

Errata for:

“From Shortest Paths to reinforcement Learning: A MATLAB-Based Introduction to Dynamic Programming” by Paolo Brandimarte, Springer, 2021

The list of errata will be periodically updated (hopefully, not too many times....).
This version is dated January 26, 2021.

To point out errors, typos, and whatnot: paolo.brandimarte@polito.it

Page 136, MATLAB code of Fig. 4.18. The following portion of code

```
% evaluate current policy
aux = matrix\rhs;
oldRelVal = [aux(1:(numStates-1));0];
oldAve = aux(numStates);
% improve policy
```

should be replaced by

```
% evaluate current policy
auxSol = matrix\rhs;
oldRelVal = [auxSol(1:(numStates-1));0];
lambda = auxSol(numStates);
% improve policy
```

In the old code there is an error, as the variable `lambda` was not assigned (so, replace `oldAve` by `lambda`). A second issue is that I have used a variable `aux` twice, in the main body of the function and in the nested functions. It is an auxiliary variable, so no harm done, but it is better to avoid the issue. The MATLAB code in the zipped archive has been corrected.

Page 156, MATLAB code of Fig. 5.5. In the following line (-4 in the code)

```
[QFactors, policy] = QLearning(systemObj,'max',discount,numSteps, ...
    startQFactors,alpha);
```

the output value `visits` is not assigned. The line should be replaced by

```
[QFactors, policy, visits] = QLearning(systemObj,'max',discount,numSteps, ...
    startQFactors,alpha);
```

The MATLAB code in the zipped archive has been corrected.

Page 176, MATLAB code of Fig. 6.9. The following portion of code should be omitted:

```
% Out-of-sample scenarios for comparing policies
rng default
numOutSample = 500;
emplOutPaths = SampleEmplPaths(empl0,transMatrix,timeHorizon,numOutSample);
retOutPaths = lognrnd(mu, sigma, numOutSample, timeHorizon) - 1;
```

No harm done, but in this first script we do not run any simulation. These out of sample scenarios are used in the next script. The MATLAB code in the zipped archive has been corrected.