In the three-day series of tutorial lectures, I describe ``learning with function approximators'' in the context of neural-network (NN) supervised learning and reinforcement learning (also known as neuro-dynamic programming). I begin with fundamental concepts of classical dynamic programming (DP), and then derive the well-known backpropagation [for multilayer-perceptron (MLP) learning] from the so-called Kelley-Bryson optimal-control gradient formula. Here, I shall demonstrate the use of a non-optimal value function in the spirit of DP. In optimal-control theory, it also plays an important role in developing a class of second-order methods that start with a given nominal non-optimal state trajectory; in particular, I spotlight stagewise Newton and differential dynamic programming (DDP). Such second-order optimal-control methods can be employed for accelerating supervised NN-learning as well as reinforcement learning with function approximators. For optimization purpose, I emphasize that it is very important to pursue a good compromise between the steepest-descent method and a Newton-type method. Besides the theoretical aspects of the posed methods, I shall illustrate in detail small-scale longest path problems in non-Markovian settings to give a clear idea on reinforcement learning approaches as well as a classical DP algorithm.

This series of talks is largely based on joint work with Stuart Dreyfus, UC Berkeley.