ΤΜΗΜΑ ΜΗΧΑΝΙΚΩΝ Η/Υ & ΠΛΗΡΟΦΟΡΙΚΗΣ
ΠΑΝΕΠΙΣΤΗΜΙΟ ΙΩΑΝΝΙΝΩΝ

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF IOANNINA

ΠΑΝΕΠΙΣΤΗΜΙΟΥΠΟΛΗ ΙΩΑΝΝΙΝΩΝ
Τ.Θ. 1186, 45110 ΙΩΑΝΝΙΝΑ
Τ: 26510 08817 - F: 26510 08890

P.O. Box 1186
GR 45110 IOANNINA, GREECE
T: +30 26510 08817 - F: +30 26510 08890

ΠΑΡΟΥΣΙΑΣΗ
ΔΙΔΑΚΤΟΡΙΚΗΣ ΔΙΑΤΡΙΒΗΣ

ΗΜΕΡΟΜΗΝΙΑ: Πέμπτη, 5 Μαρτίου 2015

ΩΡΑ: 15.00

ΑΙΘΟΥΣΑ: Αίθουσα Σεμιναρίων (ισόγειο Ι1-Ι2)
Κτήριο Τμήματος Μηχανικών Η/Υ & Πληροφορικής

ΟΜΙΛΗΤΗΣ: **Νικόλαος Τζιωρτζιώτης**

Θ έ μ α

«Machine Learning for Intelligent Agents»

Επταμελής Εξεταστική Επιτροπή:

1. ***Κωνσταντίνος Μπλέκας*** (επιβλέπων), Επίκουρος Καθηγητής του Τμήματος Μηχανικών Η/Υ & Πληροφορικής του Πανεπιστημίου Ιωαννίνων
2. ***Αριστείδης Λύκας***, Καθηγητής του Τμήματος Μηχανικών Η/Υ & Πληροφορικής του Πανεπιστημίου Ιωαννίνων
3. ***Χριστόφορος Νίκου***, Αναπληρωτής Καθηγητής του Τμήματος Μηχανικών Η/Υ & Πληροφορικής του Πανεπιστημίου Ιωαννίνων
4. ***Κωνσταντίνος Βλάχος***, Επίκουρος Καθηγητής του Τμήματος Μηχανικών Η/Υ & Πληροφορικής του Πανεπιστημίου Ιωαννίνων
5. ***Γεώργιος Βούρος***, Καθηγητής του Τμήματος Ψηφιακών Συστημάτων του Πανεπιστημίου Πειραιώς
6. ***Μιχαήλ Λαγουδάκης***, Αναπληρωτής Καθηγητής του Τμήματος Ηλεκτρονικών Μηχανικών και Μηχανικών Υπολογιστών του Πολυτεχνείου Κρήτης.
7. ***Ανδρέας-Γεώργιος Σταφυλοπάτης***, Καθηγητής της Σχολής Ηλεκτρολόγων Μηχανικών & Μηχανικών Υπολογιστών του Ε.Μ.Π.

Π ε ρ ί λ η ψ η

This dissertation studies the problem of developing intelligent agents, which are able to acquire skills in an autonomous way, simulating human behaviour. An autonomous intelligent agent acts effectively in an unknown environment, directing its activity towards achieving a specific goal based on some performance measure. Through this interaction, a rich amount of information is received, which allows the agent to perceive the consequences of its actions, identify important behavioural components, and adapt its behaviour through learning. In this direction, the present dissertation concerns the development, implementation and evaluation of machine learning techniques for building intelligent agents. Three important and very challenging tasks: i) Value function approximation on reinforcement learning, where the agent's policy is evaluated and improved by representing the value functions as a linear combination of adaptive features, ii) Bayesian reinforcement learning, where the reinforcement learning problem is modeled as a decision-theoretic problem, by placing a prior distribution over Markov Decision Processes (MDPs) that encodes the agent's belief about the true environment, and iii) Development of intelligent agents on games, which constitute a really challenging platform for developing machine learning methodologies, involving a number of issues that should be resolved, such as the appropriate choice of state representation, continuous action spaces, e.t.c..

In the first part, we focus on the problem of value function approximation suggesting two different methodologies. Firstly, we propose the Relevance Vector Machine Temporal Difference (RVMTD) algorithm, which constitutes an advanced kernelized Bayesian methodology for model-free value function approximation, employing the RVM regression framework as a generative model. The key aspect of RVMTD is the restructure of the policy evaluation problem as a linear regression problem. An online kernel sparsification technique is adopted, rendering the RVMTD practical in large scale domains. Based on this scheme, we derive recursive low-complexity formulas for the online update of the model observations. For the estimation of the unknown model coefficients a sparse Bayesian methodology is adopted that enhances model capabilities. Secondly, a model-based reinforcement learning algorithm is proposed, which is based on the online partitioning of the input space into clusters. As the data arrived sequentially to reinforcement learning, an online extension of the vanilla EM algorithm is used for clustering. In this way, a number of basis functions are created and updated automatically. Also, statistics are kept about the dynamics of the environment that are next used for policy evaluation. Finally, the least-squares solution is used for the estimation of the unknown coefficients of the value function model.

In the second part, we address the Bayesian reinforcement learning problem proposing two advanced Bayesian algorithms. Firstly, we present the Linear Bayesian Reinforcement Learning (LBRL) algorithm showing that the system dynamics can be estimated accurately by a Bayesian linear Gaussian model, which takes into account correlations in the state features. Policies are estimated by applied approximate dynamic programming on a transition model that is sampled from the current posterior. This form of approximate Thompson sampling results in good exploration in unknown MDPs. Secondly, the Cover Tree Bayesian Reinforcement Learning (CTBRL) algorithm is proposed which constitutes an online tree-based Bayesian approach for reinforcement learning. The main idea of CTBRL is the construction of a cover tree from the observations, which remains efficient in high dimensional spaces. In this way, we create a set of partitions of the state space. An efficient non-parametric Bayesian conditional density estimator is also introduced on the cover tree structure. This is a generalized context tree, endowed with a multivariate linear Bayesian model at each node and is used for the estimation of the dynamics of the underlying environment. Thus, taking a sample for the posterior, we obtain a piecewise linear Gaussian model of the dynamics. The main advantages of this approach are its flexibility and efficiency, rendering it suitable for reinforcement learning problems in continuous state spaces.

In the third part of this thesis, we consider the problem of developing intelligent agents in two challenging games, the Ms.~PacMan and the AngryBirds. Firstly, we propose the RL-PacMan agent, which is based on an abstract but informative state space representation. The adopted representation is able to encode a game scene, giving the opportunity to our agent to distinguish different situations. For discovering an optimal policy, we use the model-free SARSA(λ) reinforcement learning algorithm. In our study, we demonstrate that an efficient state representation is of central interest for the design of an intelligent agent. Finally, we propose the AngryBER agent, which is based on an efficient tree structure for representing each game screenshot. This representation has the advantage of establishing an informative feature space and modifying the task of game playing to a regression problem. A Bayesian ensemble regression framework is used for the estimation of the return of each action, where each pair of `object material' and `bird type' has its own regression model. After each shot, the regression model is incrementally updated, in a fully closed form. The AngryBER agent participated in the international AIBIRDS 2014 competition winning the 2nd price.