

Universitatea “Babeș-Bolyai”

Facultatea de Științe Economice și Gestiunea Afacerilor

**Școala doctorală Științe Economice și Gestiunea
Afacerilor**

REZUMAT TEZA DE DOCTORAT

Doctorand:

Gabriel Marius Pop

Conducător:

Prof. univ. dr. habil. Dorina Lazăr

Cluj-Napoca

2025



UNIVERSITATEA BABEȘ-BOLYAI
BABES-BOLYAI TUDOMÁNYEGYETEM
BABES-BOLYAI UNIVERSITÄT
BABES-BOLYAI UNIVERSITY
TRADITIO ET EXCELLENTIA



Facultatea de Științe Economice și Gestiunea Afacerilor

ȘCOALA DOCTORALĂ DE ȘTIINȚE
ECONOMICE ȘI GESTIUNEA AFACERILOR

CONTRIBUȚII PRIVIND ANALIZA MONTE CARLO ÎN JOCURI REPETATE: ITERAȚIILE DILEMEI PRIZIONERULUI DE TIP AXELROD ȘI AUTOCOVARIANTĂ ÎN MIXTURA ALLISON

– REZUMAT –

Doctorand:

Gabriel Marius Pop

Conducător:

Prof. univ. dr. habil. Dorina Lazăr

Cluj-Napoca

2025

Cuprinsul Tezei de Doctorat

1	Introducere	1
2	Fundamente teoretice, formale și metodologice ale luării deciziilor strategice în jocuri repetate și Dilema Prizonierului iterată	6
2.1	Concepte teoretice și notații	6
2.1.1	Jocuri în formă normală	6
2.1.2	Strategii dominante și dominate	7
2.1.3	Strategii mixte și utilități așteptate	8
2.1.4	Implicații strategice	10
2.2	Jocuri strategice și Dilema Prizonierului iterată	11
2.2.1	Jocuri strategice – Concepte	11
2.2.2	Modele statice și limitele alegerii raționale	13
2.2.3	Structura clasică a Dilemei Prizonierului și implicațiile strategice	13
2.2.4	Dilema Prizonierului iterată: fundamente strategice ale cooperării	15
2.3	Jocuri repetate: orizont finit vs. orizont infinit	19
2.3.1	Jocuri repetate cu orizont finit: o limitare teoretică	19
2.3.2	Jocuri repetate cu orizont infinit și posibilități de cooperare	21
2.3.3	Reputații și stimulente pe termen lung în medii incerte	23
2.3.4	Jocuri repetate finit vs. infinit cu reputație	25
2.3.5	Metrici de evaluare a câștigurilor pentru orizonturi finite și infinite	27
2.4	Strategii specifice din turneele lui Axelrod	29
2.4.1	Structura și obiectivele turneele lui Axelrod	29
2.4.2	Strategii specifice din turneele lui Axelrod	31
2.4.3	Strategii din al doilea turneu Axelrod	33
2.5	Metodologia simulării Monte Carlo	36
2.5.1	Fundamentare teoretică	36
2.5.2	Cadru de simulare în IPD	37
2.5.3	Prezentare generală a algoritmului Monte Carlo	39
2.5.4	Scheme de simulare	40
2.5.5	Modelare formală a strategiilor în IPD	40

2.5.6 Simularea Monte Carlo a duratelor incerte în jocuri cu interacțiuni repetate	43
2.6 Indicatori de performanță utilizați în analiza strategiilor	47
2.6.1 Indicatori descriptivi: rezultate câștig–pierdere–remiză și vectori ai câștigurilor	48
2.6.2 Indicatori de distanță între scenarii deterministe și incerte	49
2.6.3 Comparații de performanță pe bază de matrice	51
2.6.4 Interpretări și implicații	54
3 Analiza experimentală a comportamentului strategic în condiții deterministe vs. incerte	56
3.1 Considerații teoretice asupra mediilor deterministe și incerte	56
3.1.1 Jocuri cu orizont finit (medii deterministe)	56
3.1.2 Jocuri cu orizont infinit și durată incertă	57
3.1.3 Terminare probabilistică și jocuri repetate cu discount	59
3.2 Jocuri dinamice cu orizont infinit și Dilema Prizonierului iterată	61
3.3 Configurare experimentală și rezultate din primul turneu Axelrod	63
3.3.1 Descrierea bibliotecii Axelrod	65
3.3.2 Funcții de transformare a strategiilor	69
3.3.3 Utilitare și API-uri interne	71
3.4 Configurare experimentală	72
3.4.1 Implementarea turneului personalizat	72
3.4.2 Avantajele configurației personalizate	73
3.4.3 Rezultate și analiză	77
3.5 Analiză grafică a sensibilității strategiilor în primul turneu Axelrod	79
3.6 Configurare experimentală și rezultate din al doilea turneu Axelrod	87
3.6.1 Prezentare generală și motivație	87
3.6.2 Proiectarea simulării Monte Carlo	89
3.7 Rezultate și discuții	90
3.7.1 Indicatori principali și observații	92
3.7.2 Explicarea unor exemple relevante	95
3.7.3 Concluzii	96
3.8 Analiză detaliată a comportamentului strategic al agenților în condiții determinate și incerte	97
3.9 Rolul adaptabilității și flexibilității strategice în gestionarea incertitudinii	100
3.10 Prezentare generală	101

4 Autocovarianța în modelul Allison Mixture – O analiză Monte Carlo a luării deciziilor strategice în jocuri repetate	104
4.1 Introducere în Allison Mixture și paradoxul lui Parrondo	104
4.1.1 Fundamente conceptuale ale paradoxului Parrondo	105
4.1.2 Alternanță aperiodică și indicatori de structură în jocurile Parrondo ..	107
4.1.3 Extensii cuantice: paradoxul Parrondo în Quantum Walk-uri discrete	108
4.1.4 Aplicații și implicații ale paradoxului Parrondo	109
4.2 Modelul Allison Mixture	110
4.3 Aplicarea modelului Allison Mixture în jocul Entry-Deterrence	112
4.3.1 Simulare în mai multe runde a jocului de tip Entry-Deterrence	114
4.3.2 Interpretare grafică a simulării Allison Mixture	115
4.3.3 Analiză informațională a secvenței mixte	116
4.4 Metodologia experimentală pentru simulările Monte Carlo	117
4.5 Metodologie experimentală pentru simulările Monte Carlo aplicate Allison Mixture	119
4.5.1 Motivație conceptuală pentru cadrul Monte Carlo	119
4.5.2 Proiectarea experimentului Monte Carlo	120
4.6 Rezultatele experimentale obținute prin simulări Monte Carlo	126
4.6.1 Introducere în studiul simulării Monte Carlo	126
4.6.2 Analiza rezultatelor de autocovarianță și autocorelație	128
4.6.3 Discuția rezultatelor	147
4.7 Concluzii privind analiza autocovarianței în Allison Mixture prin simulări Monte Carlo pentru luarea deciziilor strategice	153
5 Concluzii generale și direcții viitoare de cercetare	155
A Anexa A: Cod sursă pentru simulările Monte Carlo din turneul Axelrod	161
B Anexa B: Cod sursă pentru experimentul de autocovarianță Allison Mixture .	168
Bibliografie	176

Cuprinsul rezumatului

Cuprinsul Tezei de Doctorat	ii
1 Introducere	1
2 Rezumat al Capitolului 2: Fundamentele teoretice, formale și metodologice ale luării deciziilor strategice în jocuri repetate și în Dilema Prizonierului iterată	5
3 Rezumatul Capitolului 3: Analiza experimentală a comportamentului strategic în condiții determinate versus incerte	14
4 Rezumatul Capitolului 4: Autocovarianța în modelul Allison Mixture – O analiză Monte Carlo a deciziei strategice în jocuri repetate	25
5 Concluzii generale și direcții viitoare de cercetare	30

1

Introducere

Luarea deciziilor strategice este esențială pentru înțelegerea interacțiunilor complexe între mai mulți agenți, iar teoria jocurilor este un instrument puternic pentru analiza structurii complexe a acestor comportamente. Dintre diversele abordări analitice, jocurile repetate se remarcă prin capacitatea lor de a explica evoluția cooperării, a conflictului și a adaptării, oferind perspective mai profunde decât modelele statice tradiționale, cu o singură rundă.

O direcție majoră a acestei teze este dedicată analizei Dilemei Prizonierului Iterate (IPD), un cadru fundamental pentru studiul interacțiunilor strategice repetate, care a fost explorat pe larg prin turneele computaționale organizate de Robert Axelrod. Contribuțiile lui Axelrod au evidențiat că deciziile strategice ale agenților nu sunt determinate exclusiv de câștigurile imediate, ci sunt influențate în mod semnificativ de istoricul interacțiunilor, de anticipările privind viitorul relației strategice și de procesele de învățare adaptivă pe termen lung.

În acest context, se conturează o întrebare teoretică fundamentală: *Pot fi considerați agenții dezvoltați în cadrul turneelor lui Axelrod capabili să distingă între implicațiile unui orizont finit și cele ale unui orizont infinit?* Deși un orizont infinit nu poate fi testat direct în practică, incertitudinea privind numărul de etape ale unui superjoc poate funcționa ca un substitut funcțional, permițând explorarea comportamentului strategic în condiții de orizont nedefinit. Acest lucru ne permite să observăm dacă generațiile succesive de agenți din turneele lui Axelrod se comportă pur mecanic din perspectiva duratei jocului sau dacă își adaptează comportamentul în funcție de certitudinea sau incertitudinea legată de numărul de etape ale superjocurilor în care sunt implicați. Această teză își propune să abordeze aceste întrebări nuanțate, contribuind la aprofundarea înțelegerii procesului decizional strategic în contexte intertemporale și oferind perspective relevante pentru teoria economică, modelarea politică, științele comportamentale și dezvoltarea agenților inteligenți.

Această teză propune un cadru nou pentru înțelegerea interacțiunii strategice,

punând accent pe *incertitudinea ca trăsătură structurală fundamentală* a jocurilor repetate. Printr-o *tranziție esențială de la modelarea cu orizont fix la cea cu orizont incert*, această lucrare introduce îmbunătățiri teoretice și operaționale semnificative, aplicabile în contexte experimentale bazate pe simulare. O contribuție majoră este *proiectarea și implementarea unui sistem inovator de simulare Monte Carlo*, dezvoltat special pentru această cercetare, care oferă o abordare metodologică solidă și testată experimental pentru evaluarea comportamentului strategic în condiții de durată necunoscută a jocului. Prin *integrarea raționamentului formal din teoria jocurilor cu simulări computaționale*, această cercetare avansează o teorie scalabilă a luării deciziilor în sisteme dinamice multi-agent. Ea oferă perspective noi asupra mecanismelor cooperării, vulnerabilităților acestora și influenței adaptabilității și volatilității mediului. Depășind ipoteza orizontului fix și validând un cadru flexibil pentru modelarea comportamentală, teza contribuie în mod semnificativ la înțelegerea inteligenței strategice în condiții de incertitudine.

Capitolul 2 oferă un cadru teoretic și metodologic solid pentru analiza comportamentului strategic în jocurile repetate, având ca model central Dilema Prizonierului Iterată (IPD). Prima parte a capitolului este consacrată fundamentele teoriei jocurilor clasice, incluzând reprezentarea în formă normală, criteriile de dominanță și echilibrul Nash — în strategii pure și mixte — ca piloni ai raționamentului strategic formalizat. Capitolul analizează apoi modul în care repetiția modifică comportamentele strategice: dacă în jocurile cu o singură rundă este anticipată trădarea, interacțiunile repetate — mai ales cele cu orizont incert sau infinit — pot susține cooperarea prin mecanisme precum reciprocitatea, pedeapsa și reputația. Teorema Folk (Folk Theorem) și formulările rafinate ale echilibrului evidențiază capacitatea agenților răbdători de a susține cooperarea fără intervenție externă. De asemenea, sunt examinate formarea convingerilor și structurile de monitorizare, evidențiind importanța acestora pentru credibilitatea strategiilor. Este introdusă o abordare originală de *simulare Monte Carlo* pentru a evalua robustețea strategiilor în condiții de incertitudine. Mediul utilizează criterii de oprire stocastice bazate pe distribuții normale, permițând evaluarea repetată a performanței agenților în scenarii probabilistice. Modelarea agenților prin reguli deterministe și structuri probabilistice finite asigură coerența metodologică necesară pentru analizarea diferențiată a strategiilor într-un mediu de simulare controlat. Capitolul 3 analizează modul în care structura temporală a jocurilor repetate influențează deciziile strategice în sisteme multi-agent, cu un accent deosebit pe rolul certitudinii sau incertitudinii privind durata interacțiunii. Analiza compară jocurile cu orizont finit — în care inducția inversă tinde să conducă

la trădare — cu cele desfășurate în orizont incert sau infinit, unde cooperarea poate fi susținută prin mecanisme precum actualizarea informației, formarea de anticipații și reciprocitate condiționată. Capitolul integrează contribuții recente din literatura dedicată jocurilor dinamice cu orizont infinit, incluzând modele de feedback anticipativ și mecanisme de actualizare informațională. Astfel, este conturat un cadru analitic coerent pentru investigarea deciziilor strategice în contexte caracterizate de incertitudine temporală și complexitate interactivă. Această cercetare extinde turneele lui Axelrod printr-un cadru de simulare care integrează durate incerte ale jocului, utilizând criterii stocastice de oprire bazate pe distribuții normale. Acest demers reflectă mai fidel scenariile reale, în care momentul finalizării interacțiunilor nu este cunoscut. Astfel, se mărește considerabil aria analitico-experimentală a studiilor clasice privind Dilema Prizonierului iterată (IPD).

Pentru susținerea acestei direcții, a fost dezvoltat un sistem personalizat de simulare Monte Carlo. *O contribuție esențială o constituie chiar platforma experimentală propusă*, care permite jocuri de lungime variabilă, monitorizarea automata a rezultatelor și exporturi largi de date. Instrumentele vizuale dezvoltate—precum diagramele circulare, matricile de cooperare sau hărțile termice—fac vizibile pattern-uri emergente și nepotriviri strategice relevante. Simularea include și reconstituirea celui de-al Doilea Turneu organizat de Axelrod, aplicând ipoteze probabilistice privind orizontul temporal. Evaluarea performanței strategiilor s-a realizat prin rulări repetate, utilizând un set extins de indicatori cantitativi. *Un element inovator de ordin metodologic constă în utilizarea unor indicatori diferențiali*, precum variațiile poziționării ierarhice și volatilitatea randamentelor, care permit o clasificare empiric-validabilă a strategiilor în funcție de robustețe, capacitate adaptativă sau instabilitate comportamentală.

Capitolul 4 oferă o analiză aprofundată a modelului Allison Mixture, evidențiindu-l ca o inovație metodologică relevantă pentru studiul comportamentului strategic în jocuri repetate sub incertitudine. Modelul este plasat în contextul proceselor stocastice și al paradoxului lui Parrondo, demonstrând cum comutarea probabilistă între strategii neutre sau suboptimale poate genera rezultate emergente favorabile. Această abordare ilustrează modul în care reguli structurale simple pot da naștere unor comportamente adaptative complexe în medii caracterizate de ambiguitate.

O contribuție majoră constă în *transpunerea modelului Allison Mixture într-un cadru strategic* pentru interacțiuni repetate, prin intermediul unui model riguros din punct de vedere matematic al dinamicii tranzițiilor și al autocovarianței. Această formalizare permite investigarea modului în care memoria strategică, variabilitatea semnalelor și asimetria decizională converg în configurarea comportamentului agenților

în condiții de incertitudine. Capitolul prezintă, de asemenea, *proiectarea unui sistem personalizat de simulare*, destinat desfășurării de experimente Monte Carlo extinse, variind probabilitățile de tranziție și lungimile interacțiunii. Acest sistem permite testarea precisă a convergenței, a varianței estimatorilor și a consistenței recompenselor în condiții dinamice, oferind soluții pentru limitările abordărilor anterioare, de scară redusă sau strict analitice.

Modelul este aplicat în continuare atât în scenarii simetrice, cât și în scenarii de joc asimetrice, evidențiindu-l *nu doar ca și construct statistic, ci și ca instrument conceptual* pentru modelarea strategiilor adaptative. Prin capacitatea de a reproduce autocorelații și dependențe structurale acolo unde strategiile clasice eșuează, modelul se conturează ca un instrument promițător pentru cercetările viitoare în inteligența artificială și teoria deciziilor.

Cercetarea propune, în cele din urmă, *o arhitectură de simulare extensibilă și reproductibilă*, deschizând calea pentru modele hibride care combină tranziții probabiliste cu euristici strategice. Acest cadru nu funcționează doar ca mecanism de validare pentru modelul Allison Mixture, ci și ca fundație pentru viitoare demersuri experimentale în domeniul mai larg al interacțiunilor strategice dinamice. În acest cadru extins, lucrarea contribuie la o reconfigurare a modului în care jocurile repetate sunt modelate teoretic și evaluate empiric, deplasând accentul analitic de la ipotezele cu orizont fix la reprezentări probabiliste ale duratei interacțiunii. Prin construcția unui cadru integrat de simulare Monte Carlo și aplicarea sa duală atât în turnee de strategii clasice, cât și în procese stocastice de tip Allison Mixture, teza testează nu doar robustețea structurală a strategiilor consacrate, ci introduce noi *direcții metodologice pentru surprinderea volatilității, adaptabilității și cooperării emergente în medii incerte*.

Tratarea incertitudinii ca element structural — și nu ca simplă condiție periferică — permite acestei lucrări să reflecte mai fidel complexitatea deciziei descentralizate în contexte caracterizate de ambiguitate temporală. Instrumentele și perspectivele dezvoltate oferă un cadru relevant pentru cercetări viitoare în sisteme multi-agent, modelare comportamentală și fundamentele computaționale ale inteligenței adaptative. Prin articularea dintre intuițiile teoriei clasice a jocurilor și experimentarea computațională scalabilă, teza propune o bază conceptuală și tehnică pentru construcția unor modele mai realiste și predictive ale comportamentului strategic în medii dinamice și incerte.

2

Rezumat al Capitolului 2: Fundamentele teoretice, formale și metodologice ale luării deciziilor strategice în jocuri repetate și în Dilema Prizonierului iterată

Capitolul 2 stabilește cadrul conceptual, formal și computațional necesar pentru analiza comportamentului strategic în jocurile repetate, utilizând Dilema Prizonierului Iterată (IPD) ca model central. Această structură este esențială pentru înțelegerea dinamicii cooperării și conflictului în sisteme multi-agent, atât în contexte teoretice, cât și aplicative.

Concepte teoretice și notații

Această secțiune introductivă fundamentează raționamentul de inspirație teoretico-jocurilor prin prezentarea structurii formale a jocurilor în formă normală, incluzând reprezentarea matematică a jucătorilor, strategiilor și câștigurilor (payoffs). Echilibrul Nash este definit riguros și discutat atât în strategii pure, cât și în strategii mixte, evidențiindu-se stabilitatea strategică pe care o oferă într-o gamă largă de medii și contexte (Nash, 1950). Discuția depășește cadrul clasic prin integrarea perspectivelor teoriei evolutive a jocurilor și a modelelor cu informație incompletă. Aceste abordări permit analiza comportamentului strategic sub raționalitate limitată și semnale ambigue, facilitând o înțelegere mai nuanțată a dinamicii agenților în sisteme complexe.

Jocuri strategice și Dilema Prizonierului Iterată

Această secțiune aprofundează fundamentele teoretice și conceptuale necesare pentru analiza interacțiunilor strategice în medii caracterizate prin decizii interdependente. Discuția debutează prin diferențierea dintre deciziile strategice și cele non-strategice, subliniind faptul că comportamentul strategic apare atunci când rezultatul alegerii unui agent depinde nu doar de preferințele proprii, ci și de acțiunile simultane sau anticipate ale celorlalți (Malook, 2024). Această interdependență introduce un nivel suplimentar de complexitate, care necesită formarea unor așteptări raționale și un raționament adaptativ.

Jocurile strategice oferă un cadru formal pentru modelarea unor astfel de scenarii, în care agenții — fie indivizi sau sisteme artificiale — trebuie să țină cont de strategiile și reacțiile probabile ale co-jucătorilor. Bazându-se atât pe surse clasice, cât și pe contribuții contemporane, această secțiune explorează modul în care raționamentul strategic se aplică diverselor tipologii de agenți: naturali (umani), corporativi sau de inteligență artificială (AI) (Floridi, 2023).

Pe acest fundament se poate construi o evaluare obiectivă a modelelor de interacțiune strategică statică, de tip unic(one-shot). Deși aceste modele sunt analitic elegante și esențiale pentru înțelegerea unor concepte fundamentale precum strategie dominantă și echilibrul Nash, ele nu surprind pe deplin complexitatea deciziilor din lumea reală. Realitatea decizională este adesea ancorată în relații recurente, structuri instituționale sau contexte strategice dinamice, aflate în continuă evoluție. O ilustrare emblematică a acestei tensiuni este oferită de problema clasică a Prizonierului, care pune în contrast raționalitatea individuală cu optimul colectiv. În lipsa unui cadru extern de constrângere sau a unui mecanism de reglementare, agenții raționali au tendința să se saboteze reciproc, chiar dacă o colaborare mutuală ar aduce beneficii maxime pentru ambele părți.

Trecerea de la interacțiuni unice la interacțiuni repetate marchează o schimbare teoretică esențială. Dilema prizonierului iterată (IPD) introduce o dimensiune temporală care modifică fundamental peisajul strategic.(Myerson, 1991) Repetiția deschide posibilitatea pentru un comportament condiționat de istoricul anterior, permițând strategiilor să își adapteze deciziile în funcție de rezultatele trecute. Această caracteristică permite implementarea unor mecanisme precum reciprocitatea, pedeapsa și iertarea , esențiale pentru susținerea cooperării chiar și în medii dominate de interesul propriu.(Axelrod, 1984) Strategii ca Tit-for-Tat, Grim Trigger sau Win-Stay Lose-Shift ilustrează modul în care simplitatea, memoria și condiționalitatea interacționează pentru a genera echilibre de cooperare rezistente .

Jocuri repetate: orizont finit versus orizont infinit

Această secțiune oferă o analiză cuprinzătoare a consecințelor strategice care decurg din structura temporală a jocurilor repetate, contrastând în mod specific jocurile cu orizont finit—în care numărul de iterații este cunoscut și fixat—cu cele care presupun un orizont infinit sau incert. Această distincție nu este una pur tehnică; ea remodelează peisajul strategic prin modificarea modului în care agenții evaluează câștigurile viitoare, compară cooperarea cu trădare și anticipează comportamentul celorlalți participanți.

În jocurile repetate finit, principiul inducției inverse impune o constrângere teoretică puternică. Deoarece se presupune că jucătorii raționali anticipează viitorul și raționează recursiv, trădarea devine strategia dominantă în ultima rundă, ceea ce declanșează un efect de cascadă de trădări în toate etapele anterioare. Această logică, fundamentată pe raționalitatea clasică din teoria jocurilor, prezice colapsul cooperării chiar și în jocurile în care cooperarea mutuală ar conduce la câștiguri cumulative mai mari. (Cressman, 1996)

Jocurile repetate cu orizont infinit sau cele cu continuare probabilistică (adică jucătorii nu știu cu certitudine dacă jocul continuă după fiecare rundă) oferă un mediu mai favorabil pentru apariția și stabilitatea cooperării. (Fudenberg & Maskin, 1986) Absența unei runde finale cunoscute elimină argumentul inducției inverse, permițând consecințele viitoare să influențeze comportamentul prezent. Conceptul de „umbra viitorului” (shadow of the future) devine central: atunci când agenții atribuie o valoare suficient de mare câștigurilor pe termen lung astfel sunt stimulați să mențină cooperarea pentru a evita declanșarea unor pedepse viitoare.

Dincolo de condițiile teoretice, rafinamente recente integrează caracteristici mai realiste, precum monitorizarea imperfectă, memorie limitată și informație asimetrică. De exemplu, în cadrul jocurilor repetate cu informație imperfectă, modelele cu semnalizare publică—în care semnalele despre acțiunile anterioare sunt comune tuturor jucătorilor—permit implementarea echilibrelor sub forma strategiilor de pedeapsă perfect credibile. În schimb, semnalizarea privată sau afectată de zgomot introduce incertitudine în actualizarea credințelor bayesiene și complică sustenabilitatea echilibrelor cooperative, deoarece jucătorii nu pot distinge cu certitudine între deviații intenționate și fluctuații stocastice. Cercetători precum Hörner și Olszewski (2009) au demonstrat că, chiar și în contexte cu monitorizare imperfectă, existența echilibrelor cooperative este posibilă, cu condiția ca jucătorii să dispună de o capacitate de memorie suficientă pentru a condiționa strategiile de istoricul interacțiunilor

și să evalueze beneficiile viitoare printr-un factor de actualizare suficient de ridicat. În astfel de medii, cooperarea poate fi susținută ca echilibru perfect bayesian sau echilibru sub joc perfect (subgame-perfect equilibrium), în funcție de structura informațională și de mecanismele de penalizare disponibile.

În concluzie, această secțiune subliniază faptul că orizontul temporal al jocurilor repetate este un determinant de ordinul întâi al posibilităților strategice. În timp ce jocurile finite constrâng cooperarea prin raționament invers, orizonturile infinite sau incerte permit strategii pe termen lung care recompensează cooperarea și sancționează defecțiunea. Această distincție nu este doar de importanță teoretică, ci este esențială și pentru proiectarea agenților robuști în medii reale multi-agent, în care durata și structura interacțiunii sunt rareori fixe sau cunoscute în avans.

Strategii din Turneele lui Axelrod

Această secțiune oferă o analiză detaliată a contribuțiilor empirice și teoretice rezultate din influențele turneele organizate de Robert Axelrod pe tema Dilemei Prizonierului Iterat (Iterated Prisoner's Dilemma, IPD). (Axelrod, 1984) Aceste experimente au marcat un punct de cotitură în studiul cooperării, demonstrând că un comportament strategic favorabil beneficiului mutual poate apărea chiar și în medii guvernate de raționalitatea individuală și interesul propriu. Mai mult decât o inovație metodologică, turneele lui Axelrod au acționat ca o punte între modelele teoretice ale jocurilor și simulările bazate pe agenți, oferind un cadru empiric solid pentru testarea ipotezelor privind reziliența și adaptabilitatea strategiilor în interacțiuni repetate.

Primul turneu organizat de Axelrod (1980) a invitat cercetători din domenii diverse să trimită strategii algoritmice pentru IPD, competiția desfășurându-se în format round-robin: fiecare strategie a jucat împotriva tuturor celorlalte—including auto-jocuri—într-o serie de meciuri a câte 200 de runde. Rezultatele au fost remarcabile. În ciuda variației considerabile în ceea ce privește complexitatea strategică, utilizarea memoriei și regulile comportamentale, cea mai simplă strategie trimisă—Tit for Tat (TFT), propusă de Anatol Rapoport—a ieșit câștigătoare. Regula TFT era elegantă: începe cu cooperare, apoi imită alegerea precedentă a adversarului. Această strategie s-a dovedit nu doar eficientă în generarea unor câștiguri ridicate, ci și robustă în fața exploatării și represaliilor.

Primul turneu a evidențiat mai multe principii strategice esențiale. Strategiile care combinau bunăvoința (niciodată nu trădează primul), capacitatea de represalii (răspund la trădare), iertarea (revenirea la cooperare) și claritatea (comportament ușor

de interpretat de către oponent) au avut performanțe superioare în contexte repetitive. (Axelrod & Hamilton, 1981) Succesul emergent al strategiei TFT l-a determinat pe Axelrod să organizeze un al doilea turneu, mai complex, care a reunit 63 de strategii provenite de la cercetători din șase țări. A doua ediție a turneului a simulat un proces evolutiv, eliminând strategiile ineficiente și propagând cele de succes prin replicare pe baza selecției performanței .

Interesant este că al doilea turneu a scos la iveală și limitele strategiei TFT, deschizând calea pentru proiectarea unor strategii mai nuanțate. Deși TFT a rămas competitivă, în anumite contexte a fost depășită de variante precum Generous Tit for Tat (GTFT), care iartă ocazional trădarea, sau de strategii mai complexe precum Tester și Tranquilizer, care utilizau mișcări de testare pentru a detecta oponenti exploatabili.

Din perspectivă teoretică, turneele lui Axelrod au demonstrat că cooperarea poate constitui un rezultat evolutiv viabil, chiar și în absența unei autorități centralizate sau a comunicării directe între agenți. Rezultatele au contrazis dominația strategiilor pur egoiste (precum Always Defect) și au arătat că regulile de comportament bazate pe reciprocitate și proporționalitate pot susține beneficiul mutual pe orizonturi temporale extinse.

În concluzie, acest subcapitol evidențiază modul în care turneele lui Axelrod au furnizat atât un model metodologic, cât și o paradigmă teoretică pentru studiul cooperării în jocurile repetate. În acest sens, cadrul conceptual oferit de Axelrod rămâne o piatră de temelie în analiza comportamentului strategic în sisteme multi-agent, unind teoria normativă cu dinamica empirică într-un mod rar egalat în literatura de specialitate.

Metodologia simulării Monte Carlo

Această secțiune introduce cadrul computațional de simulare utilizat pentru a evalua comportamentele strategice în cadrul Dilemei Prizonierului Iterate (Iterated Prisoner's Dilemma, IPD), în condiții de incertitudine. Metodologia prezentată aici extinde cadrul teoretic și abordarea experimentală propusă în articolul nostru: Axelrod First Tournament: examining certainty versus uncertainty about the end stage in repeated games. (Milencianu & Pop, 2023) Bazându-se atât pe perspective clasice, cât și pe abordări empirice recente, metodologia valorifică simularea Monte Carlo pentru a depăși limitările analitice și pentru a explora sistematic performanțele strategiilor în structuri de interacțiune stocastică.

În esență, simularea Monte Carlo implică generarea unui număr mare de eșantioane aleatorii repetate pentru a estima rezultatele așteptate. În contextul jocurilor repetate, aceasta permite aproximarea câștigurilor medii, stabilității comportamentale și tendințelor de convergență, în funcție de lungimea variabilă a interacțiunii și de combinațiile de oponenti. Spre deosebire de modelarea deterministă, care presupune o durată fixă a jocului, acest cadru utilizează reguli probabilistice de terminare, surprinzând condiții mai realiste în care agenții nu cunosc dinainte momentul încheierii interacțiunii. (Metropolis & Ulam, 1949)

Simularea urmează o structură de tip *round-robin*, în care fiecare strategie este testată în jocuri IPD cu terminare probabilistă împotriva fiecăreia dintre celelalte strategii din setul analizat. Pentru a asigura robustețea rezultatelor, fiecare interacțiune pereche este repetată de un număr mare de ori (de exemplu, 1000 de iterații), permițând estimarea fiabilă a indicatorilor de performanță. În cadrul fiecărui meci, secvența deciziilor este generată de logica internă a fiecărui agent — fie ea deterministă sau stocastică — în funcție de istoricul interacțiunii și de natura mecanismului decizional (de exemplu, automat finit, strategie de tip *one-memory* sau funcție de tranziție probabilistă). Mediul de simulare este implementat în Python, utilizând operații vectorizate eficiente prin biblioteca NumPy. Acesta suportă :

- **Simulare Secvențială:** Generează istoricul complet al jocului pe runde, permițând analiza dinamică a evoluției comportamentale în cadrul meciurilor.
- **Eșantionare Non-Secvențială:** Se concentrează asupra distribuțiilor de rezultate fără a înregistra tranzițiile pas-cu-pas, reducând astfel încărcarea computațională.

Evaluarea performanței se realizează prin generarea de indicatori cantitativi, atât pe baza câștigurilor, cât și a comportamentelor observate:

- Câștigul mediu per agent per meci;
- Vectori de frecvență pentru victorii, înfrângeri și egaluri;
- Sensibilitatea la incertitudine (măsurată prin deviația câștigurilor între orizonturi deterministe și stocastice).

(Milencianu & Pop, 2023)

Din perspectivă metodologică, cadrul de simulare dezvoltat aici reprezintă o extensie a designului experimental original al lui Axelrod. În timp ce turneele lui Axelrod presupuneau un număr fix de runde (de exemplu, 200), abordarea actuală generalizează lungimea interacțiunii, extrăgând-o din distribuții de probabilitate

(de exemplu, geometrică, normală, uniformă). Acest lucru permite compararea strategiilor identice în condiții de certitudine și incertitudine, facilitând investigații experimentale controlate asupra robusteții temporale.

Procedura de simulare urmează o succesiune structurată de pași. Pentru fiecare pereche ordonată de strategii (s_i, s_j) , simularea execută K instanțe independente. În fiecare instanță, numărul de runde T este extras dintr-o distribuție geometrică cu probabilitate de continuare ω , iar ambele strategii interacționează pe durata a T runde ale jocului Dilemei Prizonierului. Deciziile agenților sunt determinate de logica internă a fiecăruia (de exemplu, regulă deterministă, reacție bazată pe memorie, tranziție probabilistă), iar câștigurile rezultate sunt înregistrate și agregate.

În concluzie, metodologia simulării Monte Carlo descrisă în această secțiune constituie fundamentul empiric al prezentei cercetări. Ea permite testarea sistematică a unei game extinse de strategii în contexte temporale probabiliste și oferă un cadru experimental solid pentru investigarea impactului incertitudinii asupra comportamentului strategic. Dincolo de aplicabilitatea sa în cadrul Dilemei Prizonierului Iterate, această abordare este relevantă pentru o varietate de medii decizionale multi-agent, în care impredictibilitatea orizontului temporal și adaptarea descentralizată joacă un rol esențial.

Indicatori de performanță și evaluare strategică

Această secțiune prezintă cadrul evaluativ utilizat pentru analiza și compararea strategiilor în jocuri repetate, atât în condiții de orizont determinist, cât și sub incertitudine temporală. În loc să se bazeze exclusiv pe câștigul mediu—un indicator adesea insuficient al robusteții strategice—evaluarea propusă integrează un sistem metric multidimensional conceput pentru a surprinde nu doar performanța finală, ci și stabilitatea comportamentală, volatilitatea și sensibilitatea contextuală. Acest cadru a fost aplicat în mod sistematic tuturor perechilor de strategii în cadrul mediului de simulare descris în subcapitolul anterior. (Milencianu & Pop, 2023)

Pentru fiecare pereche de strategii, simularea generează doi vectori empirici principali: vectorul de frecvență al victoriilor, înfrângerilor și egalurilor, respectiv vectorul câștigurilor medii, calculat prin media rezultatelor obținute în toate rulările simulării. Acești indicatori sunt calculați separat atât pentru jocuri cu lungime fixă, cât și pentru jocuri cu orizont incert, permițând o comparație directă. Diferențele comportamentale sunt surprinse prin metrice de discrepantă, care cuantifică diferențele absolute în rezultate observate (de exemplu, modificări ale ratei de succes sau ale câștigurilor)

între cele două condiții temporale. O discrepanță mare sugerează că strategia este structural sensibilă la natura orizontului temporal, în timp ce discrepanțele mici pot indica robustețe sau consistență mecanică.

Un alt aspect esențial al sistemului evaluativ este clasificarea strategiilor în funcție de tipologii derivate din valorile de discrepanță. Rezultatele empirice arată că strategiile tind să se încadreze în trei categorii generale:

- **Strategii robuste**, precum *Tit-for-Tat* și *Grudger*, care înregistrează performanțe constante în ambele condiții temporale (determinate și probabilistice).
- **Strategii adaptive**, care manifestă discrepanțe moderate, dar își mențin poziția relativă și avantajul competitiv, sugerând o capacitate de adaptare consistentă la schimbarea condițiilor de joc.
- **Strategii volatile**, caracterizate prin variații mari ale câștigurilor, tiparelor de succes/insucces sau ale poziționării relative între cele două medii, adesea ca urmare a supra-specializării sau rigidității structurale.

Pentru identificarea acestor tipare, au fost utilizați și indicatori bazați pe rang, complementari metricilor de câștig. De exemplu, *Rank Shift* urmărește modificările poziției unei strategii în clasamentul global al performanței între condițiile de joc cu orizont fix și cele cu orizont incert. În combinație cu valorile discrepanțelor bazate pe câștiguri, acest instrument permite detectarea strategiilor care nu doar că au performanță scăzută, ci sunt și structural instabile—potențial utile în medii adversariale, dar nesigure în sisteme dinamice.

Toți indicatorii de performanță au fost structurați sub formă matricială (matrici ale câștigurilor, matrici de discrepanță, matrici de volatilitate), facilitând atât agregarea cantitativă, cât și reprezentarea calitativă.

În concluzie, sistemul de evaluare a performanței prezentat în această secțiune depășește măsurătorile tradiționale ale succesului și propune o abordare multidimensională, sensibilă la discrepanță, adaptată mediilor în care domină incertitudinea temporală. Acesta sprijină obiectivul general al tezei, acela de a înțelege reziliența strategică, de a identifica comportamente emergente și de a fundamenta dezvoltarea unor agenți adaptativi, capabili de luare robustă a deciziilor în sisteme stocastice multi-agent.

Concluzie

Capitolul 2 sintetizează modelele clasice din teoria jocurilor cu tehnici moderne de simulare, oferind o abordare unificată pentru analiza deciziilor strategice în condiții de incertitudine temporală. Prin formalism teoretic, validare empirică și inovație metodologică, capitolul fundamentează riguros analiza robusteții strategice. Cadrul de simulare bazat pe Monte Carlo, împreună cu setul de indicatori evaluativi multidimensionali, oferă instrumentele necesare pentru investigarea dinamicilor emergente, evaluarea complexității strategice și identificarea comportamentelor fiabile în sisteme multi-agent.

3

Rezumatul Capitolului 3: Analiza experimentală a comportamentului strategic în condiții determinate versus incerte

Conținutul acestui capitol a fost publicat parțial în: *The Second Axelrod Tournament: A Monte Carlo Exploration of Uncertainty About the Number of Rounds in Iterated Prisoner's Dilemma*, Studia Universitatis Babeș-Bolyai Oeconomica, 70(1), 67–82. (G. Pop et al., 2025)

Perspective teoretice asupra mediilor deterministe și incerte

Aceast subcapitol explorează modul în care strategiile evoluează în raport cu tipul orizontului temporal al jocului — determinat sau cu durată incertă.

În medii deterministe, cu orizont finit, jucătorii cunosc cu precizie numărul de runde ale jocului. Această cunoaștere comună permite aplicarea raționamentului invers, un principiu fundamental al teoriei clasice a jocurilor. (Fudenberg & Tirole, 1991) Agenții raționali, anticipând că în ultima rundă cooperarea nu poate fi susținută, deduc că trădarea reprezintă mutarea optimă în acel punct. Cunoscând acest lucru, ei proiectează trădarea înapoi prin fiecare etapă precedentă, concluzionând că trădarea mutuală este singura cale rațională încă de la început. Această logică, deși elegantă și formal riguroasă, eșuează adesea în a descrie comportamentul observat în situații reale, unde cooperarea apare și persistă chiar și în jocuri cu durată finită. Soluția prin

inducție inversă se bazează pe condiții idealizate: informație perfectă, cunoaștere comună a raționalității și absența zgomotului sau a incertitudinii. (Myerson, 1991)

În numeroase contexte strategice reale, agenții nu știu cu exactitate când se va încheia interacțiunea. Fiecare etapă este urmată de o posibilitate de continuare, a cărei probabilitate este necunoscută sau specificată stocastic. Aceasta generează ceea ce literatura numește adesea „umbra viitorului” (shadow of the future)—unde ideea este că, consecințele viitoare influențează deciziile prezente, chiar și în absența unor mecanisme formale de impunere. (Bó, 2002) În condiții de incertitudine, agenții au motive să investească în reciprocitate, să-și construiască reputații și să adopte strategii care condiționează comportamentul prezent de perspectiva unei interacțiuni viitoare. Spre deosebire de jocurile cu orizont finit, unde ultima rundă dictează așteptările strategice, incertitudinea permite susținerea cooperării prin așteptări implicite și norme informale.

Dintr-o perspectivă sistemică, mediile incerte reflectă mai fidel structura multor sisteme socio-economice. În contexte precum piețele, negocierile, alianțele sau platformele online, agenții nu cunosc, de regulă, numărul exact al interacțiunilor viitoare. Cu toate acestea, în timp, apar și persistă tipare de cooperare și sancționare. Acest fenomen sugerează că incertitudinea nu reprezintă doar un obstacol de depășit, ci o caracteristică structurală ce poate favoriza rezultate sociale mai flexibile și mai eficiente. Un orizont temporal incert stimulează apariția unor forme avansate de interacțiune, precum dezvoltarea normelor, consolidarea cooperării și stabilizarea comportamentelor colective, tocmai prin evitarea constrângerii artificiale impuse de un final prestabilit. (Leyton-Brown & Shoham, 2008)

În lumina acestor considerații teoretice, orice investigație empirică privind robustețea strategiilor trebuie să ia în serios implicațiile incertitudinii temporale. Evaluarea strategiilor exclusiv în condiții cu orizont fix ar ignora mecanismele esențiale prin care cooperarea apare și persistă. Ar subestima, de asemenea, capacitățile strategice ale agenților care își adaptează comportamentul nu doar în funcție de istoricul interacțiunilor, ci și în raport cu așteptările privind viitorul și dinamica acestuia. Acest capitol se bazează pe fundamentele conceptuale prezentate aici pentru a dezvolta o metodologie de simulare capabilă să testeze performanța strategică în medii deterministe și incerte deopotrivă.

Jocuri dinamice cu orizont infinit și Dilema Prizonierului Iterată

Această secțiune integrează contribuții recente din teoria jocurilor dinamice cu orizont infinit în analiza Dilemei Prizonierului Iterate (IPD), punând accent pe controlul anticipativ, robustețea strategică în condiții de incertitudine și învățarea adaptivă. Aceste extensii teoretice sporesc capacitatea explicativă a cadrului IPD, oferind o formalizare riguroasă a modului în care agenții pot planifica în condiții de viitor incert, își pot actualiza în mod dinamic estimările despre contextul de joc și pot elabora strategii reziliante față de variații în lungimea orizontului sau în comportamentul celorlalți participanți.

O contribuție semnificativă în această direcție este introducerea *mecanismelor de feedback anticipativ* (Chen & Zadrozny, 2002) În contrast cu logica inducției inverse caracteristică jocurilor finite sau cu strategiile invariabile în timp asociate echilibrelor clasice, structurile de tip feedback anticipativ oferă agenților — în special celor aflați în poziții de lider în cadrul jocurilor ierarhice — posibilitatea de a incorpora așteptările privind reacțiile viitoare ale oponentilor direct în procesul decizional curent. Aplicate la Dilema Prizonierului Iterate (IPD), aceste mecanisme anticipative permit modelarea unor strategii de tip leadership generos, în care un jucător optează pentru cooperare nu dintr-un interes imediat, ci pornind de la premisa că un asemenea semnal va influența optim răspunsul viitor al adversarului. O astfel de abordare reflectă dinamici observabile în sisteme descentralizate, în care inițiativele timpurii de cooperare contribuie la emergența și consolidarea normelor și așteptărilor colective pe termen lung.

În paralel, cercetările privind *strategiile de control al costurilor garantate*, dezvoltate de Gyurkovics și Takács (2005), oferă un cadru structurat pentru analiza jocului cu risc limitat în contexte incerte și adversariale. (Gyurkovics & Takács, 2005) Aceste strategii garantează că jucătorii nu vor suporta pierderi care depășesc un prag prestabilit, indiferent de acțiunile adversarului sau de prezența zgomotului sistemic. În cadrul Dilemei Prizonierului Iterate (IPD), ele exprimă o formă de cooperare precaută: agenții sunt dispuși să colaboreze, însă doar în măsura în care riscul de exploatare repetată rămâne controlabil din punct de vedere formal. Această abordare introduce un grad suplimentar de robustețe în strategiile de tip represiv, precum Grim Trigger sau Tit for Tat, reinterpretându-le ca mecanisme adaptative capabile să mențină echilibrul cooperării în medii dinamice și stocastice, chiar și în condiții de incertitudine persistentă.

O altă inovație teoretică relevantă, datorată lui D. Yeung și Petrosian, provine din cadrul *mecanismelor de actualizare informațională*, care modelează modul în care jucătorii își revizuiesc așteptările în timp, pe baza rezultatelor observate. (D. W. K. Yeung & Petrosyan, 2017) Această perspectivă se distanțează de cadrul static al convingerilor comune și al funcțiilor de utilitate fixe, modelând Dilema Prizonierului Iterate (IPD) ca un joc evolutiv, în care strategiile sunt supuse unei revizuirii endogene. În medii cu orizont stohastic și observație imperfectă, agenții își ajustează strategiile prin procese adaptative, ghidate de dinamici bazate pe informație acumulată și inferență probabilistă.

Odată cu integrarea acestor cadre teoretice, Dilema Prizonierului Iterată (IPD) își depășește rolul tradițional de reper static, devenind un instrument analitic valoros pentru explorarea inteligenței strategice dinamice. Abordarea bazată pe feedback anticipativ arată cum strategiile orientate spre viitor pot susține cooperarea pe termen lung - strategiile de control al costurilor garantate oferă robustețe în fața adversarilor imprevizibili - iar mecanismele de actualizare informațională oferă un cadru formal pentru modelarea adaptabilității comportamentale, a raționalității limitate și a plauzibilității evolutive a strategiilor. Împreună, aceste contribuții conturează un cadru integrat pentru înțelegerea modului în care agenții strategici gestionează incertitudinea, oferind fundamente solide pentru transpunerea modelelor teoretice în contexte aplicate de interacțiune strategică.

Structura experimentală și rezultate din Primul Turneu Axelrod

Această secțiune se bazează direct pe fundamentele metodologice și concluziile prezentate în articolul *Axelrod first tournament: Examining certainty versus uncertainty about the end stage in repeated games* (Milencianu & Pop, 2023), oferind o prezentare detaliată a arhitecturii experimentale, a procedurilor de evaluare și a rezultatelor obținute în ceea ce privește performanța strategiilor.

Configurația jocului a presupus un orizont determinist de 200 de runde per meci, o alegere atât pragmatică—pentru a facilita simulări limitate computațional—cât și teoretic semnificativă. Cunoașterea comună a unui număr fix de runde a generat un stimulent structural pentru trădare pe măsură ce jocul se apropia de final. Acest cadru a exercitat o presiune semnificativă asupra strategiilor cooperative, oferind un

test empiric al condițiilor în care cooperarea poate totuși să apară, în ciuda predicției inducției inverse.

Strategiile testate au variat considerabil în complexitate și metodologie de proiectare. Unele au fost bazate pe euristici simple, precum *Always Defect*, *Always Cooperate* sau *Tit for Tat* (TFT). Altele au integrat logici mai sofisticate, cu memorie, randomizare sau comportamente de testare. În ciuda acestei eterogenități, turneul a evidențiat regularități empirice robuste, care au remodelat înțelegerea cooperării.

Cel mai remarcabil rezultat a fost succesul strategiei TFT—o strategie minimalistă, dar eficientă—care începe cu cooperare și apoi imită alegerea adversarului din runda anterioară. Performanța TFT a fost notabilă nu doar prin câștigul obținut, ci și prin capacitatea sa de a promova cooperarea fără a fi vulnerabilă la exploatare.

Dincolo de succesul TFT, turneul a oferit date valoroase despre modul în care diferite tipuri de strategii performează în contexte de interacțiune repetată. Strategiile „bune” (care nu inițiază niciodată trădare) au avut, în general, performanțe mai bune, în special când au fost combinate cu oponenti similari. Strategiile agresive sau înșelătoare, deși uneori eficiente în anumite meciuri, au suferit pierderi considerabile în fața adversarilor capabili de represalii. Aceste rezultate oferă sprijin empiric pentru ipoteza conform căreia cooperarea poate fi stabilă evolutiv în medii caracterizate prin interacțiuni repetate și raționalitate limitată.

Din punct de vedere metodologic, Primul Turneu Axelrod a introdus utilizarea simulărilor ca instrument riguros de evaluare comparativă a strategiilor. În loc să se bazeze exclusiv pe echilibre analitice, formatul turneului a facilitat observarea fenomenelor emergente, cum ar fi ciclurile de cooperare și trădare, efectele de blocaj și lipsa de coordonare strategică. De asemenea, turneul a evidențiat rolul condițiilor inițiale și al tiparelor de interacțiune timpurii în determinarea rezultatelor pe termen lung—o trăsătură deosebit de relevantă în jocurile cu orizont extins.

În concluzie, Primul Turneu Axelrod a demonstrat că un comportament cooperant poate apărea și persista chiar și în medii în care, conform presupunerilor clasice, trădarea ar fi predicția de echilibru. Printr-un design experimental transparent și reproductibil, turneul a contestat dominația modelelor strict egoiste și a introdus metrici empirice—precum câștigul mediu, robustețea la zgomot și reacția la provocări—care rămân esențiali în evaluarea strategiilor din jocurile repetate. Perspectivele obținute continuă să influențeze proiectarea agenților strategici în sisteme descentralizate, de la economie până la inteligență artificială.

Arhitectura experimentală

Această parte descrie infrastructura experimentală personalizată dezvoltată pentru evaluarea comportamentului strategic în turnee de tip IPD desfășurate în condiții temporale atât determinate, cât și incerte. Biblioteca Python `Axelrod` a oferit o bază solidă pentru simularea interacțiunilor repetate între agenți, însă implementarea sa originală nu includea funcționalitățile esențiale necesare obiectivelor analitice ale acestui studiu—în special, identificarea explicită a câștigătorilor la nivel de meci și controlul dinamic al duratei meciurilor prin procese stohastice.

Pentru a depăși aceste limitări, a fost implementată în Python o extensie modulară a motorului de turneu `Axelrod`. Nucleul acestui sistem extins permite controlul aleatoriu asupra numărului de meciuri jucate (`mean_m`, `dev_m`) și asupra numărului de runde per meci (`mean_t`, `dev_t`), simulând structuri de interacțiune cu orizont infinit prin distribuții de timp probabilistice. În plus, numărul de rulări independente ale simulării a fost parametrizat pentru a asigura robustețea statistică a execuțiilor repetate ale turneului.

Implementarea integrează componente esențiale pentru specificarea strategiilor agenților, organizarea secvențială a interacțiunilor repetate și agregarea performanțelor în cadrul simulărilor stocastice. Fiecare strategie a jucat împotriva tuturor celorlalte, inclusiv împotriva sa însăși, cu lungimi ale meciurilor extrase dintr-o distribuție normală, asigurând că jucătorii s-au confruntat cu orizonturi variabile, fără a cunoaște în prealabil momentul terminării. Rezultatele au fost înregistrate sub formă de statistici cumulative (victorii/înfrângeri/egaluri) și normalizate prin rutine automate de agregare. Funcționalitatea dedicată de export în format CSV a asigurat un management structurat al datelor, iar fișierele în versiuni controlate au facilitat reproductibilitatea și monitorizarea pe termen lung a experimentelor.

Vizualizarea a constituit o componentă esențială a arhitecturii experimentale. Prin intermediul bibliotecilor `matplotlib` și `pandas`, datele generate în urma simulărilor au fost sintetizate vizual în reprezentări grafice accesibile, menite să evidențieze structuri strategice relevante. Performanța fiecărei strategii a fost sintetizată prin diagrame — reflectând distribuția rezultatelor (victorii, înfrângeri, remize) în raport cu diverși adversari — și prin matrici de cooperare, care au permis identificarea unor patternuri comportamentale recurente și a eventualelor asimetrii decizionale.

Sistemul de simulare personalizat a oferit un cadru flexibil pentru configurarea mediilor de joc, permițând integrarea duratelor stocastice ale meciurilor, a instrumentelor vizuale de analiză și a unui set extins de metrice de performanță. Acesta

a asigurat un control fin asupra parametrilor experimentali și a susținut evaluări comparative riguroase între scenarii cu orizont determinist și cele marcate de incertitudine. În consecință, arhitectura experimentală a furnizat o bază empirică robustă pentru investigațiile ulterioare privind stabilitatea strategiilor, dinamica comportamentală și adaptabilitatea agenților, în raport cu constrângerile informaționale și structurale impuse de cadrul jocului.

Analiza grafică a sensibilității strategiilor în cadrul Primului Turneu Axelrod

Această secțiune oferă o analiză grafică și statistică a sensibilității strategiilor din cadrul Primului Turneu Axelrod la modificările structurale ale orizontului temporal—în special contrastul dintre interacțiuni cu durată deterministă (fixă) și cele cu orizont aleatoriu. Analiza urmărește să evidențieze modul în care incertitudinea privind durata jocului influențează comportamentele strategice, dincolo de ceea ce reflectă simpla medie a payoff-urilor sau ratele de succes agregate.

Un indicator analitic esențial este Matricea Discrepanțelor, care măsoară variațiile absolute ale payoff-urilor obținute de fiecare strategie între scenariile cu orizont determinist și cele cu terminare stocastică. Această matrice furnizează o sinteză compactă a modului în care configurația temporală a jocului—reprezentată exclusiv prin certitudinea sau incertitudinea încheierii—afectează interacțiunile bilaterale. Strategiile care manifestă variații minime între cele două tipuri de medii sunt considerate insensibile la variații de orizont, în timp ce deviațiile accentuate indică o sensibilitate structurală sau o optimizare excesivă față de un anumit cadru.

Pe lângă variațiile de câștig, *Graficul Cooperării* vizualizează frecvența și simetria acțiunilor cooperative între toate perechile de strategii. Aceasta permite identificarea tiparelor stabile de cooperare, a regimurilor de exploatare asimetrică sau a celor dominate de trădare mutuală. Este de remarcat că anumite strategii performante în mediul determinist înregistrează scăderi semnificative ale cooperării atunci când durata meciurilor devine incertă, semnalând o schimbare comportamentală indusă de ambiguitatea orizontului temporal.

Analiza vizuală indică, per ansamblu, că strategiile variază nu doar în ceea ce privește performanța absolută, ci și în raport cu sensibilitatea la modificările structurale ale mediului de joc. Anumite strategii, precum *Tit for Tat* sau *Grudger*, demonstrează o consistență robustă între cadrele determinist și stocastic, în timp

ce altele, în special cele de natură agresivă sau orientate spre testare, înregistrează instabilitate sau declin semnificativ al performanței în condiții de incertitudine. Aceste constatări susțin conturarea unui cadru tipologic mai nuanțat, fundamentat pe dimensiuni precum volatilitatea comportamentală, capacitatea adaptivă și reacția strategică la ambiguitatea contextuală.

Structura experimentală și rezultate din Al Doilea Turneu Axelrod

Al Doilea Turneu Axelrod a introdus o serie de modificări critice în designul experimental inițial, devenind un reper în studiul empiric al comportamentului strategic în condiții de incertitudine. Bazându-se pe succesele și limitările primei competiții de tip round-robin, această a doua ediție a avut ca obiectiv testarea viabilității evolutive și a robusteții temporale a strategiilor într-un cadru de interacțiune mai dinamic și mai incert. (Axelrod, 1984)

O parte substanțială din datele și analizele comparative prezentate în această secțiune au fost publicate formal în articolul nostru *The Second Axelrod Tournament: A Monte Carlo Exploration of Uncertainty About the Number of Rounds in Iterated Prisoner's Dilemma*. (G. Pop et al., 2025)

Inovația majoră a constat în introducerea unui orizont de joc incert. În locul unui număr fix de runde, fiecare interacțiune din al doilea turneu a continuat cu o probabilitate fixă ω , introducând o terminare stocastică și simulând, în esență, un joc indefinit cu continuare probabilistă. Această ajustare a eliminat inducția inversă și a apropiat mai fidel jocul de situații reale, în care agenții nu pot anticipa exact momentul finalizării interacțiunii.

Turneul a reunit 63 de strategii, incluzând atât variante rafinate ale strategiilor anterioare, cât și propuneri complet noi. Participanții au variat de la reguli deterministe simple (de exemplu, *Always Cooperate*, *Tit for Tat*) la strategii mai complexe bazate pe testare, randomizare sau logici de automat finit. Strategiile nu aveau acces la identitatea sau codul sursă al adversarului și se bazau exclusiv pe istoricul mișcărilor, menținând astfel integritatea interacțiunilor.

Rezultatele celui de-al doilea turneu au confirmat puterea competitivă a strategiilor reciproce și iertătoare, în special *Tit for Tat* și *Generous Tit for Tat*. Acestea au demonstrat robustețe în fața incertitudinii temporale, consistență ridicată a câștigurilor și sensibilitate minimă la terminarea aleatorie. În schimb, strategiile

mai agresive sau exploatare au suferit deseori din cauza represaliilor timpurii și nu au reușit să stabilească echilibre mutual benefice.

Este important de subliniat că orizontul incert a relevat noi fațete ale robusteții strategice. Unele strategii care au obținut rezultate remarcabile în primul turneu au manifestat o performanță oscilantă în condiții de interacțiune imprevizibilă. Pe de altă parte, alte strategii s-au adaptat mai eficient la schimbarea mediului prin utilizarea iertării, a testării stocastice sau a euristicilor bazate pe memorie mixtă. Apariția acestui tip de comportament subliniază importanța evaluării strategiilor nu doar în funcție de randamentul total, ci și în raport cu capacitatea de adaptare la structura mediului.

Din punct de vedere metodologic, al doilea turneu a validat utilizarea terminării probabilistice ca mecanism pentru eliminarea comportamentului artificial din etapele finale și pentru generarea unor tipare mai bogate de cooperare și competiție. De asemenea, acesta a servit drept prototip pentru cercetări ulterioare bazate pe simulare în cadrul jocurilor repetate, inclusiv pentru cadrul Monte Carlo dezvoltat în această teză.

Rezultate și discuții

Simulările extinse de tip turneu oferă o perspectivă multifactorială asupra influenței exercitate de incertitudinea temporală și de selecția evolutivă asupra performanței strategice. Printr-o comparație sistematică între scenariile cu orizont determinist și cele cu terminare stocastică, această secțiune furnizează o evaluare integrată a robusteții, volatilității și capacității adaptative a strategiilor testate.

Un rezultat semnificativ al analizei arată că strategiile performante în condiții cu orizont determinist nu își mențin neapărat avantajul atunci când durata interacțiunii devine incertă. De exemplu, deși strategii reciproce simple, precum *Tit-for-Tat*, rămân competitive în ambele contexte, ierarhiile câștigurilor și tiparele comportamentale asociate suferă deviații măsurabile. Această observație indică faptul că incertitudinea temporală introduce o modificare structurală distinctă în peisajul strategic — una care privilegiază flexibilitatea și penalizează comportamentele excesiv de rigide sau exploatare.

Datele susțin, de asemenea, apariția unei tipologii a profilurilor strategice:

- **Strategii robuste**, caracterizate prin câștiguri medii ridicate și variație redusă, atât în condiții deterministe, cât și în cele stocastice.

- **Strategii adaptative**, care își ajustează comportamentul în funcție de schimbările temporale, menținând performanța prin decizii probabiliste sau mecanisme de iertare.
- **Strategii volatile**, sensibile la modificările orizontului de joc și adesea supra-specializate pentru un anumit context.

Indicatorii comportamentali relevanți—precum rapoartele între succese și eșecuri, frecvențele cooperării și variațiile între câștigurile obținute—au fost integrați în reprezentări matriciale vizuale, menite să evidențieze structuri strategice latente. Aceste instrumente analitice au permis identificarea unor asimetrii comportamentale emergente care rămân ascunse în analiza convențională, agregată, a performanțelor individuale. De exemplu, anumite strategii au demonstrat o predispoziție către cooperare în contexte cu orizont temporal determinat, dar au adoptat conduite predominant defectuoase în medii caracterizate de incertitudine privind finalul jocului. Această tranziție strategică sugerează o dependență marcată de transparența temporală a jocului pentru menținerea cooperării sau pentru exercitarea unor forme de control strategic asupra adversarilor. Astfel, predictibilitatea orizontului de joc devine un parametru critic în modelarea comportamentului rațional și în structurarea echilibrelor dinamice în jocurile repetate.

Dintr-o perspectivă evolutivă, strategiile care integrează componente stocastice sau reguli adaptative de complexitate redusă au demonstrat o durabilitate superioară în simulările ce au inclus mecanisme de feedback evolutiv. Această tendință evidențiază un avantaj adaptativ asociat plasticității comportamentale, în special în medii fără orizont temporal bine definit. Relevant este și faptul că introducerea unui orizont probabilist a suspendat logica inducției inverse, facilitând emergența spontană a cooperării condiționate și a încrederii strategice.

Discuția analizează și contribuțiile metodologice ale cadrului de simulare. Adaptarea modelului lui Axelrod pentru a include durate aleatorii ale meciurilor, repetiție automată și urmărirea extinsă a datelor a permis o analiză mult mai detaliată în comparație cu studiile originale. Aceste inovații au îmbunătățit înțelegerea modului în care incertitudinea structurală influențează succesul strategic, atât în ceea ce privește rezultatele imediate, cât și sustenabilitatea pe termen lung din punct de vedere evolutiv.

Rezultatele evidențiază necesitatea integrării incertitudinii în proiectarea și evaluarea strategiilor pentru jocuri repetate. Agenții de succes în medii cu ambiguitate temporală integrează, în mod tipic, reciprocitate, adaptabilitate și toleranță la zgomot.

Această concluzie are implicații semnificative pentru teoria jocurilor, economia comportamentală și dezvoltarea sistemelor multi-agent reziliente în domeniul inteligenței artificiale și al coordonării descentralizate.

4

Rezumatul Capitolului 4: Autocovarianța în modelul Allison Mixture – O analiză Monte Carlo a deciziei strategice în jocuri repetate

Introducere în modelul Allison Mixture și paradoxul lui Parrondo

Rezultatele prezentate în această secțiune se bazează pe fundamentele empirice introduse în articolul nostru *Autocovariance in the Allison Mixture: A Monte Carlo Analysis of Strategic Decision-Making in Repeated Games* (G. M. Pop et al., 2025), care a formulat motivațiile teoretice și a demonstrat aplicabilitatea modelului Allison Mixture printr-o serie de rezultate obținute prin simulare.

Această secțiune introduce modelul *Allison Mixture* ca un cadru probabilist pentru investigarea comportamentului strategic emergent în jocuri repetate desfășurate sub incertitudine. Modelul se află la intersecția dintre teoria proceselor stocastice și dinamica decizională în jocuri repetate, fiind inspirat din paradoxul lui Parrondo. Acesta arată că alternanța între strategii individual dezavantajoase poate conduce, în mod contraintuitiv, la un rezultat agregat favorabil. Aplicat în context strategic, modelul sugerează că succesiunea controlată a unor politici aparent ineficiente poate exploata instabilități dinamice și genera avantaje pe termen lung. Allison Mixture transpune această logică paradoxală în context strategic, ilustrând cum combinațiile adaptative de strategii neutre sau suboptime pot produce, la nivel agregat, performanțe superioare în termeni evolutivi sau comportamentali.

Dintr-o perspectivă teoretică, Allison Mixture contrazice accentul tradițional pus

pe optimizarea statică și raționalitatea de tip echilibru. În schimb, evidențiază modul în care tranzițiile probabiliste între moduri comportamentale—fiecare suboptim în mod izolat—pot produce comportamente robuste și sensibile la context. (Harmer & Abbott, 2002) Acest mecanism dinamic de comutare este modelat printr-un amestec de procese cu stare finită, guvernat de probabilități de tranziție care codifică gradul de incertitudine structurală și dependență de memorie în sistem.

Paradoxul lui Parrondo constituie logica fundamentală a acestui mecanism. Inițial derivat din studiul sistemelor fizice care prezintă comportament de tip „Brownian ratchet”, paradoxul a fost aplicat în diverse domenii, de la biologie evolutivă la economie. În contexte de tip joc strategic, acesta sugerează că alternarea strategică între reguli deterministe și stocastice poate genera rezultate imposibil de atins printr-o strategie pură. (Parrondo et al., 2000) Modelul Allison Mixture valorifică această intuiție prin incorporarea comutării stocastice în structura operațională a comportamentului strategic în jocuri repetate. Rezultatul constă în strategii capabile să se adapteze dinamic la variabilitatea contextului local, la reacțiile adversarului și la incertitudinea indusă de zgomotul informațional.

Relevanța acestei abordări rezidă în capacitatea sa de a modela raționalitatea limitată, adaptarea indusă de semnale și dependențele autocorelate ale comportamentului, fără a presupune acces la informație completă sau memorie nelimitată. Combinând logica probabilistă a sistemelor de tip Parrondo cu formalismul interacțiunilor strategice, modelul Allison Mixture constituie un cadru inovator pentru explorarea emergenței cooperării, coordonării și competiției în medii marcate de ambiguitate, asimetrie informațională și tranziții stocastice. Acest cadru teoretic fundamentează dezvoltarea unor metode de simulare menite să evalueze empiric comportamentele strategice în jocuri repetate complexe.

Modelul Allison Mixture

Această secțiune definește modelul Allison Mixture ca un cadru formal de modelare care permite alternanța controlată între două structuri strategice complementare. Esențial este faptul că această alternanță nu este de natură mecanică, ci este proiectată să inducă autocorelație în comportamentul agenților—reflectând o dependență statistică între deciziile succesive, ce imprimă o structură temporală coerentă asupra traiectoriei decizionale. (Gunn et al., 2014)

Modelul Allison Mixture este definit prin două proprietăți fundamentale. În primul rând, acesta integrează două strategii de bază într-o secvență comportamentală coerentă printr-un mecanism probabilist care determină tranzițiile dintre ele. (Harmer & Abbott, 1999) În al doilea rând, modelul generează secvențe comportamentale structurate, cu proprietăți care nu pot fi deduse din analiza individuală a strategiilor componente. Mecanismul de alternanță conduce la apariția unor tipare emergente, a căror structură este determinată de logica internă a comutării.

Secțiunea subliniază, de asemenea, că această construcție nu presupune existența unui algoritm de învățare, adaptare prin feedback sau acces la istoricul interacțiunilor. În schimb, corelația comportamentală emergentă este indusă direct de arhitectura internă a mecanismului de alternanță. Prin această structură, modelul Allison Mixture se dovedește adecvat pentru simulări menite să investigheze în mod controlat influența asimetriilor structurale interne asupra comportamentului strategic în jocuri repetate.

Aplicarea modelului Allison Mixture în jocul Entry - Deterrence

Această parte explorează aplicabilitatea modelului *Allison Mixture* în cadrul jocului de tip Entry-Deterrence, cu scopul de a evalua efectele alternanței probabiliste între strategii elementare asupra comportamentului de echilibru într-un mediu strategic caracterizat prin asimetrie. Configurația experimentală modelează interacțiunea dintre un Incumbent (actor deja prezent pe piață) și un potențial Entrant (agent cu intenție de intrare), într-un cadru de recompensare care amplifică tensiunile strategice dintre acțiunile de descurajare și cele de intrare efectivă. (G. M. Pop et al., 2025)

Rezultatele simulării se concentrează pe modul în care variațiile parametrului de comutare din cadrul modelului influențează eficiența de descurajare a jucătorului Incumbent. Analiza documentează fluctuații în structura câștigurilor, frecvența intrării și răspunsurile condiționate, confirmând că alternanța de tip Allison introduce tipare comportamentale sistematice. Aceste tipare diferă substanțial de cele generate de strategii pur aleatoare sau deterministe, sugerând că modelul încorporează dependențe structurale care se propagă de-a lungul interacțiunii.

De asemenea, secțiunea evidențiază rolul jocului Entry-Deterrence ca platformă de testare pentru investigarea expresivității strategice a modelului Allison Mixture în contexte nesimetrice. Prin aplicarea mecanismului de alternanță într-un cadru

cu payoff-uri asimetrice, analiza oferă perspective relevante asupra modului în care comutarea structurată intern poate influența anticipațiile adversarului, poate induce forme de nealiniere strategică sau poate genera efecte reputaționale emergente, chiar în absența unor mecanisme explicite de semnalizare.

Metodologia experimentală pentru simulările Monte Carlo aplicate modelului Allison Mixture

Această secțiune descrie metodologia de simulare utilizată pentru a analiza comportamentul strategic al agenților influențați de modelul Allison Mixture în cadrul jocurilor repetate. Simularea are la bază un cadru experimental de tip *Monte Carlo* care permite evaluarea proprietăților statistice și dinamice ale strategiilor bazate pe alternanță, în condiții variabile de interacțiune strategică. Sistemul implementat permite rulări repetate ale jocului între perechi de agenți, ale căror decizii sunt guvernate de mecanisme interne de comutare strategică.

Cadrul de simulare este implementat în `Python` și configurat pentru a permite experimente la scară largă, permițând extragerea unor concluzii statistice robuste prin eșantionare repetată. Comportamentul agenților este modelat prin sisteme de tranziție care reflectă dinamica de alternanță specificată de modelul Allison Mixture. Rezultatele obținute în urma simulărilor sunt sintetizate prin statistici rezumative, care includ câștigurile medii, secvențele decizionale și indicatori structurali ai tiparelor de corelație comportamentală.

Metodologia experimentală pentru simulările Monte Carlo aplicate modelului Allison Mixture

Această parte a lucrării prezintă procedura utilizată pentru rularea simulărilor Monte Carlo, având ca scop testarea comportamentului agenților construiți conform modelului Allison Mixture. Metodologia este concepută pentru a examina impactul regulilor interne de alternanță asupra dinamicii strategice a agenților în multiple scenarii simulate. Analiza urmărește identificarea trăsăturilor statistice emergente, atât la nivelul secvențelor decizionale, cât și al distribuțiilor de câștig asociate acestora.

Designul experimental se bazează pe o schemă de eșantionare aleatorie repetată, în care fiecare rundă de simulare constă într-un număr extins de iterații ale jocurilor repetate, desfășurate sub aceleași reguli de alternanță structurală. Agenții care

implementează modelul Allison Mixture comută între două strategii de bază, iar comportamentul emergent este monitorizat de-a lungul meciurilor cu durate variabile. Simulările astfel obținute generează distribuții robuste, din care pot fi extrase tendințe centrale, variații statistice și proprietăți de autocorelație. Comportamentul agenților este implementat printr-o logică programabilă, capabilă să aplice mecanismul de comutare specificat de modelul Allison Mixture. Starea internă a fiecărui agent evoluează în funcție de reguli de tranziție probabilistice, iar deciziile rezultate sunt înregistrate rundă cu rundă. Datele din simulare sunt apoi analizate pentru a identifica regularități precum tipare de fază, periodicitate și intervale de fluctuație.

Rezultatele experimentale obținute prin simulările Monte Carlo

Etapa metodologică descrisă în continuare prezintă constatările empirice obținute din simulările Monte Carlo care utilizează agenți ghidați de modelul Allison Mixture. Analiza se concentrează asupra metricilor statistice esențiale rezultate din experimente extinse cu jocuri repetate, având ca scop evaluarea modului în care comutarea probabilistă influențează structural comportamentul agenților și stabilitatea câștigurilor.

Rezultatele indică faptul că, într-o varietate largă de configurații parametrice, agenții guvernați de modelul Allison Mixture generează secvențe comportamentale caracterizate prin regularitate structurală, distincte de un comportament aleatoriu pur. Simulările relevă apariția autocorelației în tiparele de acțiune, sugerând că mecanismul intern de alternanță funcționează ca un substitut al memoriei, chiar și în absența unei componente explicite de memorie în strategiile de bază. (G. M. Pop et al., 2025)

Secțiunea prezintă, de asemenea, diagrame și statistici rezumative care evidențiază abateri sistematice ale comportamentului generat de modelul de amestec față de strategiile pur aleatorii sau complet deterministe. Aceste tipare sugerează că modelul reușește să integreze impredictibilitatea funcțională, menținând totodată o coerență comportamentală stabilă în timp. Rezultatele experimentale confirmă faptul că modelul Allison Mixture generează caracteristici comportamentale distincte, cum ar fi autocorelația, consistența câștigurilor și adaptabilitatea în funcție de mediu—attribute care îl diferențiază de strategiile tradiționale din scenariile de joc repetat.

5

Concluzii generale și direcții viitoare de cercetare

Această teză oferă un cadru analitic cuprinzător pentru examinarea comportamentului strategic în jocuri repetate, în contexte caracterizate prin diferite grade de incertitudine asupra duratei interacțiunii. Lucrarea își propune să conecteze principiile clasice ale teoriei jocurilor cu perspective contemporane asupra deciziei bazate pe agenți și robusteții evolutive, prin integrarea fundamentelor teoretice formale cu modelarea algoritmică și simularea de tip Monte Carlo.

Primul capitol a furnizat cadrul conceptual și formal necesar pentru replicarea interacțiunii strategice în medii repetitive. Bazându-se pe lucrări fundamentale din teoria jocurilor non-cooperative, capitolul a subliniat importanța echilibrului Nash, a strategiilor mixte și a utilității așteptate ca instrumente centrale pentru modelarea comportamentului rațional. Aplicarea acestor concepte în medii dinamice, precum Dilema Prizonierului Iterată (IPD), a arătat că anticiparea rezultatelor viitoare modifică semnificativ logica decizională strategică. Folk Theorem demonstrează că, în jocurile repetate pe termen nelimitat, cooperarea poate apărea ca echilibru stabil, cu condiția ca agenții să fie suficient de orientați spre viitor.

Investigația asupra metricilor de discrepanță oferă o bază principală pentru identificarea strategiilor structurale solide, facilitând nu doar compararea performanței în condiții stabile, ci și evaluarea flexibilității strategice în contexte cu incertitudine temporală. Tranziția de la jocuri cu orizont determinist la jocuri cu orizont incert evidențiază faptul că reziliența comportamentală nu constituie o proprietate intrinsec superioară, ci depinde de capacitatea unei strategii de a-și menține performanța în condiții variabile. Această concluzie susține ideea conform căreia sustenabilitatea deciziilor strategice pe termen lung este strâns condiționată de sensibilitatea față de structura temporală a interacțiunii.

Înțelegerea mai profundă a adaptării agenților în condiții de incertitudine este

posibilă prin integrarea măsurătorilor de performanță structurală în analiza interacțiunilor recurente. Această abordare nu doar interpretează rezultatele simulărilor, ci fundamentează un cadru principal pentru proiectarea agenților adaptivi capabili să își reconfigureze strategiile ca răspuns la volatilitate, asimetrie și dinamica jocului. Aceste constatări fundamentează extinderea cercetării către designul deciziilor robuste, dinamica evolutivă a strategiilor și algoritmi de învățare adaptivă (reinforcement learning).

Modelele teoretice dezvoltate au fost investigate empiric în al doilea capitol, prin experimente bazate pe turneele celebre organizate de Axelrod. Simulările comparative ale performanței strategiilor în contexte cu număr cunoscut și necunoscut de runde au evidențiat diferențe comportamentale esențiale. În anumite cazuri, variante mai adaptative sau iertătoare au depășit strategii precum Tit-for-Tat, care s-au dovedit constante în medii deterministe, dar mai puțin eficiente în condiții de incertitudine. În contrast, strategiile rigide, fără componente probabilistice, au înregistrat performanțe slabe în condiții de incertitudine privind orizontul temporal. Utilizarea metricilor de discrepanță—precum volatilitatea câștigurilor și variațiile în raportul victorii-înfrângeri— a permis identificarea agenților sensibili la modificări structurale și a celor cu performanță stabilă în toate condițiile.

Analiza a demonstrat că agenții cu comportament determinist pot manifesta sensibilitate emergentă la durata jocului, chiar și în absența unui raționament temporal explicit. Această observație susține ideea conform căreia caracteristicile structurale ale interacțiunii repetate pot genera rezultate complexe în comportamentul agenților, mai ales atunci când aceștia se bazează pe cooperare condiționată sau răspunsuri bazate pe memorie.

Lucrarea subliniază, de asemenea, importanța construirii unor cadre de evaluare care să reflecte natura stocastică a proceselor decizionale reale. Contrastul dintre simulările cu orizont fix și cele cu orizont incert evidențiază modul în care ipotezele experimentale rigide pot distorsiona înțelegerea comportamentului strategic optim.

Un alt rezultat notabil constă în evidențierea potențialului metodologiilor hibride, care integrează sisteme adaptive de învățare — capabile să gestioneze zgomotul informațional, ambiguitatea intențiilor și variațiile neanticipate ale mediului — cu robustețea euristicilor reactive. Astfel de arhitecturi hibride combină capacitatea de colaborare strategică cu mecanisme eficiente de apărare, oferind modele mai realiste ale procesului decizional în medii conflictuale sau incerte.

Această cercetare reconfirmă valoarea educațională și metodologică a turneele simulate ca instrumente de evaluare a interacțiunii strategice. Inspirare de designul

original al lui Axelrod, astfel de competiții rămân relevante ca medii experimentale controlate în care ipotezele despre comportamentul agenților—umani sau artificiali—pot fi testate, rafinate și generalizate.

Rezultatele indică faptul că gestionarea incertitudinii în contexte strategice necesită mai mult decât sofisticare algoritmică; ea presupune o înțelegere conceptuală profundă a trăsăturilor comportamentale fundamentale — precum adaptabilitatea, iertarea și capacitatea de învățare — care susțin succesul pe termen lung. Această constatare subliniază că, deși predictibilitatea poate fi avantajoasă în medii deterministe, ea devine o vulnerabilitate în contexte volatile, în care flexibilitatea decizională este esențială. Agenții care urmează rigid reguli prestabilite pot obține performanțe satisfăcătoare în medii idealizate, dar tind să eșueze în fața complexității și variabilității inerente interacțiunilor reale.

Proiectarea eficientă a strategiilor viitoare ar trebui să ia în considerare relația dinamică dintre profunzimea memoriei și variabilitatea răspunsului. Agenții cu memorie redusă pot reacționa prompt la schimbări locale, însă riscă să ignore regularitățile pe termen lung; în schimb, cei cu memorie extinsă și capacitate evaluativă pot identifica tipare comportamentale recurente și ajusta deciziile în mod contextualizat. Echilibrarea adaptivă a acestui compromis — între reactivitate imediată și integrare istorică — reprezintă o direcție promițătoare pentru dezvoltarea unor arhitecturi decizionale mai flexibile, context-sensibile și robuste la variații ambientale.

Capitolul 4 s-a concentrat pe investigarea formulei Allison Mixture, pornind de la premisa teoretică potrivit căreia, deși elegantă în structura sa probabilistă, aceasta necesită o validare empirică sistematică în medii complexe și dinamice. S-a acordat o atenție specială contextelor cu tranziții stocastice, unde incertitudinea inherentă provoacă ipotezele tradiționale de modelare. Obiectivul principal a fost evaluarea viabilității formulei în cadrul simulărilor Monte Carlo și explorarea potențialului său pentru dezvoltarea unor agenți inteligenți capabili să anticipeze și să răspundă comportamentului strategic al adversarilor în interacțiuni repetate.

Pentru realizarea acestui obiectiv, a fost construit un design experimental solid, în care simulările au fost rulate pe o gamă largă de parametri și orizonturi de interacțiune. Pentru a asigura fiabilitatea statistică a rezultatelor, fiecare configurație de joc a fost rulată de sute de ori, reducând impactul fluctuațiilor aleatorii și îmbunătățind acuratețea măsurărilor de performanță. Această abordare a permis estimarea consistentă a indicatorilor cheie, precum media și abaterea standard a autocovarianței și autocorelației în diverse regimuri.

Dincolo de validarea formulei teoretice, analiza a relevat că modelul Allison Mixture deține un potențial considerabil pentru dezvoltarea agenților strategici adaptivi. În medii în care tehnicile convenționale eșuează adesea în captarea interdependențelor subtile sau a semnalelor comportamentale evolutive, Allison Mixture facilitează apariția unor structuri de ordin superior prin utilizarea aleatorietății controlate. Capacitatea sa de a sintetiza tipare prin alternanță probabilistă îl poziționează ca un instrument valoros pentru modelarea și navigarea incertitudinii în sisteme multi-agent. În această direcție, studiul aduce o contribuție duală — metodologică și conceptuală — la literatura privind modelarea strategică, oferind totodată o bază empirică solidă pentru aplicații viitoare în inteligența artificială și teoria jocurilor computaționale.

Aplicarea metodelor Monte Carlo pe parcursul întregii lucrări s-a dovedit crucială pentru gestionarea caracterului stocastic și a complexității dimensionale ridicate caracteristice jocurilor repetate. Această abordare computațională a permis estimarea valorilor așteptate și analiza volatilității strategice prin rularea extensivă a scenariilor de interacțiune între agenți, desfășurate în contexte parametrice variabile. Constatările evidențiază că simularea nu reprezintă doar un instrument ilustrativ, ci un mijloc indispensabil de explorare computațională în studiul dinamicii strategice, mai ales atunci când soluțiile analitice devin intractabile.

Rigoarea metodologică ce stă la baza simulărilor Allison Mixture—bazată pe iterație la scară largă și calcul de înaltă performanță—garantează că rezultatele observate nu sunt artefacte ale zgomotului, ci reflecții statistic fundamentate ale dinamicii modelate, consolidând astfel validitatea empirică a abordării.

În ciuda rezultatelor încurajatoare, investigația a evidențiat limite privind generalizabilitatea modelului Allison Mixture. Performanța formulei s-a dovedit sensibilă la caracteristicile structurale ale jocului de bază, ceea ce limitează aplicabilitatea imediată în contexte strategice diverse. Deși multe scenarii s-au aliniat predicțiilor teoretice, absența rezultatelor consecvente în toate condițiile indică necesitatea unei explorări suplimentare. Viitoare studiile ar trebui să exploreze variabilitatea comportamentală a modelului în funcție de setările inițiale și de diversitatea arhitecturii jocurilor, pentru a clarifica domeniul său de validitate.

Rezultatele indică faptul că modelul Allison Mixture are potențialul de a contribui la îmbunătățirea procesului decizional adaptiv în medii dinamice, oferind o direcție promițătoare pentru dezvoltarea unor strategii capabile să anticipeze și să răspundă eficient la tipare complexe de interacțiune.

În ansamblu, teza aduce contribuții semnificative pe trei direcții esențiale în

înțelegerea jocurilor repetate: (1) oferă o demonstrație a modului în care cooperarea și complexitatea strategică se pot manifesta chiar și în absența ipotezelor stricte privind raționalitatea perfectă și informația completă, consolidând relevanța teoriei clasice; (2) dezvoltă un cadru de simulare destinat evaluării performanței agenților în contexte cu orizonturi deterministe și stocastice; și (3) introduce o metodologie pentru analiza robusteții structurale și a adaptabilității strategice, fundamentată pe date empirice generate în cadrul interacțiunilor simulate.

Una dintre contribuțiile majore ale acestei cercetări constă în integrarea formală a incertitudinii în evaluarea strategică a jocurilor repetate — un aspect adesea neglijat în analiza tradițională. Prin dezvoltarea și aplicarea unui cadru robust de simulare bazat pe metode Monte Carlo, atât în turneele de tip Axelrod, cât și în modelul Allison Mixture, studiul evidențiază modul în care structurile probabilistice ale orizontului temporal transformă comportamentul agenților, expun vulnerabilități și relevă capacități adaptative emergente. Această abordare duală consolidează conexiunea dintre teoria experimentală a jocurilor și modelarea stocastică, oferind totodată o metodologie replicabilă pentru testarea robusteții strategiilor în condiții de incertitudine realistă.

Cadrul propus introduce un protocol flexibil de simulare, capabil să acomodeze atât structuri deterministe, cât și stocastice, permițând evaluări mai nuanțate și sensibile la context ale performanței agenților. În cele din urmă, cercetarea deschide noi direcții pentru proiectarea unor agenți inteligenți capabili să învețe, să se adapteze și să mențină cooperarea în medii dinamice și incerte.

Prezenta lucrare propune un cadru experimental unificat pentru compararea sistematică a comportamentului agenților în funcție de natura orizontului temporal — determinist sau stocastic — oferind perspective noi asupra adaptabilității strategice și volatilității decizionale. Extins atât asupra strategiilor clasice testate în turneele Axelrod, cât și asupra dinamicii statistice a modelului Allison Mixture, acest cadru furnizează o bază metodologică pentru analizarea modului în care incertitudinea afectează structura stimulentei și rezultatele de coordonare. Această dublă aplicare evidențiază originalitatea abordării, care se desprinde de analiza statică a echilibrului și subliniază influența trăsăturilor structurale ale interacțiunii — precum incertitudinea privind numărul de rundă — asupra apariției cooperării sustenabile, a modificării graduale a comportamentului și a tendinței către tipare comportamentale persistente.

Bibliografie Selectivă

- Abbott, D. (2010). Asymmetry and disorder: A decade of parrondo's paradox. *Fluctuation and Noise Letters*, 9(1), 129–156.
- Abreu, D., & Pearce, D. (2007). Bargaining, reputation, and equilibrium selection in repeated games with contracts. *Econometrica*, 75(3), 653–710. <https://doi.org/10.1111/j.1468-0262.2007.00762.x>
- Akin, E. (2016). The iterated prisoner's dilemma: Good strategies and their dynamics. In A. Idris (Ed.), *Ergodic theory* (pp. 77–107). De Gruyter.
- Allison, A., Pearce, C., & Abbott, D. (2007). Finding keywords amongst noise: Automatic text classification without parsing. *Proceedings of SPIE - The International Society for Optical Engineering*, 6601. <https://doi.org/10.1117/12.724655>
- Asmussen, S., & Glynn, P. W. (2007). *Stochastic simulation: Algorithms and analysis* (Vol. 57). Springer. <https://doi.org/10.1007/978-0-387-69033-4>
- Axelrod, R. (1984). *The evolution of cooperation*. Basic Books.
- Axelrod, R., & Hamilton, W. D. (1981). The evolution of cooperation. *Science*, 211(4489), 1390–1396. <https://doi.org/10.1126/science.7466396>
- Ben-Porath, E. (1990). The complexity of computing a best response automaton in repeated games with mixed strategies. *Games and Economic Behavior*, 2(1), 1–12.
- Binmore, K. (1992). *Fun and games: A text on game theory*. D. C. Heath; Company.
- Bó, P. D. (2002). Cooperation under the shadow of the future: Experimental evidence from infinitely repeated games. *Working Paper, Brown University*, No. 2002-21.
- Brunauer, R., Löcker, A., Mayer, H. A., Mitterlechner, G., & Payer, H. (2007). Evolution of iterated prisoner's dilemma strategies with different history lengths in static and cultural environments. *Proceedings of the 2007 ACM Symposium on Applied Computing (SAC)*, 2058–2065. <https://doi.org/10.1145/1244002.1244163>
- Bukhori, S., Hariadi, M., Purnama, I. K. E., & Purnomo, M. H. (2013). Parrondo's paradox based strategies in the serious game of rtgs using forest fire model. *The International Arab Journal of Information Technology*, 10(5), 428–435.
- Chen, B., & Zadrozny, P. A. (2002). An anticipative feedback solution for the infinite-horizon, linear-quadratic, dynamic, stackelberg game. *Journal of Economic*

- Dynamics and Control*, 26(9–10), 1397–1416. [https://doi.org/10.1016/S0165-1889\(01\)00077-X](https://doi.org/10.1016/S0165-1889(01)00077-X)
- Cheong, K. H. (2019). Adaptive therapy and parrondo's paradox: A model-based perspective. *BioSystems*, 178, 10–17.
- Cressman, R. (1996). Evolutionary stability in the finitely repeated prisoner's dilemma game. *Journal of Economic Theory*, 68, 234–248.
- Doucet, A., de Freitas, N., & Gordon, N. (Eds.). (2001). *Sequential monte carlo methods in practice*. Springer. <https://doi.org/10.1007/978-1-4757-3437-9>
- Eckhardt, R. (1987). Stan ulam, john von neumann, and the monte carlo method. *Los Alamos Science*, 15, 131–137.
- Floridi, L. (2023). *The ethics of artificial intelligence: Principles, challenges, and opportunities*. Oxford University Press.
- Fudenberg, D., & Maskin, E. (1986). The folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, 54(3), 533–554.
- Fudenberg, D., & Tirole, J. (1991). *Game theory*. MIT Press.
- Gunn, L. J., Allison, A., & Abbott, D. (2014). Allison mixtures: Where random digits obey thermodynamic principles. *Fluctuation and Noise Letters*, 13(03), 1450015. <https://doi.org/10.1142/S0219477514500152>
- Gunn, L. J., Chapeau-Blondeau, F., Allison, A., & Abbott, D. (2016). Towards an information-theoretic model of the allison mixture stochastic process. *Journal of Statistical Mechanics: Theory and Experiment*, 2016(5), 054041. <https://doi.org/10.1088/1742-5468/2016/05/054041>
- Gyurkovics, É., & Takács, T. (2005). Guaranteeing cost strategies for infinite horizon difference games with uncertain dynamics. *International Journal of Control*, 78(8), 587–599. <https://doi.org/10.1080/00207170500093978>
- Harmer, G. P., & Abbott, D. (1999). Parrondo's paradox. *Statistical Science*, 14(2), 206–213. <http://www.jstor.org/stable/2676740>
- Harmer, G. P., & Abbott, D. (2002). A review of parrondo's paradox. *Fluctuation and Noise Letters*, 2(02), R71–R107.
- Imhof, L. A., & Nowak, M. A. (2007). Stochastic evolutionary dynamics of direct reciprocity. *Proceedings of the Royal Society B: Biological Sciences*, 274(1621), 173–179. <https://doi.org/10.1098/rspb.2006.3720>
- Koh, J. M., & Cheong, K. H. (2020). Generalized solutions of parrondo's games. *Advanced Science*, 7(24), 2001126.
- Lesne, A. (2014). Shannon entropy: A rigorous notion at the crossroads between probability, information theory, dynamical systems and statistical physics.

- Mathematical Structures in Computer Science*, 24(3), e240311. <https://doi.org/10.1017/S0960129512000783>
- Leyton-Brown, K., & Shoham, Y. (2008). *Essentials of game theory: A concise multidisciplinary introduction*. Morgan; Claypool.
- Lindsten, F., & Schön, T. B. (2013). Backward simulation methods for monte carlo statistical inference. *Foundations and Trends® in Machine Learning*, 6(1), 1–143.
- Mailath, G. J., & Samuelson, L. (2006). *Repeated games and reputations: Long-run relationships*. Oxford University Press.
- Malook, S. (2024). The prisoner's versus pardoner's dilemmas: A juxtaposition of two strategic decision-game theoretic approaches in social sciences. *Journal of Social & Organizational Matters*, 3(3), 52–74. <https://doi.org/10.56976/jsom.v3i3.21>
- Metropolis, N., & Ulam, S. (1949). The monte carlo method. *Journal of the American Statistical Association*, 44(247), 335–341.
- Milencianu, M., & Pop, G. (2023). Axelrod first tournament: Examining certainty versus uncertainty about the end stage in repeated games. *Virgil Madgearu Review of Economic Studies and Research*, 16, 57–73. <https://doi.org/10.24193/RVM.2023.16.99>
- Mittal, V., & Huang, Y.-P. (2024). Parrondo's paradox in quantum walks with inhomogeneous coins. *arXiv preprint arXiv:2407.16558*. <https://arxiv.org/abs/2407.16558>
- Myerson, R. B. (1991). *Game theory: Analysis of conflict*. Harvard University Press.
- Nachbar, J. H. (2005). Beliefs and learning in repeated games. *Econometrica*, 73(2), 459–480. <https://doi.org/10.1111/j.1468-0262.2005.00578.x>
- Nash, J. (1950). Equilibrium points in n-person games. *Proc. Natl. Acad. Sci. USA*, 36, 48–49.
- Nowak, M. A., & Sigmund, K. (1993). A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, 364, 56–58. <https://doi.org/10.1038/364056a0>
- Parrondo, J. M. R., Harmer, G. P., & Abbott, D. (2000). New paradoxical games based on brownian ratchets. *Physical Review Letters*, 85(24), 5226–5229.
- Pop, G., Milencianu, M., & Pop, A. (2025). The second axelrod tournament: A monte carlo exploration of uncertainty about the number of rounds in iterated prisoner's dilemma. *Studia Universitatis Babeş-Bolyai Oeconomica*, 70(1), 67–82. <https://doi.org/10.2478/subboec-2025-0004>
- Pop, G. M., Tătar, F.-A., Vlădărean, M.-M., & Marian, C. G. (2025). Autocovariance in the allison mixture: A monte carlo analysis of strategic decision-making in

- repeated games. *Virgil Madgearu Review of Economic Studies and Research*, 18(1), 143–164. <https://doi.org/10.24193/RVM.2025.18.130>
- Rubinstein, A. (1998). *Modeling bounded rationality*. MIT Press. <https://mitpress.mit.edu/9780262681001/modeling-bounded-rationality/>
- Schelling, T. C. (1980). *The strategy of conflict*. Harvard University Press. <https://www.hup.harvard.edu/books/9780674840317>
- Stewart, A. J., & Plotkin, J. B. (2012). Extortion and cooperation in the prisoner's dilemma. *Proceedings of the National Academy of Sciences*, 109(26), 10134–10135. <https://doi.org/10.1073/pnas.1208087109>
- Stewart, A. J., & Plotkin, J. B. (2013). From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proceedings of the National Academy of Sciences*, 110(38), 15348–15353.
- von Neumann, J., & Morgenstern, O. (1974). *Theory of games and economic behavior* (3rd). Princeton University Press.
- Weinstein, J., & Yildiz, M. (2016). Reputation without commitment in finitely repeated games. *Theoretical Economics*, 11(1), 157–185.
- Yeung, D., & Petrosian, O. (2017). Infinite horizon dynamic games: A new approach via information updating. *International Game Theory Review*, 19(4), 1750026. <https://doi.org/10.1142/S0219198917500268>
- Yeung, D. W. K., & Petrosyan, L. A. (2017). Infinite horizon dynamic games: A new approach via information updating. *Dynamic Games and Applications*, 7(4), 581–608. <https://doi.org/10.1007/s13235-017-0212-7>