Reinforcement Learning China Summer School

RLChina 2020

Introduction and Opening

夏令营官网： **rlchina.org**
夏令营对于所有同学、学者公开开放，
扫描下方二维码进入报名入口：

Prof. Jun Wang

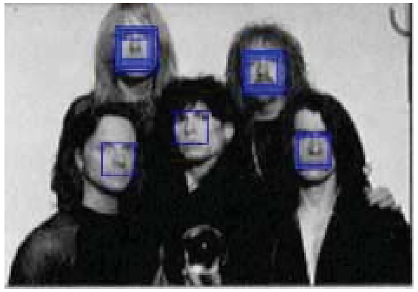
Computer Science, UCL

jun.wang@cs.ucl.ac.uk

http://www0.cs.ucl.ac.uk/staff/Jun.Wang/

July 27, 2020

Progress of Artificial Intelligence



Face detection (Computer Vision)

Recognition
(识别)

Play Go (Reinforcement Learning)

Decision Making
(单个智体决策)

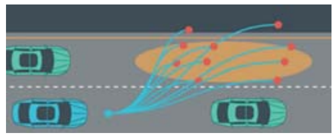
DeepMind
(rooted from UCL)

OpenAI

EMBODIED INTELLIGENCE
(rooted from UC Berkeley)

Multi-agent AI and Distributed Decision Making
(多智体决策)

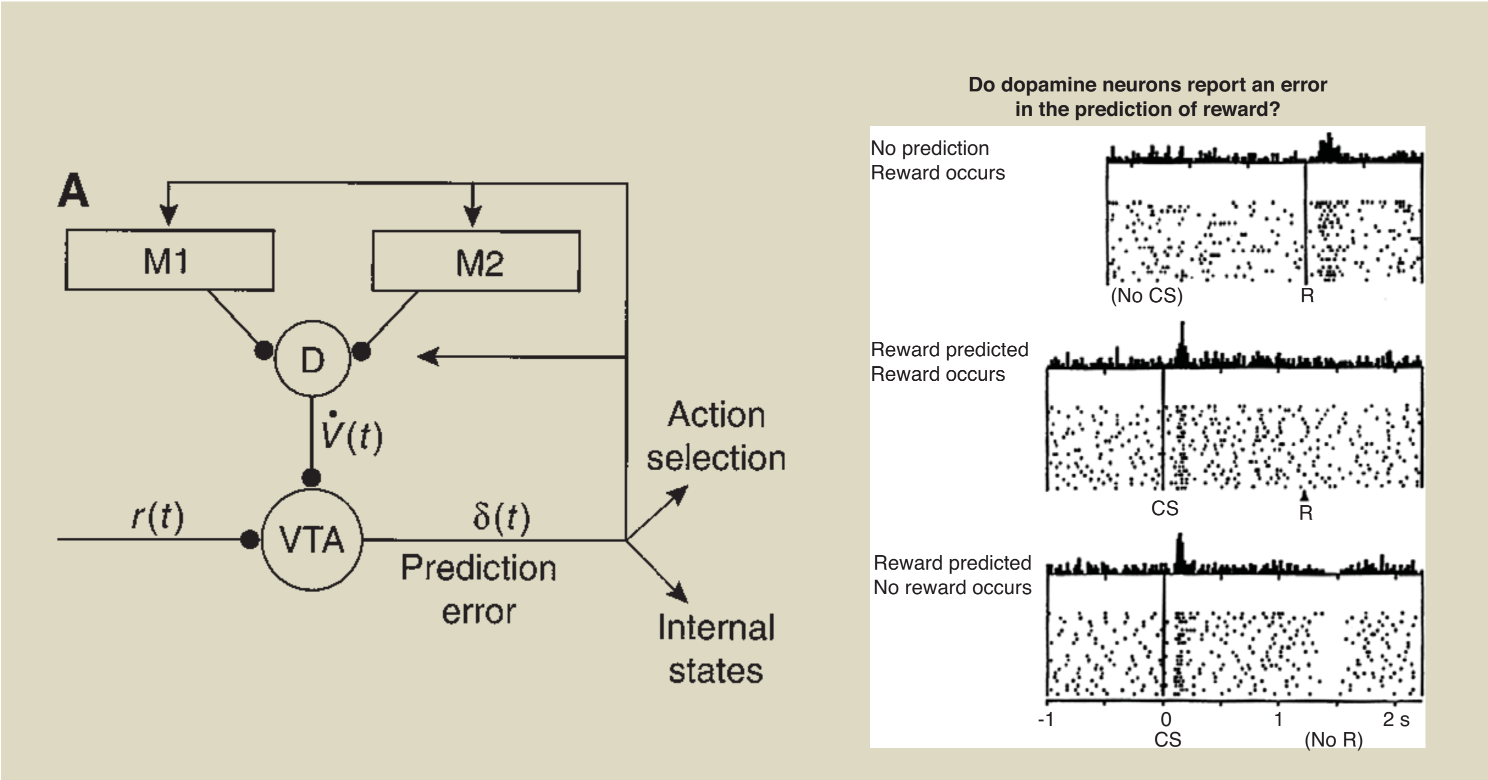
Bots Routine and Navigation



Distributed Optimisation

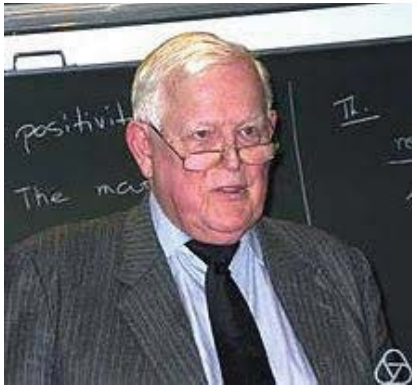
Multi-objectives optimisation

Classical conditioning (Pavlov 1906) & reinforcement Learning



[Schultz, Wolfram, Peter Dayan, and P. Read Montague, 1997]

The Duality Between Estimation and (Single) Control



R. E. KALMAN
Research Institute for Advanced Study,[2]
Baltimore, Md.

A New Approach to Linear Filtering and Prediction Problems[1]

The classical filtering and prediction problem is re-examined using the Bode-Shannon representation of random processes and the "state transition" method of analysis of dynamic systems. New results are:

(1) The formulation and methods of solution of the problem apply without modification to stationary and nonstationary statistics and to growing-memory and infinite-memory filters.

(2) A nonlinear difference (or differential) equation is derived for the covariance matrix of the optimal estimation error. From the solution of this equation the coefficients of the difference (or differential) equation of the optimal linear filter are obtained without further calculations.

(3) The filtering problem is shown to be the dual of the noise-free regulator problem.

The new method developed here is applied to two well-known problems, confirming and extending earlier results.

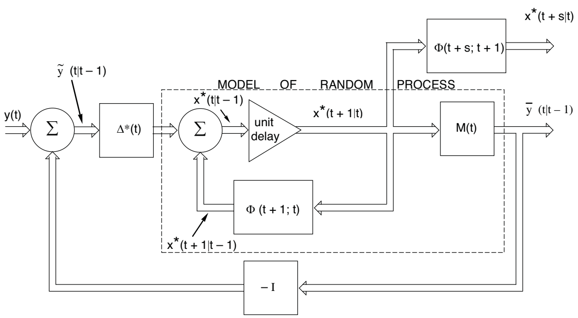
The discussion is largely self-contained and proceeds from first principles; basic concepts of the theory of random processes are reviewed in the Appendix.

Introduction

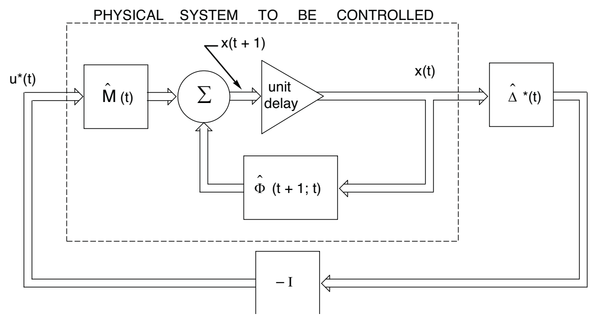
AN IMPORTANT class of theoretical and practical

Present methods for solving the Wiener problem are subject to a number of limitations which seriously curtail their practical

Optimal estimation



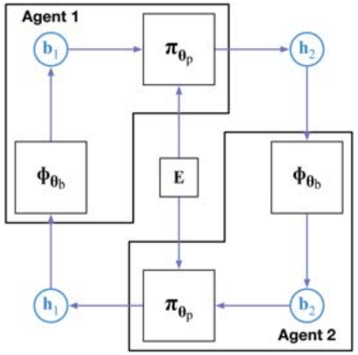
Optimal control

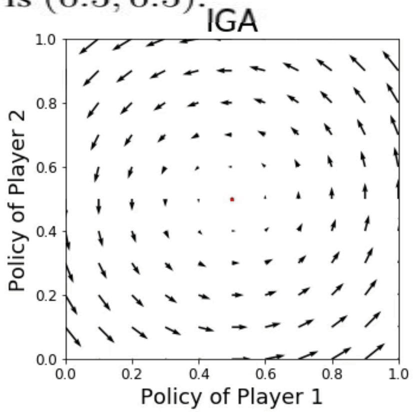


Kalman, Rudolph Emil. "A new approach to linear filtering and prediction problems." (1960): 35-45.

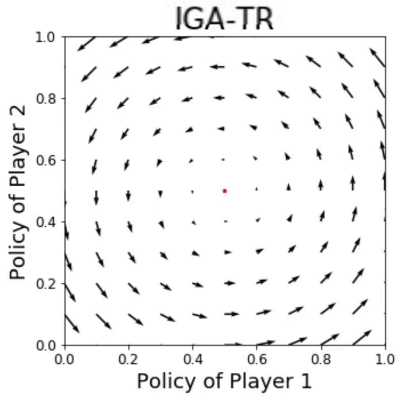
Multiagent gradient ascent

An example matrix game is $R_r = \begin{bmatrix} 0 & 3 \\ 1 & 2 \end{bmatrix}$, $R_c = \begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix}$ with the initial strategy $(\alpha, \beta) = (0.0, 0.9)$. The only Nash Equilibrium is $(0.5, 0.5)$.

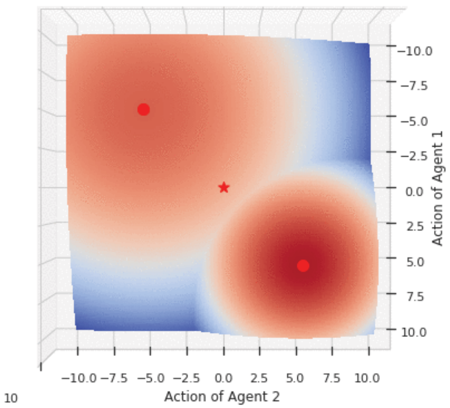




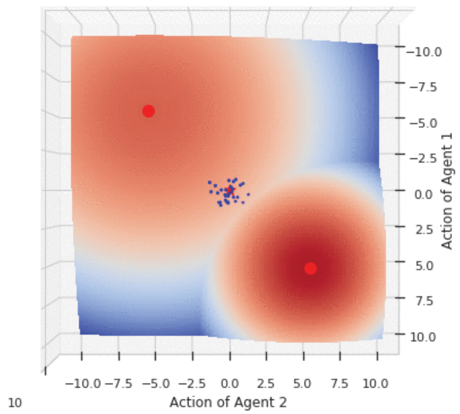
Independent learner



Aware of other agent



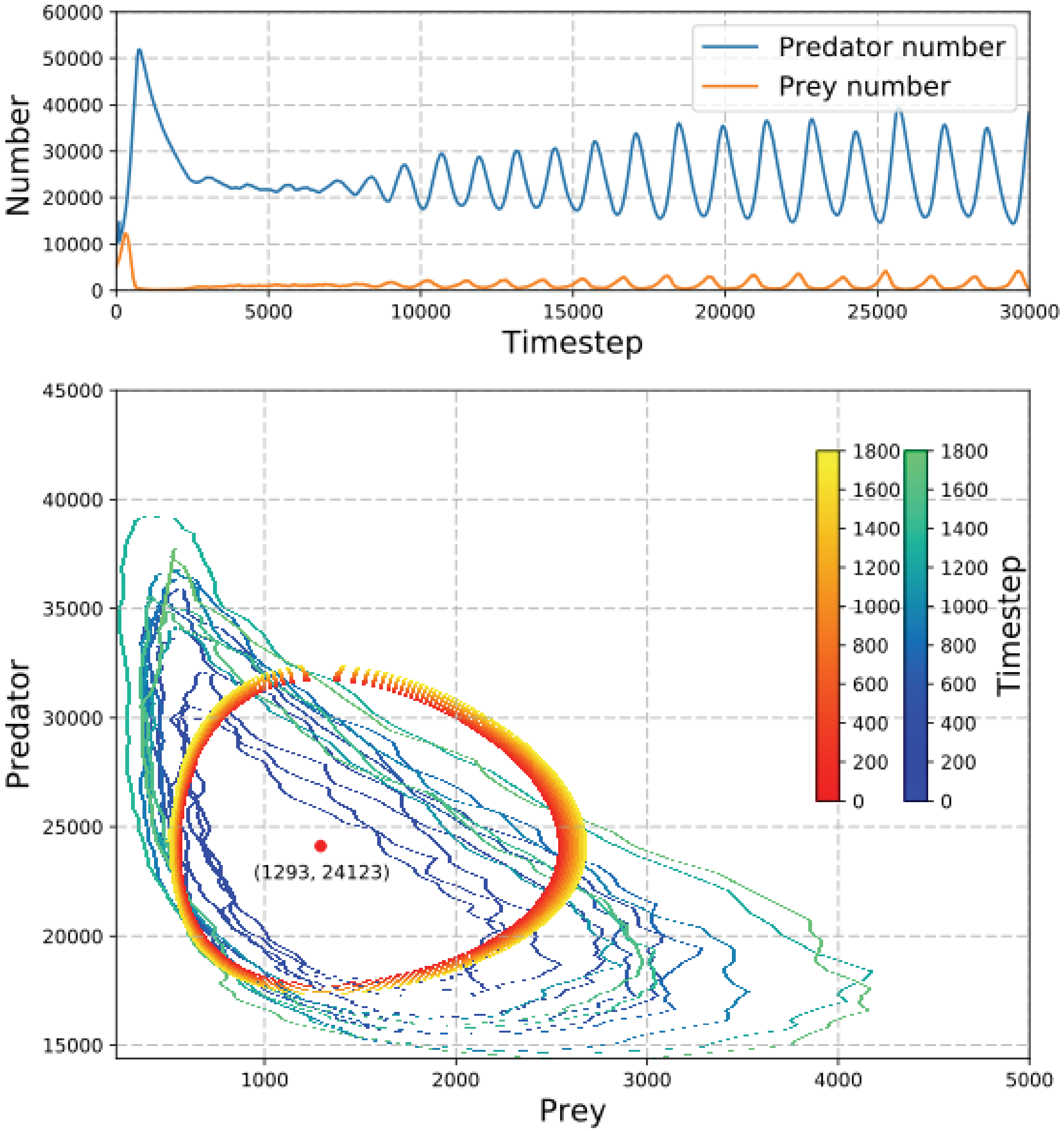
PR2-AC



DDPG-OM

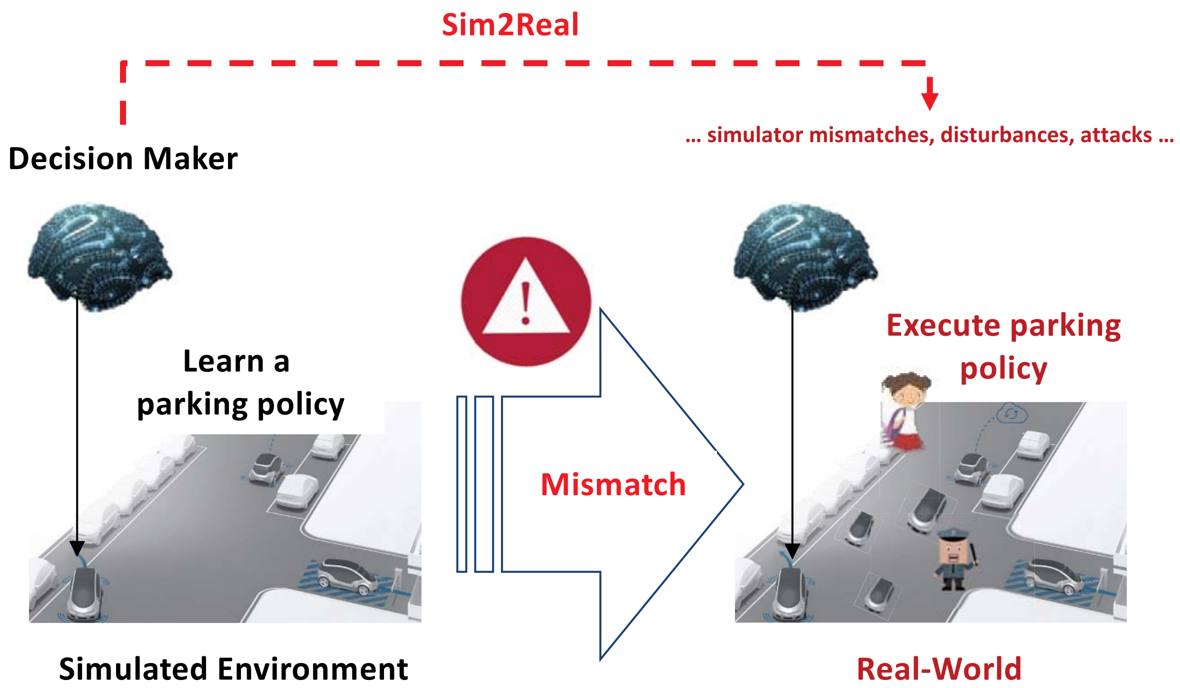
Wen, Ying, et al. "Probabilistic recursive reasoning for multi-agent reinforcement learning." ICLR19

The Dynamics of the Artificial Population



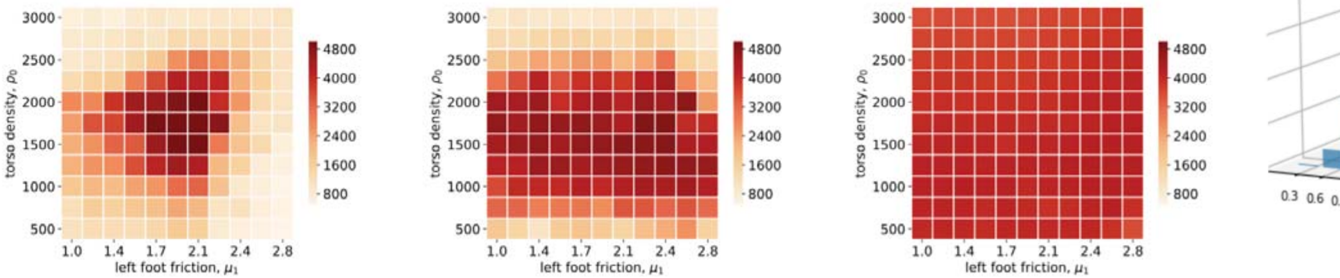
Yaodong Yang , Lantao Yu , Yiwei Bai , Jun Wang , Weinan Zhang , Ying Wen , Yong Yu, , Dynamics of Artificial Populations by Million-agent Reinforcement Learning, 2017

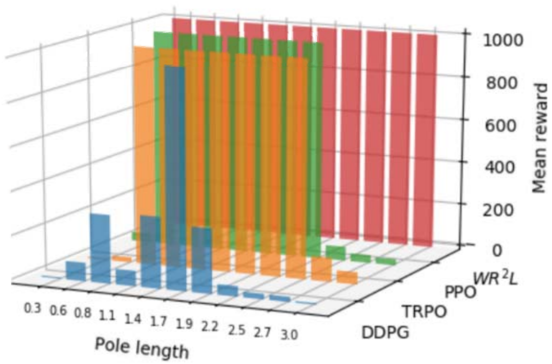
Robustness in Machine Decision Making



Abdullah MA, Ren H, Ammar HB, Milenkovic V, Luo R, Zhang M, Wang J. Wasserstein Robust Reinforcement Learning. arXiv preprint arXiv:1907.13196.

- **Robustness** against the changes in data environments
- **Safety** of the executed decisions to catastrophic failures





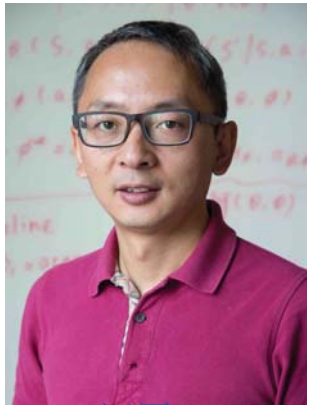
Lecturers

安波

Haitham Bou Ammar

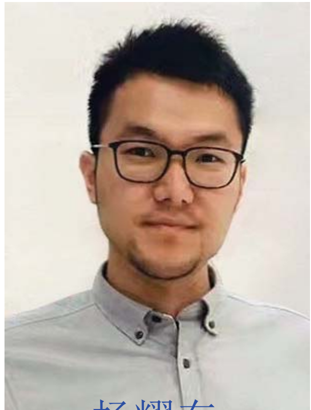
郝建业

卢宗青



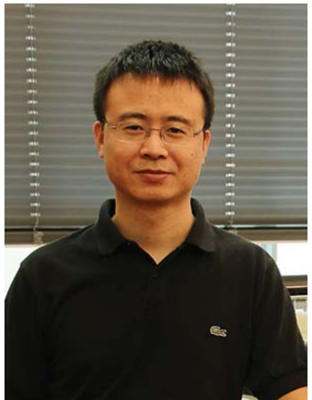
汪军

徐任远

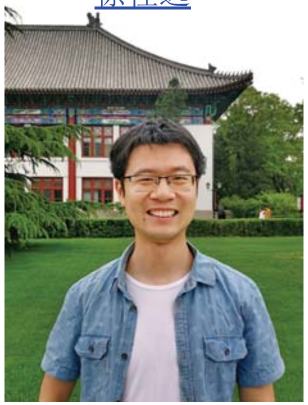


杨耀东

俞扬



张崇洁



张海峰

张伟楠

朱占星

DateTime	Course	Teacher
2020-07-27 19:00-19:10	Openning and Introduction	汪军
2020-07-27 19:10-20:50	Value-based Reinforcement Learning	卢宗青
2020-07-28 19:00-20:40	Policy-based RL and RL Theory	汪军
2020-07-29 19:00-20:40	Optimisation in Learning	Haitham
2020-07-30 19:00-20:40	Model-based Reinforcement Learning	张伟楠
2020-07-31 19:00-20:40	Control as Inference	朱占星
2020-08-01 19:00-20:40	Imitation Learning	俞扬
2020-08-03 19:00-20:40	Hierarchical Reinforcement Learning	郝建业
2020-08-04 19:00-20:40	Game Theory Basic	张海峰
2020-08-05 19:00-20:40	Multi-agent Systems	安波
2020-08-06 19:00-20:40	Deep Multi-agent Learning	张崇洁
2020-08-07 19:00-20:40	Advances in Multi-agent Learning	杨耀东
2020-08-08 19:00-20:40	Mean-field Games and Controls	徐任远
2020-08-08 20:40-21:10	Panel Discussion	全体导师

Today’s lecture: introduction of RL and value-based methods

卢宗青
北京大学计算机科学系“博雅”助理教授。在2017年9月加入北京大学之前，他在美国宾夕法尼亚州立大学计算机系从事博士后工作。他于2014年4月获得了新加坡南洋理工大学计算机博士学位，并获得了东南大学的硕士学位和学士学位。他的主要研究领域包括（多智能体）强化学习、移动/边缘智能系统等。