One-week intensive course on

``Introduction to Dynamic Programming and Optimal Control''

Hino Campus, Tokyo Metropolitan University（首都大学東京　日野キャンパス）

for **9:40 am – 11:40 am**

July 31 - August 4, 2017.

Room: 2-305

Instructor: 國立台灣科技大學 助教授 水谷英二 先生 (Prof. Eiji Mizutani)

Dynamic programming (DP) is an optimization procedure that is particularly useful in solving a sequential decision-making problem. We consider a general discrete-time optimal-control problem, where we wish to find a sequence of optimal decisions so as to minimize a given criterion, the sum of stage costs plus the terminal-state cost. In this framework, we begin with a minimum-cost path problem, and its several variants, showing a standard *four-step* DP formulation:

(1) Define the optimal value function;
(2) Write down its recurrence relation;
(3) Give appropriate boundary conditions;
(4) State ``Answer is given by what?''

In (2), we introduce a basic concept of *consultant questions* in order to choose the proper state description for the *Bellman's principle of optimality* to hold.

Next, we consider optimal-control problems having linear dynamics and quadratic cost (or, LQ problem in short), and then more general problems with nonlinear dynamics. In the general nonlinear situations, the so-called optimal-control gradient procedure is widely employed for iterative optimization. Its computational convenience is enormous in any multistage setting. In machine learning, the procedure is also known as back-propagation (BP). Due to the current breakthroughs of deep (neural network) learning, BP has gained great popularity. We derive BP in the spirit of DP using nominal cost-to-go value functions. We also describe an optimal-control problem corresponding to recurrent neural network learning, for which we derive BP through time (BPTT) from a DP standpoint.