

Multiagent Learning: dynamic games & applications

Citation for published version (APA):

Hennes, D. (2013). Multiagent Learning: dynamic games & applications. Maastricht: Maastricht University.

Document status and date:

Published: 01/01/2013

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

Summary

More and more contemporary technological challenges require decentralized solutions. Examples include the coordination of multiple autonomous vehicles and automated control systems for air traffic. The demand for multiagent systems is widespread and originates from a collection of significant advantages. Multiagent systems have built-in redundancies that allow for robustness, i.e., if one or a few components fail the system is still able to operate properly. In addition, different tasks can be assigned to independent agents pursuing these goals in parallel. Multiagent systems are modular per definition and components can be easily added or removed. While a multiagent approach offers an elegant solution paradigm to issues like fault tolerance and load balancing, it also introduces new challenges. The presence of multiple agents results in a highly dynamic and nondeterministic environment. Interactions between agents and with the environment inevitably cause very complex systems, which are not trivial to control. In order to facilitate adaptation to changes in the environment and other agents, *learning* is crucial.

This dissertation investigates the dynamics of learning in multiagent systems and is organized in two parts. The first part investigates evolutionary game theory as a formal framework to model the dynamics of multiagent learning. The second part focuses on the application of these models to real world problems. Part I commences with the formal concepts and key results of multiagent learning. We define the fundamental models of single- and multiagent learning and give an introduction to two well-known reinforcement learning algorithms, *Q*-learning and learning automata. Subsequently, we survey the most important multiagent learning algorithms. We continue by explaining the relation between evolutionary game theory and multiagent reinforcement learning. In particular, we show the formal relation between games of learning automata and the asymmetric continuous time replicator dynamics, as well as the link between *Q*-learning and the selection-mutation dynamics. This dissertation makes two important contributions to the methodology of modeling multiagent reinforcement learning by evolutionary processes. First, we prove that a large collective of learning agents can be approximated by the symmetric continuous time replicator dynamics. Second, we extend the link between multiagent reinforcement learning and evolutionary game theory to multi-state games.

Part II of this dissertation focuses on the dynamics of multiagent learning in real world domains. In particular, we investigate three domains: complex negotiation scenarios, markets, and collision avoidance. Complex negotiations involve strategic interactions in settings that include multiple agents and multiple actions as well as interrelated goals, tasks and resources. As the complexity of these large-scale strategic inter-

actions grows exponentially with the number of agents, we introduce metastrategies as a method to reduce the strategy space significantly. We show that we can preserve many of the strategic characteristics in the reduced setting, which makes it possible to use population games to model the dynamics of such interactions. The competitive advantage of price signal information for traders in auctions is the subject of the second domain study. More information does not guarantee higher performance, in particular, traders with limited information perform below market average and are outperformed by random traders; only insiders beat the market. We analyze the market dynamics with an evolutionary model of agents with competing information levels. Results show that the highest information level will dominate if information comes for free. If information is costly, less-informed traders may prevail reflecting a more realistic distribution over information levels. In the third domain, we perform an evolutionary analysis of collision avoidance in multiagent systems based on the velocity obstacle paradigm. The velocity obstacle is a geometric representation of all velocities that will eventually result in a collision given that all dynamic obstacles maintain their observed velocities. Three variants of the velocity obstacle description are compared and the most robust technique under evolutionary pressure is identified.

To conclude, this dissertation contributes to the understanding of multiagent learning dynamics in several ways. First, we provide an alternative interpretation of the single population replicator dynamics. In particular, we derive that if fitness proportionate selection is used, the overall population average of a population of learners moves independently of the adaptation process of individuals; it follows the symmetric continuous time replicator dynamics. Second, we derive a dynamical model of multi-state multiagent learning. The state-coupled replicator dynamics extend the existing evolutionary framework of multiagent learning and allow to predict learning dynamics in stochastic games. Finally, we show how evolutionary models of learning can be used to analyze complex scenarios and predict behavior, thereby linking the theoretical concepts to everyday problems.

About the author: Daniel Hennes was born on July 4th, 1984 in Duisburg, Germany. He graduated *summa cum laude* from Maastricht University with a B.Sc. in Knowledge Engineering and Computer Science (August 2007) and a M.Sc. in Artificial Intelligence (June 2008). In July 2008, he started his Ph.D. research at Eindhoven University of Technology. From September 2009 onwards, he continued his research at the Department of Knowledge Engineering at Maastricht University. From September 2007 to January 2008 and February 2010 to June 2010, he worked at Oregon State University as a visiting research scholar in the area of Dynamics and Control. From September 2010 to December 2010, he was a research intern at Willow Garage, a robotics research lab and technology incubator in Menlo Park, California.

Samenvatting

Steeds meer hedendaagse technologische uitdagingen vereisen decentrale oplossingen. Voorbeelden hiervan zijn onder andere de coördinatie van meerdere autonome voertuigen en geautomatiseerde regelsystemen voor het luchtverkeer. De vraag naar multi-agentsystemen is wijdverbreid en komt voort uit enkele belangrijke voordelen. Multi-agentsystemen hebben ingebouwde redundanties die leiden tot robuustheid; als één of enkele onderdelen falen is het systeem nog steeds in staat om goed te functioneren. Daarnaast kunnen verschillende taken worden toegewezen aan individuele agenten die deze doelen parallel nastreven. Multi-agentsystemen zijn per definitie modulair en componenten kunnen eenvoudig worden toegevoegd of verwijderd. Naast het feit dat multi-agentsystemen een elegante oplossingsmethode bieden voor problemen zoals fouttolerantie en load balancing, worden er ook nieuwe uitdagingen geïntroduceerd. De aanwezigheid van meerdere agenten leidt tot een zeer dynamische en complex omgeving. Interacties tussen agenten en met de omgeving leiden onvermijdelijk tot zeer complexe systemen die niet triviaal te controleren zijn. Om aanpassing aan veranderingen in de omgeving en andere agenten te vergemakkelijken, is *leren* van groot belang.

Dit proefschrift onderzoekt de dynamiek van leergedrag in multi-agentsystemen en is georganiseerd in twee delen. In het eerste deel onderzoeken we evolutionaire speltheorie als een formeel kader om de dynamica van *multiagent learning* te modelleren. In het tweede deel richten we ons op de toepassing van deze modellen binnen praktische domeinen. Deel I begint met de formele definities en de belangrijkste resultaten met betrekking tot multiagent learning. We introduceren de fundamentele modellen van single- en multiagent learning en geven een inleiding tot twee bekende *reinforcement learning* algoritmen, *Q-learning* en *learning automata*. Hierna geven we een overzicht van de belangrijkste algoritmen voor multiagent learning. Vervolgens zetten we de relatie tussen evolutionaire speltheorie en multiagent reinforcement learning uiteen. In het bijzonder tonen we de formele relatie aan tussen het leergedrag van learning automata en de *asymmetric continuous time replicator dynamics*, alsmede het verband tussen *Q-learning* en selectie-mutatie dynamica. Dit proefschrift draagt op twee manieren bij aan de methodiek van het modelleren van multiagent reinforcement learning doormiddel van evolutionaire processen. Ten eerste bewijzen we dat een grote populatie van lerende agenten kan worden benaderd door de symmetric continuous time replicator dynamics. Ten tweede breiden we de relatie tussen multiagent reinforcement learning en evolutionaire speltheorie uit naar *multi-state games*.

Deel II van dit proefschrift richt zich op de toepassing van deze dynamische modellen binnen praktische domeinen. We onderzoeken drie domeinen: complexe on-

derhandelingsscenario's, aandelenmarkten, en *collision avoidance*. Complexe onderhandelingen zijn strategische interacties in scenario's die meerdere agenten, vele mogelijke acties, en verbanden tussen doelen, taken en middelen bevatten. Omdat de complexiteit van deze strategische interacties exponentieel groeit met het aantal agenten, introduceren we *meta-strategies* als een methode om de complexiteit beduidend in te perken. We laten zien dat veel van de strategische kenmerken van het originele scenario behouden blijven binnen deze gereduceerde representatie, wat ons in staat stelt om de dynamiek van deze interacties te modelleren met behulp van evolutionaire spellen. Het competitief voordeel van prijssignaal-informatie voor handelaren op de aandelenmarkt is het onderwerp van de tweede praktische toepassing. Het hebben van meer informatie is geen garantie voor betere prestaties, zo presteren met name handelaren met beperkte informatie onder het marktgemiddelde en worden zij overtroffen door handelaren die willekeurig opereren; alleen insiders zijn in staat om de markt te verslaan. We analyseren de dynamiek van de aandelenmarkt met een evolutionair model van agenten met verschillende informatieniveaus. Uit de resultaten blijkt dat het hoogste informatieniveau zal domineren zolang informatie vrij beschikbaar is. Als informatie kostbaar is, kunnen ook minder geïnformeerde handelaren voortbestaan, met een meer realistische verdeling over informatieniveaus tot gevolg. Tot slot voeren we een evolutionaire analyse uit van *collision avoidance* in multi-agentsystemen op basis van de *velocity obstacle* methode. Het *velocity obstacle* is een geometrische weergave van alle snelheden die uiteindelijk tot een botsing zullen leiden, aangenomen dat het dynamische obstakel zijn waargenomen snelheid handhaaft. Drie varianten van de *velocity obstacle* methode worden vergeleken en de meest robuuste aanpak wordt bepaald middels een aantal experimenten.

Concluderend, dit proefschrift draagt op meerdere manieren bij aan een beter begrip van de dynamiek van multiagent learning. Ten eerste bieden we een alternatieve interpretatie van de *single population replicator dynamics*. In het bijzonder leiden we af dat, wanneer selectie proportioneel aan prestatie plaatsvindt, het gemiddelde gedrag van een populatie lerende agenten niet afhankelijk is van het gedrag van de individuele agenten, maar de *single population replicator dynamics* volgt. Ten tweede leiden we een dynamisch model af van *multi-state multi-agent learning*. De *state-coupled replicator dynamics* verbreden de evolutionaire methodiek van multiagent learning en stellen ons in staat om leerdynamica in stochastische spellen te voorspellen. Tot slot laten we zien hoe evolutionaire modellen van leergedrag kunnen worden gebruikt om complexe systemen te analyseren en gedrag te voorspellen, waarmee we een brug slaan tussen de theoretische concepten en problemen uit het dagelijks leven.