

# CS4200/CS5200, On-line Machine Learning

## Class 9: Reinforcement Learning

Yuri Kalnishkan

Department of Computer Science  
Royal Holloway, University of London

2018/19

# Class Outline

## 1. Motivation and Preliminaries

# References

- [SB] R. S. Sutton and A. G. Barto, “Reinforcement Learning: An Introduction”, 2nd edition, The MIT Press, 2018
- [CS] C. Szepesvári “Algorithms for Reinforcement Learning”, Morgan & Claypool, 2010
- [JT] J. N. Tsitsiklis, On the Convergence of Optimistic Policy Iteration, JMLR 3 (2002) 59-72
- [WD] C. J. C. H. Watkins and P. Dayan, Technical Note: Q-Learning, Machine Learning, 8, 279-292 (1992)

## 1. Motivation and Preliminaries

## Example: USPS

- the problem is to label an image, which is a  $16 \times 16$  matrix of pixels
  - it is known that an image represents a hand-written digit, from 0 to 9
- we are given a training set containing a large number of labelled images
  - USPS dataset: scanned zip codes from envelopes