

Seminar in Statistics: Learning Blackjack

ETH Zurich, Spring Semester 2016

Jonas Peters, Christina Heinze, Gian Thanei, Nicolai Meinshausen

Version: April 25, 2016

§ 1 Stationary Environments

I Python and Blackjack (22.02.2016, C. Heinze)

Jonas Erni

introduction to object-oriented programming and python, quick intro to blackjack, present python blackjack framework, implement "HumanPlayer"

II Optimal Strategy to play Blackjack (29.02.2016, G. Thanei)

Arturo Vivas, Ben Molitor

rules and optimal strategy of blackjack (with proof), implement an "OptimalPlayer", present the "StrategicPlayer"

III Multi-Armed Bandits (07.03.2016, C. Heinze)

Elena Asoni, Valentin Stalder

problem statement (regret), different strategies (epsilon-greedy, Thompson sampling, UCB, ...), Chapter 2 of Sutton, Barto: "Reinforcement Learning An Introduction" (2nd edition), contextual bandits (as single bandits, LinUCB: Chu et al 2011: "Contextual Bandits with Linear Payoff Functions", nonlin. variants (short))

IV Statistics of Multi-Armed Bandits (14.03.2016, C. Heinze)

Orellano, Mauchle

Chapter 2 in Bubeck and Cesa-Bianchi: "Survey" (compute upper bound, present lower bound), Gittins index (talk from Richard Weber)

V CODING SESSION I (21.03.2016)

everyone

to finish (until 20.03.2016): random strategy and explore-exploit

to code: tictactoe

§ 2 Reinforcement Learning

VI Reinforcement Learning and Bellman Equations (04.04.2016, G. Thanei)

Buck, Hertrich

Chapters 3 and 4 of Sutton, Barto: “Reinforcement Learning An Introduction” (2nd edition)

VII Monte Carlo Methods (11.04.2016, G. Thanei)

Heinzer, Profumo

Chapter 5 of Sutton, Barto: “Reinforcement Learning An Introduction” (2nd edition), implementation for blackjack

VIII Temporal Differences, Q-Learning (25.04.2016, J. Peters)

Abegglen, Zoechbauer

Chapters 6 and 7 of Sutton, Barto: “Reinforcement Learning An Introduction” (2nd edition)

§ 3 Causality

IX Causal Inference (02.05.2016, J. Peters)

Bunkin, Steffen

Structural Equation Models, Interventions, Simpson’s Paradox

X Inverse Probability Weighting (09.05.2016, J. Peters)

Gianolli, Scholtes

soft interventions, expectations, gradients, finite sample estimators, confidence intervals; distribution of weights. ask Jonas for material

§ 4 Challenge

XI AlphaGo and CODING SESSION II (23.05.2016, Gian Thanei)

Bausch

HOMEWORK (until 25.05.2016): strategy different from explore-exploit

David Silver et al: “Mastering the game of Go with deep neural networks and tree search”, Nature 529:484-489, 2016. For Neural Networks: Chapter 11 from Hastie, Tibshirani: “The Elements of Statistical Learning”, Springer, 2008.

XII CHALLENGE and SUMMARY (30.05.2016)

everyone, Peters

to bring: random strategy