

UNIVERSITATEA BABEȘ-BOLYAI
FACULTATEA DE MATEMATICĂ ȘI INFORMATICĂ



Modele computaționale cu aplicații în domeniul sistemelor complexe

Rezumatul tezei de doctorat

Doctorand: Florentin BOTA
Conducător Științific: Prof. Dr. Horia F. POP

Cluj-Napoca 2021

Cuprinsul tezei de doctorat

Lista figurilor	4
Lista tabelelor	5
Lista publicațiilor	6
1 Introducere	8
2 Informații generale	17
2.1 Modele computaționale inspirate de natură	18
2.1.1 Modele bazate pe comportamentul coloniilor de viespi	19
2.1.2 Optimizarea bazată pe furnici și <i>Visual Agents</i>	20
2.2 Comportamentul uman în procesele economice	24
2.2.1 Contextul	24
2.2.2 Preferința temporală pentru agenții economici	27
2.2.3 Jocul ultimatumului.	31
2.3 Detectarea structurilor de comunități	33
2.4 Rețele neuronale	35
3 Noi modele computaționale bazate pe agenți	37
3.1 Algoritmi inspirați de coloniile de viespi în învățarea prin joc: Sotirios .	38
3.1.1 Metodologie	38
3.1.2 Modelul propus	40
3.1.3 Discuție și dezvoltări viitoare	43
3.2 Preferința temporală pentru agenții economici multidimensionali	44
3.2.1 Experimente privind preferința temporală	46
3.2.2 Metodologie	49
3.2.4 Concluzii și etape următoare	54
3.3 Integrarea comportamentului uman folosind agenți economici multidi-	
mensionali	55
3.3.1 Metodologie	55
3.3.2 Modelul propus	56
3.3.3 Rezultate experimentale	62
3.3.4 Concluzii	67
4 Structuri comunitare: un nou criteriu de fitness	68
4.1 Detectarea structurilor de comunități în rețelele multipartite: o nouă	
măsură de fitness	68
4.1.1 Metodologie	69

4.1.2	Soluția propusă	71
4.1.3	Rezultate experimentale	71
4.1.4	Concluzii	79
4.2	O analiză teoretică a rețelelor de autori în domeniul elaborării lucrărilor academice	80
4.2.1	Metodologie	80
4.2.2	Rezultate	84
4.2.3	Concluzii	86
5	Noi aspecte educaționale și de calitate a codului	88
5.1	Un cadru conceptual pentru predicția defectelor software prin utilizarea rețelelor neuronale	88
5.1.1	Metodologie	90
5.1.2	Modelul propus și experimente	92
5.1.3	Rezultate și validare	99
5.1.4	Concluzii și lucrări viitoare	99
5.2	Aspecte educaționale în testarea programelor utilizând platforme colab- orative	100
5.2.1	Obiective educaționale	101
5.2.2	Metodologia cercetării	103
5.2.3	Discuție	104
5.2.4	Concluzii	109
5.3	Componenta emoțională umană în contextul COVID-19	110
6	Concluzii și dezvoltări viitoare	113
	Bibliografie	115

Capitolul 1

Introducere

Înțelegerea naturii umane este un subiect fascinant ce i-a provocat pe filosofi, istorici și oamenii de știință încă de la începuturile timpului. De obicei, ne gândim la noi înșine ca având nevoi și dorințe de bază, însă interacțiunile noastre sunt unele dintre cele mai dinamice procese pe care le-am experimentat până acum [GDDG⁺03]. Atât liberul arbitru, cât și constrângerile naturale produc rezultate neașteptate, ce au crescut în complexitate de-a lungul timpului.

Teza mea de doctorat „Modele computaționale cu aplicații în domeniul sistemelor complexe” își propune să aducă contribuții în acest domeniu, cu accent pe modelarea comportamentului uman în sisteme complexe precum economia. De asemenea, am realizat contribuții în discipline precum educația, sisteme multi-agent și învățare automată.

În [CCC⁺18] Calder et al. aseamănă modelele computaționale niște instrumente ce ne pot ajuta să transformăm observațiile în predicții pentru evenimente viitoare, să experimentăm teorii și să înțelegem semnificația rezultatelor obținute.

Simplu spus, în modelarea computațională, modelele matematice sunt utilizate pentru a simula și analiza sisteme complexe cu ajutorul informaticii [TTT96]. Un sistem poate fi definit ca un grup de elemente (interdependente) ce interacționează și care formează un întreg. În acest context, un sistem poate fi definit de limitele sale, structura internă și funcția declarată sau scopul [Bac00, Boc10].

Astfel de sisteme, ce pot fi naturale sau create de către om, fac obiectul studiului interdisciplinar al *teoriei sistemelor*. Sistemele complexe prezintă proprietăți specifice, cum ar fi *emergența, neliniaritatea și auto-organizarea*. Termenul este de asemenea utilizat pentru a descrie studiul unor astfel de structuri, interacțiunile și comportamentul colectiv al acestora.

Sistemele complexe reprezintă un subset ce este considerat dificil de modelat și compus dintr-o multitudine de elemente care interacționează între ele [Har87]. Economia, clima, internetul [Par05], organismele și universul în sine sunt toate exemple de sisteme complexe.

Este important să înțelegem diferența dintre complex și complicat. Un mecanism complicat (precum un motor de mașină de exemplu) este extrem de previzibil, chiar dacă descrierea tehnică este dificilă. Acest lucru se datorează faptului că relațiile componentelor sunt previzibile. Un sistem complex este exact opusul, unde acțiuni ne semnificative pot avea consecințe neliniare. În rândurile următoare voi oferi o scurtă analiză a studiului complexității și a modului în care rezultatele din cadrul tezei mele se integrează în acest cadru.

Pornind de la știința sistemelor și cibernetică, studiul sistemelor complexe a evoluat pe parcursul a câtorva decenii. În 2010, Hiroki Sayama a propus o structură organizațională ce împarte sistemele complexe în șapte subgrupuri: *teoria jocurilor, comportament colectiv, rețele, evoluția și adaptarea, formarea modelelor, teoria sistemelor și dinamica neliniară*.

Acesta reprezintă cadrul în care vom începe investigația noastră, un mediu în care, așa cum a menționat Holland în 1995, *adaptarea în sine creează complexitate* [Hol96].

Probleme abordate

În cadrul acestei teze, am identificat mai multe obiective pe care le-am urmărit pe parcursul cercetării:

- Evaluarea modelelor computaționale inspirate de natură
- Investigarea comportamentului uman și a interacțiunilor sociale.
- Dezvoltarea unor metode noi de determinare a comunităților ascunse.
- Examinarea mecanismelor de evoluție ale agenților
- Utilizarea algoritmilor stocastici pentru a simula comportamentul imprevizibil.
- Analizarea componentei de corectitudine a naturii umane
- Testarea agenților dezvoltați și comparația rezultatelor cu datele experimentale.
- Dezvoltarea unor noi instrumente educaționale

Una dintre primele teme pe care le-am analizat a fost dezvoltarea unui model alternativ pentru așa-numitul “Homo Economicus” [MT00, Tha00]. Acești agenți, denumiți modelul standard în economie, iau decizii perfect raționale, nu au impulsuri de moment și prezintă o înțelegere absolută a mediului în care interacționează.

Modelul nou propus se bazează pe comportamentul uman, iar mai multe aspecte ale acestuia pot fi găsite în *Economia comportamentală* [MT00]. Am sugerat noi aspecte ce ar putea înlocui sau îmbunătăți sau înlocui modelul economic standard și care permit agenților să ia decizii impulsive, iraționale sau influențate social.

Structura tezei

Această lucrare este împărțită în șase capitole, după cum urmează:

- **Capitolul 1:** Am introdus contextul investigației noastre și temele de cercetare propuse. Această secțiune oferă o descriere a studiului sistemelor complexe.
- **Capitolul 2:** Analiza modelelor computaționale inspirate de natură și detectarea structurilor de comunități. De asemenea, este descris primul proiect: *Visual Agents*, pe care l-am folosit pentru a simula o colonie de furnici și pentru a demonstra proprietățile cheie ale unui sistem complex [BD15]. În acest capitol am prezentat de asemenea studii existente privind comportamentul uman în procesele economice și concepte teoretice legate de învățarea automată.
- **Capitolul 3:** Acest capitol prezintă investigația modelelor computaționale bazate pe agenți, inclusiv un sistem multi-agent adaptiv inspirat de comportamentul coloniilor de viespi, alături de conceptul propus, de agenți economici multidimensionali. Am realizat dezvoltarea agenților economici cu o distribuție realistă a componentei de echitate, prezentă în multe procese economice. Aceste rezultate au fost publicate în mai multe lucrări [SB18, BS21, Bot21b].
- **Capitolul 4:** În acest capitol este prezentată munca privind detectarea structurilor de comunități. Am conceput o nouă metodă de detectare a acestora în rețele multipartite, care ia în considerare relațiile de tip “*shadow*” dintre nodurile ce au un vecin comun. Am propus și validat o funcție de fitness care ia în considerare aceste relații și poate fi utilizată pentru a identifica populațiile din rețele sintetice, precum și cu aplicație în context real, pentru seturi de date masive. În plus, am descris modul în care am folosit teoria jocurilor pentru a identifica rețelele de coautor în elaborarea lucrărilor științifice. Rezultatele noastre au fost publicate în două articole [GBSL17, GBSL20].
- **Capitolul 5:** Pe parcursul acestui capitol am explorat modelul propus de învățare automată pentru predicția defectelor software și care a fost construit folosind o combinație de metrici CK. Apoi am prezentat utilizarea instrumentelor și activităților colaborative în cadrul cursului de Verificare și Validare a Sistemelor Software (VVSS) [CCBP21, SB20]. În plus, am discutat despre activitatea în curs de desfășurare privind gestionarea emoțiilor pentru adolescenții afectați de pandemia COVID-19.
- **Capitolul 6:** Acest capitol rezumă concluziile studiului curent și identifică cele mai relevante direcții pentru investigații viitoare.

Contribuții originale

Obiectivul principal al tezei mele este de a dezvolta noi modele computaționale și algoritmi ce pot fi utilizați pentru o înțelegere și predicție mai bună a comportamentului agenților în procesele complexe din viața reală. Principalele contribuții se concentrează în următoarele direcții.

1. Din punct de vedere economic, am conceput și implementat un nou agent economic multidimensional (MEA).
2. Am folosit Teoria Jocurilor și Algoritmi Genetici (GA) pentru a detecta structuri de comunități.
3. Din perspectivă tehnică, am dezvoltat instrumente web și 3D ce pot fi utilizate ca resurse educaționale sau pentru a ajuta la recuperarea sănătății mintale.
4. În plus, rezultatele au fost utilizate pentru a îmbunătăți componente educaționale legate de învățarea colaborativă și calitatea codului.

Contribuțiile sunt descrise în capitolele 3,4 și 5 și au fost publicate în opt lucrări de cercetare [GBSL17, SB18, GBSL20, Bot21b, BS21, CCBP21, BD15, SB20]:

- Am creat o platformă experimentală ce modelează procesele naturale dintr-o colonie de furnici, pentru a indica felul în care algoritmul de optimizare bazat pe comportamentul coloniilor de furnici (ACO) poate fi utilizat pentru a demonstra principiul emergenței într-un sistem complex [BD15]. Ca și aplicație am dezvoltat framework-ul *Visual Agents*, care folosește componente 3D și realitatea augmentată pentru a simula o colonie de furnici într-un mod realist. Acest proiect a reprezentat primul pas în investigația noastră.
- Am propus un nou model de alocare al quest-urilor bazat pe comportamentul coloniilor de viespi [SB18] pentru jocuri virtuale educaționale. Modelul se bazează pe ierarhia de dominație a viespilelor pentru a construi o componentă nouă de învățare colaborativă în jocul video Sotirios [Bot12], unde fiecare participant are un proces unic de învățare, în funcție de performanța sa. Am analizat comportamentul de viespe ca și soluție pentru menținerea unui joc multiplayer echilibrat și pentru a maximiza performanța echipei în interiorul jocului, cu rezultate promițătoare.
- În studiile noastre, am remarcat semnificația dimensiunii sociale și am folosit teoria jocurilor cooperative pentru a detecta rețelele de coautori în mediul academic de elaborare al lucrărilor științifice. Această metodă poate fi utilizată cu succes pentru a identifica autori interdisciplinari proeminenți și modele de publicare din

comunitatea științifică prin calcularea contribuției marginale a autorilor la efortul colectiv [GBSL20].

- Am dezvoltat o nouă soluție pentru problema detectării structurilor comunitare în rețelele multipartite, luând în considerare legăturile de tip “*shadow*” dintre nodurile care au un vecin comun. Am propus o funcție de fitness care ia în considerare aceste relații și a demonstrat că poate fi utilizată pentru a clasifica populațiile relevante într-un context de date sintetice, precum și într-o aplicație din lumea reală [GBSL17].
- Am studiat fenomenul de preferințe temporale regăsit în componenta Rațională a agenților noștri economici multidimensionali și am evaluat soluțiile existente din literatura științifică. Am propus apoi o nouă abordare hibridă, în care rezultatele comparative (pe baza datelor experimentale) au arătat că metoda noastră funcționează mai bine în majoritatea cazurilor [Bot21b].
- Am prezentat un nou algoritm evolutiv care a demonstrat viabilitatea dezvoltării de agenți inteligenți cu comportament uman. Am creat cu succes agenți economici cu o distribuție realistă a aspectului de echitate găsit în procesele economice. Datele experimentale ce utilizează agenți umani și sintetici au fost folosite pentru a testa modelul și am integrat o rețea neuronală în acest context, cu rezultate promițătoare [BS21]. Ca și aplicație am dezvoltat două versiuni distincte de *Ultimatum Game Simulator*.
- Am proiectat un model de predicție al erorilor software, folosind metode de învățare automată. Principala noastră provocare a fost să definim un model de rețea neuronală bazat pe combinații de metrici software CK [SB20] și pentru a stabili care dintre parametrii sporesc acuratețea predicției defectelor. Am descoperit un set de măsurători performante și am obținut rezultate pozitive prin această abordare.
- Am analizat utilizarea instrumentelor și activităților colaborative în cadrul materiei de verificare și validare a sistemelor software (VVSS) [CCBP21]. Această strategie a fost propusă înainte de pandemia de COVID-19 și s-a dovedit a fi o abordare eficientă atunci când studenților li s-a cerut să se implice în noul mediu online. Rezultatele noastre au arătat că elevii au preferat activitățile de colaborare, iar notele lor s-au îmbunătățit semnificativ.
- Am publicat o carte despre eficiența în dezvoltarea aplicațiilor software [Bot21a]. Am folosit cunoștințele pe care le-am dobândit prin cercetarea noastră despre natura umană pentru a determina cele mai eficiente abordări pentru a motiva studenții în contextul disciplinelor de Informatică.

Sunt optimist că modelele propuse sunt relevante pentru domeniile discutate în acest studiu și că vor contribui la o cunoaștere mai bună atunci când modelăm comportamente complexe.

Cuvinte cheie: modele computaționale, agenți multidimensionali, sisteme complexe, teoria jocurilor, algoritmi evolutivi, jocul ultimatumului, învățare automată, jocuri educaționale

Aș dori să imi exprim recunoștința față de Prof. dr. Horia F. Pop, pentru coordonarea științifică din cadrul studiilor de doctorat și față de soția mea, Alexandra și fiica noastră, Eva, pentru sprijinul lor continuu.

Aprecierea mea sinceră se îndreaptă de asemenea și către colegii cu care am colaborat: Noémi Gaskó, Mihai Suciu, Rodica Ioana Lung, Dana Simian, Camelia Șerban, Camelia Chisăliță-Crețu și Andreea Pop. Le mulțumesc foarte mult colegilor mei de la CodexWorks pentru sprijinul important oferit de-a lungul timpului.

Mulțumesc Susan Shapiro, Dr. Sally Baas de la Universitatea Concordia, St Paul și Nadia Crișan de la Universitatea Princeton pentru includerea contribuțiilor mele în proiectul *6 Feet Apart*.

In memoriam prof. Dumitru Dan Dumitrescu, cel care a făcut această cercetare posibilă.

Capitolul 2

Informații generale

Modelarea computațională poate fi definită ca studiul sistemelor complexe prin utilizarea informaticii și a altor discipline precum matematică, fizică și statistică. Prin efectuarea simulărilor computerizate, putem utiliza ieșirea unui sistem de calcul C pentru a descrie comportamentul unui alt sistem S în anumite condiții [Pic07].

În acest capitol am descris contextul unor astfel de modele inspirate de natură, incluzând coloniile de furnici, comportamentul coloniilor de viespi, comunități și comportamentul uman în cadrul proceselor economice.

2.1 Modele computaționale inspirate de natură

Modelele bazate pe agenți (ABM) [NH11] pot fi utilizate pentru a studia agenții autonomi și interacțiunile acestora. Originile teoretice se regăsesc în așa-numita *mașină Von Neumann*, un concept de auto-replicare propus în 1940 de John von Neumann [Pes95, NB⁺66] ce formează baza *automatelor celulare*. Teoria centrală este că legile de bază pot produce un comportament complex și că putem înțelege fenomene complexe prin studiul interacțiunilor dintre componente [GR13].

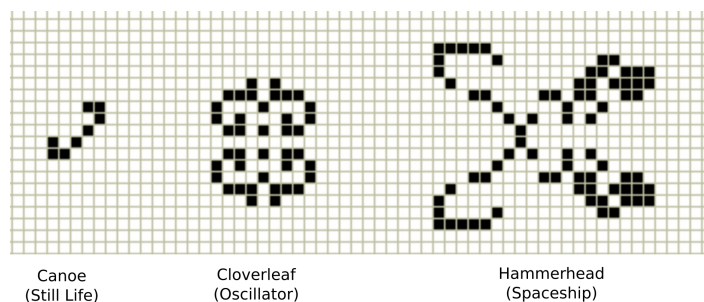


Figura 2.1: Modele stabile în “Jocul vieții” al lui Conway [MSJM16]

Un cunoscut automat celular este “*Jocul vieții*” al lui Conway, dezvoltat la sfârșitul anilor 1960 și bazat pe o matrice simplă de celule bi-dimensionale (Figura 2.1).

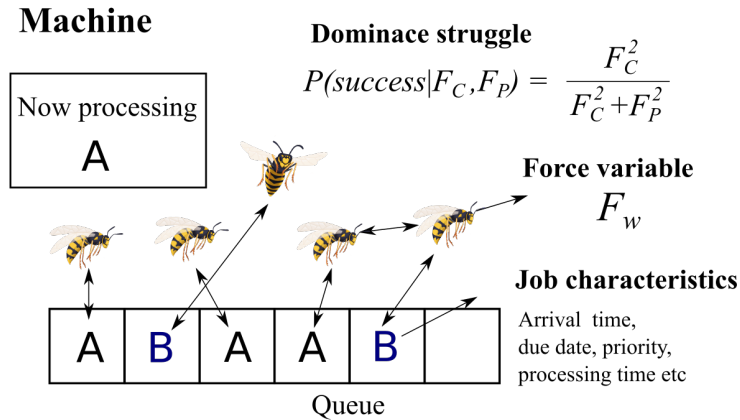


Figura 2.2: Planificarea bazată pe viespi, Cicirello și Smith, 2001 [CS01]

2.1.1 Modele bazate pe comportamentul coloniilor de viespi

Modelul de autoorganizare într-o colonie de viespi a fost propus de Theraulaz et al. în [TBGD91]. Comportamentul de viespe se caracterizează prin interacțiunea dintre indivizi și mediu folosind un mecanism de răspuns-stimul distribuit ce controlează alocarea sarcinilor.

După cum am menționat în [SB18], algoritmi care simulează comportamentul coloniilor de viespi au fost folosiți pentru a crea structuri multi-agent ce rezolvă probleme complexe: roboți autonomi [TBGD91], distribuirea resurselor și coordonarea într-un mediu de producție [CS04], distribuirea sarcinilor și proiectelor în medii de învățare virtuale [SSMP08], probleme de optimizare [SSS10], probleme de planificare dinamică [SM11, Sim09], extragerea cunoștințelor [DC10], etc. Aceste probleme pot fi asimilate problemei de alocare a sarcinilor în sistemele distribuite propusă de Cicirello în [CS04]. Un concurs de dominație poate apărea dacă două sau mai multe viespi concurează pentru aceeași sarcină cu aceeași probabilitate, așa cum se vede în Figura 2.2. Ierarhia dominanței se bazează pe o funcție de „forță” care este invers proporțională cu probabilitatea de dominație [CS04]. Au fost dezvoltate diferite modele pentru actualizarea pragurilor de răspuns. În [BTD98] pragul de răspuns rămâne fix în timp ce în [TBD98] este prezentată ideea de consolidare a pragului de răspuns [SB18].

2.1.2 Optimizarea bazată pe furnici și *Visual Agents*

Marco Dorigo a propus modelul coloniilor de furnici [Dor92] (ACO) în care agenții sunt denumiți furnici artificiale iar comportamentul lor este folosit pentru a găsi soluții relevante pentru o anumită problemă de optimizare. ACO reduce de obicei problema de optimizare în problema găsirii celui mai bun traseu într-un graf ponderat.

Am dezvoltat un instrument grafic modern pentru simularea și redarea modelelor

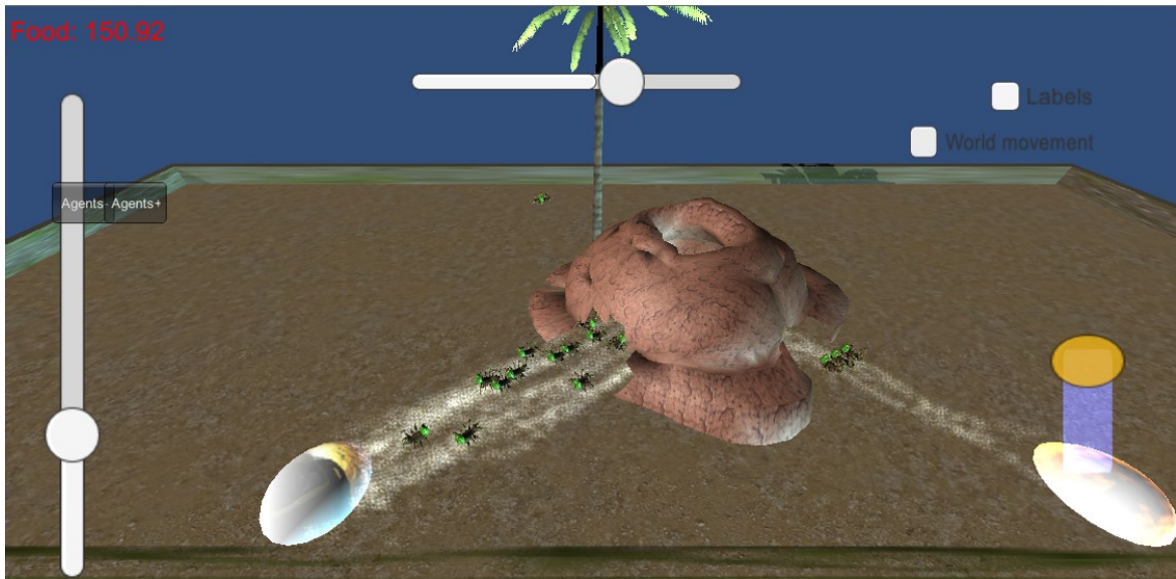


Figura 2.3: Visual Agents - Ant Colony Optimization [BD15]

compuționale în 3D. Utilizând Unity3D ca motor grafic și C# ca limbaj de programare principal, am putut observa procese precum comportamentele emergente într-o nouă perspectivă [BD15]. Am început cu comportamentul coloniei de furnici și am dezvoltat un instrument de simulare care poate fi utilizat în cercetări viitoare, dar și în scopuri educaționale. Principala subiect de studiu din [BD15] a fost următorul.

Cum s-ar împărți o colonie de furnici între două surse de hrană?

După cum se vede în 2.3, am început cu un mediu simplu, un spațiu limitat cu un strat de sol pentru navigație și mai mulți pereți pentru a împiedica agenții noștri să cadă în timpul interacțiunilor.

Am folosit mușuroiul ca punct central de generare, unde agenții se pot aduna pentru a-și lăsa hrana, dacă o descoperă pe hartă. De asemenea, utilizatorii își pot regla unghiul camerei, similar cu un joc video.

2.2 Comportamentul uman în procesele economice

2.2.1 Contextul

În [F⁺12] Farmer susține faptul că ar trebui să tratăm sistemele economice în mod realist ca fiind complexe. Lucrarea sa începe prin compararea termenului de “Economie Complexă” cu jena pe care matematicianul Stan Ulam a simțit-o în 1977 cu privire la titlul cursului său de “Matematică neliniară”. Ulam a afirmat că, din moment ce aproape toată matematica și cele mai fascinante fenomene sunt neliniare, termenul este incorect și asemănător cu referirea la un animal ca „animal non-elefant”.

Realist vorbind, întreprinderile economice sunt în mare parte coordonate de agenți umani ce sunt vulnerabili la greșeli, mai ales datorită așa-numitelor *prejudecăți cognitive*. Biasurile cognitive [HNM15] sunt anomalii predictibile ale raționamentului uman. Din observațiile noastre menționate în [Bot21b], modelul standard este cel mai util pentru sistemul macro-economic, în timp ce modelul comportamental este mai potrivit pentru micro-decizii.

2.2.2 Preferința temporală pentru agenții economici

În această secțiune am discutat despre concepte precum satisfacția imediată sau întârziată și modelele existente privind modul în care se comportă agenții umani atunci când li se oferă opțiunea de a consuma acum sau mai târziu. Aceste aspecte se bazează pe cercetările noastre din [Bot21b].

Am analizat anterior conceptul de *utilitate*, propus de Daniel Bernoulli în 1738 și care poate fi utilizat pentru a cuantifica satisfacția obținută din utilizarea unui produs sau serviciu. În acest context, **preferința temporală** descrie fenomenul în care agenții apreciază același bun (orice produs dorit care oferă utilitate unui consumator) în mod diferit, în funcție de momentul în care este consumat.

Acest *factor de discount*, să îi spunem δ , este valoarea necesară de multiplicare pentru suma din viitor prin care să obținem valoarea curentă și se poate calcula prin

$$\delta = \frac{1}{1 + \rho} \quad (2.1)$$

unde *rata preferinței temporale* ρ poate fi calculată:

$$\rho = \frac{1 - \delta}{\delta} \quad (2.2)$$

Rata ρ se referă la rata de discount al dobânzii. Am identificat câteva exemple în cel de-al patrulea capitol din “Economia Comportamentală” de Edward Cartwright [Car18].

2.2.3 Jocul ultimatumului

Pentru a studia *erorile sistematice de judecată*, metode experimentale au fost aplicate ca parte a așa-numitei “*Experimental economics*”. Multiple studii [CLS06, BDW03, BdW14, BDGW11, WB20] au demonstrat prezența prejudecăților la subiecții non-umani <https://youtu.be/xot4z1CKFMo>.

Teoria jocurilor abordează unele dintre aceste comportamente prin analiza interacțiunilor strategice dintre jucători. Matematica jocurilor a fost studiată mult timp. De exemplu, Blaise Pascal și Christiaan Huygens au analizat structura jocurilor de noroc în 1657, iar Gerolamo Cardano a scris “*Book on Games of Chance*” în 1564 [Car15].

Lui John von Neumann și Oskar Morgenstern li se atribuie faptul că au pus bazele teoriei moderne a jocurilor în 1944 prin analiza jocurilor cooperative și a celor cu sumă nulă [VNM07].

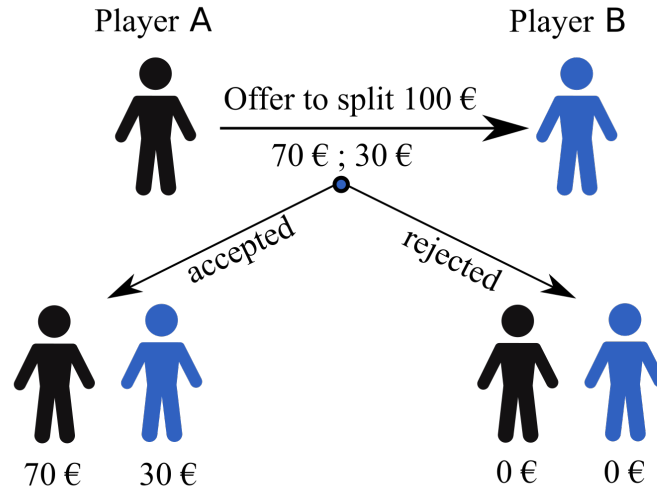


Figura 2.4: Exemplu joc “Ultimatum Game”, suma = 100 EUR

Jocul Ultimatum (UG) este unul dintre cele mai cunoscute jocuri economice și servește ca un bun exemplu de psihologie umană cu un comportament aparent irațional [Gua08, TLS⁺13], iar regulile jocului sunt simple, așa cum se arată în Figura 2.4.

2.3 Detectarea structurilor de comunități

Detectarea grupurilor în sisteme complexe este o provocare importantă, ce prezintă o gamă largă de aplicații în domenii precum criminologia (găsirea grupărilor potențial periculoase), sănătatea publică (grupuri vulnerabile de indivizi), marketing (identificarea setului de potențiali clienți), rețele sociale (identificarea conturilor false) și multe altele [KŞ18].

Una dintre principalele abordări ale detectării structurii comunității pentru rețelele unipartite transformă problema căutării într-o problemă de optimizare prin utilizarea unei funcții de fitness care este menită să demonstreze structura modulară a rețelei. Una dintre cele mai cunoscute și dezbătute funcții de fitness este *modularitatea* [NG04, New06]. Aceasta Q_N compară o structură comunitară pentru o rețea cu structura corespunzătoare atunci când se ia în considerare o rețea aleatorie. Poate fi calculat ca:

$$Q_N = \sum_{C \in \mathcal{C}} \left(\frac{m_C}{m} - \left(\frac{D_C}{2m} \right)^2 \right), \quad (2.3)$$

unde m reprezintă muchiile, m_C reprezintă marginile din comunitatea C , iar D_C este calculat ca suma gradelor nodurilor din C . În mod normal, o valoare ridicată pentru

funcția de modularitate înseamnă o soluție mai bună. În [GBSL17] am investigat acest aspect și am descoperit diferite abordări în funcție de tipul rețelei.

2.4 Rețele neuronale

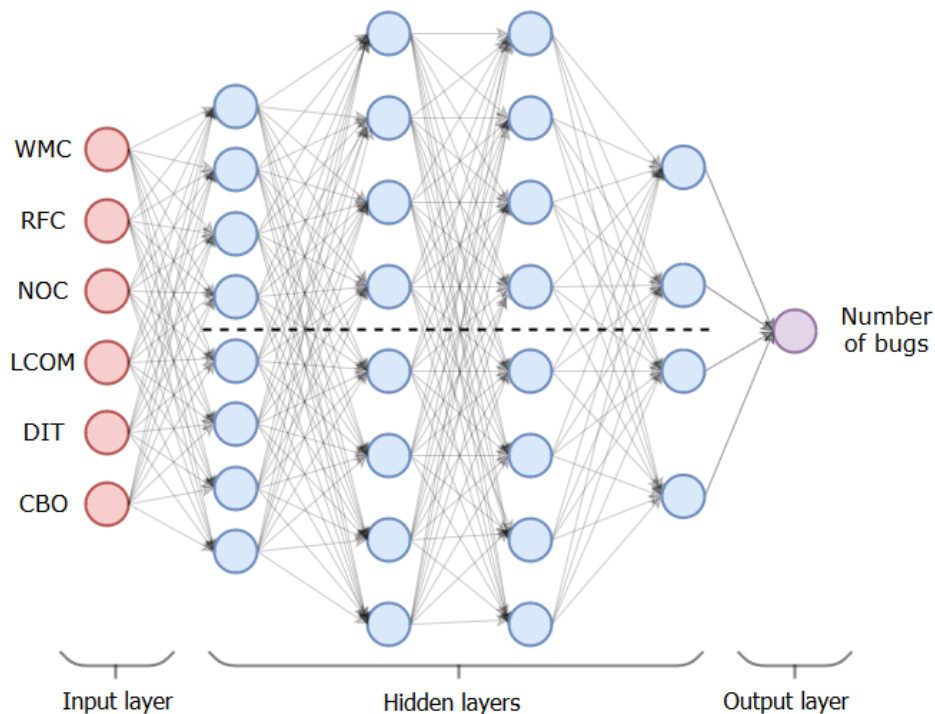


Figura 2.5: Feed-forward Neural Network [SB20]

Rețelele neuronale artificiale (ANN) [RN95] simulează comportamentul creierului natural. Ele sunt sisteme de calcul formate din noduri interconectate, la fel ca neuronii din creierul uman și pot fi organizate în straturi specializate.

În figura 2.5 putem vedea un exemplu de astfel de rețea neuronală stratificată (LFFNN) din [SB20], unde primul strat este folosit ca punct de intrare iar stratul final ca ieșire. Nodurile dintre ele fac parte din așa-numitele “straturi ascunde”. În această rețea, un nod poate fi conectat doar la un alt nod din stratul următor și avem mai multe straturi ascunde. Procesul de învățare este direct: intrarea este dată prin primul strat și apoi rezultatul este comparat cu valoarea așteptată. Diferența se numește “eroare” și poate fi utilizată pentru a regla greutatea rețelei.

Capitolul 3

Noi modele computaționale bazate pe agenți

Natura prezintă o mulțime de fenomene interesante, atât probleme cât și soluții, iar oamenii au folosit aceste informații pentru a progresa, încă de la începuturile timpului. Trebuie să recunosc că observarea sistemelor naturale este o activitate fascinantă pentru mine și îmi folosesc întotdeauna imaginația pentru a înțelege modul în care acestea funcționează, atât la nivel macro, cât și la nivel micro. În secțiunile următoare am discutat despre munca mea în cadrul mai multor proiecte, unde am folosit agenți artificiali pentru a simula procese și diverse comportamente.

3.1 Algoritmi inspirați de coloniile de viespi în învățarea prin joc: Sotirios

Cercetarea mea a început cu algoritmi inspirați din natură și unul dintre primele sisteme pe care le-am studiat a fost cel al cuibului de viespe. Pe baza lucrărilor mele anterioare de creare a unui joc video educațional, am integrat un nou model de alocare a puzzle-urilor bazat pe comportamentul coloniilor de viespi. *Sotirios* [Bot12] este un joc de e-learning (Digital Game Based Learning) ce îmbină grafica 3D cu puzzle-uri educative. Naratiunea jocului gravitează în jurul unui robot ce se opune liderilor planetei iar jucătorul trebuie să rezolve diferite puzzle-uri (Figura 3.1) pentru a progresa [SB18].

3.1.1 Metodologie

Abordarea mea a fost inspirată de către comportamentul uman. Mai exact, modul în care copiii și chiar adulții *învață prin joacă*. Am observat că studenții se plictisesc sau își pierd concentrarea în timpul unei lecții, dar nu au o problemă să petreacă cu jocurile video zile întregi.

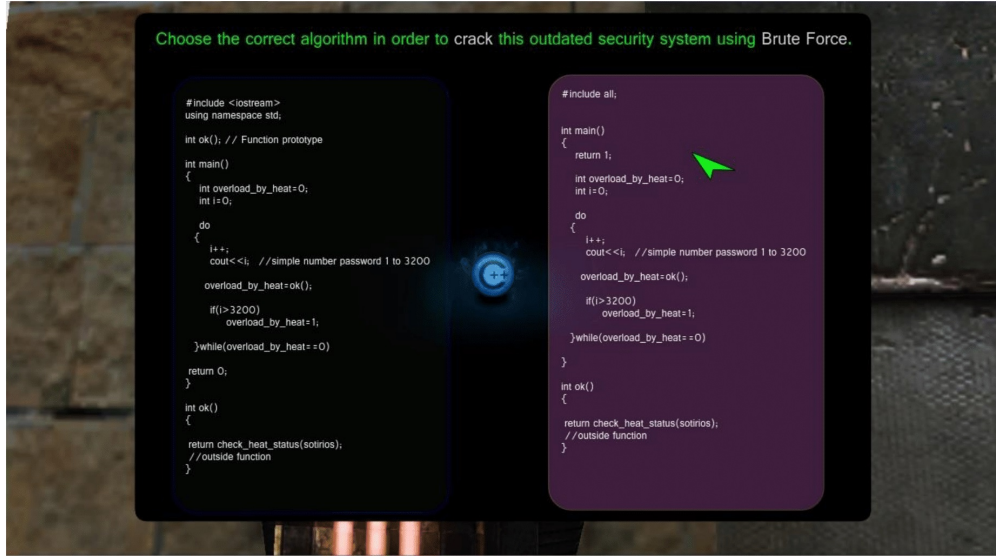


Figura 3.1: Decizii efectuate de către user, Project Sotirios [Bot12]

3.1.2 Modelul propus

În [SB18] am prezentat soluția noastră pentru problema distribuirii dinamice a puzzle-urilor pentru jucători multipli. Ne-am adresat componentei multiplayer din SOTIRIOS pentru a crește potențialul educațional al jocului, pe baza mai multor constatări privind efectul pozitiv al învățării în grup [Col80]. Specializarea unui jucător i pentru rezolvarea unui puzzle k - $PS_{i,k}$

$$PS_{i,k} = \max \left\{ 0, \frac{ns_{i,k} \cdot l_{i,k} - nus_{i,k} \cdot l_{i,k}}{ns_{i,k} + nus_{i,k}} \right\} \quad (3.1)$$

Unde $ns_{i,k}$ este numărul de puzzle-uri de tip k , rezolvate de către jucătorul i ; $nus_{i,k}$ - numărul de puzzle-uri de tip k , greșite de către jucătorul i ; $l_{i,k}$ - nivelul unui puzzle din $ns_{i,k}$ și $nus_{i,k}$.

Pragurile de răspuns au valori cuprinse între $[w_{min}, w_{max}]$. Prin $w_{i,k}$ înțelegem pragul de răspuns al viespei în raport cu jucătorul i pentru puzzle-ul k . Inițial, toate pragurile de răspuns sunt inițiate cu valoarea medie $w_{med} = (w_{min} + w_{max})/2$. Puzzle-ul k transmite un stimul S_k pe baza complexității puzzle-ului C_k . Agentul de viespe i licitează pentru task-ul k dacă

$$MD_i \leq MAXMT \quad (3.2)$$

cu probabilitatea

$$P_{i,k} = \frac{(S_k)^\gamma}{(S_k)^\gamma + (w_{i,j})^\gamma} \quad (3.3)$$

Când două viespi i și s intră într-un concurs de dominație, probabilitatea viespei i

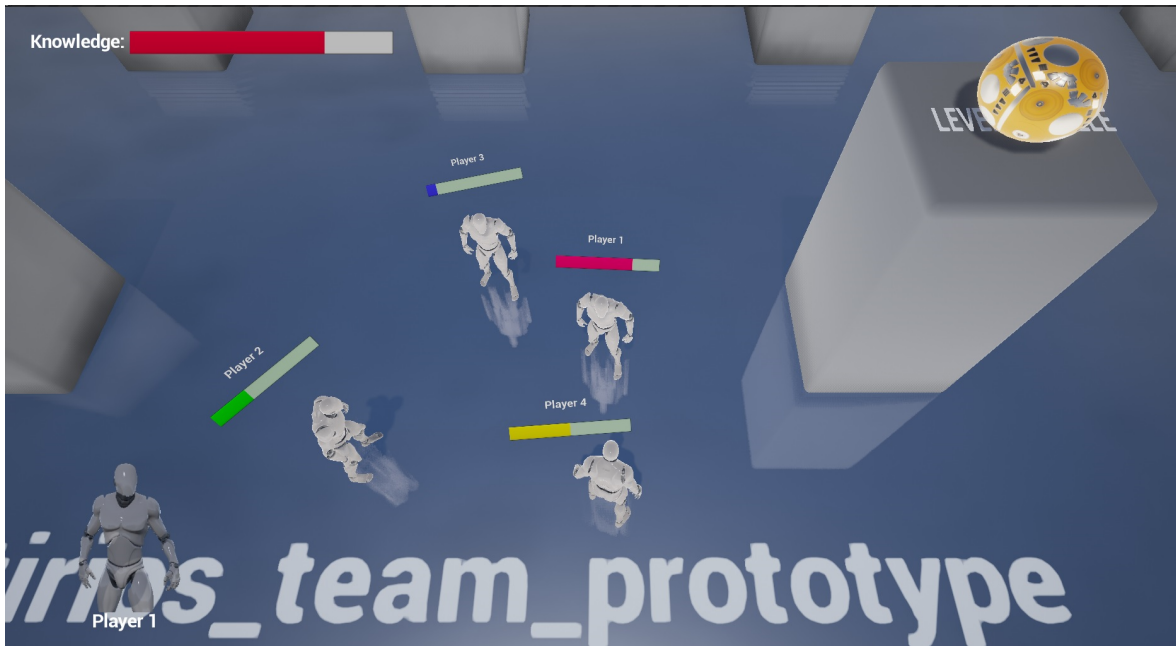


Figura 3.2: Versiunea multiplayer, Sotirios [SB18]

să obțină acel puzzle este:

$$P_c(i, s) = P(i \text{ win} | F_i, F_s) = \frac{F_s^2}{F_s^2 + F_i^2} \quad (3.4)$$

3.1.3 Discuție și dezvoltări viitoare

Viespile de învățare alocate jucătorilor cu specializare similară și cu un număr egal de puzzle-uri rezolvate în cadrul jocului au șanse egale de a obține puzzle-ul.

După cum am menționat în [SB18], o creștere adaptivă a pragurilor din modelul de viespe menține un echilibru între componentele unei echipe și evită blocarea jocului cauzată de introducerea unui nou tip de puzzle. Abordarea noastră este adaptabilă și oferă o varietate de strategii pentru sistemul multi-agent propus. Parametrii pot fi reglați manual sau ar putea fi dezvoltate alte abordări în acest scop, precum algoritmi genetici (GA) sau rețele neuronale. Cu toate acestea, este necesară comparația mai multor modele pentru a determina cea mai bună strategie.

Una dintre direcțiile viitoare ale activității noastre constă în validarea și analiza modelului cu valori diferite pentru parametrii și strategiile din model. O altă direcție ar fi implementarea structurilor de învățare în alte jocuri virtuale și compararea rezultatelor. În Figura 3.2 putem vedea imagini din ultima versiune a jocului Sotirios, dezvoltat folosind Unreal Engine.

3.2 Preferința temporală pentru agenții economici multidimensionali

În această secțiune, am discutat despre preferința de timp în procesele economice. Am validat modelele standard folosind date empirice din studii conexe și am efectuat propriile noastre experimente [Bot21b]. De asemenea, am comparat predicția modelelor și am sugerat noi abordări hibride, care au fost validate. Inspirația pentru acest subiect își are originea din interesul meu de a înțelege de ce oamenii acționează în anumite moduri și uneori iau decizii aparent iraționale într-un context financiar. Am propus un nou model computațional unificat, ce reflectă deciziile umane în sisteme complexe precum economia. Noul model este proiectat folosind o metodologie bazată pe date empirice [Bot21b] și poate fi folosit ca un instrument eficient pentru dezvoltarea agenților economici multidimensionali (MEA). În acest proiect ne-am concentrat pe componenta *rațională* a modelului (paradigma standard), în special preferințele temporale ce pot fi observate în economie.

3.2.1 Metodologie

Am realizat un sondaj cu studenții de la Universitatea Babeș-Bolyai din Cluj-Napoca și alți participanți dintr-o comunitate online (reddit) și am folosit Google Forms ca platformă de sondaj.

Tabelul 3.1: Exemple de răspunsuri la sondaj, [Bot21b]

50\$					5000\$				
3 mo	6 mo	1 yr	2yr	4yr	3 mo	6 mo	1 yr	2yr	4yr
200	400	600	1000	2000	8000	10000	15000	20000	35000
75	100	300	500	1000	5999	7420	8499	9999	14000
60	75	100	200	400	6000	7500	8000	15000	25000
65	80	100	150	300	6000	6666	8000	10000	12500
100	300	1000	2000	5000	10000	10000	12000	15000	20000

Șablonul era:

Sunteți indiferent față de Y \$ acum față de X \$ în t ani. Scrieți suma X mai jos

și am întrebat subiecții cu privire la indiferența de a primi bani pe perioade de 6 luni, 1 an, 2 ani și 4 ani. Suma a variat între 50 \$ și 5000 \$. În comparație cu datele experimentale mai vechi discutate în lucrarea noastră, rezultatele prezintă unele variații, însă sunt în concordanță cu cercetările mai recente [EMT⁺18].

Modelul propus

În acest context am sugerat două funcții noi ce pot fi considerate hibridzi ale modelelor existente, simplificate. Credem că prin utilizarea mai multor parametri, acestea pot modela mai bine uman modern, mai ales având în vedere anomaliile identificate în literatură și în propriile noastre experimente.

Ecuția 3.5 descrie funcția **Hybrid Exponential-Hyperbolic**:

$$hyb_{EH}(x) = \frac{\delta^x}{1 + (\alpha \times x)} \quad (3.5)$$

iar ecuația 3.6 descrie funcția temporală **Hybrid Quasi-Exponential-Hyperbolic**.

$$hyb_{QEH}(x) = \beta \times \frac{\delta^x}{1 + (\alpha \times x)} \quad (3.6)$$

În aceste funcții, δ este *factorul de discount* iar α reprezintă parametrul de discount ce scalează gradul de nevoie pentru satisfacție imediată, iar β este folosit pentru a reprezenta bias-ul pentru prezent [Bot21b].

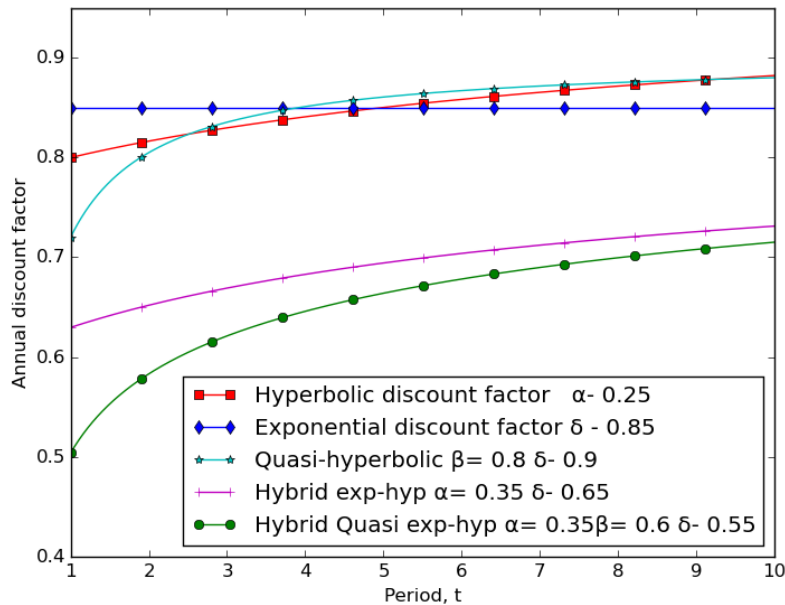


Figura 3.3: Factorul de discount anual, comparație funcții, [Bot21b]

Figura 3.3 prezintă o comparație a abordărilor actuale și a funcțiilor noastre hibride sugerate pentru factorul de discount temporal anual. Putem observa similitudini între funcția cvasi-exponențială-hiperbolică și cvasi-hiperbolică, și între hiperbolică exponențială cu funcția de discount hiperbolică. După cum am arătat în secțiunea următoare, variațiile sunt importante pentru analiza noastră.

3.2.2 Rezultate experimentale

Rezultatele arată o precizie de **95.17%** pentru suma de 50 \$ și **91.15%** pentru întrebările cu 5000 \$. Un eșantion de răspunsuri din sondajul nostru din [Bot21b] poate fi văzut în Tabelul 3.1. Rezultatele finale ce includ acuratețea tuturor funcțiilor se află în Tabelul 3.2.

Tabelul 3.2: Acuratețea funcțiilor pentru datele experimentale [Bot21b]

Discount function	ACC
Hybrid Quasi-Exponential-Hyperbolic	99.39 %
Hybrid Exponential-Hyperbolic	98.65 %
Quasi-Hyperbolic	96.24 %
Hyperbolic	96.07 %
Exponential	89.02 %

3.2.3 Concluzii și etape următoare

În concluzie, în [Bot21b] am furnizat o analiză a fenomenelor de preferință temporală găsite în procesele economice. Folosind studii din literatura de specialitate, am evaluat funcțiile de discount existente și am propus noi soluții hibride bazate pe propriile noastre date experimentale.

Funcțiile noastre propuse au funcționat bine și am determinat rezultate foarte bune pentru funcția Quasi-Exponențială-Hiperbolică, cu o precizie de până la 99,39%.

Conform rezultatelor noastre, funcțiile standard au funcționat bine pe seturile de date mai vechi și au avut diferențe în modelarea comportamentului din date mai noi, în timp ce soluțiile noastre propuse arată un potențial foarte bun de modelare al consumatorilor [Bot21b].

În plus, am obținut rezultate promițătoare cu învățarea zero-shot și GPT-3 sub formă de preferință inter-temporală, cu un studiu aprofundat propus ca următoarea etapă a cercetării noastre.

3.3 Integrarea comportamentului uman folosind agenți economici multidimensionali

Această secțiune descrie contribuțiile realizate în proiectarea agenților autonomi ce pot simula comportamentul uman și pot răspunde la factori multipli precum impulsuri, emoții, influențe sociale [BS21].

În [BS21] am propus un algoritm evolutiv stocastic pentru dezvoltarea MEA într-un mediu dinamic ce ia în considerare comportamentul echitabil și influența comunității. Am investigat evoluția acestor agenți într-un mediu dinamic.

Am examinat jocul Ultimatum “one-shot” și modelarea strategiilor corespunzătoare. În următoarele paragrafe am descris rezultatele inițiale, folosind funcții existente. Pentru a ne asigura că modelul propus este valid, am realizat propriul nostru experiment cu voluntari și am coroborat rezultatele empirice [BS21]. Următoarele secțiuni descriu algoritmul evolutiv propus.

3.3.1 Metodologie

În [BS21] am discutat despre obiectivul nostru de a proiecta și implementa agenți economici multidimensionali artificiali (MEA), bazați pe un model computațional capabil să depășească dezavantajele modelelor menționate mai sus și să emuleze comportamentul uman într-un mod util atât micro, cât și macroeconomiei. Am propus patru componente, ce sunt corelate cu patru dimensiuni ale MEA: *cognitiv*, *social*, *emoțional* și *rațional*.

Fiecare componentă are o anumită pondere (semnificație în alegerea finală), bias (posibilitatea de a accelera o decizie - creșterea sau reducerea ponderii cu o anumită magnitudine) și o funcție de calcul. În această investigație, m-am concentrat asupra examinării și evaluării următoarelor întrebări de cercetare:

RQ1: Poate fi simulat comportamentul uman folosind algoritmi evolutivi?

RQ2: Cum putem integra rețelele neuronale în conceptul nostru MEA?

3.3.2 Modelul

Teoria propusă este că acest comportamentul uman perceput este de fapt rezultatul evoluției pe parcursul mai multor generații, într-un mediu dinamic, favorizând apariția strategiilor câștigătoare economice pe termen lung. Aceste strategii sunt strâns legate de caracteristicile sociale ale omului, cum ar fi corectitudinea și împărtășirea resurselor, precum și legate de restricțiile de mediu [BS21].

În acest sens am propus o soluție pentru modificarea comportamentului social și cognitiv al actorilor umani ce se joacă experimentul economic numit „Ultimatum Game” (UG) [Cam11, RTON13]. UG a fost utilizat în studiul nostru MEA, deoarece oferă informații despre interacțiunile agenților și asupra consecințelor alegerilor echitabile sau nedrepte. Dezvoltarea agenților capabili să reproducă rezultatele experimentelor cu subiecți umani în timp ce joacă UG poate avea loc numai atunci când este luată în considerare dimensiunea socială a modelului de calcul.

3.3.3 Rezultate experimentale

Ultimatum Game Evolution Algorithm

Rezultatele obținute în [BS21] folosind algoritmul nostru au dovedit că evoluția naturală favorizează corectitudinea, dar distribuția este foarte dependentă de parametrii stochastici.

A fost necesar un alt filtru comunitar pentru a negocia oferte între indivizii aceleiași comunități. În aceste experimente am folosit o singură comunitate, pentru a simula mai bine majoritatea participanților noștri reali (colegi)

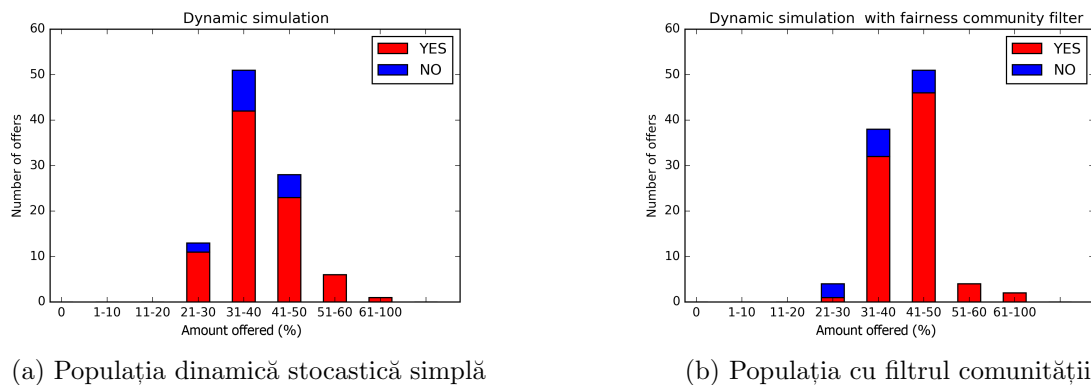


Figura 3.4: Rezultatele finale ale agenților, [BS21]

Ultimatum Game Simulator

Pe baza lucrărilor anterioare [BD15] am implementat și două aplicații 3D, pentru a ilustra mai bine conceptele implicate. Prima simulare a constatat în versiunea simplificată a jocului ultimatumului, unde 2 agenți se joacă conform regulilor.

Unele răspunsuri emoționale au fost concepute și integrate în jucătorii noștri artificiali, așa cum se vede în Figura 3.5. În această versiune, primul jucător primește sume aleatoare de bani din gama 0 – 100 și un factor de *încredere* este calculat, bazat pe jocurile anterioare. Dacă acest factor este peste 0.5 (valoarea maximă este 1 și valoarea minimă este 0), primul jucător face o ofertă celui de-al doilea jucător [BS21].

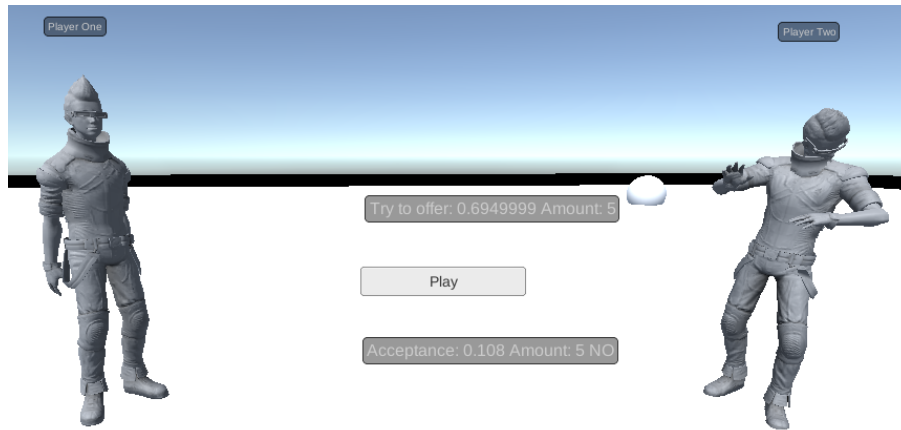


Figura 3.5: Ultimatum Game Simulator, versiunea 0.7, [BS21]

Utilizarea rețelelor neuronale pentru a juca jocul Ultimatum

Pentru a simula jocul Ultimatum, am creat o rețea neuronală simplă cu următoarea structură: un nod pe stratul de intrare, două straturi ascunse (cu 16 și 32 de noduri) și un singur nod de ieșire. Am folosit funcția de activare ReLU [NH10], metoda de optimizare *ADAM* și *eroarea medie pătratică* (MSE) ca și “loss function”. Rețeaua a determinat corect intervalul de 40% pe baza datelor de instruire și a prezentat o distribuție limitată a valorilor.

3.3.4 Concluzii

În [BS21] am propus o metodă evolutivă de creare a unor agenți inteligenți cu comportament similar uman și am validat algoritmul folosind experimente atât cu agenți umani cât și cu agenți sintetici. Rezultatele (Figura 3.4) demonstrează faptul că algoritmul nostru ne permite să modelăm comportamentul uman.

În modelul nostru, mediul dinamic (schimbarea sumelor de bani pe durata evoluției) a contribuit la o distribuție mai realistă a corectitudinii. Am simulat rezultatele din experimentul real cu ajutorul teoriei jocurilor evolutive stocastice (Figura 3.4) [BS21]. De asemenea, am creat aplicații de simulare 3D care pot fi utilizate pentru a descrie conceptele implicate în acest studiu. Simulatorul UG poate fi accesat la <https://codeworks.com/dev/ug/>.

Experimentul utilizării rețelei neuronale a avut rezultate încurajatoare în estimarea ofertei medii și poate fi încorporat în modelul nostru MEA o dată cu abordarea problemei de distribuție a valorilor.

Capitolul 4

Structuri comunitare: un nou criteriu de fitness

4.1 Detectarea structurilor de comunități în rețelele multipartite: o nouă măsură de fitness

În [GBSL17] am abordat problema detectării structurilor de comunități într-o rețea, care constă în identificarea grupurilor de noduri ce sunt mai conectate între ele (mai dens) decât cu alte noduri [For10], problemă studiată pe larg pentru rețelele unipartite și bipartite.

Cu toate acestea, există foarte puține rezultate care se extind la rețelele multipartite, în ciuda faptului că numeroase aplicații ar putea beneficia de astfel de metode, deoarece multe sisteme sau fenomene pot fi modelate ca și rețele multipartite [LLMW16, AT05].

Am propus o extindere a fitnessului comunității definită în [Lan09] pentru identificarea comunităților din rețelele multipartite, considerând că nodurile care au un vecin comun sunt conectate între ele și astfel se creează legături de tip *shadow*, ajutând la descoperirea structurilor de comunități [GBSL17].

4.1.1 Metodologie

Particularitățile rețelelor multipartite necesită luarea în considerare a faptului că, în timp ce nodurile aparținând aceleiași partiții nu pot fi conectate direct între ele, ele pot fi conectate indirect de vecinii comuni din alte partiții. Astfel de conexiuni induse pot fi, de asemenea, responsabile pentru structura comunitară și ar trebui luate în considerare la evaluarea potențialelor soluții.

În acest context, am propus “Multipartite Community Fitness” M prin extensia *fitnessului comunitar* din [Lan09] pentru a include astfel de link-uri [GBSL17]. Deoarece

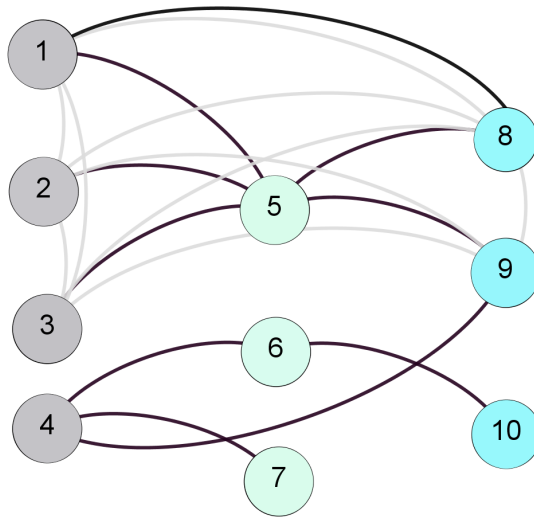


Figura 4.1: O rețea simplă tripartită. [GBSL17]

astfel de legături nu pot fi considerate conexiuni 'reale' în rețea (dacă ar fi fost cazul, acestea ar fi fost adăugate la construcția rețelei) le numim legături *shadow* și le putem adăuga la gradul unui nod folosind *factor de sensibilitate multipartit* α .

Fitness-ul comunității C este calculat astfel:

$$f_{\alpha}(C) = \frac{\sum_{i \in C} k_{\alpha}(i|C)}{\sum_{i \in C} d_{i,\alpha}}, \quad (4.1)$$

ca și raportul dintre gradul total modificat al nodurilor în comunitatea C și gradul total al nodurilor din C . O valoare de fitness mai mare poate fi considerată un indicator al unei comunități mai mari. Având în vedere aceeași comunitate $C = \{1, 2, 3, 5\}$ din Figura 4.1, fitness-ul ei $f_{0.05}(C) = \frac{1+1+1+3 \cdot 15}{2 \cdot 05+1+1+5 \cdot 5} = 0.64$.

Fitness-ul unei *structuri comunitare* \mathcal{C} , adică o partiție a setului de noduri, se calculează ca medie a fitness-ului tuturor comunităților $C \in \mathcal{C}$:

$$M_{\alpha}(\mathcal{C}) = \frac{1}{|\mathcal{C}|} \sum_{C \in \mathcal{C}} f_{\alpha}(C), \quad (4.2)$$

unde $|\mathcal{C}|$ este numărul de comunități.

Pentru rețelele ponderate folosim aceeași abordare, luând în considerare în toate formulele gradul ponderat sau gradul interior ponderat al nodurilor și suma fitness-ului fiecărei comunități din (4.2).

4.1.2 Soluția propusă

Maximizarea fitness-ului M ar trebui să descopere o structură comunitară multipartită într-o manieră similară precum o funcție de fitness, ca și modularitatea, o explorează pentru rețelele unipartite. Pentru a valida abordarea noastră am utilizat o metodă eficientă bazată pe optimizarea extremală pentru detectarea structurii comunității, numită *NoisyEO* [ISG17] pentru a maximiza fitness-ul M .

NoisyEO evoluează perechi de indivizi s și s_{best} prin reatribuirea valorilor aleatorii celor mai slabe componente din s și înlocuind s_{best} ori de câte ori a fost găsită o soluție mai bună. s și s_{best} encodează vectori de dimensiune N ce reprezintă dimensiunea rețelei. $s(i)$ reprezintă comunitatea pentru nodul i . Fitness-ul unui nod este calculat ca și contribuția pe care nodul o are în comunitatea sa (Măsurăm fitness-ul comunității cu și fără nodul dat). Modificarea față de algoritmul original *NoisyEO* o reprezintă noua funcție de fitness comunitară M din (4.2).

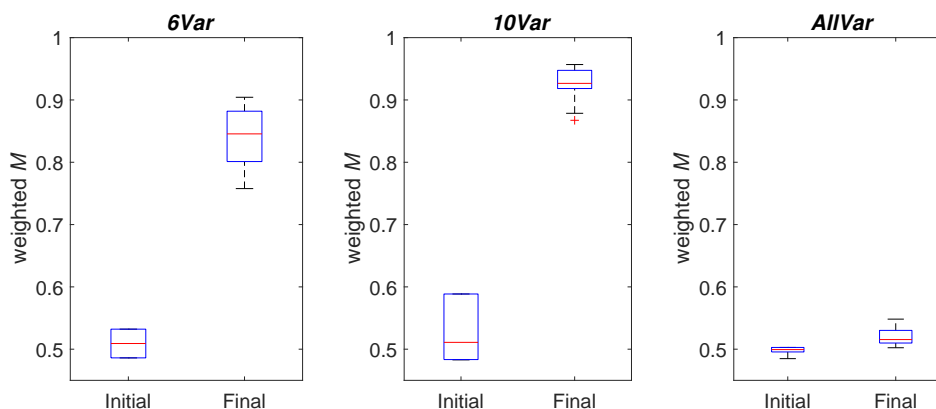


Figura 4.2: Rezultate obținute pentru cele trei rețele de date financiare [GBSL17].

4.1.3 Rezultate experimentale

Rezultate

Matricea de corelație a valorilor NMI ale indivizilor care au cele mai bune valori de fitness în populația finală arată că funcțiile de modularitate sunt puternic corelate, cu 0,92 și distribuții similare, în timp ce M prezintă un comportament diferit, dar cu mai multe valori NMI de 1. Aceste grafice arată că, în timp ce experimentele efectuate pe date sintetice arată rezultate mai bune obținute cu fitnessul nostru comunitar multipartit, există încă loc de îmbunătățire, deoarece există situații în care valoarea maximă a lui M nu identifică individul care are cea mai bună valoare NMI din populație.

Date de marketing bancare Pentru a ilustra analiza datelor bancare am construit trei rețele ponderate: una care conține toate variabilele din setul de date (*AllVar*,

785 noduri, 31549 muchii), una care conține doar date referitoare la client (“age-job-marital-education-default-balance-housing-loan-duration-y” *10 Var*, 87 noduri, 2845 muchii) și una cu variabilele binare eliminate (“age-job-marital-education-balance-duration” *6Var*, 79 noduri, 2217 muchii) (Figura 4.3).

Pentru a interpreta rezultatele, de exemplu, pentru rețeaua *6Var*, algoritmul a separat persoanele cu vârsta sub 53 de ani de celelalte. La categoria de locuri de muncă, algoritmul a plasat în aceeași comunitate de *gospodine*, *pensionari* și pe cei cu locuri de muncă *necunoscute*. După starea civilă, cei divorțați sunt separați de celibatari și căsătoriți. Educația *primară* este separată de toate celelalte niveluri de educație. Aceste rezultate arată cum structura comunităților poate dezvălui conexiuni despre categorii și poate ajuta la analiza seturilor mari de date.

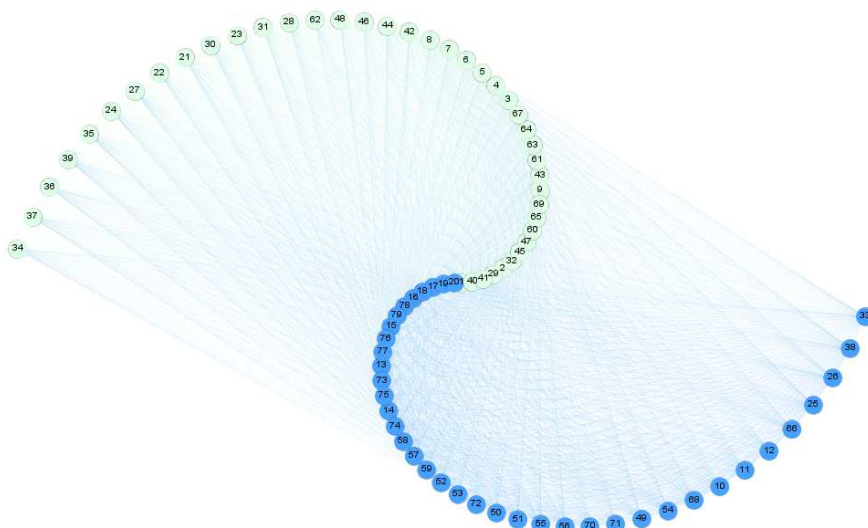


Figura 4.3: Structura comunitară detectată în setul de date *6Var* [GBSL17].

4.1.4 Concluzii

Problema detecției structurilor de comunități în rețelele multipartite poate fi abordată luând în considerare conexiuni de tip “shadow” între nodurile care au un vecin comun, chiar dacă aparțin aceleiași partiții.

În [GBSL17] am propus o funcție de fitness care să ia în considerare aceste conexiuni și să arate că poate fi utilizată pentru a identifica comunități pentru reperele sintetice și cu aplicații din lumea reală. Putem compara rezultatele cu abordările existente pentru rețelele bipartite; rezultatele numerice ilustrează eficiența acestei abordări și identifică un nou mod de a analiza seturi mari de date.

Acest proiect a furnizat cele mai provocatoare aspecte ale cercetării mele, precum și cele mai satisfăcătoare contribuții; lucrarea completă a fost aprobată și prezentată la Conferința de Calcul Genetic și Evolutiv (GECCO).

4.2 O analiză teoretică a rețelelor de autori în domeniul elaborării lucrărilor academice

În [GBSL20] am analizat domeniul elaborărilor lucrărilor științifice din perspectiva rețelelor de autori, combinată cu o abordare din teoria jocurilor. Rețelele de coautori sunt construite din datele publice de de publicare. Se calculează valoarea Shapley a fiecărui autor, indicând potențialul mediu de colaborare marginal al autorilor. Rezultatele oferă informații interesante despre comportamentele de publicare ale comunității academice.

Motivația din spatele studiului nostru a fost următoarea. Acest tip de cercetare poate oferi o analiză al publicațiilor dintr-un domeniu emergent în ceea ce privește publicarea și poate oferi studenților o imagine de ansamblu atât de necesară asupra propriului domeniu de studiu. Rețelele de coautori au devenit un element de bază în analiza scientometrică iar examinarea lor poate oferi informații importante unui începător în această disciplină.

4.2.1 Metodologie

Perspectivile teoretice ale jocurilor

Teoria jocurilor este un domeniu important de cercetare, având un caracter interdisciplinar cu aplicații principale în economie, biologie, inginerie, politică etc. Jocurile pot fi clasificate în diferite clase: jocuri cu informații perfecte și jocuri cu informații imperfecte, cooperativ vs. jocuri necooperante etc. Valoarea *Shapley* [Har89] definește cât de important este fiecare jucător pentru coalitie și modul în care recompensa ar trebui împărțită între jucători.

Statistici rețele

Deoarece rețelele mari sunt greu de analizat și de comparat vizual cu alte rețele, datele statistice sunt utilizate pentru a caracteriza și a oferi informații despre structura și proprietățile lor. Modularitatea măsoară cât de bine se descompun în grupuri de noduri ("clusters") [BGLL08]. Valorile ridicate indică o structură complexă a rețelei [GBSL20].

Coeficientul de grupare [WS98], aplicat unui singur nod din rețea, măsoară cât de complet este vecinătatea nodului, adică proporția dintre numărul de noduri învecinate și numărul total de vecini posibili. Într-o rețea, două noduri sunt vecine dacă există o legătură (conexiune directă) între ele. Pentru rețea, coeficientul de *clusterizare* reprezintă media tuturor coeficienților de clusterizare. Este o valoare cuprinsă între 0 și 1 și indică tendința nodurilor de a se grupa împreună. Un nod izolat din rețea are valoarea 0, iar un nod legat de toate celelalte noduri va avea valoarea 1.

Rețele de coautori și valoarea Shapley

În [GBSL20] am discutat despre modul în care centralitatea rețelei este un indicator important, cu mai multe metode de măsură (“betweenness, closeness, PageRank centrality”) care încearcă să capteze anumite proprietăți ale rețelelor studiate. În [MAS⁺13] este descris un algoritm liniar ce obține valorile Shapley ale jocului cooperativ introdus, care reprezintă contribuția marginală medie a fiecărui nod la fiecare coalțiie formată din restul nodurilor. Am folosit acest algoritm pentru a calcula valoarea Shapley a autorilor care au publicat articole pe tema “academic writing” în perioada 2015 – 2019.

Rețele de coautori pentru “academic writing”

Am colectat datele folosind API-ul web al bazei de date Scopus (www.scopus.com) și am filtrat articolele care conțineau termenii cheie “academic writing” publicate între 2015 și 2019. Am folosit următoarea interogare:

```
TITLE-ABS-KEY("academic writing") AND ( LIMIT-TO (
DOCTYPE,"ar")) AND ( LIMIT-TO ( PUBYEAR,x) ),
```

unde pentru x am folosit 5 ani, din 2015 până în 2019.

Am descoperit că rețelele cresc ca și dimensiune pe parcursul investigației noastre de cinci ani. Primii patru indicatori sunt în creștere: numărul de noduri, care reprezintă numărul de autori ce au publicat în acest domeniu, numărul de conexiuni și gradul mediu, care indică faptul că articole succesive sunt publicate colaborativ.

4.2.2 Rezultate

În [GBSL20], valoarea Shapley pentru autorii dintr-o rețea este calculată ca și contribuția marginală medie la valoarea tuturor coalțițiilor de noduri care pot fi formate cu nodurile din rețea. Valoarea unei coalțiții este calculată ca și dimensiunea coalțiției la care se adaugă numărul de noduri care pot fi accesate la distanța 1 de nodurile din coalțiție.

Valoarea Shapley relevă autori care au cea mai mare contribuție marginală din rețeaua lor - în acest caz prin publicațiile dintr-un anumit an. În domeniul “academic writing”, o astfel de informație relevă cine sunt cei mai influenți autori dintr-un an. O caracteristică interesantă a acestui domeniu este că lucrările publicate sunt adresate unui câmp specific de cercetare iar un astfel de studiu ar indica alte discipline ce se concentrează în mod formal pe prezentarea cercetării în articole.

Figura 4.4 prezintă rețeaua de coautori pentru anul 2016 (am ales anul 2016 deoarece este cea mai mică dintre cele cinci rețele studiate și poate fi afișată mai ușor), fără nodurile izolate, iar cele mai mari valori Shapley sunt evidențiate cu culoare roșie.

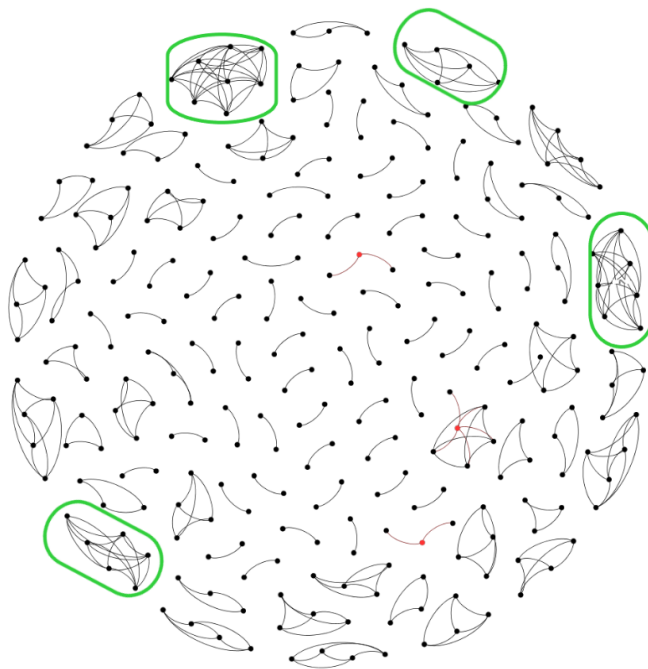


Figura 4.4: Rețea de coautori, [GBSL17]

Observăm că autorii cu cea mai mare valoare Shapley nu sunt neapărat autorii cu cele mai mari grade (componentele conectate marcate în verde reflectă de fapt o lucrare cu mulți autori), ci sunt cei care au contribuit cu mai multe lucrări scrise cu diferiți colaboratori, contribuind astfel la dezvoltarea domeniului.

4.2.3 Concluzii

Scrierea academică ca și domeniu de cercetare este unul în curs de dezvoltare, cu diverse particularități, ceea ce face ca datele și tendințele publicațiilor să justifice efortul unor studii mai aprofundate. Rețelele de comunități din din domeniul elaborării lucrărilor științifice sunt analizate cu ajutorul teoriei jocurilor cooperative.

Valoarea Shapley este un concept care măsoară contribuția marginală medie a unui jucător la un câștig colectiv. În cazul rețelelor de coautori, acesta indică potențialul mediu de colaborare ale acestora. Cei mai influenți autori în domeniul scrierii academice în perioada 2015-2019 sunt identificați pentru fiecare an. În [GBSL20] am constatat că domeniile lor de cercetare variază, indicând un interes pentru acest domeniu. Utilizarea valorii Shapley pentru rețelele de autori și implementarea tehnicii noastre s-au dovedit a fi o abordare interesantă a problemei, urmând ca pe viitor să continuăm investigația în acest domeniu.

Capitolul 5

Noi aspecte educaționale și de calitate a codului

5.1 Un cadru conceptual pentru predicția defectelor software prin utilizarea rețelelor neuronale

Testarea software-ului poate fi definită ca o investigație privind calitatea produselor dezvoltate. Identificarea defectelor software este o sarcină dificilă care necesită de obicei mult timp și resurse. Astfel, un mecanism care poate fi utilizat pentru a prezice defecțiunile software ar oferi un avantaj în crearea unor aplicații mai bune, utilizând mai puține resurse.

Contribuția din [SB20] este dublă: în primul rând, ne propunem să abordăm problema predicției defectelor software, definind un cadru conceptual care încorporează componente precum: modelul de proiectare orientat pe obiecte, metrici și metodele utilizate pentru predicție. În al doilea rând, investigația își propune să stabilească cea mai performantă combinație de indicatori CK pentru predicția defectelor software.

Indicatori folosiți

Indicatorii software sunt foarte utili atunci când încercăm să monitorizăm și să îmbunătățim calitatea software-ului. Aceștia sunt din ce în ce mai utilizați în domeniul ingineriei software. Unele dintre cele mai cunoscute și studiate metrici sunt cele propuse de Chidamber și Kemerer în [CK94]: *Depth of Inheritance Tree* (DIT), *Weighted Methods per Class* (WMC), *Coupling Between Objects* (CBO), *Response for a Class* (RFC), *Lack of Cohesion in Methods* (LCOM), *Number of children* (NOC)

5.1.1 Metodologie

În [SB20] am definit problema predicției defectelor software ca o problemă de regresie, scopul fiind de a prezice numărul de defecte (bug-uri) identificate pentru o anumită aplicație software. Am creat setul de instruire utilizând o parte din datele existente. Apoi am început procesul de instruire cu caracteristici selectate din setul de metrice software CK [CK94]. Obiectivul nostru a fost cel de a prezice, utilizând un model de rețea neuronală, numărul de defecte (bug-uri) al unei clase date.

Obiective și întrebări de cercetare

În [SB20] am investigat următoarele două întrebări de cercetare.

RQ1: Poate un subset al indicatorilor CK să performeze mai bine decât toate valorile existente în sarcina de a prezice numărul de probleme dintr-o clasă?

RQ2: Modelul de predicție bazat pe rețele neuronale funcționează mai bine folosind un set de date definit de un singur proiect decât luând în considerare setul de date obținut prin reunirea tuturor proiectelor?

5.1.2 Modelul propus și experimente

Setul de date de referință

Am folosit “Bug prediction dataset” din [DLR10], care conține metrice și date istorice din mai multe proiecte software: Java Development Tool (JDT) - Eclipse, Plug-in Development Environment (PDE), Equinox, Lucene și Mylyn. Setul de date conține informații despre versiuni, valorile CK, hot-fix-uri și numărul de erori clasificate după importanță. Mylyn, de exemplu, prezintă 701 de clase fără bug-uri și 1862 de clase cu bug-uri. În același proiect au fost 14577 de vulnerabilități, dintre care 235 au fost critice. Există 997 de clase în Java Development Tool care au bug-uri, dintre care 432 sunt importante. Lucene nu are erori critice sau majore, însă prezintă 1714 erori non-triviale.

Descrierea modelului de predicție bazat pe rețele neuronale

Pentru a prezice numărul de defecte pentru o anumită clasă, am folosit o rețea neuronală *feed-forward* cu “backpropagation” și structura din Figura 2.5: șase noduri pe stratul de intrare, unul nod pe stratul de ieșire și patru straturi ascunse, fiecare dintre ele având între 64 și 500 de noduri. Fiecare nod folosește funcția de activare *ReLU* [NH10]. Condiția de finalizare a instruirii este fie o lipsă a îmbunătățirii rezultatelor timp de 100 de epoci, fie la atingerea numărului maxim de epoci 1200. După

instruirea acestei rețele neuronale, am obținut un model de predicție pentru numărul de erori dintr-o clasă.

Descrierea experimentelor

Investigația noastră efectuată în [SB20] a folosit patru experimente, bazate pe cele cinci proiecte / seturi de date. În fiecare experiment, am instruit un model de predicție utilizând 65% din date eșantionate aleatoriu și am folosit diferența pentru validarea modelului, o metodă numită “holdout”, o versiune de validare încrucișată.

5.1.3 Rezultate și validare

În [SB20], modelul inițial nu a performat, cu variații substanțiale în acuratețe. Pentru al doilea experiment, diferențele dintre instruire și datele de testare au fost nesemnificative. În al treilea experiment, am constatat o creștere de 42,86% a preciziei între combinațiile CK, unde (DIT, NOC, LCOM) au o precizie de 0,35, comparativ cu (WMC, DIT, LCOM) care prezintă o acuratețe de 0,5.

Am constatat că există combinații de metrici ((WMC, DIT, LCOM) și (WMC, RFC, LCOM)) care depășesc toți indicatorii la prezicerea numărului de erori găsite într-o clasă. Am identificat o creștere de 21,95% a preciziei prin utilizarea unor combinații performante de valori în loc de utilizarea tuturor indicatorilor CK. Valorile CK performante au prezentat o precizie de 0,5 în testele noastre, în timp ce utilizarea tuturor valorilor a furnizat o precizie de 0,41. Aceasta răspunde la prima noastră întrebare de cercetare din secțiunea 5.1.1. Un răspuns interesant pentru a doua întrebare a fost că atunci când se utilizează toate proiectele din setul de date, predicțiile au performat mai slab decât prin utilizarea unui singur proiect. În experimentele noastre, precizia a fost mai slabă cu 46,34%, de la 0,41 precizie pentru proiectul JTD (Eclipse) la 0,22 pentru toate proiectele.

5.1.4 Concluzii și studii viitoare

În [SB20], am propus un cadru integrat pentru predicția defectelor software, definind un model bazat pe rețele neuronale pentru a stabili empiric o combinație de metrici software CK care sporesc acuratețea predicției. Rezultatele experimentale au arătat că există o creștere de 21,95% a preciziei prin utilizarea unei combinații performante de indicatori în loc să se utilizeze toate valorile CK și că o combinație a tuturor proiectelor a avut o predicție mai slabă decât utilizând un singur proiect. Ca și pași viitori, ne propunem să folosim alte metode de învățare automată pentru instruire, sau un model hibrid.

5.2 Aspecte educaționale în testarea programelor utilizând platforme colaborative

Colaborarea și accesul la informații, conform observațiilor mele, sunt esențiale atunci când abordăm sisteme multi-agent. Acest lucru mi-a permis să propun mai multe aspecte ale procesului de învățare și, împreună cu colegii mei din [CCBP21] am beneficiat de oportunitatea de a testa impactul pe care îl au instrumentele digitale de lucru în echipă și digitale asupra proceselor de învățare ale studenților noștri.

În [CCBP21] am discutat despre modul în care testarea software-ului este un element important în procesul de dezvoltare software. Principalele contribuții din [CCBP21] se referă la percepția elevilor asupra activităților de învățare pe durata materiei, îmbunătățirea eficienței pentru sarcinile de laborator și consecințele adoptării instrumentelor online în activitatea de predare.

5.2.1 Metodologia cercetării

În [CCBP21] am investigat următoarele întrebări de cercetare:

RQ1: Ce activități sunt considerate utile de către studenți în timp ce învață testarea software?

RQ2: Poate fi îmbunătățită eficiența pe parcursul temelor de laborator?

RQ3: Care sunt consecințele integrării instrumentelor online în cadrul VVSS?

Pentru a răspunde la întrebările noastre, am luat în considerare un sondaj completat de către studenți și feedback-ul profesorilor cu privire la activitățile de învățare. Chestionarul a avut atât întrebări deschise, cât și întrebări închise, necesitând părerea studenților cu privire la subiectele discutate pe parcursul cursului, experiențele de învățare și instrumentele utilizate pentru a comunica și colabora. Am analizat 32 de răspunsuri complete din 202 de studenți înscriși la VVSS în anul universitar 2019-2020.

5.2.2 Discuție

RQ1. Ce activități le sunt utile studenților în timp ce învață testarea software?

Elevii consideră relevante pentru experiența lor de învățare acele activități care implică interacțiune și sprijin. **Demo-urile live** au fost apreciate de 63% dintre studenți ca fiind **foarte importante** în învățarea testării software-ului. De asemenea, **sprijinul**

din partea profesorilor pentru diverse activități este perceput ca fiind foarte important de 67% dintre elevi. Rezultatele sugerează un grup de studenți axat pe rezolvarea problemelor, mai degrabă decât pe aspectele teoretice. Acest lucru este confirmat de 37% și 25% dintre studenți care au apreciat drept foarte importante **cursurile** și **bibliografia**, de exemplu, lucrări sau cărți privind testarea software-ului.

RQ2. Poate fi îmbunătățită eficiența pe parcursul temelor specifice de laborator?

În cadrul cursului de *Verificare și Validare al Sistemelor Software* (VVSS) 2019-2020, un total de 202 de studenți au creat 206 planuri de testare și cu 1019 cazuri de testare în Testlink. Fiecare student a implementat o integrare continuă (CI) în proiectele sale, fie singur, fie ca parte a unei echipe. Majoritatea studenților au ales să lucreze în echipă, formându-se un număr de 79 de echipe unice. O echipă poate conține maximum 3 persoane, dimensiunea minimă a unei echipe Agile [FH⁺01].

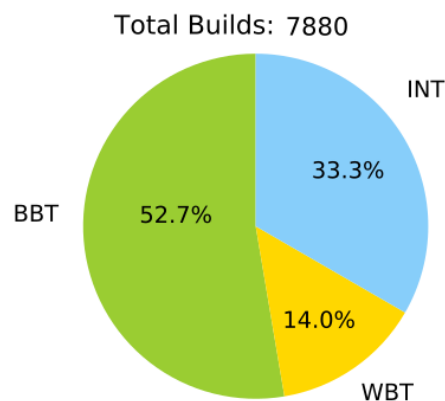


Figura 5.1: Distribuția execuțiilor exportate din Testlink

În figura 5.1 putem observa distribuția build-urilor Jenkins pentru fiecare activitate de laborator. Au existat un total de 7880 de versiuni, cu o distribuție mai mare pentru primul laborator (BBT) și mai mică pentru următoarele două.

În primul laborator al acestei serii (BBT), mai multe compilări înseamnă că studenții depun mai mult efort pentru a finaliza un scenariu de caz de testare reușit, petrecând mai mult timp pentru remedierea erorilor și configurarea mediului [CCBP21]. Prezentarea generală pentru toate cazurile de test, distribuită pe laboratoare: BBT avea 69.26% teste cu succes, WBT 76.64% și INT 80.23%, sugerând mici îmbunătățiri pe parcursul temelor de laborator.

Tabelul 5.1: Rezultate statistice între 2017 și 2020. [CCBP21]

Year	#Students	Slack	Skype	Labwork in groups	Lab grades mean	Final grades mean
2017	180	no	no	no	5.86	6.88
2018	202	yes	no	no	7.76	7.47
2019	201	yes	no	no	7.71	7.46
2020	202	yes	yes	yes	8.53	8.98
2020	self assessment	yes	yes	yes		7.13

RQ3. Care sunt consecințele integrării instrumentelor online în cadrul VVSS?

Rezultatele din 5.1 indică o creștere generală a notelor pe măsură ce a fost utilizat instrumentul Slack. În primii trei ani investigați, studenții au lucrat în mod individual la sarcinile lor. Sprijinul îmbunătățit acordat studenților pe parcursul activităților cursului este reflectat de notele superioare din anul 2018 până în 2020. În cursul anului 2020 au fost atribuite sarcini în echipă pentru activitățile de laborator, pentru a îmbunătăți abilitățile sociale ale participanților. Pentru media claselor de laborator, am observat o creștere de la 7,46 la 8,98. Până în prezent, analiza noastră cu privire la această creștere a notelor este asociată activităților de lucru în echipă, deoarece proiectele desemnate să lucreze în timpul activităților de laborator au avut niveluri similare de complexitate în toți anii investigați. Sunt necesare investigații suplimentare pentru a verifica dacă și alți factori au contribuit la această creștere [CCBP21].

5.2.3 Concluzii

În [CCBP21] am abordat aspecte importante legate de testarea software și învățarea acesteia. Cercetările pe această temă indică necesitatea unei abordări cuprinzătoare care să gestioneze obiectivele de învățare, activitățile asociate, instrumentele utilizate în testarea software-ului și pregătirea adecvată a profesorilor. Îmbunătățirea eficienței de la o sarcină de laborator la alta este posibilă și ajută la dobândirea unor abilități noi. Integrarea instrumentelor de comunicare și colaborare precum Slack și Skype în timpul activităților cursului a îmbunătățit notele studenților la evaluarea formativă și sumativă.

Lucrările viitoare cu privire la problemele investigate își propun să aprofundeze cercetarea privind creșterea diversității instrumentelor de testare și a instrumentelor de învățare în timpul activităților colaborative.

5.3 Componenta emoțională în contextul COVID-19

Acest segment prezintă o nouă inițiativă pentru cercetarea noastră cu privire la natura umană.

Munca noastră este realizată în colaborare cu Noah Shapiro, în vârstă de 14 ani, care a început proiectul 6 Feet Apart cu ajutorul altor șapte adolescenți și a lui Susan Shapiro, coautor al “The Teacher Within” *The Teacher Within* [SB19]. Proiectul a debutat la începutul lunii martie 2020, la scurt timp după ce noul coronavirus a apărut în mai multe țări din lume, inclusiv în SUA. Acest studiu este motivat de nevoia de a înțelege pandemia și de a cultiva un sentiment de comunitate în rândul celor afectați.



Figura 5.2: Muzeul Virtual, 6 Feet Apart

Contribuțiile noastre în acest proiect includ aplicația web “Stories”, platforma “6 Feet Apart” și *Muzeul virtual* (Unity 3D) din Figura 5.2. Ele reflectă amprenta digitală a proiectului și o soluție tehnică care poate ajuta adolescenții, profesorii și părinții să își gestioneze emoțiile în noua paradigmă a coronavirusului. Implementarea folosește Python (Django), Angular și Unity 3D pentru Muzeul Virtual.

Acest proiect oferă oportunitatea de a perfecționa agenții multidimensionali și de a studia în continuare comportamentul emoțional pe care l-am observat în timpul cercetării noastre. Pe baza acestor rezultate putem decide direcția adecvată de utilizare a platformelor digitale dezvoltate. Sperăm că Muzeul Virtual va rămâne un instrument istoric ce va păstra în siguranță experiențele actuale. Muzeul poate fi vizitat aici: <http://museum.sharemycovidstory.com>

Capitolul 6

Concluzii și dezvoltări viitoare

Concluzii

În această teză am investigat modelarea comportamentelor în sisteme complexe, cu obiectivul general de a contribui la o mai bună înțelegere a problemelor existente în acest domeniu. Am prezentat soluții de ultimă generație și am dezvoltat noi modele computaționale axate pe algoritmi evolutivi, teoria jocurilor și învățare automată.

Am construit un model experimental ce demonstrează mecanismul natural de optimizare a coloniilor de furnici. Ca și aplicație, am creat o colonie de furnici virtuale 3D și am evidențiat conceptul de emergență.

Pentru dezvoltarea de jocuri educaționale captivante, am propus un nou model de alocare al quest-urilor, bazat pe comportamentul coloniilor de viespi. Am integrat acest model în jocul video Sotirios pentru a optimiza performanța echipelor. Introducerea strategiilor de “gamificare” pentru îmbunătățirea procesului educațional a avut rezultate bune iar unele elemente au fost utilizate pentru a rafina cursul VVSS și pentru a evalua diferite metode de învățare bazată pe jocuri.

Am folosit teoria jocurilor cooperative pentru a detecta rețele de coautori din comunitatea științifică, ceea ce poate ajuta la descoperirea scriitorilor interdisciplinari și a tendințelor de publicare. De asemenea, am dezvoltat o soluție de ultimă generație pentru identificarea clusterelor de noduri în rețelele multipartite. Abordarea noastră descoperă structuri comunitare multipartite folosind o nouă funcție de fitness și optimizare *NoisyEO*. Rezultatele sunt relevante și în alte domenii precum ecologia și etnobiologia [BHBD20].

Folosind o nouă metodă evolutivă, am creat agenți autonomi care pot juca jocul ultimatumului într-un mediu dinamic, prin cele două versiuni de *Ultimatum Game Simulator*, iar pentru a-i ajuta pe studenți să-și îmbunătățească abilitățile tehnice și sociale, am dezvoltat instrumente și activități noi de colaborare.

Dezvoltări viitoare

Îmi propun să extindem cercetarea în dimensiunile emoționale și sociale ale agenților noștri și să modelăm comportamentul social în timpul evenimentelor haotice. Modelele noastre pot ajuta la predicția crizelor economice sau a răspunsurilor emoționale, precum cele observate în timpul pandemiei sau recenta criză bursieră GameStop [Bur21].

Am descoperit noi oportunități în progresul modelelor generatoare, auto-regresive din perspectiva învățării automate și intenționez să extindem tehnici similare și în alte domenii, ar fi psihologia și sănătatea mintală. O posibilă abordare ar putea fi reprezentată de *jocul intern asimetric pentru descoperirea obiectivelor* [OPS⁺21] sau *rețele liniare închise* [VLB⁺19], unde fiecare neuron încearcă să prezică datele de ieșire.

Contribuțiile mele la proiectul *6 Feet Apart* vor continua, iar munca noastră poate fi de un real beneficiu pentru copiii din întreaga lume. Platforma va fi utilizată în școli din mai multe țări, începând cu România și SUA ca programe pilot. Profesorii și grupurile de studenți vor își vor exprima emoțiile pe care le-au trăit în timpul pandemiei, iar impresiile lor vor fi evidențiate în Muzeul Virtual și platforma Mynde.

Bibliography

- [AT05] Gediminas Adomavicius and Alexander Tuzhilin. Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE transactions on knowledge and data engineering*, 17(6):734–749, 2005.
- [Bac00] Alexander Backlund. The definition of system. *Kybernetes*, 2000.
- [BD15] Florentin Bota and Dan Dumitrescu. Agent-based computational models implemented in 3D space. In *Proceedings of Modelling and Development of Intelligent Systems*, pages 13–19, Sibiu, 2015.
- [BDGW11] Melissa Bateson, Suzanne Desire, Sarah E Gartside, and Geraldine A Wright. Agitated honeybees exhibit pessimistic cognitive biases. *Current biology*, 21(12):1070–1073, 2011.
- [BDW03] Sarah F Brosnan and Frans BM De Waal. Monkeys reject unequal pay. *Nature*, 425(6955):297–299, 2003.
- [BdW14] Sarah F Brosnan and Frans BM de Waal. Evolution of responses to (un) fairness. *Science*, 346(6207), 2014.
- [BGLL08] Vincent D Blondel, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. Fast unfolding of communities in large networks. *Journal of statistical mechanics: theory and experiment*, 2008(10):P10008, 2008.
- [BHBD20] Avner Bar-Hen, Pierre Barbillon, and Sophie Donnet. Block models for generalized multipartite networks: Applications in ecology and ethnobiology. *Statistical Modelling*, page 1471082X20963254, 2020.
- [Boc10] Nino Boccara. *Modeling complex systems*. Springer Science & Business Media, 2010.
- [Bot12] Florentin Bota. Game based learning, project Sotirios. In *Proceedings of the second International Students Conference on Informatics*, pages 44–50, Sibiu, Romania, 2012. ICDD 2012.

- [Bot21a] Florentin Bota. *Software Booklet: How to Become a 10X Developer*. Col-orama Printing House, Cluj-Napoca, Romania, 1st edition, 2021.
- [Bot21b] Florentin Bota. Temporal discounting for multidimensional economic agents. *Studia Universitatis Babeş-Bolyai Informatica*, 66(1):86–103, 2021.
- [BS21] Florentin Bota and Dana Simian. Embedding human behavior using multi-dimensional economic agents. In Dana Simian and Laura Florentina Stoica, editors, *Modelling and Development of Intelligent Systems*, pages 3–19, Cham, 2021. Springer International Publishing.
- [BTD98] Eric Bonabeau, Guy Theraulaz, and Jean-Louis Deneubourg. Fixed re-sponse thresholds and the regulation of division of labor in insect societies. *Bulletin of Mathematical Biology*, 60(4):753–807, 1998.
- [Bur21] Riley Burnette. What were the factors that led to the gamestop short squeeze? 2021.
- [Cam11] Colin F Camerer. *Behavioral game theory: Experiments in strategic in-teraction*. Princeton University Press, 2011.
- [Car15] Gerolamo Cardano. *The book on games of Chance: the 16th-century treatise on probability*. Courier Dover Publications, 2015.
- [Car18] Edward Cartwright. *Behavioral economics*. Routledge, 2018.
- [CCBP21] Camelia Chisăliță-Crețu, Florentin Bota, and Andreea-Diana Pop. Soft-ware testing education experiences using collaborative platforms. In *Smart Education and e-Learning 2021*, pages 149–159, Singapore, 2021. Springer Singapore.
- [CCC⁺18] Muffy Calder, Claire Craig, Dave Culley, Richard de Cani, Christl A Donnelly, Rowan Douglas, Bruce Edmonds, Jonathon Gascoigne, Nigel Gilbert, Caroline Hargrove, et al. Computational modelling for decision-making: where, why, what, who and how. *Royal Society open science*, 5(6):172096, 2018.
- [CK94] Shyam R Chidamber and Chris F Kemerer. A metrics suite for object oriented design. *IEEE Transactions on software engineering*, 20(6):476–493, 1994.
- [CLS06] M Keith Chen, Venkat Lakshminarayanan, and Laurie R Santos. How basic are behavioral biases? evidence from capuchin monkey trading be-havior. *Journal of political economy*, 114(3):517–537, 2006.

- [Col80] Kenneth Gerald Collier. Peer-group learning in higher education: The development of higher order skills. *Studies in Higher Education*, 5(1):55–62, 1980.
- [CS01] Vincent A Cicirello and Stephen F Smith. Insect societies and manufacturing. In *The IJCAI-01 Workshop on Artificial Intelligence and Manufacturing, Working Notes*, pages 33–38, 2001.
- [CS04] Vincent A Cicirello and Stephen F Smith. Wasp-like agents for distributed factory coordination. *Autonomous Agents and Multi-agent systems*, 8(3):237–266, 2004.
- [DC10] Satchidananda Dehuri and Sung-Bae Cho. *Knowledge mining using intelligent agents*, volume 6. World Scientific, 2010.
- [DLR10] Marco D’Ambros, Michele Lanza, and Romain Robbes. An extensive comparison of bug prediction approaches. In *2010 7th IEEE Working Conference on Mining Software Repositories (MSR 2010)*, pages 31–41. IEEE, 2010.
- [Dor92] Marco Dorigo. Optimization, learning and natural algorithms. *Ph. D. Thesis, Politecnico di Milano*, 1992.
- [EMT⁺18] Kristina Esopo, Daniel Mellow, Catherine Thomas, Hannah Uckat, Justin Abraham, Prachi Jain, Channing Jang, Nicholas Otis, Michala Riis-Vestergaard, Amanda Starcev, et al. Measuring self-efficacy, executive function, and temporal discounting in kenya. *Behaviour Research and Therapy*, 101:30–45, 2018.
- [F⁺12] J Doyne Farmer et al. Economics needs to treat the economy as a complex system. In *Paper for the INET Conference ‘Rethinking Economics and Politics*, volume 14, 2012.
- [FH⁺01] Martin Fowler, Jim Highsmith, et al. The Agile manifesto. *Software Development*, 9(8):28–35, 2001.
- [For10] Santo Fortunato. Community detection in graphs. *Physics Reports*, 486(3-5):75–174, feb 2010.
- [GBSL17] Noémi Gaskó, Florentin Bota, Mihai Suciu, and Rodica Ioana Lung. Community structure detection in multipartite networks: a new fitness measure. In *Proceedings of the Genetic and Evolutionary Computation Conference, GECCO ’17*, pages 259–265, New York, NY, USA, 2017. Association for Computing Machinery.

- [GBSL20] Noémi Gaskó, Florentin Bota, Mihai Alexandru Suciuc, and Rodica Ioana Lung. A game theoretical analysis of academic writing co-authorship networks. *Journal of Scientometric Research*, 9:319–325, December 2020.
- [GDDG⁺03] Roger Guimera, Leon Danon, Albert Diaz-Guilera, Francesc Giralt, and Alex Arenas. Self-similar community structure in a network of human interactions. *Physical review E*, 68(6):065103, 2003.
- [GR13] Volker Grimm and Steven F Railsback. *Individual-based modeling and ecology*. Princeton university press, 2013.
- [Gua08] Francesco Guala. Paradigmatic experiments: The ultimatum game from testing to measurement device. *Philosophy of Science*, 75(5):658–669, 2008.
- [Har87] David Harel. Statecharts: A visual formalism for complex systems. *Science of computer programming*, 8(3):231–274, 1987.
- [Har89] Sergiu Hart. Shapley value. In *Game Theory*, pages 210–216. Springer, 1989.
- [HNM15] Martie G Haselton, Daniel Nettle, and Damian R Murray. The evolution of cognitive bias. *The handbook of evolutionary psychology*, pages 1–20, 2015.
- [Hol96] John H Holland. *Hidden order: How adaptation builds complexity*. Addison Wesley Longman Publishing Co., Inc., 1996.
- [ISG17] Lung Rodica Ioana, Mihai Suciuc, and Noémi Gaskó. Noisy extremal optimization. *Soft Computing*, 21(5):1253–1270, 2017.
- [KŞ18] Arzum Karataş and Serap Şahin. Application areas of community detection: A review. In *2018 International congress on big data, deep learning and fighting cyber terrorism (IBIGDELFT)*, pages 65–70. Ieee, 2018.
- [Lan09] A. Lancichinetti. Detecting the overlapping and hierarchical community structure in complex networks. *New Journal of Physics*, 11:3, 2009.
- [LLMW16] Xin Liu, Weichu Liu, Tsuyoshi Murata, and Ken Wakita. Community detection in multi-partite multi-relational networks based on information compression. *New Generation Computing*, 34(1-2):153–176, 2016.
- [MAS⁺13] Tomasz P Michalak, Karthik V Aadithya, Piotr L Szczepanski, Balaraman Ravindran, and Nicholas R Jennings. Efficient computation of the

- shapley value for game-theoretic network centrality. *Journal of Artificial Intelligence Research*, 46:607–650, 2013.
- [MSJM16] Daniel Marolt, Jürgen Scheible, Göran Jerke, and Vinko Marolt. Swarm: a self-organization approach for layout automation in analog ic design. *Int. Journal of Electronics and Electrical Engineering (IJEED)*, 4(5):374–385, 2016.
- [MT00] Sendhil Mullainathan and Richard H Thaler. Behavioral economics. Technical report, National Bureau of Economic Research, 2000.
- [NB⁺66] János Neumann, Arthur W Burks, et al. *Theory of self-reproducing automata*, volume 1102024. University of Illinois press Urbana, 1966.
- [New06] Mark EJ Newman. Modularity and community structure in networks. *Proceedings of the national academy of sciences*, 103(23):8577–8582, 2006.
- [NG04] Mark EJ Newman and Michelle Girvan. Finding and evaluating community structure in networks. *Physical review E*, 69(2):026113, 2004.
- [NH10] Vinod Nair and Geoffrey E Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pages 807–814, 2010.
- [NH11] Muaz Niazi and Amir Hussain. Agent-based computing from multi-agent systems to agent-based models: a visual survey. *Scientometrics*, 89(2):479–499, 2011.
- [OPS⁺21] OpenAI OpenAI, Matthias Plappert, Raul Sampedro, Tao Xu, Ilge Akkaya, Vineet Kosaraju, Peter Welinder, Ruben D’Sa, Arthur Petron, Henrique Ponde de Oliveira Pinto, et al. Asymmetric self-play for automatic goal discovery in robotic manipulation. *arXiv preprint arXiv:2101.04882*, 2021.
- [Par05] Kihong Park. The Internet as a complex system., 2005.
- [Pes95] Umberto Pesavento. An implementation of von neumann’s self-reproducing machine. *Artificial Life*, 2(4):337–354, 1995.
- [Pic07] Gualtiero Piccinini. Computational modelling vs. computational explanation: Is everything a turing machine, and does it matter to the philosophy of mind? *Australasian Journal of Philosophy*, 85(1):93–115, 2007.
- [RN95] Stuart J Russell and Peter Norvig. *Artificial intelligence: a modern approach*. Prentice Hall, New Jersey, 1995.

- [RTON13] David G Rand, Corina E Tarnita, Hisashi Ohtsuki, and Martin A Nowak. Evolution of fairness in the one-shot anonymous ultimatum game. *Proceedings of the National Academy of Sciences*, 110(7):2581–2586, 2013.
- [SB18] Dana Simian and Florentin Bota. Adaptive multi-agent system based on wasp-like behaviour for the virtual learning game sotirios. In Ivan Lirkov and Svetozar Margenov, editors, *Large-Scale Scientific Computing*, pages 416–424, Cham, 2018. Springer International Publishing.
- [SB19] Susan Shapiro and Simona Baciu. *The Teacher Within: A Mindful Journey Toward Well-Being For Teachers In The 21St Century*. Studio Impress Design, 2019.
- [SB20] Camelia Serban and Florentin Bota. A conceptual framework for software fault prediction using neural networks. In Dana Simian and Laura Florentina Stoica, editors, *Modelling and Development of Intelligent Systems*, pages 171–186, Cham, 2020. Springer International Publishing.
- [Sim09] Dana Simian. Wasp based algorithms and applications. *Proceedings of Int. Conf. Modelling and Development of Intelligent Systems, Sibiu*, pages 229–235, 2009.
- [SM11] Marco Santos and Carlos Martinho. Wasp-like scheduling for unit training in real-time strategy games. In *Proceedings of the AAAI Conference on Artificial Intelligence and Interactive Digital Entertainment*, volume 6, pages 195–200, 2011.
- [SSMP08] Dana Simian, Corina Simian, Ioana Moisil, and Iulian Pah. Computer mediated communication and collaboration in a virtual learning environment based on a multi-agent system with wasp-like behavior. In *International Conference on Large-Scale Scientific Computing*, pages 618–625. Springer, 2008.
- [SSS10] Dana Simian, Florin Stoica, and Corina Simian. Optimization of complex svm kernels using a hybrid algorithm based on wasp behaviour. In Ivan Lirkov, Svetozar Margenov, and Jerzy Waśniewski, editors, *Large-Scale Scientific Computing*, pages 361–368, Berlin, Heidelberg, 2010. Springer Berlin Heidelberg.
- [TBD98] Guy Theraulaz, Eric Bonabeau, and Jean-Louis Deneubourg. Response threshold reinforcement and division of labour in insect societies. *Proceedings: Biological Sciences*, 265(1393):327–332, 1998.

- [TBGD91] G. Theraulaz, E. Bonabeau, J. Gervet, and J. I. Demeubourg. Task differentiation in policies wasp colonies. a model for self-organizing groups of robots. from animals to animats. In *Proceedings of the First International Conference on Simulation of Adaptive behavior*, pages 346–355, 1991.
- [Tha00] Richard H Thaler. From homo economicus to homo sapiens. *Journal of economic perspectives*, 14(1):133–141, 2000.
- [TLS⁺13] Barbara Tomasino, Lorella Lotto, Michela Sarlo, Claudia Civai, Rino Rumiati, and Raffaella Ida Rumiati. Framing the ultimatum game: the contribution of simulation. *Frontiers in human neuroscience*, 7:337, 2013.
- [TTT96] Charles S Taber, Richard J Timpone, and Richard J Timpone. *Computational modeling*, volume 113. Sage, 1996.
- [VLB⁺19] Joel Veness, Tor Lattimore, David Budden, Avishkar Bhoopchand, Christopher Mattern, Agnieszka Grabska-Barwinska, Eren Sezener, Jianan Wang, Peter Toth, Simon Schmitt, et al. Gated linear networks. *arXiv preprint arXiv:1910.01526*, 2019.
- [VNM07] John Von Neumann and Oskar Morgenstern. *Theory of games and economic behavior (commemorative edition)*. Princeton university press, 2007.
- [WB20] Julia Watzek and Sarah Brosnan. Decision-making biases in animals: A critical review. 2020.
- [WS98] Duncan J Watts and Steven H Strogatz. Collective dynamics of ‘small-world’ networks. *nature*, 393(6684):440–442, 1998.