# More Q-Learning

* Complete assignments p3\_rl\_q6, p3\_rl\_q7 (but don’t work on the crawler yet), and p3\_rl\_q8 in the edX version of [Berkeley’s CS188](https://edge.edx.org/courses/course-v1:BerkeleyX+CS188+2018_SP/courseware/817468c509ef4b64bcb74cd38766ac44/41c847d3794141a7acb4f6b1e5904215/8?activate_block_id=block-v1%3ABerkeleyX%2BCS188%2B2018_SP%2Btype%40vertical%2Bblock%40789af2adfb5442468b3ed7d2dc6cfb5a).
* Download [cliffwalker.py](https://drive.google.com/file/d/1c7zjUMxKreTnHRmU5sht6EOE_wSJNZ5v/view?usp=sharing), [cliffwalking\_env.py](https://drive.google.com/file/d/1ghkxdfTAFIYIgQ80CGTmMk0nx2ZPM3Em/view?usp=sharing), and [discrete.py](https://drive.google.com/file/d/15BJL_Nv6ff3mCAys8CDfXQo1ekgZfhPM/view?usp=sharing) and run cliffwalker.py. This is another classic reinforcement learning example. It is similar to the BridgeGrid. See a description at the top of the file.
* Notice that if the noise parameter (to do\_a\_series()) is set to 0.2 or higher the agent learns to avoid the cliff by taking the top route. If noise is set to 0.0, the agent learns to take the bottom route adjacent to the cliff. The bottom route is shorter and there is no danger that the agent will fall over the cliff accidentally. Is there a noise parameter setting that will get the agent to decide that the middle route is best: shorter than the top route and safe enough to risk taking?
* Notice also that if the noise parameter is set to 0.0, the agent takes a distinctive strange-looking path on its first episode. Can you explain why? Think about the q-values the agent generates.
* Be aware that the example runs in ten parts. Each part has a reduced epsilon and a reduced alpha. At the end of each part, epsilon is set to zero and an episode is displayed, which shows how the q-table directs the agent at that time.