

Aalto-yliopisto
Perustieteiden korkeakoulu
Teknillisen fysiikan ja matematiikan tutkinto-ohjelma

Vangin dilemma häiriöisessä ympäristössä Markov-prosessina

kandidaatintyö
11.11.2012

Lasse Lindqvist

Työn saa tallentaa ja julkistaa Aalto-yliopiston avoimilla verkkosivuilla.
Muilta osin kaikki oikeudet pidätetään.

AALTO-YLIOPISTO PERUSTIETEIDEN KORKEAKOULU PL 11000, 00076 Aalto http://www.aalto.fi	KANDIDAATINTYÖN TIIVISTELMÄ	
Tekijä: Lasse Lindqvist		
Työn nimi: Vangin dilemma häiriöisessä ympäristössä Markov-prosessina		
Tutkinto-ohjelma: Teknillisen fysiikan ja matematiikan tutkinto-ohjelma		
Pääaine: Systeemitieteet	Pääaineen koodi: F3010	
Vastuopettaja(t): Harri Ehtamo		
Ohjaaja(t): Kimmo Berg		
Tiivistelmä: <p>Työssä tutkitaan erästä toistetun vangin dilemman häiriöistä mallia, jossa pelaajat havaitsevat vain omat siirtonsa ja saavuttamiensa pisteiden realisaation. Pelaajat eivät siten havaitse suoraan vastapelaajan siirtoa, vaan tämän ja normaalijakautuneen häiriön summan.</p> <p>Työn tuloksena saadaan selville, että yksinkertainen tavoitearvoon perustuva strategia pystyy pitämään yllä yhteistyötä pettävää strategiaa vastaan tietyin ehdoin. Työssä esitetään ehdot, joilla yhteistyön ylläpito on mahdollista ja tutkitaan miten tämä riippuu parametreista.</p> <p>Lisäksi työssä selvitetään, kuinka tavoitestrategia kehittyy yksinkertaisen evoluution seurauksena. Työssä esitetään ehdot, joista nähdään, milloin tavoitestrategia luisuu hiljalleen pettäväksi strategiaksi, ja milloin ei. Lisäksi löydetään tilanteet, jolloin tavoitestrategia kehittyy sopiviin tavoitearvoihin, jotka kykenevät pitämään yllä yhteistyötä satunnaisia pettäviä yksilöitä vastaan.</p>		
Päivämäärä: 12.12.2012	Kieli: suomi	Sivumäärä: 22+2
Avainsanat: evoluutio, häiriö, Markov-prosessi, peliteoria, vangin dilemma		

Sisältö

1	Johdanto	1
2	Aikaisempi tutkimus	2
3	Tutkimusongelma ja -menetelmät	5
3.1	Tavoite-strategia	6
3.2	Tilanteiden luokittelu	6
3.3	Toistettu peli Markov-prosessina	8
3.3.1	Tasapainojakauma	9
4	Tulokset	10
4.1	Käypä alue	10
4.2	Käyvän alueen muoto ja paikka	12
4.2.1	Käyvän alueen äärikohtat	12
4.3	Tavoitearvon evoluutio	15
5	Yhteenveto	16
A	Tasapainojakauma erityistapauksessa	21

1 Johdanto

Tässä työssä tutkitaan toistetun vangin dilemman käyttäytymistä simuloinnin avulla. Toistetussa vangin dilemmassa pelaajat voivat pelata vangin dilemma -nimistä peliä toistuvasti toisiaan vastaan $n:n$ pelaajan populaatiossa sekä päättää valinnoistaan edellisiin lopputuloksiin perustuen. Toistetun vangin dilemman tutkimuksen perusteos on Robert Axelrodin *The Evolution of Cooperation* vuodelta 1984. Kirjassaan Axelrod laajensi kolme vuotta aiemmin julkaisemaansa artikkeliaan sekä järjesti ohjelmointikilpailun, jossa kuka tahansa saattoi lähettää kilpailuun vangin dilemmaa pelaavan ohjelman. Kilpailun voitti silmä silmästä -strategia. Tämän jälkeen Axelrod järjesti myös toisen kilpailun, johon pystyttiin ottamaan oppia ensimmäisen turnauksen tuloksista. Myös tämän voitti silmä silmästä -strategia.[1]

Axelrod esitti tutkimustensa perusteella periaatteet, joita menestyvän strategian piti noudattaa. Strategia ei saa olla ilkeä, mikä tarkoittaa sitä, että se ei saa koskaan pettää ensimmäisenä, jos molemmat tekevät yhteistyötä. Strategia ei saa olla kateellinen, eli se ei saa yrittää pärjätä vastustajaansa paremmin. Vastustajan ylittäminen vangin dilemmassa johtaa usein vain turhiin kostonkierteisiin ja vähentää molempien osapuolien pisteitä. Strategian pitää olla vastavuoroinen. Strategian täytyy kostaa yhteistyöstä poikkeaminen sekä vastata yhteistyöhön yhteistyöllä. Jos strategia ei kosta, sitä voidaan käyttää hyväksi pettämällä. Strategia ei saa olla liian älykäs, sillä usein liika monimutkaisuus saavuttaa käytännössä heikompiä tuloksia kuin yksinkertaiset säännöt, jotka johtavat molemminpuoliseen yhteistyöhön.[1]

Tässä työssä tutkitaan toistettua vangin ongelmaa häiriöisessä ympäristössä. Pelaajien valintojen lisäksi mukana on satunnainen häiriö. Pelaaja ei voi olla varma vastapelaajan valinnasta, sillä pelaajat havannoivat vain lopputuloksen häiriön lisäämisen jälkeen. Vaikka häiriö onkin jatkuva, pelaajan valinta on yhä diskreetti valinta yhteistyön ja pettämisen väliltä.

Hypoteesina on, että hyvin pienellä häiriöllä ongelma muistuttaa häiriötöntä vangin dilemmaa. Kun häiriön suuruus kasvaa, pelaajan mahdollisuus tunnistaa vastapelaajan käytös pitäisi heikentyä, jolloin yhteistyön tulisi hypoteesin mukaan kadota.

Lopuksi tutkitaan erilaisten populaatioiden kehittymistä evoluution kautta simuloimalla. Tutkitaan, voiko pieni pettävä klusteri tuhota yhteistyötä harastavan populaation tai toisinpäin. Aikaisempien tutkimusten perusteella on syytä olettaa, että populaatiot käyttäytyvät usein siten, että silmä silmästä -strategia valtaa pettävän populaation. Yhä anteeksiantavammat stra-

	C	D
C	R, R	S, T
D	T, S	P, P

Taulukko 1: Tulomatriisi tavalliselle vangin dilemmalle.

tegiat taas valtaavat tiukemmat strategiat, kunnes pettävät stragiat lopulta valtaavat nämä kivat strategiat, jotka eivät osaa kostaa. Tällaisen kivi-paperi-sakset dynamiikan olemassaolosta on saatu havaintoja aikaisemmissa töissä.[17]

2 Aikaisempi tutkimus

Vangin dilemman kaltaista peliä peliteoreettisena pulmana käsittelevät ensimmäisenä Merrill Flood ja Melvin Dresher RAND:ssa. Nimen vangin dilemma ja version, jossa pisteet esitetään vankilatuomioina ovat peräisin Albert Tuckerilta, joka halusi tuoda ajatukset helpommin ymmärrettäviksi psykologeille.[2]

Vangin dilemma on tavallisesti kahden pelaajan peli, joka voidaan esittää taulukon 1 mukaisen tulomatriisin avulla.

Tässä pätee $T > R > P > S$. Usein asetetaan lisäksi ehto $2R > S + T$, mikä aiheuttaa sen, että jatkuva yhteistyö on parempi vaihtoehto kuin S ja T vuorotellen. Vaihtoehto C tarkoittaa tavallisesti *cooperation* eli yhteistyön tekemistä ja vaihtoehto D *defection* pettämistä. Koska $T > R$ ja $P > S$ vaihtoehto D dominoi C :tä, joten molempien on järkevintä pettää, jos peliä pelataan yhden kierroksen verran. Kuitenkin molemmille olisi parempi, jos he tekisivät yhteistyötä, sillä $R > P$.

Yhden kierroksen pelissä paras strategia on siten pettäminen. Toistetussa eli usean kierroksen pelissä tämä ei välttämättä pidä paikkaansa. Asiaa tutki muun muassa Robert Axelrod järjestämässään turnauksissa. Hän järjesti turnauksen, jossa erilaiset strategiat pelasivat toisiaan vastaan useiden kierrosten ajan. Eniten pisteitä ensimmäisessä turnauksessa keräsi Anatol Rapoportin lähettämä tit-for-tat-strategia (TFT), joka tekee yhteistyötä ensimmäisellä kierroksella ja toistaa jatkossa vastustajan edellisen siirron. Axelrod julkaisi turnauksen tulokset ja järjesti toisen turnauksen selvittääkseen, onko joku löytänyt paremman strategian. Tit-for-tat voitti tämänkin. Sen menestys perustuu siihen, että se saavuttaa suuria pistemääriä yhteistyöhalukkai-

den strategioiden kanssa molemminpuolisella yhteistyöllä, muttei jää ilkeiden strategioiden jalkoihin, koska vastaa pettämiseen pettämisellä.[1]

Vaikka tit-for-tat vaikuttaa hyvältä kandidaatilta parhaaksi strategiaksi toistetussa vangin dilemmassa, voidaan osoittaa, että sekään ei ole evolutiivisesti stabiili strategia. Voidaan myös osoittaa, että mikään puhdas strategia, eli strategia joka ei käytä todennäköisyyksiä siirron valinnassa, ei voi olla evolutiivisesti stabiili.[3]

Jos ympäristöön lisätään jonkinlainen häiriö, TFT voi helposti joutua itseään toistavaan kostonkierteeseen niin itsensä kuin muidenkin kanssa yhden kierroksen virheen takia. Tästä voidaan parantaa erilaisilla TFT:n muunnoksilla. Häiriöön tottumattomassa populaatiossa anteeksiantava tit-for-tat (T2FT) on osoittautunut toimivaksi. Anteeksiantava TFT voi versiosta riippuen antaa aina yhden pettämisen anteeksi tai minkä tahansa pettämisen anteeksi ennalta määrätyllä todennäköisyydellä. Häiriöön sopeutuneessa populaatiossa taas menestystä on saavuttanut katuva tit-for-tat (CTFT). Katuva tit-for-tat aloittaa yhteistyöllä ja jatkaa sitä, elleivät molemmat pelaajat petä samanaikaisesti. Jos vastapelaaja pettää kerran CTFT kostaa pettämisen takaisin, ellei itse aloittanut pettämistä edellisellä kierroksella. Tällöin CTFT antaa pettämisen anteeksi, jolloin on mahdollisuus päästä irti pettämisen kierteestä.[4]

Vallitseva näkemys on, että häiriön ollessa mahdollinen menestyvän strategian täytyy olla jonkin verran anteeksiantava vastapelaajan pettämiselle. Per Molander osoitti vuonna 1985, että TFT:n ja C:n sekoitus lähestyy pistekeskisarvoltaan R:ää, kun virheen todennäköisyys lähestyy nollaa. Keskenään strategiat menestyvät suurilla C:n määrillä, mutta tulevat samalla haavoittuvaisiksi hyväksikäytölle. Molanderin laskelmien mukaan, kun pettämisen jälkeen tekee yhteistyötä todennäköisyydellä $\min\{1 - \frac{T-R}{R-S}, \frac{R-P}{T-P}\}$, anteliaat strategiat saavat korkeimmat pistemäärät ilman, että tiukat strategiat menestyvät näitä vastaan paremmin kuin tavallinen TFT. Tätä strategiaa merkitään joskus GTFT.[5]

Kun esimerkiksi vastapelaajan siirron tulkinnan virheen todennäköisyys lähenee 50 prosenttia, TFT lähenee satunnaista strategiaa. Tällaiseen strategiaan taas pelkkä pettäminen on paras vastaus. Peter Kollockin simulaatio näyttäisi osoittavan että suurilla häiriötodennäköisyyksillä tiukkuus olisi anteeksiantavuutta parempi strategia.[6] Kuten Bendor, Kramer ja Swistak vastauksessaan kuitenkin huomauttavat, simulaatiossa käytetyt eivät ole kaikkia strategioita edustavia, ja siksi tuloksiin tulee suhtautua varauksella.[7]

Toistettuja pelejä epätäydellisillä havainnoilla on tutkittu viime aikoina pal-

jon. Niille on myös omat folk-teoreemansa. Folk-teoreemaa toistetuille peleille yksityisellä ja epätäydellisellä havainnoinnilla ovat tutkineet muun muassa Hörner ja Olszewski sekä Matsushima.[8, 9]

Toinen merkittävä suuntaus häiriöllisille ympäristöille ovat Pavlov-pohjaiset strategiat. Yksinkertainen Pavlov aloittaa yhteistyöllä ja pitää aina saman siirron kuin edellisellä, jos on saavuttanut pistemääräksi R tai T . Muutoin Pavlov vaihtaa siirtoa.[10] Monimutkaisemmat n -Pavlovit tekevät yhteistyötä todennäköisyydellä k/n , missä n on ennalta valittu taso. Arvoa k korjataan aina ylöspäin, kun on saavutettu R tai T ja muutoin alaspäin. Pavlov-pohjaiset strategiat menestyvät hyvin kivoja strategioita vastaan ja jopa hyväksikäyttävät liian avokätisiä. Toisaalta todella ilkeät strategiat, kuten pelkkä D , hyväksikäyttävät Pavlovia. Kun $n > 2$, Pavlov on vakaa useimpien strategioiden kanssa häiriöisessäkin ympäristössä. Kraines ja Kraines ehdottavat, että n -Pavlov arvoilla 3 tai 4 alkutiloilla $3/3$ ja $4/4$ olisivat usein ideaalisia, koska ne oppivat tarpeeksi nopeasti ollen samalla tarpeeksi laajoja.[11]

Usein kohtaaminen toisen kanssa ei ole täysin satunnaista, vaan ihminen tai eläin vuorovaikuttaa samojen yksilöiden kanssa useammin kuin muiden. Tällaisen kuvaamiseen voidaan jossain määrin käyttää spatiaalista vangin dilemmaa. Siinä pelaajat ovat jonkilaisessa avaruudellisessa järjestyksessä ja pelaavat vangin dilemmaa naapurustonsa kanssa. Jokaisen kierroksen jälkeen kaikki pelaajat valitsevat seuraavaksi strategiakseen aina sen, jota tämän menestyksekkäin naapuruston jäsen on käyttänyt. Tämä voidaan ajatella joko niin, että menestyksekkäin pelaaja valloittaa populaatiota tai siten, että jokainen pelaaja valitsee aina strategian, jota pitää parhaana. Avaruudellisella vangin dilemmalla voidaan koettaa selittää tapahtumaa, jossa pelaajat valitsevat todennäköisemmin vastapelaajikseen samaa strategiaa pelaavia pelaajia, kun pelaajien paritus ei ole satunnaista. Tällöin yhteistyön olisi helpompi syntyä.[12, 13]

Käytännössä vangin dilemma ei yleensä ole diskreetti tapaus pettämisen ja yhteistyön väliltä. Jatkuvassa vangin dilemmassa pelaava voi valita vapaasti yhteistyön asteen. Kun pelaaja korottaa panostaan, hänen pisteensä laskevat, mutta parin pisteiden keskiarvo kasvaa. Toinen pelaaja saa siis annetun panoksen jollain yhtä suuremmalla kertoimella itselleen. Tämä tapaus on vähemmän tutkittu kuin diskreetti, vaikka päteekin paremmin useampaan tosielämän tilanteeseen. Jatkuvassa tapauksessa pettävä tasapainopiste on evolutiivisesti stabiili, joten tavallisesti sitä ei voida poistaa. On myös olemassa yhteistyöstrategia, joka saavuttaa korkeimman pistemäärän pelatessaan itseään vastaan. Tämä sama strategia on myös attraktori, tosin epästabiili sellainen. Liian avokätisten ja pettävien strategioiden yhdistelmä voi saada

tämän tasapainopisteen epästabiiliksi.[14]

Lindi M. Wahl ja Martin A. Nowak (1999) totesivat, että häiriöttömässä tapauksessa ensimmäisellä siirroilla oli suuri vaikutus strategian pitkäaikaiseen menestykseen. Tutkimuksen mukaan strategiat, jotka vastaavat yhteistyöhön herkästi, mutta eivät siedä hyväksikäyttöä, pystyvät liikkumaan kohti yhteistyöhaluista Nashin tasapainoa.[15] Häiriöisessä tapauksessa havainnoitiin erityisesti, että yhteistyö on evolutiivisesti epästabiilia. Yhteistyötä harrastavat strategiat menettävät asemansa yhä anteliaammille strategioille, kunnes pettävät strategiat valtaavat nämä liian anteliaat strategiat. Tästä systeemi voi taas palautua alkuperäisen yhteistyön asteelle, jolloin sykli alkaa alusta. Vaihtoehtoisesti systeemi voi saavuttaa vakaan tasapainotilan, jossa kaikki kolme strategiaryhmää ovat tasapainossa.[16] Samanlaisen kivi-sakset-paperi-dynamiikkaan ovat päätyneet myös Nowak ja Sigmund diskreetissä tapauksessa.[17]

3 Tutkimusongelma ja -menetelmät

Tutkimusongelmana on diskreetti vangin dilemma häiriöisessä tapauksessa yhdessä mahdollisista malleista. Pelaajat valitsevat tavallisesti, tekevätkö yhteistyötä vai pettävätkö. Pelaajat saavat pisteensä taulukon 1 mukaan.

Yleensä pelaajia rankaistetaan pettämiskiarteeseen joutumisesta, sillä vangin dilemassa tavallisesti on ehtona myös $T + S < R + R$. Näin ollen vuorottaisesta T:n ja S:n pistemäärän saavuttamisesta ei saa yhtä suurta hyötyä kuin jatkuvasta R:stä.

Näihin pistemääriin lisätään normaalijakautunut virhetermi. Pelaajat havaitsevat vasta tämän virhetermin ja alkuperäisen pistemäärän summan. Normaalijakauman käyttö on luonnollista, koska sillä on taipumus esiintyä virheiden jakaumana. Normaalijakauman keskiarvoparametriksi on valittu nolla, jolloin virheet eivät ole suuntautuneet keskimäärin kumpaankaan suuntaan. Jos keskiarvo olisi muu kuin nolla, kaikkien historiaan pohjautuvien sääntöjen parametrit voisi siirtää virheen keskiarvoa vastaavalla määrällä, jolloin tapaus palautuisi nollakeskiarvoisen virheen tapaukseen. Nolla keskiarvona on muutenkin luonteva valinta. Häiriön suuruutta säädellään muuttamalla virheen varianssia. Eri pelaajille annettavat virhetermit ovat simulaatioissa myös toisistaan riippumattomia, sillä riippuvuudelle ei ole perusteita. Malli on eräänlainen välimuoto jatkuvan ja diskreetin mallin väliltä.

3.1 Tavoite-strategia

Tutkitaan pelkkää pettämistä pelaavan D- ja molempia pelaavan tavoite-strategian menestystä toisiaan vastaan populaatiossa, jossa ei ole muita strategioita. D-strategia ei muutu tavallisesta vangin dilemmasta. Se pelaa koko ajan vaihtoehtoa D. Tavoite-strategia toteutetaan siten, että se tekee yhteistyötä ensimmäisellä kierroksella ja jatkossa silloin, jos on edellisellä kierroksella havainnut ennalta valittua tavoitetta k suuremman arvon. Strategiaa voidaan täten merkitä G_k (Goal).

D saavuttaa itseään vastaan aina P määrän pisteitä kierrokselta riippumatta häiriön suuruudesta. Sen sijaan G:n menestys itseään ja D:tä vastaan riippuu häiriön suuruudesta ja G:n valitsemasta rajasta. Näitä voidaan tutkia esimerkiksi simuloimalla peliä suurella kierrosmäärällä. Koska virhetermin keskiarvo on nolla, G saavuttaa D:tä vastaan aina korkeintaan pistemäärän P, koska saa virheen takaa vain pistemääriä S ja P. Samoin D saavuttaa vain pistemääriä P ja T. Itseään vastaan G taas saavuttaa kaikkia pistemääriä, mutta keskimäärin kuitenkin P:n ja R:n väliltä.

3.2 Tilanteiden luokittelu

Kahden strategian menestys toisiaan vastaan voidaan jakaa neljään luokkaan. Tehdään tämä nyt G- ja D-strategioille. Merkitään G:n D:tä vastaan saavuttamaa pistemäärää symbolilla P_{G-D} . Merkitään myös muita pistemääriä samoin.

Tapauksessa yksi pätee $P_{G-D} < P_{D-D}$ ja $P_{G-G} < P_{D-G}$. Tällöin strategia D voittaa G:n jokaisella populaatiokombinaatiolla, sillä se saa paremmat pistemäärät niin itseään kuin TFT:täkin vastaan. Toisessa tapauksessa $P_{G-D} > P_{D-D}$ ja $P_{G-G} > P_{D-G}$. Nyt G voittaa jokaisella kombinaatiolla. Tätä tilannetta ei kuitenkaan strategioilla G ja D pitäisi ilmetä, sillä G ei voi koskaan pärjätä D-strategiaa vastaan tätä itseään paremmin, jos häiriötä ei lasketa mukaan loppupisteisiin. Jos häiriön antamat pisteet lasketaan mukaan, voi G saavuttaa kuitenkin tämän tilanteen.

Kolmannessa tapauksessa pätee $P_{G-D} < P_{D-D}$ ja $P_{G-G} > P_{D-G}$. Nyt molemmat strategiat pärjäävät itsensä kanssa parhaiten ja strategioille löytyy jokin populaatiotiheys p , jota suuremmilla arvoilla strategia pärjää paremmin ja jota pienemmilla huonommin. Neljännessä tilanteessa pätee $P_{G-D} > P_{D-D}$ ja $P_{G-G} < P_{D-G}$. Nyt strategiat pärjäävät paremmin toisiaan vastaan. Tällaisessa tilanteessa löytyy jokin populaatiotiheys p , jota suuremmil-

la arvoilla strategia pärjää huonommin ja jota pienemmällä paremmin. Kuten toista tilannetta, tätäkään ei voi esiintyä kovin usein, sillä ensimmäinen ehto saavutetaan vain, kun häiriö antaa G:lle sattumalta enemmän pisteitä kuin ottaa pois.

Kolmannelle ja neljännelle tapaukselle on yhteistä tietty tasapainopiste p , joka voidaan määrittää strategioiden saavuttamista pistemääristä. Kolmannessa tapauksessa tasapainopiste on epästabiili satulapiste, josta populaatio hakeutuu pois. Neljännessä tapauksessa tasapainopiste on stabiili ja populaatiot hakeutuvat siihen. Tasapainopiste voidaan määrittää mille tahansa kahdelle, kun merkitään, että molempien strategioiden saavuttama pistemäärä on sama. Jos G:n populaatiotiheys on p , ja D:n siten $1-p$, saadaan

$$pP_{G-G} + (1-p)P_{G-D} = pP_{D-G} + (1-p)P_{D-D}. \quad (1)$$

Yhtälöstä 1 voidaan ratkaista p , jolloin saadaan

$$p = \frac{P_{G-D} + P_{D-D}}{P_{G-G} + P_{D-D} - P_{D-G} - P_{G-D}}. \quad (2)$$

Jos menestyksen mittana pidetään saavutettuja pistemääriä ja populaatiotiheydet muuttuvat menestyksen mukaan, tapausten luokittelu määrää täysin, mihin lopputulokseen populaatio päätyy, kun lähdetään tietystä populaatiosta. Tämä voidaan esittää yhtälömuodossa seuraavasti

$$\frac{\dot{n}_D}{n_D} = k(p(P_{D-G} - P_{G-G}) + (1-p)(P_{D-D} - P_{G-D})), \quad (3)$$

missä p on strategian G osuus populaatiosta, k jokin positiivinen kerroin ja n_D strategian D tiheys populaatiossa. Voidaan merkitä strategioiden saavuttamien pistemäärien erotusta funktiona p :stä

$$\frac{\dot{n}_D}{n_D} = k(P_{dif,D-G}(p)). \quad (4)$$

Funktion määritelmästä nähdään, että arvo riippuu lineaarisesti muuttujasta p . Tapauksessa 1 erotusfunktio on aina positiivinen, joten nähdään, että D:n muutos on aina positiivinen. Tällöin strategia G katoaa. Tapauksessa 2 taas erotus on aina negatiivinen, strategia G valtaa populaation ja D häviää pois. Tapauksessa 3 funktio on vähenevä. Strategia D pärjää tällöin sitä paremmin, mitä pienempi p on. Voidaan myös etsiä funktion nollakohta, jonka jälkeen

\dot{n}_D on negatiivinen. Tämä nollakohta on epästabiili tasapainopiste. Tämä tarkoittaa myös sitä, että jos populaatio koostuu aluksi vain yhdestä strategiasta ja siihen lisätään mielivaltaisen pienen populaatiotiheyden omaava toinen strategia, uusi strategia kuolee pois. Tämä tarkoittaa sitä, että populaation perustaja valtaa aina populaation takaisin, jos oletetaan että uuden strategian tiheys on mielivaltaisen pieni. Tapauksessa 4 funktio on kasvava, jolloin strategia D pärjää sitä paremmin, mitä suurempi p on. Voidaan myös etsiä funktion nollakohta, jonka jälkeen \dot{n}_D on positiivinen. Tämä nollakohta on stabiili tasapainopiste. Tässä tapauksessa populaatio päättyy aina tasapainopisteeseen riippumatta lähtötiheyksistä.

3.3 Toistettu peli Markov-prosessina

Koska käsiteltävien strategioiden muisti on korkeintaan yhden kierroksen pituinen, peli voidaan mallintaa Markov-prosessina. Prosessin tilat ovat toteutuneet tilanteet, C-C, D-C, C-D ja D-D. Tilat voidaan nimetä pelaajan 1 näkökulmasta R, T, S ja P saavutettujen pistemäärien perusteella. Johdetaan nyt tilansiirtomatriisi.

Merkitään strategian i tavoitetasoa H_i . Todennäköisyys, että pelaaja havaitsee arvoa H_i suuremman pistemäärän, riippuu sen todennäköisyydestä, että toteutunut häiriö on pienempi kuin erotus $H_i - K$, missä K on alkuperäinen pistemäärä pelaajalle. Häiriö on normaalijakautunut odotusarvolla 0, jolloin todennäköisyys riippuu vain erotuksesta ja jakauman varianssista. Esimerkiksi kiinteällä varianssilla todennäköisyys, että pelaaja havaitsee tavoitearvoaan pienemmän arvon, kun tila on ollut R, on $P_R = P(X \leq H_i - R)$. Nyt todennäköisyys, että tilasta R siirrytään tilaan P on $P(X \leq H_1 - R)P(X \leq H_2 - R) = P_{R1}P_{R2}$, kun molemmat strategiat ovat tavoitestrategioita. Jos häiriön varianssia merkitään σ :lla, saadaan matriisi

$$A_\sigma = \begin{pmatrix} (1 - P_{R1})(1 - P_{R2}) & P_{R1}(1 - P_{R2}) & P_{R2}(1 - P_{R1}) & P_{R1}P_{R2} \\ (1 - P_{T1})(1 - P_{S2}) & P_{T1}(1 - P_{S2}) & P_{S2}(1 - P_{T1}) & P_{S2}P_{T1} \\ (1 - P_{T2})(1 - P_{S1}) & P_{S1}(1 - P_{T2}) & P_{T2}(1 - P_{S1}) & P_{S1}P_{T2} \\ (1 - P_{P1})(1 - P_{P2}) & P_{P1}(1 - P_{P2}) & P_{P2}(1 - P_{P1}) & P_{P1}P_{P2} \end{pmatrix}.$$

Tilansiirtomatriisista nähdään suoraan muutama asia. Jos häiriön varianssi kasvaa rajatta, kaikki todennäköisyydet lähestyvät arvoa $\frac{1}{4}$ ja tavoiterajan valinta muuttuu merkityksettömäksi. Matriisi siis lähestyy matriisia

$$A_\infty = \begin{pmatrix} 0,25 & 0,25 & 0,25 & 0,25 \\ 0,25 & 0,25 & 0,25 & 0,25 \\ 0,25 & 0,25 & 0,25 & 0,25 \\ 0,25 & 0,25 & 0,25 & 0,25 \end{pmatrix}.$$

Matriisista nähdään suoraan, että jokaisen tilan todennäköisyys on 0,25. Toisaalta D-strategiaa voidaan mallintaa tavoite-strategiana, jonka tavoite on äärettömän suuri. Jos pelaaja 2 pelaa D-strategiaa, saadaan kaikille tiloille $P_2 = 1$ ja matriisi muuttuu muotoon

$$A_\sigma = \begin{pmatrix} 0 & 0 & (1 - P_{R1}) & P_{R1} \\ 0 & 0 & (1 - P_{T1}) & P_{T1} \\ 0 & 0 & (1 - P_{S1}) & P_{S1} \\ 0 & 0 & (1 - P_{P1}) & P_{P1} \end{pmatrix}.$$

Nähdään, että vain tilat S ja P toteutuvat, kuten kuuluukin olla.

Viimeisenä huomiona todetaan, että jos tavoitestategia pelaa itseään vastaan, tilansiirtomatriisiksi saadaan

$$A_\sigma = \begin{pmatrix} (1 - P_R)^2 & P_R(1 - P_R) & P_R(1 - P_R) & P_R^2 \\ (1 - P_T)(1 - P_S) & P_T(1 - P_S) & P_S(1 - P_T) & P_S P_T \\ (1 - P_T)(1 - P_S) & P_S(1 - P_T) & P_T(1 - P_S) & P_S P_T \\ (1 - P_P)^2 & P_P(1 - P_P) & P_P(1 - P_P) & P_P^2 \end{pmatrix}.$$

Matriisista nähdään, että jos tavoite on välillä $T > R > H > P > S$ ja häiriön varianssi lähestyy nollaa, todennäköisyys poistua tilasta R lähenee nollaa. Samalla $P_R, P_T \rightarrow 1$ ja $P_S, P_P \rightarrow 0$. Nähdään, että todennäköisyys poistua tiloista S ja T tiloihin P ja R pienenee nopeammin kuin todennäköisyys päätyä tiloihin S ja T. Jos tila kuitenkin sattuu muuttumaan S:ksi tai T:ksi, niissä vietetään käytännössä koko loppuaika todennäköisyyksillä 0,5. Samanlaisella päättelyllä voidaan todeta, että jos pätee $P > H$, kaikki aika vietetään tilassa R. Vastaavasti, jos $R > H$, kaikki aika vietetään tilassa P, kun häiriön varianssi lähenee nollaa.

3.3.1 Tasapainojakauma

Tutkitaan, miten eri tilojen todennäköisyydet käyttäytyvät, kun kierrosten lukumäärä lähestyy ääretöntä. Markov-ketjujen konvergenssiteoreeman mu-

kaan, kun tilansiirtomatriisi A on pelkistymätön ja jaksoton ja sillä on tasapainojakauma π , pätee seuraava lause.[19]

Kun $n \rightarrow \infty$, $p^{n(x,y)} \rightarrow \pi(y)$.

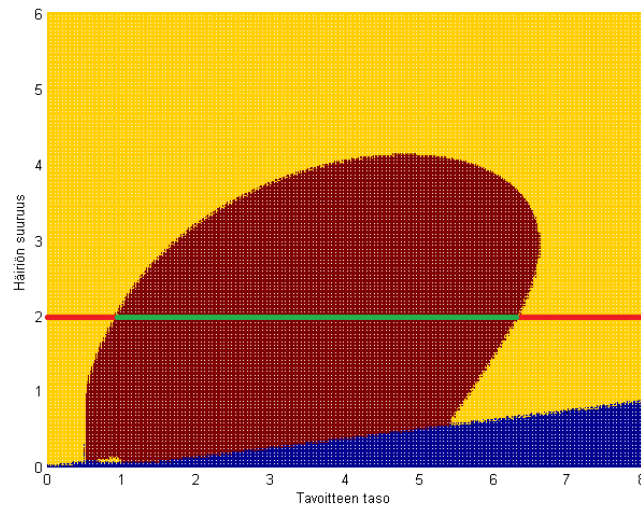
Toisin sanottuna tilatodennäköisyydet lähestyvät tasapainojakaumaa. Nyt tarvitsee vain määrittää tilansiirtomatriisin tasapainojakauma. Kun tilansiirtomatriisi on aikariippumaton, pelkistymätön ja jaksoton, tasapainojakauma saadaan seuraavasti.[19] Vähennetään tilansiirtomatriisista ensin yksikkömatriisi ja muutetaan kaikki viimeisen sarakkeen alkiot ykkösiksi. Tämän jälkeen otetaan saadusta matriisista käänteismatriisi ja tilojen todennäköisyydet voidaan lukea käänteismatriisin alimmalta riviltä. Matriisin kääntäminen johtaa hyvin monimutkaiseen lausekkeeseen, joten matriiseja ei ole mielekästä esittää valmiissa muodossa, vaan kääntämisen voi toteuttaa numeroarvoilla. Liitteessä A on esitetty kuitenkin erityistapaus, kun toisen tavoitestrategiaa pelaavan tavoite on mielivaltaisen korkea tai mielivaltaisen pieni. Nämä tilanteet vastaisivat tapausta, jossa vastapelaajan strategia on pelkkä D tai C.

4 Tulokset

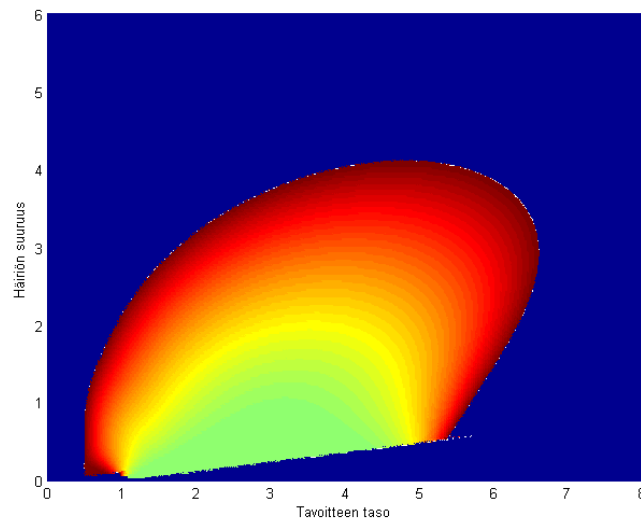
4.1 Käypä alue

Tutkitaan tavoitestrategian menestystä D-strategiaa vastaan. Valitaan aluksi pisteille jotkin arvot. Käytetään nyt arvoja $T=5$, $R=3$, $P=1$, $S=0$.

Kuvassa 1 nähdään tilanne, kun tasapainojakaumien avulla on laskettu tavoitestrategian ja D-strategian pisteiden odotusarvot ja luokiteltu tilanteet edellä esiteltyihin luokkiin. Nämä on laskettu erilaisilla tavoitteen ja häiriön arvoilla, jolloin saadaan kaksiulotteinen kuva tilanteesta. Nähdään, että normaalisti D-strategia saavuttaa paremmat pisteet kaikilla populaatiosuhteen p arvoilla. Tätä esittää keltainen alue. On kuitenkin tietty alue, punaisella merkitty, jossa tavoitestrategia päihittää D:n, kun tavoitestrategian osuus populaatiosta on tarpeeksi suuri. Kutsutaan tätä aluetta nimellä *käypä alue*. Kuvassa on myös sininen alue, joka johtuu laskentatarkkuuden rajoista. Huomioitavaa on, että tiettyä häiriön suuruutta suuremmilla arvoilla niin sanottua käypää aluetta ei enää ole. Myös tavoitesuunnassa alue ulottuu vain oikeiden ilman häiriötä saatavien pistemäärien läheisyyteen. Kuvaan 1 on merkitty häiriön suuruuden 2 kohdalle viiva. Jos nyt peliä pelataan näillä parametreilla ja häiriön suuruus on 2, tavoitestrategia pystyy selviytymään



Kuva 1: Tavoitestrategian menestys. Punaiset alueet ovat tilanteen 3 ja keltaiset tilanteen 2 mukaisia. Siniset alueet ovat seurausta numeerisen laskennan rajoitteista.



Kuva 2: Suureen p arvo. Vihreät alueet ovat lähellä nollaa. Punaiset alueet ovat lähellä arvoa 1. Sinisten alueiden kohdalla ei ole käypää aluetta lukuunottamatta alareunan numeerisen epätarkkuuden kohtaa.

D-strategiaa vastaan vain silloin, jos tavoitteen arvo on vihreän viivan kohdalla. Tässä tapauksessa nähdään, että sopivat arvot tavoitteelle rajoittuvat suurin piirtein välille $(1, 6\frac{1}{2})$.

Raja tavoitestrategian populaatio-osuudelle, jonka jälkeen tavoitestrategia päihittää D-strategian, voidaan laskea kaavan 1 mukaisesti. Nämä tulokset on esitetty kuvassa 2. Kuvasta nähdään, että alueen reunoilla tarvitaan suurempi määrä tavoitestrategiaa suhteessa D-strategiaan, jotta tavoitestrategia voisi menestyä. Alueen sisällä taas riittää hyvinkin lähellä nollaa oleva arvo.

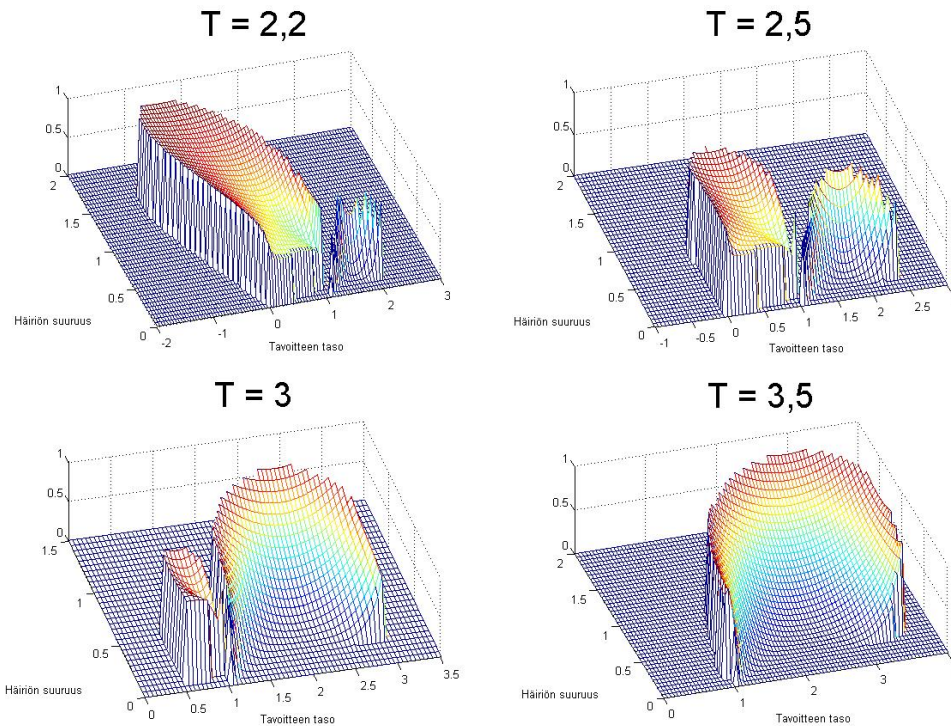
4.2 Käyvän alueen muoto ja paikka

Mielenkiinnon kohteena on heti edellisen tuloksen selvittyä tämän käyvän alueen ominaisuudet. Halutaan esimerkiksi selvittää, löytyykö alue kaikilla $R:n$, $T:n$, $S:n$ ja $P:n$ arvoilla. Koska ensimmäisen arvon valinta asettaa vain nollatason ja toisen arvon valinta skaalauksen, jäljelle jää kahden arvon vaihtelu. Asetetaan $S=0$ ja $P=1$. Tutkitaan nyt käyvän alueen ominaisuuksia vaihtelemalla $R:n$ ja $T:n$ arvoja.

Kuvassa 3 on esitetty käyvän alueen muoto erilaisilla $T:n$ arvoilla, kun $R:n$ arvoksi on valittu 2. Kuvista nähdään, että alueen yleinen muoto säilyy samantyyppisenä parametrin muutoksista huolimatta. Alueen kärki kuitenkin vaihtaa paikkaa. Kun $T:n$ arvo on lähellä $R:n$ arvoa, kärki on vasemmalla ja kun arvo on paljon suurempi kuin R , kärki on oikealla. $T:n$ ja $R:n$ arvojen erotus kertoo käytännössä siitä, kuinka houkuttelevaa on poiketa jatkuvasta yhteistyöstä pettämiseen korkeampien pisteiden toivossa. Kun erotus on pieni, houkutuskin on pieni. Tällöin suurilla häiriön arvoilla käyvän alueen tavoitearvot ovat matalimmat. Tämän voi käsittää niin, että houkutuksen ollessa pieni, on kannattavaa varmistaa, että yhteistyöstä ei vahingossa poiketa pettämiseen. Suurilla erotuksen arvoilla asia on taas päinvastoin. Näillä tietyillä parametreilla nähdään myös ilmiö, jossa alue surkastuu pois samalla kun tilalle kasvaa toinen alue.

4.2.1 Käyvän alueen äärikohtat

Voidaan myös tutkia, miten käyvän alueen maksimikohtat riippuvat $T:n$ ja $R:n$ arvoista. Lasketaan käypä alue erinäisillä arvoilla ja otetaan talteen suurin häiriön arvo sekä maksimi- ja minimiarvo tavoitteen tasolle. Kuvissa 4, 5 ja 6 on esitetty tulokset.



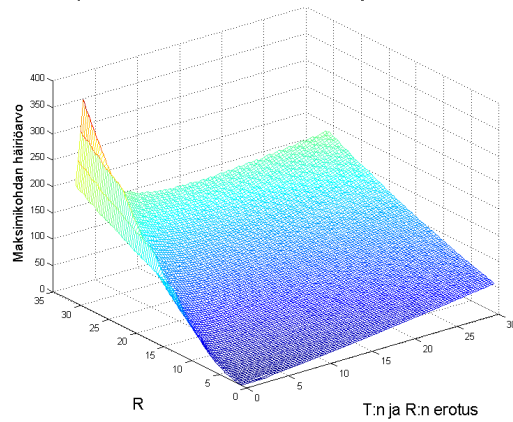
Kuva 3: Käyvän alueen muoto eri T :n arvoilla, kun $R = 2$.

Kuvasta 4 voidaan havaita, että käyvän alueen koko häiriösuunnassa riippuu eniten R :n suuruudesta, mutta myös T :n ja R :n erotuksen arvolla on merkitystä. Suurilla R :n arvoilla havaitaan, että häiriön maksimiarvo kasvaa eksponentiaalisen nopeasti, kun erotus pienenee.

Kuvasta 5 nähdään, että käyvän alueen oikean reunan ulottuminen tavoitteen suunnassa riippuu lähes samalla lailla T :n ja R :n erotuksesta ja R :n suuruudesta. Tavoitteen maksimiarvo on molempien muuttujien suhteen kasvava. Huomattavissa on myös tavallista korkeampi kasvu suurilla R :n ja pienillä erotuksen arvoilla.

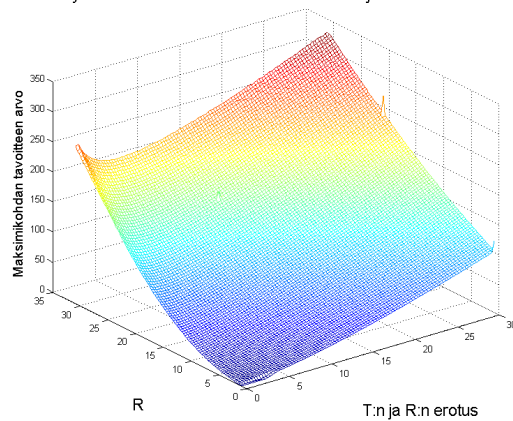
Käyvän alueen tavoitteen miniarvon riippuvuus on kuvassa 6. Tässä riippuvuus on erilainen. Suuressa osassa aluetta arvo on nollan tuntumassa. Tämän lisäksi kuvaaajassa on kuilu, jossa minimiarvo on pieni ja pienenee kuilun perää kohden. Tämä kuilu on kuitenkin pieni verrattuna tavoitteen maksimiarvoihin, jotka nähdään kuvassa 5. Yhdistämällä kuvien 5 ja 6 informaatio voidaan todeta, että käyvän alueen suuruus tavoitesuunnassa kasvaa likimain kuten häiriösuunnassakin.

Käyvän alueen suurin häiriöarvo T:n ja R:n funktiona



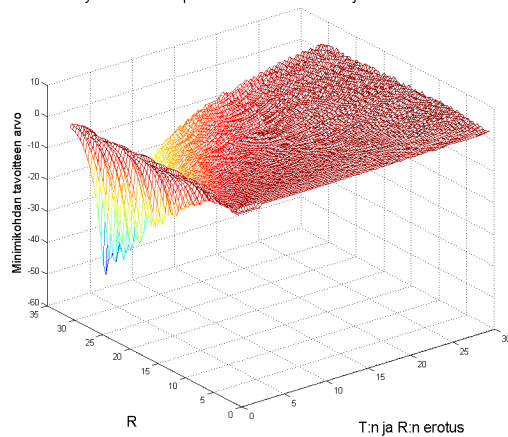
Kuva 4: Käyvän alueen maksimi-arvo häiriön suunnassa riippuu eniten T:n ja R:n erotuksen arvosta.

Käyvän alueen suurin tavoitearvo T:n ja R:n funktiona



Kuva 5: Käyvän alueen maksimi-arvo tavoitteen suunnassa riippuu vahvasti sekä T:n että T:n ja R:n erotuksen arvosta.

Käyvän alueen pienin tavoitearvo T:n ja R:n funktiona



Kuva 6: Käyvän alueen minimi-arvo tavoitteen suunnassa on yhtä uomaa lukuunottamatta lähes samassa kohdassa.

4.3 Tavoitearvon evoluutio

Käyvän alueen lisäksi mielenkiinnon kohteena voi olla tavoitearvon muutos. Voidaan tutkia, mitkä käyvän alueen valinnat tavoitearvolle pärjäävät muille strategioille. Analysoinnin mahdollistamiseksi tilannetta täytyy rajata. Rajoitutaan tarkastelemaan pelkkiä tavoitestratgioita.

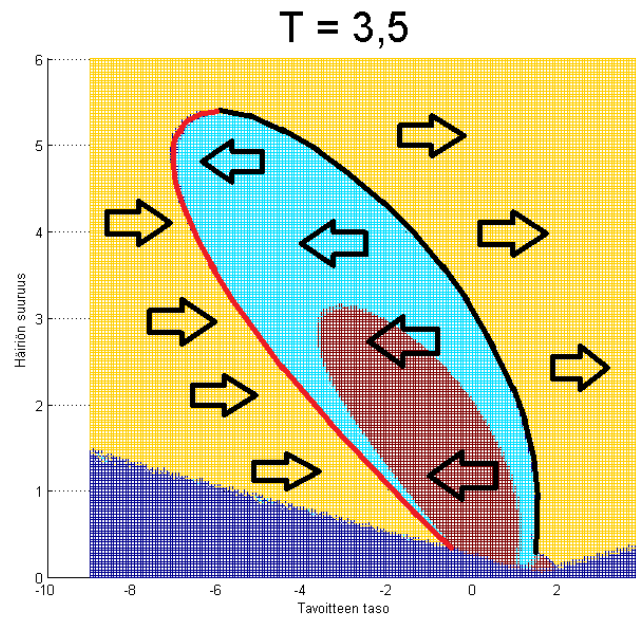
Sopeutuvan dynamiikan aksioomien mukaisesti voidaan lisätä muutama muu oletus. Oletetaan, että mutaatioita tapahtuu suhteellisesti niin harvoin, että uuden mutantin ilmaantuessa hallitseva populaatio on jo tasapainossa. Toiseksi oletetaan, että fenotyyppiin vaikuttavat mutaatiot ovat pieniä, mutta satunnaisia. Siispä riittää tarkastella strategian menestystä vain naapuristrategioita vastaan.[18] Tässä tapauksessa naapuristrategia on strategia, jonka rajan arvo poikkeaa vain vähän hallitsevan strategian arvosta.

Tarkastellaan siis tilannetta, jossa strategian tavoitearvo on jokin k . Tutkitaan, miten strategia menestyy suhteessa toiseen strategiaan, jonka tavoite on $k + \epsilon$. Jos jompi kumpi strategia voittaa toisen kaikilla populaatiosuhteilla, voidaan ajatella, että evoluutio kulkisi siinä pisteessä voittavan strategian suuntaan. Tutkitaan arvossa k muutoksen suuntaa vertaamalla tavoitestratgijan menestystä arvoilla $k - \frac{\epsilon}{2}$ ja $k + \frac{\epsilon}{2}$. Näin saadaan jokaiselle arvolle selvitettyä muutoksen suunta tai suunnan puute. Käytännössä joudutaan valita jokin äärellinen ϵ numeerisia laskutoimituksia varten.

Vertaamalla saatuja suuntia laskettuihin käypiin alueisiin, voidaan tutkia, mitä tapahtuu käyvän alueen tavoitestratgioille, jos ne alistetaan kilpailulle naapuristrategioidensa kanssa. Voidaan tutkia, ajautuuko tavoitearvo ulos käyvältä alueelta ylä- tai alapuolelta, vai onko alueella jokin alue, johon tavoitearvo pysähtyy. Jos käyvällä alueella on jokin kohta, johon tavoitearvo ajautuisi niin oikealta kuin vasemmaltakin puolelta, kutsutaan tätä tasapainokohdaksi. Jos evoluutio käyttäytyisi niin, että uusi valloittaja olisi aina tavoitestratgiaa pelaava hieman eri tavoitteella kuin edellinen, tilanne ajautuisi lopulta juurikin tähän tasapainopisteeseen.

Kuvassa 7 on esitetty eräillä parametreilla, miten tavoitestratgia käyttäytyy. Kuvassa keltainen alue on ei-käypää aluetta, jolla evoluution suunta on korkeampia tavoitearvoja kohden. Vaaleansinisellä ja punaisella alueella evoluutio kulkee pienempiä tavoitearvoja kohden, mutta vain punainen alue on käypää. Aloituspisteestä riippuen evoluutio johtaa siten tässä tapauksessa joko mielivaltaisen suuriin tavoitearvoihin tai punaisella merkittyyn tasapainoalueeseen.

Kuvissa 8 ja 9 on esitetty tasapainoalueen esiintyminen $R:n$ ja erotuksen



Kuva 7: Tavoitestrategian evoluutio.

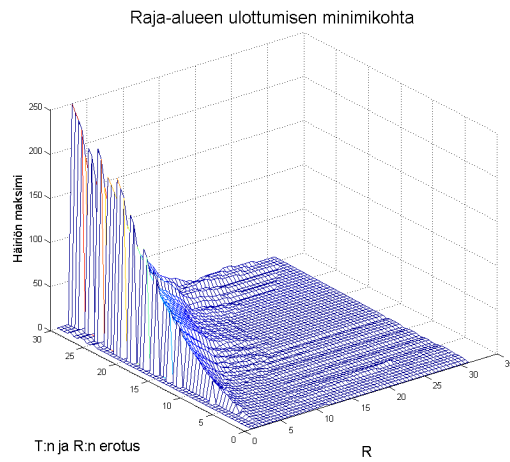
funktiona. Kuvista nähdään se, että maksimiarvo kasvaa samalla lailla kuten käyvän alueen arvo kuvassa 4. Tämä on odotettu tulos.

Mielenkiintoista on pienillä arvoilla nähtävä alue, jolla raja-alueita ei esiinny. Tämä alue rajoittuu 1:tä pienemmille erotuksen arvoille. Kun tämä pieni raja-arvo ylittyy, raja-alueen pituus kasvaa hypähtäen. Ilmiö seuraa käyvän alueen koon riippuvuutta, joka on esitetty kuvassa 4. Kokonaisuudessaan raja-alueen koko riippuu tietysti käyvän alueen koosta. Raja-alue ei voi olla tätä pidempi. Siten kuvan 10 havainnollistama riippuvuus muuttujista R ja T :n ja R :n erotus on helppo ymmärtää. Käyvän alueen koko kasvaa nimittäin samansuuntaisen riippuvuuden mukaan.

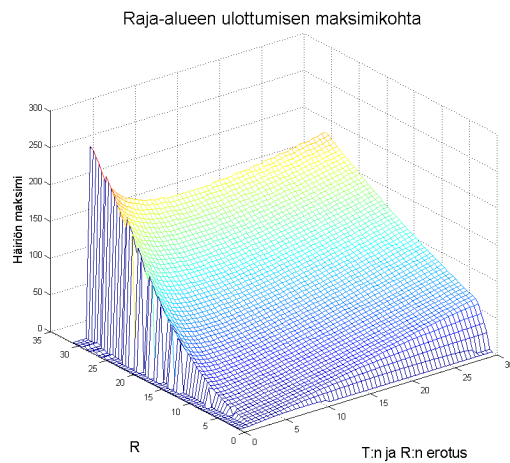
Minimikohdan kuvaajassa nähdään pieniä heilahteluja, jotka johtuvat numeerisen laskennan epätarkkuudesta.

5 Yhteenveto

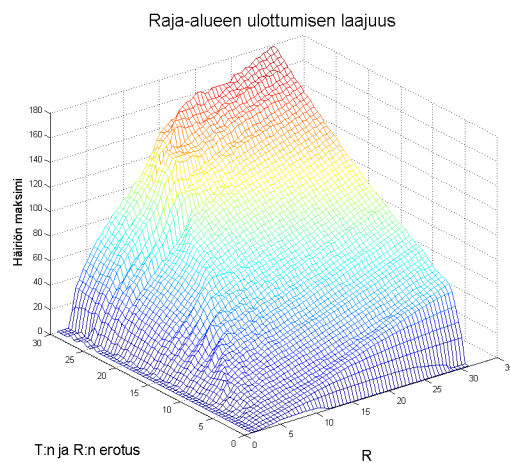
Vangin dilemma häiriöisessä ympäristössä pystyttiin mallintamaan Markov-prosessina ja tutkittiin tavoitestrategian menestystä D-strategiaa vastaan. Huomattavaa on, että kaikilla tutkituilla parametrijhdistelmillä löytyy alue,



Kuva 8: Tasapainoalueen pienimmän arvon kohta häiriön suunnassa R:n ja T:n ja R:n erotuksen funktiona.



Kuva 9: Tasapainoalueen suurimman arvon kohta häiriön suunnassa R:n ja T:n ja R:n erotuksen funktiona.



Kuva 10: Tasapainoalueen laajuus häiriön suunnassa R:n ja T:n ja R:n erotuksen funktiona.

jolla tavoite-strategia menestyy D-strategiaa vastaan, kun tavoitestategian osuuspopulaatiosta on tarpeeksi suuri tai häiriö tarpeeksi pieni. Samoin nähdään, että kaikilla tutkituilla parametriyhdistelmillä löytyy maksimaalinen häiriö, jota suuremmilla häiriöillä tavoitestategia ei voi menestyä.

Tutkittiin myös tavoite-strategian tavoitearvon mahdollista evoluutiota. Tilannetta rajaamalla voitiin osoittaa, että useimmilla parametriyhdistelmillä tavoitearvo ajautuu sopivasta alkuarvosta lähtiessä alueelle, jolla se pystyy vastustamaan pettävää strategiaa. Löydettiin myös parametrien arvot, joilla tavoitestategia ei ole evoluutio huomioon ottaen kestävä vaihtoehto, vaan ajautuu sellaisiin tavoitearvoihin, joilla pettävä strategia valloittaa lopulta populaation.

Jatkossa tutkimuksen kohteena voisi olla hieman hienostuneempi strategia, joka ottaisi huomioon myös omat valintansa. Tässä tutkittu tavoitestategia hyödyntää vain saamansa pisteinformaation, muttei ollenkaan omaa valintahistoriaansa. Voitaisiin tutkia strategiaa, joka oman siirtonsa huomioiden valitsisi vaatimansa rajan kahdesta vaihtoehdosta. Tällainen strategia olisi yhä hyvin yksinkertainen, mutta luultavasti robustimpi kuin nyt tutkittu strategia.

Viitteet

- [1] Axelrod, Robert. The Evolution of Cooperation. Basic Books, 1984. New York.
- [2] Kuhn, Steven. Prisoner's Dilemma. The Stanford Encyclopedia of Philosophy (Spring 2009 Edition), Edward N. Zalta (ed.), <http://plato.stanford.edu/archives/spr2009/entries/prisoner-dilemma/>.
- [3] Boyd, Robert and Lorberbaum, Jeffrey P. (1987) No pure strategy is evolutionarily stable in the repeated Prisoner's Dilemma game *Nature* 327, 58-59.
- [4] Axelrod, Robert, Wu, Jiangzhong. (1995) How To Cope With Noise in the Iterated Prisoner's Dilemma *Journal of Conflict Resolution* 39, Issue 1, 183-189
- [5] Molander, Per. (1987) The optimal level of generosity in a selfish, uncertain environment *Journal of Conflict Resolution* 31, Issue 4, 692-724
- [6] Kollock, Peter. (1993) 'An Eye for an Eye Leaves Everyone Blind': Cooperation and Accounting Systems. *American Sociological Review* 58, 768-86.
- [7] Bendor, Jonathan and Kramer, Roderick and Swistak, Piotr, (1996) Cooperation Under Uncertainty: What is New, What is True and What is Important? *American Sociological Review* 61, 333-338.
- [8] Matsushima, Hitoshi, (2004) Repeated Games with Private Monitoring: Two Players *Econometrica*, Vol. 72, No. 3, 823-852
- [9] Hörner, Johannes and Olszewski, Wojciech, (2006) The Folk Theorem for Games with Private Almost-Perfect Monitoring *Econometrica*, Vol. 74, No. 6, 1499-1544
- [10] Nowak, Martin and Karl Sigmund, (1993) A Strategy of Win-stay, Lose-shift that Outperforms Tit-for-tat in the Prisoner's Dilemma Game *Nature* 364, 56-58.
- [11] Kraines, David and Kraines, Vivian, (1993) Learning to Cooperate with Pavlov: an Adaptive Strategy for the Iterated Prisoner's Dilemma with Noise *Theory and Decision* 35, 107-150.
- [12] Szabo, György and Toke, Csaba, (1997) Evolutionary prisoner's dilemma game on a square lattice *Physical Review E* 58, Issue 1, 69-73

- [13] Szabo, György and Vukov, Jeromos and Szolnoki, Attila (2005) Phase diagrams for an evolutionary prisoner's dilemma game on two-dimensional lattices *Physical Review E* 72, 69-73.
- [14] Le, Stephen and Boyd, Robert, (2007) Evolutionary Dynamics of the Continuous Iterated Prisoners' Dilemma *Journal of Theoretical Biology* 245, 258–267.
- [15] Wahl, Lindi M., Nowak, Martin A., (1999a.) The continuous Prisoner's dilemma: I. linear reactive strategies *Journal of Theoretical Biology* 200, 307–321.
- [16] Wahl, Lindi M., Nowak, Martin A., (1999b.) The continuous Prisoner's dilemma: II. linear reactive strategies with noise *Journal of Theoretical Biology* 200, 323–338.
- [17] Nowak, Martin and Sigmund, Karl (1989) Oscillations in the Evolution of Reciprocity *Journal of Theoretical Biology* 137, 21-26
- [18] Geritz, S.A.H. et al (1996) Evolutionarily Singular Strategies and the Adaptive Growth and Branching of the Evolutionary Tree *Evolutionary Ecology* 12, 35-57
- [19] Durrett, Rick. *Essentials of Stochastic Processes*, 2010.

A Tasapainojakauma erityistapauksessa

Lasketaan tasapainojakauma, kun toisen pelaajan tavoitearvo on mielivaltaisen suuri. Tällöin saadaan, että $P_{R2}, P_{T2}, P_{S2}, P_{P2} = 1$. Tilansiirtomatriisi muuttuu muotoon

$$A_\sigma = \begin{pmatrix} 0 & 0 & (1 - P_{R1}) & P_{R1} \\ 0 & 0 & (1 - P_{T1}) & P_{T1} \\ 0 & 0 & (1 - P_{S1}) & P_{S1} \\ 0 & 0 & (1 - P_{P1}) & P_{P1} \end{pmatrix}.$$

Vähennetään tästä yksikkömatriisi, muutetaan neljännen sarakkeen alkion ykkösiksi ja otetaan käänteismatriisi. Saadaan tasapainomatriisi

$$\pi = \begin{pmatrix} -1 & 0 & \frac{-P_{P1}+P_{R1}}{1-P_{P1}+P_{S1}} & \frac{1-P_{R1}+P_{S1}}{1-P_{P1}+P_{S1}} \\ 0 & -1 & \frac{P_{P1}-P_{T1}}{1+P_{P1}-P_{S1}} & \frac{1+P_{S1}-P_{T1}}{1-P_{P1}+P_{S1}} \\ 0 & 0 & \frac{-1+P_{P1}-P_{S1}}{-1+P_{P1}} & \frac{1-P_{P1}+P_{S1}}{P_{S1}} \\ 0 & 0 & \frac{-1+P_{P1}-P_{S1}}{-1+P_{P1}-P_{S1}} & \frac{1-P_{P1}+P_{S1}}{1-P_{P1}+P_{S1}} \end{pmatrix}.$$

Nyt alimmalta riviltä nähdään, että tilojen tasapainotodennäköisyydet ovat $S = \frac{-1+P_{P1}}{-1+P_{P1}-P_{S1}}$ ja $P = \frac{P_{S1}}{1-P_{P1}+P_{S1}}$.

Samoin saadaan, kun toisen pelaajan tavoitearvo on mielivaltaisen pieni. Nyt $P_{R2}, P_{T2}, P_{S2}, P_{P2} = 0$. Tilansiirtomatriisi muuttuu muotoon

$$A_\sigma = \begin{pmatrix} (1 - P_{R1}) & P_{R1} & 0 & 0 \\ (1 - P_{T1}) & P_{T1} & 0 & 0 \\ (1 - P_{S1}) & P_{S1} & 0 & 0 \\ (1 - P_{P1}) & P_{P1} & 0 & 0 \end{pmatrix}.$$

Vähennetään tästä yksikkömatriisi, muutetaan neljännen sarakkeen alkion ykkösiksi ja otetaan käänteismatriisi. Saadaan tasapainomatriisi

$$\pi = \begin{pmatrix} \frac{-1-P_{P1}+P_{T1}}{1+P_{R1}-P_{T1}} & \frac{P_{P1}-P_{R1}}{1+P_{R1}-P_{T1}} & 0 & 1 \\ \frac{-P_{P1}+P_{T1}}{-P_{P1}+P_{T1}} & \frac{-1+P_{P1}-P_{R1}}{-1+P_{P1}-P_{R1}} & 0 & 1 \\ \frac{1+P_{R1}-P_{T1}}{-P_{P1}+P_{S1}} & \frac{1+P_{R1}-P_{T1}}{P_{P1}-P_{S1}} & -1 & 1 \\ 1 + \frac{P_{R1}}{-1-P_{R1}+P_{T1}} & \frac{P_{R1}}{1+P_{R1}-P_{T1}} & 0 & 0 \end{pmatrix}.$$

Nyt alimmalta riviltä nähdään, että tilojen tasapainotodennäköisyydet ovat $R = 1 + \frac{P_{R1}}{-1-P_{R1}+P_{T1}}$ ja $T = \frac{P_{R1}}{1+P_{R1}-P_{T1}}$.