] Jonathan Gardner. Implementing cfr for kuhn poker (in python). <https://www.youtube.com/watch?v=NE7V8e77vg4>, 2024.
- [5] John C. Harsanyi. Games with incomplete information played by “bayesian” players, part i: The basic model. *Management Science*, 14(3):159–182, 1967. doi: 10.1287/mnsc.14.3.159.
- [6] John C. Harsanyi. Games with incomplete information played by “bayesian” players, part ii: Bayesian equilibrium points. *Management Science*, 14(5):320–334, 1968. doi: 10.1287/mnsc.14.5.320.
- [7] John C. Harsanyi. Games with incomplete information played by “bayesian” players, part iii: The basic probability distribution of the game. *Management Science*, 14(7):486–502, 1968. doi: 10.1287/mnsc.14.7.486.
- [8] David M. Kreps and Robert Wilson. Sequential equilibria. *Econometrica*, 50(4):863–894, 1982. doi: 10.2307/1912767.
- [9] H. W. Kuhn. *Simplified Two-Person Poker*, volume 1. Princeton University Press, 1950.
- [10] Michael Maschler, Eilon Solan, and Shmuel Zamir. *Game Theory*. Cambridge University Press, 2020. ISBN 9781108493451.
- [11] Martin J. Osborne and Ariel Rubinstein. *A Course in Game Theory*. MIT Press, 1994.
- [12] Jonathan Weinstein. Two elementary proofs of the minimax theorem. *Working Paper*, 2022. URL <https://bpb-us-w2.wpmucdn.com/sites.wustl.edu/dist/d/2034/files/2022/05/minimax.pdf>.
- [13] Xinyu Zhang. Learn ai game playing algorithm (part iii): Counterfactual regret minimization. <https://xyzml.medium.com/learn-ai-game-playing-algorithm-part-iii-counterfactual-regret-minimization-b182a7ec85f1>, 2020.
- [14] Martin Zinkevich, Michael Johanson, Michael Bowling, and Carmelo Piccione. Regret minimization in games with incomplete information. <http://cs.ualberta.ca>, 2007. University of Alberta, Computing Science Department.