

Reinforcement Learning

MDP (markov decision process)

Adapted from: <http://www7.inra.fr/mia/T/MDPtoolbox/QuickStart.pdf>

```
library(MDPtoolbox)

## Warning: package 'MDPtoolbox' was built under R version 4.0.2
## Loading required package: Matrix
## Loading required package: linprog
## Warning: package 'linprog' was built under R version 4.0.2
## Loading required package: lpSolve
## Warning: package 'lpSolve' was built under R version 4.0.2
x <- mdp_example_forest()
mdp_check(x$P, x$R)

## [1] ""
discount <- 0.95

y <- mdp_policy_iteration(x$P, x$R, discount)
V <- y$V
policy <- y$policy

z <- mdp_LP(x$P, x$R, discount)
V <- z$V
policy <- z$policy

Vpolicy <- mdp_eval_policy_matrix(x$P, x$R, discount, policy)
Vpolicy

## [1] 58.482 61.902 65.902
Vpolicy <- mdp_eval_policy_iterative(x$P, x$R, discount, policy)

## [1] "MDP Toolbox: iterations stopped, epsilon-optimal value function"
Vpolicy

##           [,1]
## [1,] 58.48015
## [2,] 61.90015
## [3,] 65.90015
Vpolicy <- mdp_eval_policy_TD_0(x$P, x$R, discount, policy)
Vpolicy

## $V
## [1] 54.98117 58.65873 63.64552
##
## $mean_discrepancy
##      [1]      NA 0.43060205 0.26766950 0.20059361 0.16992243 0.15021071
##      [7] 0.13469194 0.11936998 0.11063406 0.09975291 0.13640458 0.09269907
##     [13] 0.09946345 0.08202312 0.07599025 0.06576575 0.06290057 0.05670672
##     [19] 0.07683651 0.06427114 0.06446001 0.04685300 0.04845936 0.04962113
```

```
## [25] 0.04874008 0.04609876 0.03660481 0.04208108 0.04775155 0.04469552
## [31] 0.03889769 0.03701329 0.03769780 0.02903419 0.04404198 0.03307655
## [37] 0.03091932 0.03314946 0.03385592 0.02982490 0.03269933 0.02154674
## [43] 0.02830598 0.03088086 0.02741372 0.03220929 0.02841705 0.02655718
## [49] 0.02845145 0.02875893 0.02359960 0.02235703 0.02749650 0.02101332
## [55] 0.02353669 0.02599887 0.02506887 0.02156636 0.01909237 0.02266701
## [61] 0.01957554 0.01967844 0.02368531 0.02473205 0.02223086 0.02081371
## [67] 0.02227978 0.01931021 0.01900899 0.01965489 0.02186374 0.01850046
## [73] 0.01719466 0.02026194 0.01502661 0.01777795 0.01740360 0.01911117
## [79] 0.01723470 0.01638249 0.01561226 0.01535405 0.01428444 0.01955766
## [85] 0.01857572 0.01725339 0.01712832 0.01662187 0.01660273 0.01445423
## [91] 0.01465620 0.01610518 0.01606698 0.01331878 0.01489407 0.01369575
## [97] 0.01326000 0.01256156 0.01378656 0.01352426
```

Model parameter change

change...

didn't finish: <http://www7.inra.fr/mia/T/MDPtoolbox/QuickStart.pdf>

```
q <- mdp_example_forest(3, 0.4, 2, .2)
# didn't finish w <- mdp_
```

wildfire high