



Iran University
of Science and
Technology

به نام خدا

یادگیری تقویتی در کنترل

دکتر سعید شمقدری

دانشکده مهندسی برق
گروه کنترل

نیمسال اول ۱۴۰۵-۱۴۰۴

Motivation

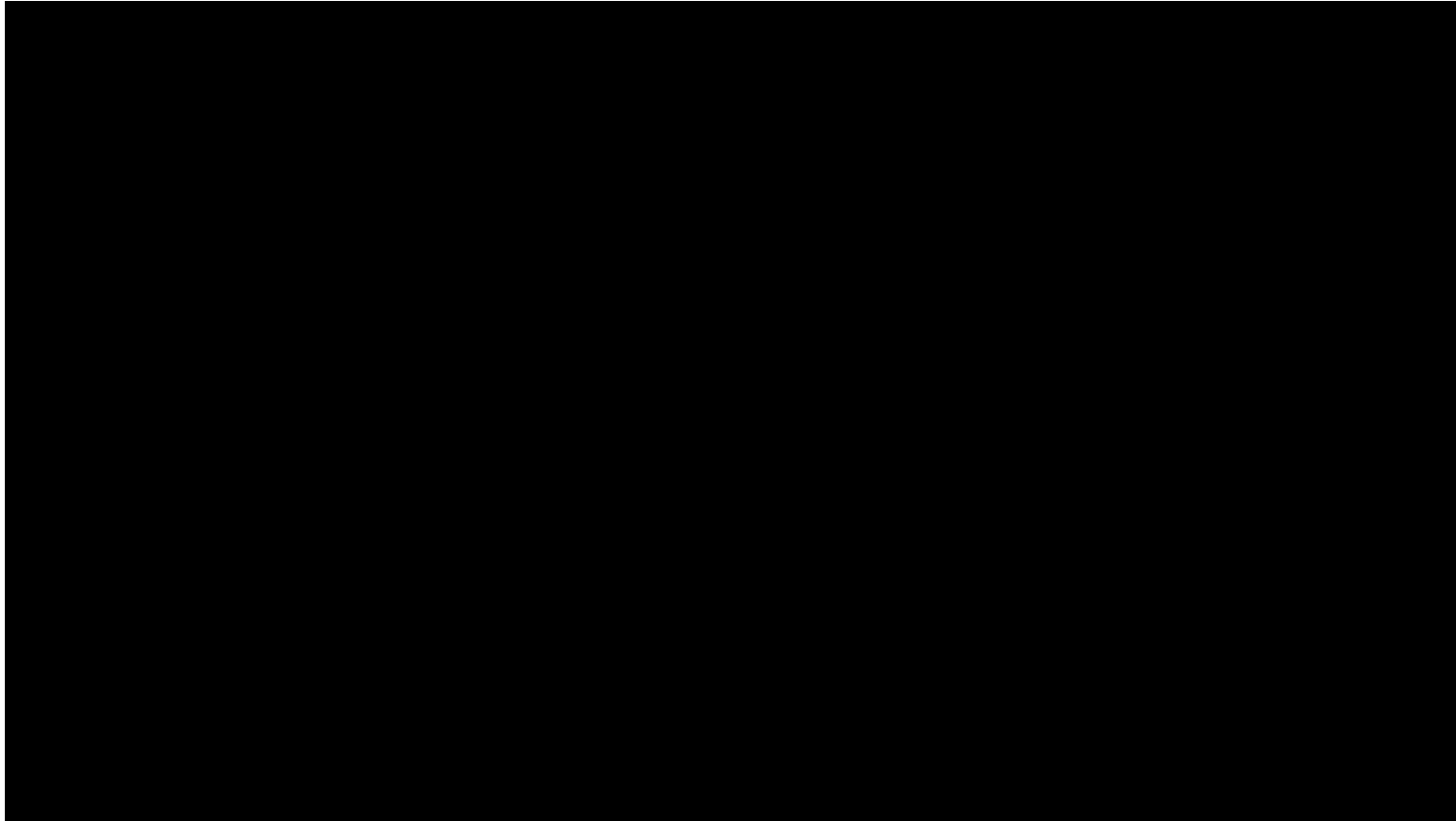
Robot Motor Skill Coordination with EM-based Reinforcement Learning

**Petar Kormushev, Sylvain Calinon,
and Darwin G. Caldwell**

Italian Institute of Technology

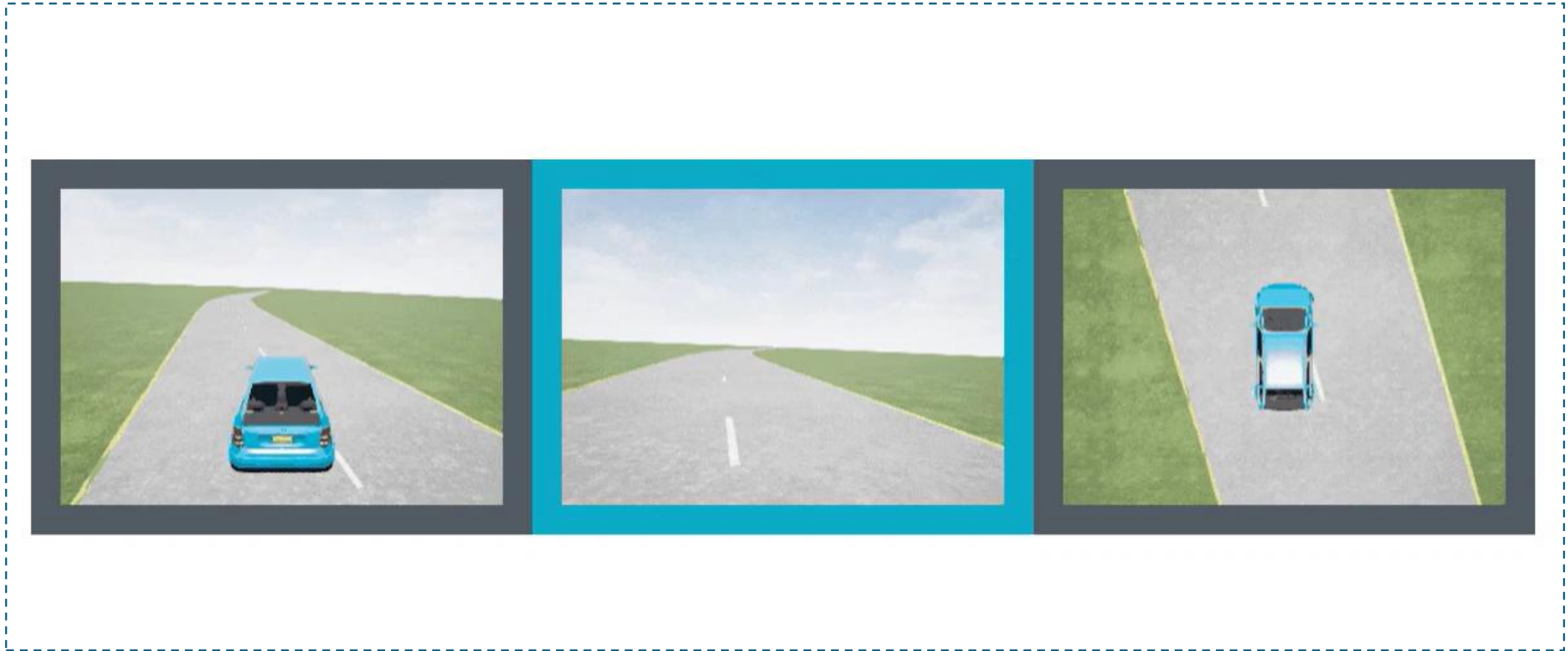
Robot Motor Skill Coordination with EM-based Reinforcement Learning
([Paper](#))

I Motivation



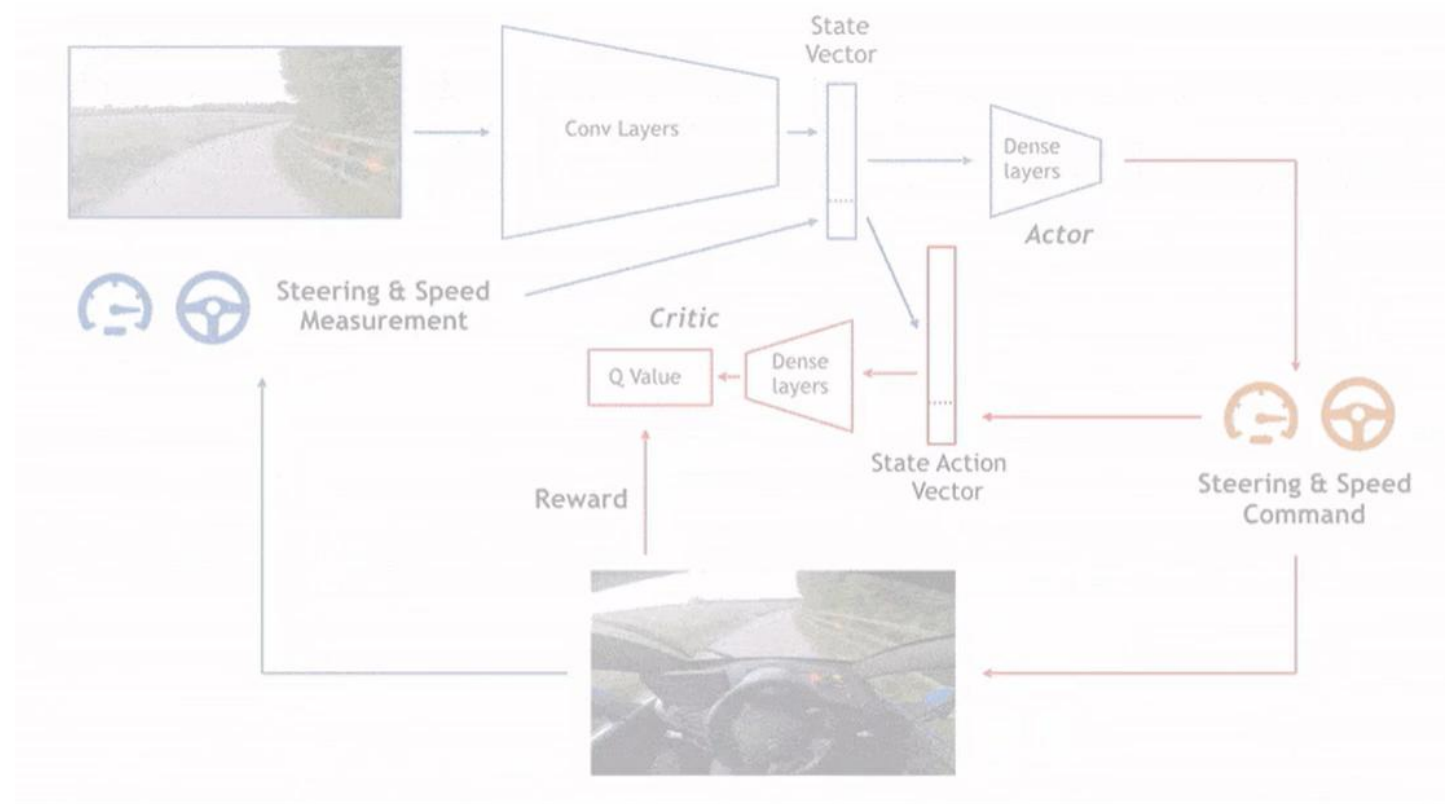
Learning to Drive in a Day
([Paper](#))

I Motivation



Learning to Drive in a Day
([Paper](#))

I Motivation

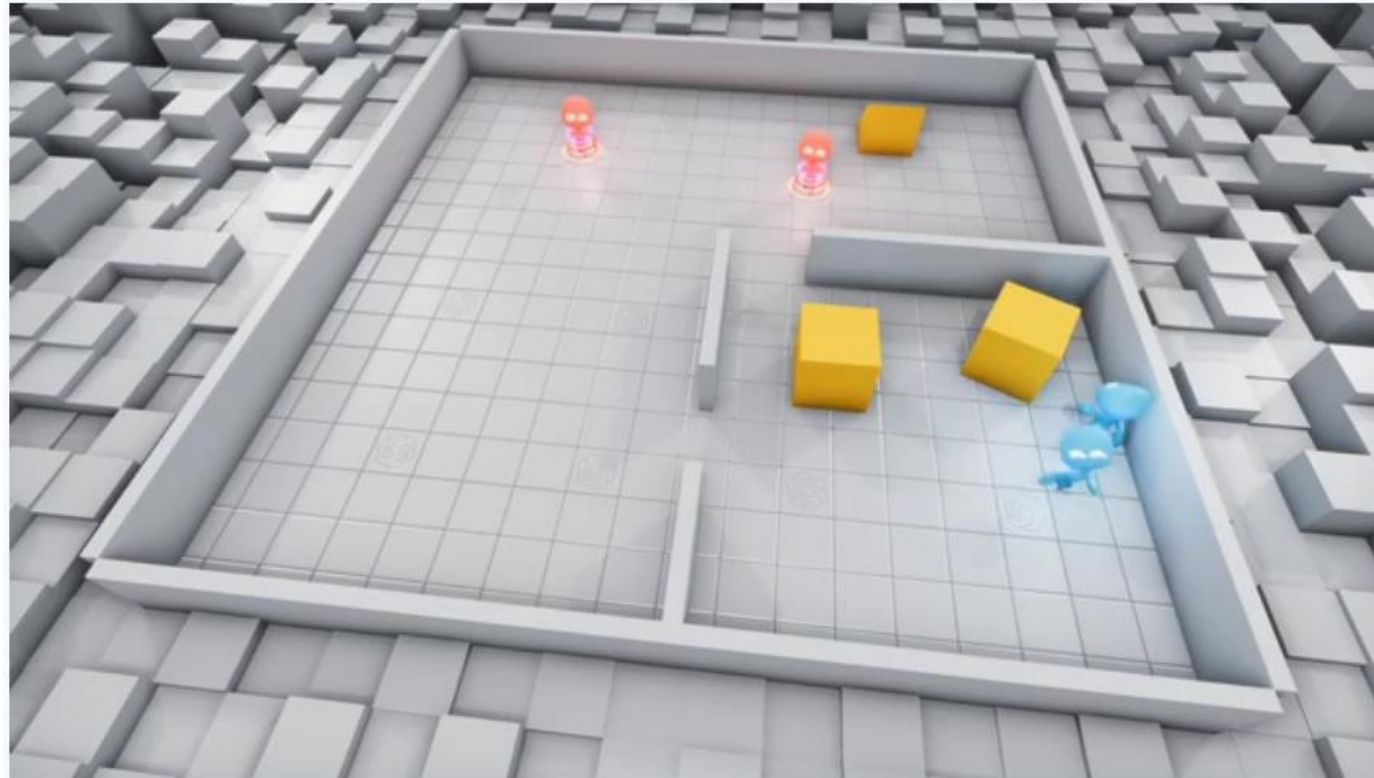


Learning to Drive in a Day
([Paper](#))

Multi-Agent Hide and Seek

EMERGENT TOOL USE FROM MULTI-AGENT AUTOCURRICULA
([Paper](#)) ([Webpage](#))

I Motivation



Episode 0

Random The agents move randomly.

EMERGENT TOOL USE FROM MULTI-AGENT AUTOCURRICULA ([Paper](#)) ([Webpage](#))

Table of Contents

بخش اول: یادگیری تقویتی در علوم کامپیوتر

مقدمه یادگیری تقویتی و یادگیری ماشین

مفاهیم یادگیری ماشین و یادگیری تقویتی

مقایسه روش‌های یادگیری

بخش اول: یادگیری تقویتی در علوم کامپیوتر

مقدمه یادگیری تقویتی و یادگیری ماشین

مفاهیم پایه‌ای یادگیری تقویتی

مطالعه سیستم تک حالت

بخش اول: یادگیری تقویتی در علوم کامپیوتر

مقدمه یادگیری تقویتی و یادگیری ماشین

مفاهیم پایه‌ای یادگیری تقویتی

فرایندهای تصمیم مارکوف

بخش اول: یادگیری تقویتی در علوم کامپیوتر

مقدمه یادگیری تقویتی و یادگیری ماشین

مفاهیم پایه‌ای یادگیری تقویتی

فرایندهای تصمیم مارکوف

برنامه ریزی پویا

بخش اول: یادگیری تقویتی در علوم کامپیوتر

مقدمه یادگیری تقویتی و یادگیری ماشین

مفاهیم پایه‌ای یادگیری تقویتی

فرایندهای تصمیم مارکوف

برنامه ریزی پویا

روش‌های مونت کارلو

یادگیری (TD) Temporal Difference

TD(0) and SARSA

Q-Learning

یادگیری (TD) Temporal Difference

مبانی یادگیری تقویتی عمیق

مفاهیم شبکه‌های عصبی

Policy Gradient Methods and Actor Critic

DDPG methods

کنترل بهینه سیستم‌های پیوسته و گسسته

بخش اول: یادگیری تقویتی در مهندسی کنترل

کنترل بهینه سیستم‌های پیوسته و گسسته

مبانی یادگیری تقویتی در کنترل

Policy Iteration and Value Iteration

Q-Learning

بخش اول: یادگیری تقویتی در مهندسی کنترل

کنترل بهینه سیستم‌های پیوسته و گسسته

مبانی یادگیری تقویتی در کنترل

کنترل فرایندهای تصمیم مارکوف مبتنی بر یادگیری تقویتی

بخش اول: یادگیری تقویتی در مهندسی کنترل

کنترل بهینه سیستم‌های پیوسته و گسسته

مبانی یادگیری تقویتی در کنترل

کنترل فرایندهای تصمیم مارکوف مبتنی بر یادگیری تقویتی

کنترل مبتنی بر یادگیری تقویتی برای سیستم‌های خطی پیوسته و گسسته

بخش اول: یادگیری تقویتی در مهندسی کنترل

کنترل بهینه سیستم‌های پیوسته و گسسته

مبانی یادگیری تقویتی در کنترل

کنترل فرایندهای تصمیم مارکوف مبتنی بر یادگیری تقویتی

کنترل مبتنی بر یادگیری تقویتی برای سیستم‌های خطی پیوسته و گسسته

حل مسئله LQT با استفاده از الگوریتم‌های یادگیری تقویتی

الگوریتم‌های یادگیری تقویتی برای سیستم‌های غیر خطی

بخش اول: یادگیری تقویتی در مهندسی کنترل

الگوریتم‌های یادگیری تقویتی برای سیستم‌های غیرخطی

پیاده سازی الگوریتم‌های غیرخطی با استفاده از شبکه‌های عصبی

بخش اول: یادگیری تقویتی در مهندسی کنترل

الگوریتم‌های یادگیری تقویتی برای سیستم‌های غیرخطی

پیاده سازی الگوریتم‌های غیرخطی با استفاده از شبکه‌های عصبی

معرفی مسئله Zero-sum-game و طراحی کنترل کننده حذف اغتشاش

Class Format

وبسایت درس (برای اطلاع از تاریخ انتشار تمرین‌ها، مهلت آن‌ها و همچنین تاریخ برگزاری کوییزها حتما تقویم درس مشاهده شود)

تکالیف و شبیه سازی:

- ۸ تمرین (۳)
- تمرین کلاسی (۱)
- تمرین جامع (بخش ۱ کامپیوتری) (۱)

کوئیز:

- ۴ کوئیز تئوری (۶)

آزمون پایانی (۶)

سمینارها:

- **سمینار کوتاه: پلتفرم‌های شبیه سازی RL (۰.۷۵)**

در گروه‌های دونفره

ارائه باید شامل معرفی کامل پلتفرم و همچنین نتایج شبیه سازی و اجرای الگوریتم نمونه باشد.
مثال: کتابخانه‌های PettingZoo، Gymnasium، ...

- **سمینار عمومی: روش‌های پیشرفته RL (۰.۷۵)**

در گروه‌های دونفره

مثال‌هایی از موضوعات: یادگیری تقویتی معکوس، سیستم‌های چندعاملی، کاربردهای یادگیری تقویتی در سایر حوزه‌های هوش مصنوعی (پردازش زبان طبیعی، مدل‌های مولد و ...)

پروژه: (۱.۵)

در گروه‌های دونفره

- انتخاب مقاله کنترلی مبتنی بر یادگیری تقویتی و دارای نوآوری تئوری
- مقاله‌ای انتخاب کنید که توانایی شبیه سازی آن را داشته باشید!
- ارائه مقاله مانند سمینارهای قبلی
- شبیه سازی مقاله و سعی در بازتولید نتایج آن (یا تولید نتایج نزدیک به مقاله)
- تحویل کامل گزارش و کد

پروژه: (۱.۵)

در گروه‌های دونفره

- انتخاب مقاله کنترلی مبتنی بر یادگیری تقویتی و دارای نوآوری تئوری
- مقاله‌ای انتخاب کنید که توانایی شبیه سازی آن را داشته باشید!
- ارائه مقاله مانند سمینارهای قبلی
- شبیه سازی مقاله و سعی در بازتولید نتایج آن (یا تولید نتایج نزدیک به مقاله)
- تحویل کامل گزارش و کد

فعالیت در کلاس

بهترین جزوه کلاس شامل مطالب درسی و تمرین‌ها و مطالب تکمیلی

پروژه: (۱.۵)

در گروه‌های دونفره

- انتخاب مقاله کنترلی مبتنی بر یادگیری تقویتی و دارای نوآوری تئوری
- مقاله‌ای انتخاب کنید که توانایی شبیه سازی آن را داشته باشید!
- ارائه مقاله مانند سمینارهای قبلی
- شبیه سازی مقاله و سعی در بازتولید نتایج آن (یا تولید نتایج نزدیک به مقاله)
- تحویل کامل گزارش و کد

فعالیت در کلاس
بهترین جزوه کلاس شامل مطالب درسی

Bonus Points !

ب تکمیلی

References

Books:

1. Sutton, R.S and Barto, A.G., Reinforcement Learning: An Introduction 2018

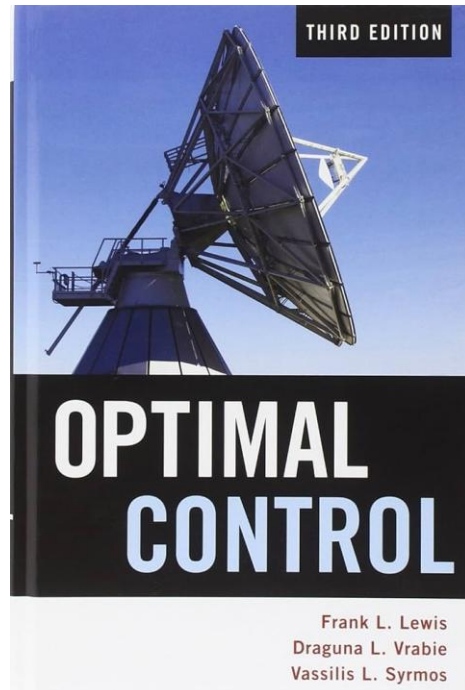


Books:

1. Sutton, R.S and Barto, A.G., Reinforcement Learning: An Introduction 2018
2. Szepesvari, Csaba., Algorithms for Reinforcement Learning

Books:

1. Sutton, R.S and Barto, A.G., Reinforcement Learning: An Introduction 2018
2. Szepesvari, Csaba., Algorithms for Reinforcement Learning
3. Lewis, F.L., Vrabie, D. and Syrmos, V.L., Optimal Control 2012



Articles:

4. Lewis, F.L. and Vrabie, D., Reinforcement Learning and Adaptive Dynamic Programming for Feedback Control, 2009
5. Lewis, F.L., Vrabie, D. and Vamvoudakis, K.G., reinforcement learning and feedback control, 2012
6. Lewis, F.L., Vrabie, D. and Syrmos, V.L., Optimal Control, 2012
7. Vrabie, D., Pastravanu, O., Abu-Khalaf, M. and Lewis, F.L., adaptive optimal control for continuous-time linear systems based on policy iteration, 2009
8. Jiang, Y. and Jiang, Z.P., Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamic, 2012
9. Kiumarsi, B., Lewis, F.L., Modares, H., Karimpour, A. and Naghibi-Sistani, M.B, Reinforcement Q-learning for optimal tracking control of linear discrete-time systems with unknown dynamics, 2014

Articles:

10. Wang, T., Zhang, H. and Luo, Y., Stochastic linear quadratic optimal control for model-free discrete-time systems based on Q-learning algorithm, 2018
11. Liu, D. and Wei, Q., policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems, 2014
12. Al-Tamimi, A., Lewis, F.L. and Abu-Khalaf, M., Discrete-Time Nonlinear HJB Solution Using Approximate Dynamic Programming: Convergence Proof, 2008
13. Liu, D. and Wei, Q., policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems, 2014
14. Al-Tamimi, A., Lewis, F.L. and Abu-Khalaf, M., Discrete-Time Nonlinear HJB Solution Using Approximate Dynamic Programming: Convergence Proof, 2008
15. Liu, D. and Wei, Q., policy iteration adaptive dynamic programming algorithm for discrete-time nonlinear systems, 2014

Articles:

16. Rizvi, S.A.A. and Lin, Z., Output feedback Q-learning for discrete-time linear zero-sum games with application to the H-infinity control, 2018 – Cited by 12
17. Kim, J.H. and Lewis, F.L., Model-free H^∞ control design for unknown linear discrete-time systems via Q-learning with LMI, 2010
18. Kiumarsi, B., Lewis, F.L. and Jiang, Z.P., H^∞ control of linear discrete-time systems: Off-policy reinforcement learning, 2017
19. Liu, D., Li, H. and Wang, D., Neural-network-based zero-sum game for discrete-time nonlinear systems via iterative adaptive dynamic programming algorithm, 2013
20. Zhu, Y., Zhao, D. and Li, X., Iterative Adaptive Dynamic Programming for Solving Unknown Nonlinear Zero-Sum Game Based on Online Data, 2016

Prerequisites

Prerequisites

کنترل مدرن

آمار و احتمال

Teaching Assistants

- **سجاد رفیعی**
• [Email](#) | [ID](#) (Telegram)

- **سیده ستاره خسروی**
• [Email](#) | [ID](#) (Telegram)