Appendix A: Basic Notation for MDP

S	State space
s/s	A state vector/scalar, given a continuous/discrete state space
i	A state, given a countable state space
r	Reward
$r^a(\mathbf{s})$	Expected one step reward/reward rate in state **s**, under action a
c	Costs
$c^a(\mathbf{s})$	Expected one step cost/cost rate in state **s**, under action a
a	Action
$A(\mathbf{s})$	Set of actions available in state **s**, given a continuous/discrete state space
$A(i)$	Set of actions available in state i, given a countable state space
α	Discount factor
δ	Decision rule
δ_t	Decision rule at time t
$\delta_t(\mathbf{s})$	Action at time t, when in state **s**, given a continuous/discrete state space
$\pi = (\delta, \delta, \ldots)$	Stationary Policy
$\pi = (\delta_0, \delta_1, \delta_2, \ldots)$	Policy with decision rule at time $t = 0, 1, 2, \ldots$
P^π	One step transition probability distribution/matrix under policy π
$P(dy\|(\mathbf{s}, a))$	Transition probability/distribution under action $\pi(\mathbf{s}) = a$, in state **s**
$P^\pi(dy\|\mathbf{s})$	Transition probability distribution/matrix under policy π
$p(j\|i, a)$	Transition probability into state j, when in state i, under action a
Q^π	Transition rate (infinitesimal generator) matrix under policy π
$Q^a_{i,j}$	Transition rate from state i into j (countable) under action a
$q(\mathbf{s'}\|(s, a))$	Transition rate from a state s into a state **s'** under action a, given a continuous/discrete state space
V^π_t	Value function under policy π of expected cumulative reward/costs over t steps (up to time t)
$V^\pi_t(\mathbf{s})$	Value function under policy π of expected cumulative reward /costs over t steps (up to time t) starting in state **s** at time 0
$V^*_t(\mathbf{s})$ or $V_t(\mathbf{s})$	Optimal value function of expected cumulative reward/costs over t steps up to time t, starting in state **s** at time 0
V^π_α	Discounted value function under policy π
V^*_α or V_α	Optimal discounted value function
$G^\pi(\mathbf{s}), g^\pi$ (if ergodic)	Average expected reward/cost function/value under policy π
$G^*(\mathbf{s})$ or $G(\mathbf{s})$, g^* or g (if ergodic)	Optimal average expected reward/cost function/value
$W^\pi(\mathbf{s})/W^*(\mathbf{s})$ or $W(\mathbf{s})$	Expected total reward/costs, $\lim_{t\to\infty} V^\pi_t(\mathbf{s})$, given that the limit exists, under policy π/ an optimal policy
$H^\pi(\mathbf{s})/H^*(\mathbf{s})$ or $H(\mathbf{s})$	Bias of the policy π/ an optimal policy

R.J. Boucherie, N.M. van Dijk (eds.), *Markov Decision Processes in Practice*,
International Series in Operations Research & Management Science 248,
DOI 10.1007/978-3-319-47766-4

Appendix B: Dichotomy and Criteria

The table below gives a compact overview of the dichotomy on Discrete or continuous (time and state) modeling aspects. Here the distinction is made based upon the natural or primary description, i.e. not on the used solution procedure e.g. as by uniformization. It also states the performance measure of interest and the optimization criterion used.

R.J. Boucherie, N.M. van Dijk (eds.), *Markov Decision Processes in Practice*,
International Series in Operations Research & Management Science 248,
DOI 10.1007/978-3-319-47766-4

Ch.	Topic	Measure	Time	State	Criteria
		R: Rewards C: Costs O: Other	DT: Discrete CT: Continuous	DS: Discrete CS: Continuous	Time Horizon: ITH: Infinite FTH: Finite Costs: AC: Average DC: Discounted
General theory					
1	One-step improvements	R/C O: Delay/Payoff	DT	DS	ITH AC
2	Value function approximation in queueing	O: Delay/Loss	DT	DS	ITH AC
3	ADP: approximate dynamic programming	R: Revenues C: Routing	DT	DS	FTH (ITH: DC)
4	Infinite state queueing	C O: Delay	CT	DS	ITH AC
5	Infinite state structural properties	C O: Delay	DT/CT	DS	ITH AC/DC
Healthcare					
6	Screening and treatment of diseases	O: QALY (see chapter)	DT	DS + CS	FTH - ITH
7	Breast cancer	O: QALY (see chapter)	DT	DS	FTH - ITH
8	Patient appointment scheduling	O: Service level/ Overtime	DT	DS	ITH DC
9	Ambulance dispatching	O: Late arrivals/ Response time	CT	DS	ITH AC
10	Blood supply	O: Outdating	DT	DS	FTH (+ ITH: AC)
Transportation					
11	Airports: noise load management	O: Noise Load	DT	CS	FTH
12	Car park	O: Imbalance	CT	DS	FTH
13	Traffic lights	O: Delays/ Queues	CT	DS	ITH AC
14	Electric vehicles	C: Charging	DT	DS	FTH
Production					
15	Lot scheduling	R: Order Acceptances	DT	DS	ITH AC
16	Fisheries	R: Welfare/profit	DT	CS	ITH DC
17	Flow controllable service rates	O: Delays/workload	CT	DS	ITH AC
Communications					
18	Wireless channel selection	O: Throughput	DT	DS + CS	ITH AC
19	Call center staffing	C: Staffing/ Service level	CT	CS	FTH
20	Query wireless sensoring	O: Freshness/ Response times	CT	DS + CS	ITH AC
Financial modelling					
21	Financial derivatives	R/C: Utility and costs	DT	CS	FTH DC