Contents

Preface v

1. Finite-Horizon Models 1

1.1 Preliminaries 1
1.2 Model Description 3
1.3 Dynamic Programming Approach 5
1.4 Examples 8
1.4.1 Non-transitivity of the correlation 8
1.4.2 The more frequently used control is not better . . 9
1.4.3 Voting 11
1.4.4 The secretary problem 13
1.4.5 Constrained optimization 14
1.4.6 Equivalent Markov selectors in non-atomic MDPs 17
1.4.7 Strongly equivalent Markov selectors in non-atomic MDPs 20
1.4.8 Stock exchange 25
1.4.9 Markov or non-Markov strategy? Randomized or not? When is the Bellman principle violated? . . 27
1.4.10 Uniformly optimal, but not optimal strategy . . . 31
1.4.11 Martingales and the Bellman principle 32
1.4.12 Conventions on expectation and infinities 34
1.4.13 Nowhere-differentiable function $v_t(x)$; discontinuous function $v_t(x)$ 38
1.4.14 The non-measurable Bellman function 43
1.4.15 No one strategy is uniformly ε-optimal 44
1.4.16 Semi-continuous model 46

Piunovskiy, A.B.
Examples in Markov decision processes
2013

digitalisiert durch:
IDS Basel Bern

2. Homogeneous Infinite-Horizon Models: Expected Total Loss 51
2.1 Homogeneous Non-discounted Model 51
2.2 Examples 54
2.2.1 Mixed Strategies 54
2.2.2 Multiple solutions to the optimality equation . . . 56
2.2.3 Finite model: multiple solutions to the optimality equation; conserving but not equalizing strategy . 58
2.2.4 The single conserving strategy is not equalizing and not optimal 58
2.2.5 When strategy iteration is not successful 61
2.2.6 When value iteration is not successful 63
2.2.7 When value iteration is not successful: positive model I 67
2.2.8 When value iteration is not successful: positive model II 69
2.2.9 Value iteration and stability in optimal stopping problems 71
2.2.10 A non-equalizing strategy is uniformly optimal . . 73
2.2.11 A stationary uniformly ε-optimal selector does not exist (positive model) 75
2.2.12 A stationary uniformly ε-optimal selector does not exist (negative model) 77
2.2.13 Finite-action negative model where a stationary uniformly ε-optimal selector does not exist 80
2.2.14 Nearly uniformly optimal selectors in negative models 83
2.2.15 Semi-continuous models and the blackmailer's dilemma 85
2.2.16 Not a semi-continuous model 88
2.2.17 The Bellman function is non-measurable and no one strategy is uniformly ε-optimal 91
2.2.18 A randomized strategy is better than any selector (finite action space) 92
2.2.19 The fluid approximation does not work 94
2.2.20 The fluid approximation: refined model 97
2.2.21 Occupation measures: phantom solutions 101
2.2.22 Occupation measures in transient models 104
2.2.23 Occupation measures and duality 107

2.2.24 Occupation measures: compactness 109
2.2.25 The bold strategy in gambling is not optimal (house limit) 112
2.2.26 The bold strategy in gambling is not optimal (inflation) 115
2.2.27 Search strategy for a moving target 119
2.2.28 The three-way duel ("Truel") 122

3. Homogeneous Infinite-Horizon Models: Discounted Loss 127

3.1 Preliminaries 127
3.2 Examples 128
3.2.1 Phantom solutions of the optimality equation . . 128
3.2.2 When value iteration is not successful: positive model 130
3.2.3 A non-optimal strategy $\hat{\pi}$ for which $v_x^{\hat{\pi}}$ solves the optimality equation 132
3.2.4 The single conserving strategy is not equalizing and not optimal 134
3.2.5 Value iteration and convergence of strategies . . . 135
3.2.6 Value iteration in countable models 137
3.2.7 The Bellman function is non-measurable and no one strategy is uniformly ε-optimal 140
3.2.8 No one selector is uniformly ε-optimal 141
3.2.9 Myopic strategies 141
3.2.10 Stable and unstable controllers for linear systems 143
3.2.11 Incorrect optimal actions in the model with partial information 146
3.2.12 Occupation measures and stationary strategies . . 149
3.2.13 Constrained optimization and the Bellman principle 152
3.2.14 Constrained optimization and Lagrange multipliers 153
3.2.15 Constrained optimization: multiple solutions . . . 157
3.2.16 Weighted discounted loss and (N, ∞)-stationary selectors 158
3.2.17 Non-constant discounting 160
3.2.18 The nearly optimal strategy is not Blackwell optimal 163
3.2.19 Blackwell optimal strategies and opportunity loss 164

3.2.20 Blackwell optimal and n-discount optimal strategies . . . 165
3.2.21 No Blackwell (Maitra) optimal strategies . . . 168
3.2.22 Optimal strategies as $\beta \to 1-$ and MDPs with the average loss – I . . . 171
3.2.23 Optimal strategies as $\beta \to 1-$ and MDPs with the average loss – II . . . 172

4. Homogeneous Infinite-Horizon Models: Average Loss and Other Criteria 177

4.1 Preliminaries . . . 177
4.2 Examples . . . 179
4.2.1 Why lim sup? . . . 179
4.2.2 AC-optimal non-canonical strategies . . . 181
4.2.3 Canonical triplets and canonical equations . . . 183
4.2.4 Multiple solutions to the canonical equations in finite models . . . 186
4.2.5 No AC-optimal strategies . . . 187
4.2.6 Canonical equations have no solutions: the finite action space . . . 188
4.2.7 No AC-ε-optimal stationary strategies in a finite state model . . . 191
4.2.8 No AC-optimal strategies in a finite-state semi-continuous model . . . 192
4.2.9 Semi-continuous models and the sufficiency of stationary selectors . . . 194
4.2.10 No AC-optimal stationary strategies in a unichain model with a finite action space . . . 195
4.2.11 No AC-ε-optimal stationary strategies in a finite action model . . . 198
4.2.12 No AC-ε-optimal Markov strategies . . . 199
4.2.13 Singular perturbation of an MDP . . . 201
4.2.14 Blackwell optimal strategies and AC-optimality . . . 203
4.2.15 Strategy iteration in a unichain model . . . 204
4.2.16 Unichain strategy iteration in a finite communicating model . . . 207
4.2.17 Strategy iteration in semi-continuous models . . . 208
4.2.18 When value iteration is not successful . . . 211
4.2.19 The finite-horizon approximation does not work . . . 213

4.2.20 The linear programming approach to finite models 215
4.2.21 Linear programming for infinite models 219
4.2.22 Linear programs and expected frequencies in finite models 223
4.2.23 Constrained optimization 225
4.2.24 AC-optimal, bias optimal, overtaking optimal and opportunity-cost optimal strategies: periodic model 229
4.2.25 AC-optimal and average-overtaking optimal strategies 232
4.2.26 Blackwell optimal, bias optimal, average-overtaking optimal and AC-optimal strategies . . 235
4.2.27 Nearly optimal and average-overtaking optimal strategies 238
4.2.28 Strong-overtaking/average optimal, overtaking optimal, AC-optimal strategies and minimal opportunity loss 239
4.2.29 Strong-overtaking optimal and strong*-overtaking optimal strategies 242
4.2.30 Parrondo's paradox 247
4.2.31 An optimal service strategy in a queueing system 249

Afterword 253

Appendix A Borel Spaces and Other Theoretical Issues 257
A.1 Main Concepts 257
A.2 Probability Measures on Borel Spaces 260
A.3 Semi-continuous Functions and Measurable Selection . . . 263
A.4 Abelian (Tauberian) Theorem 265

Appendix B Proofs of Auxiliary Statements 267

Notation 281

List of the Main Statements 283

Bibliography 285

Index 291