Learning-Based Control for Mobility Systems

Lecture 1: Introduction

Kyunghwan Choi

kh.choi@kaist.ac.kr

September 2, 2025

Outline

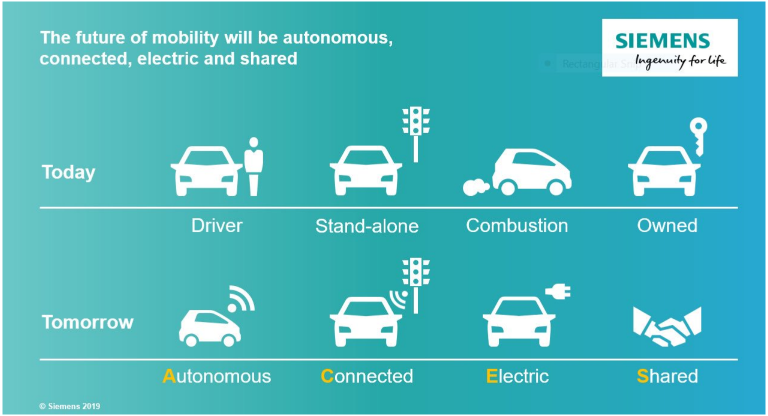
Motivation

Mobility

Definition by MLIT

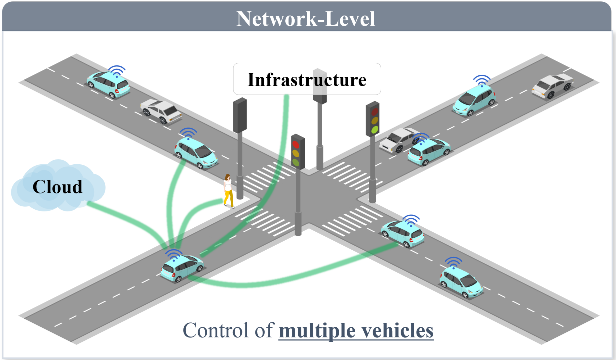
- A comprehensive concept of mobility from the user's perspective, referring to the act, function, or process of moving or transporting people or goods from one location to another, which is enabled through related means, infrastructure, and a series of services.

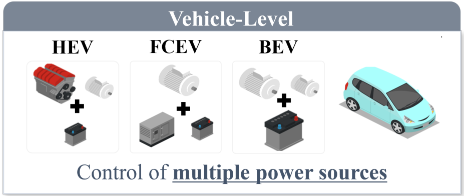
Key Technologies in Mobility

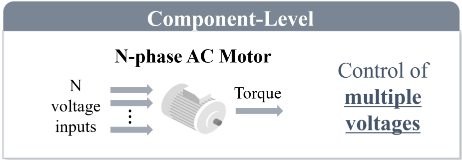


Connected, Automated, and Electrified Vehicles (CAEVs)

Multiple Control Levels in CAEVs

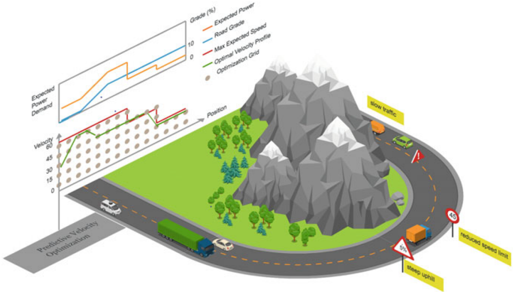




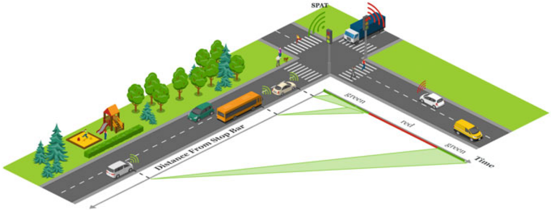


- How can we effectively leverage the multiple degrees of freedom across different control levels in CAEVs?

Examples of Network-level Control [1]



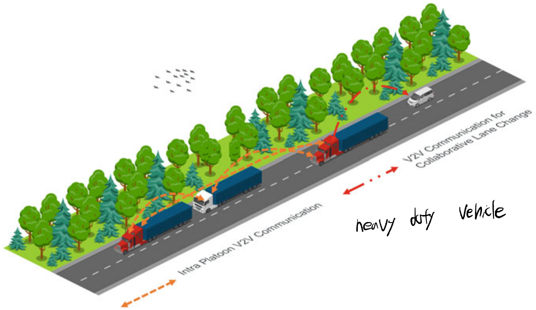
Eco driving in anticipation of upcoming hills



Eco-driving with SPaT preview (Example)

Signal phase and timing

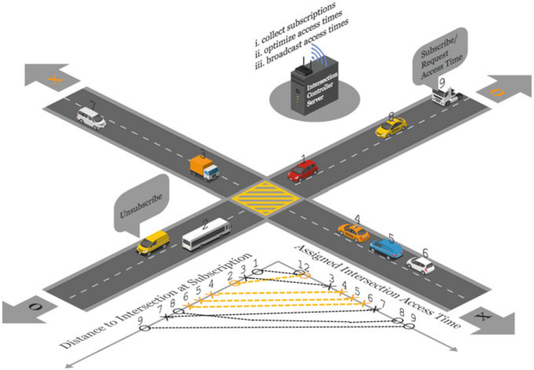
Examples of Network-level Control [1]



heavy duty vehicle

Collaborative car following and lane selection

(Example)



Cooperative intersection

- To maximize CAEVs' potential (e.g., energy efficiency), optimal control is essential for managing multiple degrees of freedom.

Optimal Control of CAEVs

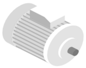
Optimal Control

Find Control input
To minimize Performance Index
Subject to Control constraints
System model
Environmental conditions

Applications

Component-Level

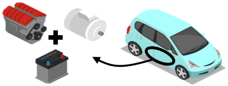
Optimal torque control of electric motor



Torque ↑
Energy ↓

Vehicle-Level

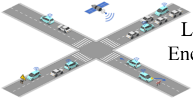
Optimal control of hybrid powertrain



Energy ↓

Network-Level

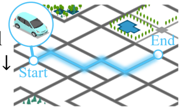
Optimal cooperative driving control



Local
Energy ↓

Optimal eco-driving & energy management

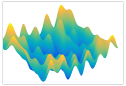
Global
Energy ↓



Challenges in Optimal Control of CAEVs

Main challenge

C1: Mathematical complexity



Optimal Control

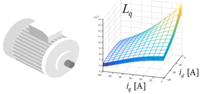
Find Control input
To minimize Energy consumption
Subject to Control constraints
System model ?
Environmental conditions ?

No analytic solution

Applications

Component-Level

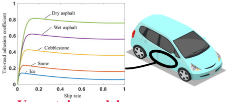
Optimal torque control of electric motor



Uncertain model

Vehicle-Level

Optimal control of hybrid powertrain



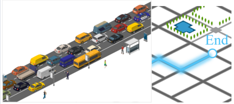
Uncertain model

Network-Level

Optimal cooperative driving control

Uncertain intentions

Optimal eco-driving & energy management



Uncertain future traffic

C2: System model uncertainties

key param depending operation systems manufacturing

C2: Environmental uncertainties

human driver

Challenges in Optimal Control of CAEVs

Examples: Optimal Torque Control of Electric Motors

Torque command T_e^{cmd}

Optimal torque control problem

Objective 1: $\min \sum_{k=1}^{N} \left(T_{e,k}^{cmd} - T_{e,k} \right)^2$

Objective 2: $\min J_p = \sum_{k=1}^{N} P_{loss,k}$

Constraints: System model, ?

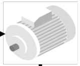
$i_{d,k}^2 + i_{q,k}^2 \le I_{\max}^2, v_{d,k}^2 + v_{q,k}^2 \le V_{\max}^2.$

v_d^*, v_q^*

Inverter

v_d, v_q

Synchronous motor



Torque T_e

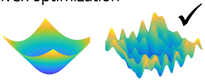
i_d, i_q, w_r

System model

$$\dot{\lambda}_d = -R_s i_d + w_r \lambda_q + v_d,$$
$$\dot{\lambda}_q = -R_s i_q - w_r \lambda_d + v_q,$$
$$T_e = 1.5P(\lambda_d i_q - \lambda_q i_d).$$

Challenge 1: Mathematical complexity

- Nonconvex optimization

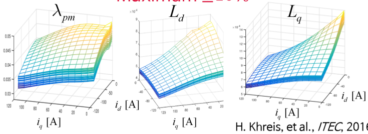


- Need to be solved within T_s ($\le 100\ \mu s$)

Challenge 2: Uncertainty

Manufacturing tolerances

Maximum $\pm 10\%$



H. Khreis, et al., *ITEC*, 2016.

- $\lambda_d = L_d i_d + \lambda_{pm}, \ \lambda_q = L_q i_q$

Challenges in Optimal Control of CAEVs

Examples: Optimal Control of Hybrid Powertrain

Optimal control problem

Torque demand
T_{dem}

$$\min_{T_{eng}(t)} \int_{t_0}^{t_N} \dot{m}_f(T_{eng}(t),t)\,dt$$

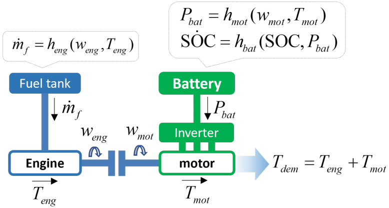
subject to

$$\dot{\mathrm{SOC}}(t) = h_{bat}(\mathrm{SOC}(t), T_{mot}(t), t),$$
$$\mathrm{SOC}_{\min} \le \mathrm{SOC}(t) \le \mathrm{SOC}_{\max},$$
$$\mathrm{SOC}(t_0) = \mathrm{SOC}_0,\ \mathrm{SOC}(t_N) = \mathrm{SOC}_N,$$
$$T_{dem}(t) = T_{eng}(t) + T_{mot}(t),$$
$$T_{mot,\min}(t) \le T_{mot}(t) \le T_{mot,\max}(t).$$

Motor/Engine torque command

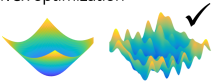
Status

Parallel Hybrid powertrain



Challenge 1: Mathematical complexity

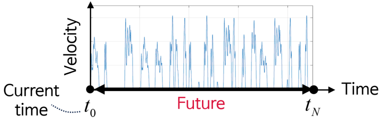
- Nonconvex optimization



- Long horizon ($t_0 \le t \le t_N$; tens of mins.)

Challenge 2: Uncertainty

- Uncertain future information is needed.



Challenges in Optimal Control of CAEVs

Examples: Optimal Cooperative Driving Control

Optimal coordination problem

$$\min_{u_i(t),\forall i\in A} \sum_{i\in N} TTD_i^2$$

subject to

$$\dot{p}_i(t) = v_i(t), \dot{v}_i(t) = u_i(t),$$

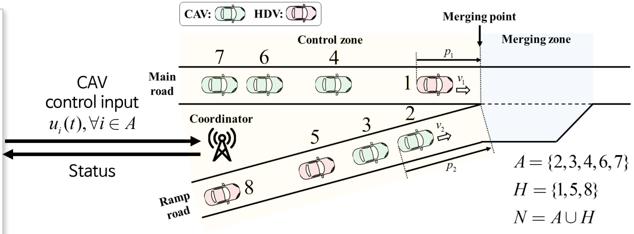
$$u_{\min} \le u_i(t) \le u_{\max}, 0 \le v_i(t) \le v_{\max},$$

$$|T_i - T_j| \ge T_{safe}, \forall j \text{ such that } r_j \ne r_i,$$

$$p_j(t) - p_i(t) \ge d_{safe} + T_{safe} v_i(t),$$

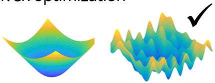
$$j = \max\{k \mid r_k = r_i, k < i\}.$$

Merging scenario in mixed traffic



Challenge 1: Mathematical complexity

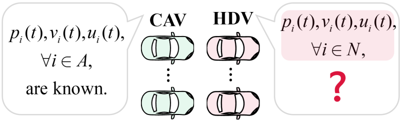
- Nonconvex optimization



- Large number of vehicles

Challenge 2: Uncertainty

- Uncertain intentions of the HDVs



Motivation for Learning-based Control

From Optimization to **Policy Learning**

MPC

We need to solve the problem forever
not efficient

- Numerical optimization: solve problem repeatedly $\rightarrow$ computationally heavy
- Policy learning: store and improve policies $\rightarrow$ more efficient, potentially more accurate

From Traditional Robustness to **Model Learning**

nominal system model

conventional approach

- Robust control: treats uncertainty in a lumped way $\rightarrow$ difficult to achieve optimality
- Adaptive control / estimation: limited transient accuracy DOB (Disturbance observer)
- Model learning: directly learn and use system model $\rightarrow$ Utilize information for policy learning
  define parameter / state
- Beyond: model-free policy learning

↳ model is not explicitly express

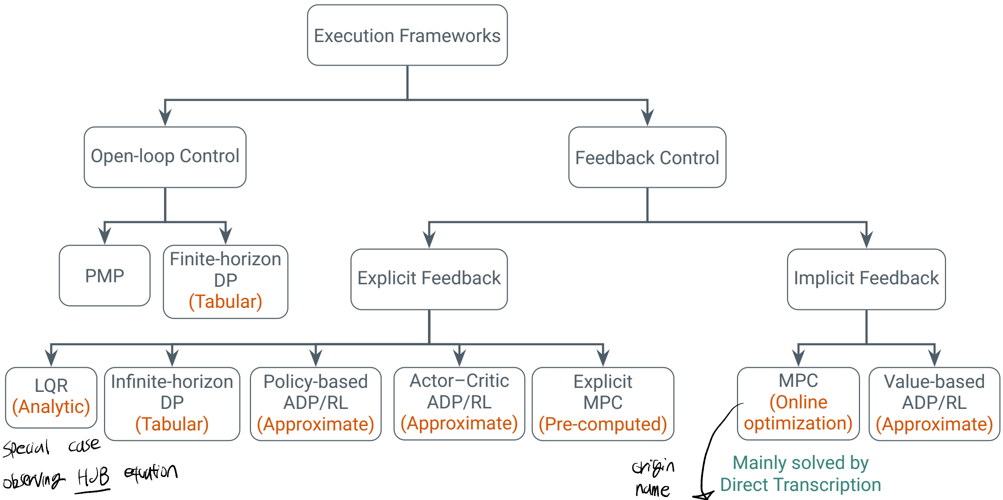
About this Course

Optimal Control

Solution Methods

- **PMP**
  - Necessary (local) → It cannot provide global solution
  - Mainly finite horizon
  - TPBVP / indirect
  - Model-based
- **Principle of Optimality**
  - Hamilton Jacobian
  - **Bellman (disc.) / HJB (cont.) Equation**
    - Global (under std. assumptions)
    - Finite / infinite horizon
  - **DP**
    - Model-based
  - **ADP** (control society)
    - Approximate
    - Often model-based
  - **RL** (AI community)
    - Often model-free
  - Control lens
  - AI lens
  - origin은 다르지만 비슷해서 ADP/RL 이라고도 함
- **Direct Transcription**
  - Discretize problem
  - parameter optimization problem
  - Typically local (global if convex)
  - Finite horizon
  - Constraints-friendly useful to handle constraints
  - Model-based

Optimal Control



special case
observing HJB equation
Hamilton Jacobian Bellman

origin name
Receiving Horizon Control (RHC)

DP/ADP/RL

Exact DP applies (in principle) to a very broad range of optimization problems

- Deterministic $\longleftrightarrow$ Stochastic
- Combinatorial optimization $\longleftrightarrow$ Optimal control w/ infinite state/control spaces
- One decision maker $\longleftrightarrow$ Two player games
- ... BUT is plagued by the curse of dimensionality and need for a math model

target model

ADP/RL overcomes the difficulties of exact DP by:

Starting different community but key component
quite similar

- Approximation (use neural nets and other architectures to reduce dimension)
- Simulation (use a computer model in place of a math model)

From the textbook

"In this chapter, we provide some background on exact dynamic programming (DP for short), with a view towards the suboptimal solution methods that are the main subject of this book. These methods are known by several essentially equivalent names: reinforcement learning, approximate dynamic programming, and neuro-dynamic programming. In this book, we will use primarily the most popular name: reinforcement learning (RL for short)."

Terminology in RL (AI) and DP (Control)

RL (AI)	DP (Control)
Agent	Controller (or decision maker)
Action	Control (or decision)
Environment	Dynamical system
Reward of a stage	(Opposite of) Cost of a state
Value (or reward, or state-value) function	(Opposite of) Cost function
Maximizing the value function	Minimizing the cost function
Planning	Solving a DP problem with model-based simulation
Learning	Solving a DP problem with model-free simulation
Learning a model	System identification
Deep RL	ADP using value and/or policy approximation with DNNs
⋮	⋮

Aims and Requirements

Principal aim:

- To explore the state of the art of ADP/RL at a graduate level
- To explore the common boundary between AI and optimal control
- To provide a bridge that workers with background in either field find it accessible (modest math)

Requirements:

- Homework (20%): A total of 4
- Midterm (40%)
- Term project (40%): An individual project applying RL to your research interests.
    - Report (30%)
    - Presentation (10%)

Office Hours:

- After class or by appointment

Mathematical Requirements

Math requirements for this course are modest.

- Calculus, elementary probability, and minimal use of vector-matrix algebra.
- Our objective is to use math to the extent needed to develop insight into the mechanism of various methods, and to be able to start research.

However a math framework is critically important.

- A math framework is essential for DP problem formulation, understanding, and solution.
- DP relies on substantial math theory, particularly for infinite horizon problems.

Textbook and References

Textbook:

- D. Bertsekas, "*Reinforcement Learning and Optimal Control*," Athena Scientific, 2019.

Supplementary Materials:

- D. Bertsekas, "*A Course in Reinforcement Learning*," 2nd Edition, Athena Scientific, 2024 (online).
- R. S. Sutton and A. G. Barto, "*Reinforcement Learning: An Introduction*," 2nd Edition, MIT Press, 2018 (online; a valuable resource that approaches the subject from the AI point of view).
- D. Liu, et al., "*Adaptive Dynamic Programming with Applications in Optimal Control*," Springer International Publishing, 2017.
- S. L. Brunton and J. N. Kutz, "*Data-Driven Science and Engineering: Machine Learning, Dynamical Systems, and Control*," 2nd Edition, Cambridge University Press, 2022.

Course Schedule

Week	Topic	Remark
1	Introduction	
2	Finite-horizon Exact DP	HW1
3	Approximation in Value Space	
4	Rollout	
5	Parametric Approximation	HW2
6	No class – Chuseok holiday	
7	Model-based and Model-free Parametric ADP	
8	Midterm exam	
9	Infinite-horizon RL	
10	Value Iteration	HW3
11	Policy Iteration	
12	Approximation in Policy Space	HW4
13	Aggregation	
14	Model Learning I	
15	Model Learning II	Term project report due
16	Term project presentations	

About the Next Lecture

We will cover:

- Overview of Optimal Control (PMP, Principle of Optimality, DP, ADP, RL, MPC)

PLEASE READ CHAPTER 1.1

References I

[1] A. Sciarretta, A. Vahidi, et al.
Energy-efficient driving of road vehicles.
Springer, 2020.