Special Topics in
Learning, Games and Optimization

SUTD 40.616 – Fall 2025

[Introductory Lecture and Course Overview]

Co-instructors:
Anas Barakat
John Lazarsfeld
Joseph Sakos

Course website: learning-in-games.github.io

What is this Course?

- New special topics course on **online learning** and **learning in games**.

Optimization-based paradigm for *sequential decision-making*
in *multi-agent, game-theoretic settings*.

- Co-designed and instructed by three Postdoctoral Research Fellows (ESD Pillar) working in these areas:

John Lazarsfeld

Anas Barakat

Joseph Sakos

- **Target audience:** graduate students broadly interested in **ML theory** *(game theory, multi-agent learning, RL, optimization, economics, …)*
- **Objective**: introduce core algorithms, results, proof techniques; bring students to frontier of research in these fields (very active area of ML).

Today's Class

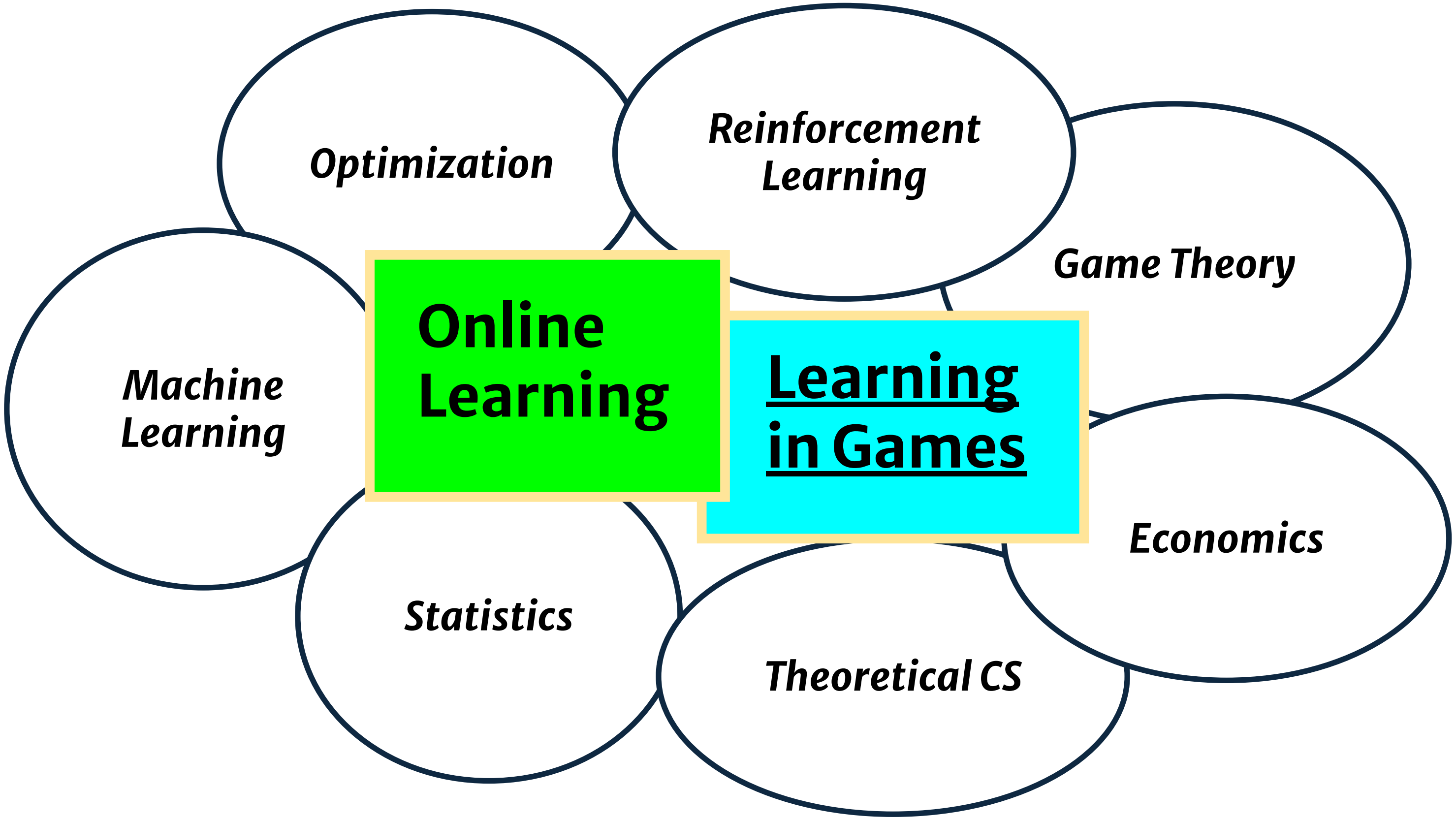
- (1) Overview of topics and motivating examples.
- (2) Course logistics (*schedule*, *assignments*, *project*)

- (3) **Lecture 1: Introduction to Online Learning**

(These slides will be posted on eDimension / course website)

Overview of Topics

Overview of Topics



This course:
1. Introduce fundamental algorithms and concepts in online learning.
2. Study (non)-equilibria arising from simultaneous use of learning algorithms in multi-agent settings.

Online Learning

Online Learning

Sequential decision-making setting:

Every round/day….

1. ***Learner***: chooses an action from a fixed set.
2. ***Nature***: chooses a *loss function* over actions.
3. Learner incurs loss and observes feedback.

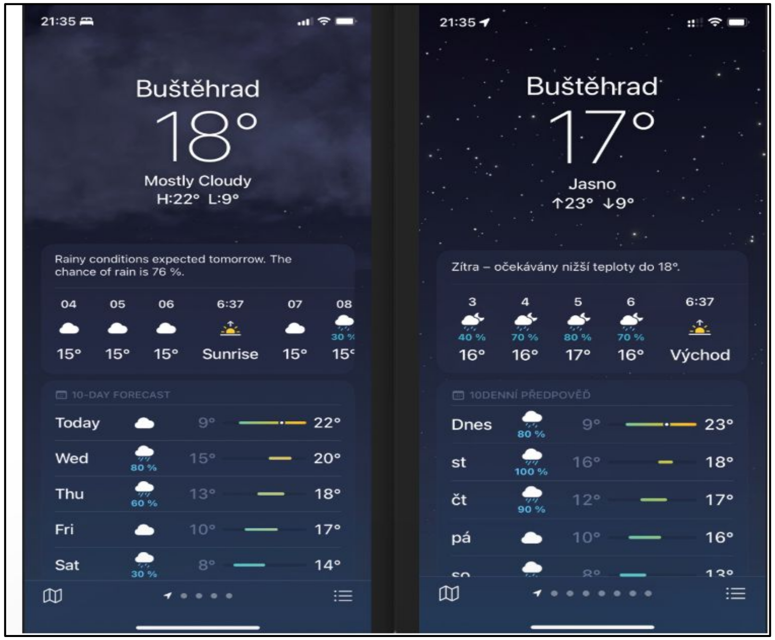
Goal of learner: minimize total incurred loss over time.

Learning = choosing a sequence of actions with total loss not much worse than that of a fixed benchmark sequence.

Challenge: *non-stationarity* **–** the future is unknown, and losses can be adversarially changing with time!

Online Learning

Example 1: Weather forecast publishing



Data sources

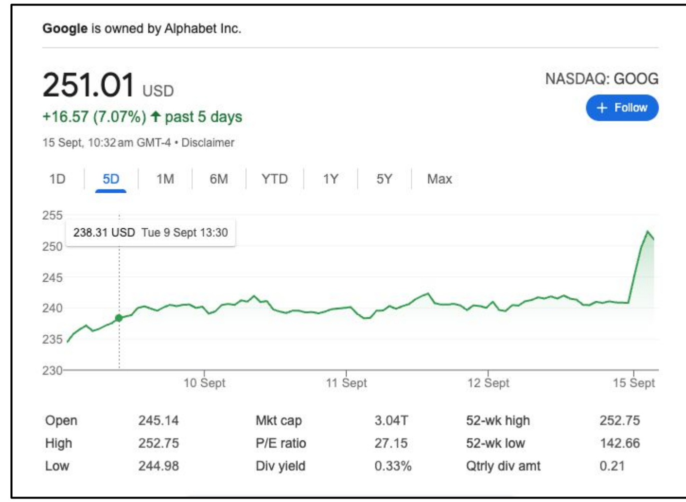
- Australian Government Bureau of Meteorology
- BreezoMeter
- CMA Public Meteorological Service Centre
- Deutscher Wetterdienst
- Environment and Climate Change Canada
- EUMETNET - MeteoAlarm
- European Centre for Medium-Range Weather Forecasts (ECMWF)
- India Meteorological Department
- Instituto Nacional de Meteorologia
- Japan Meteorological Agency
- National Weather Service/National Oceanic and Atmospheric Administration (NOAA)
- QWeather
- Servicio Meteorológico Nacional
- Thai Meteorological Department
- The Met Office
- The Weather Channel

Setting: weather app needs to display current conditions and hourly/weekly forecast, but many possible data sources to choose from.

After each day, can evaluate whether each source was accurate.

Online Learning

Example 2: Portfolio Selection

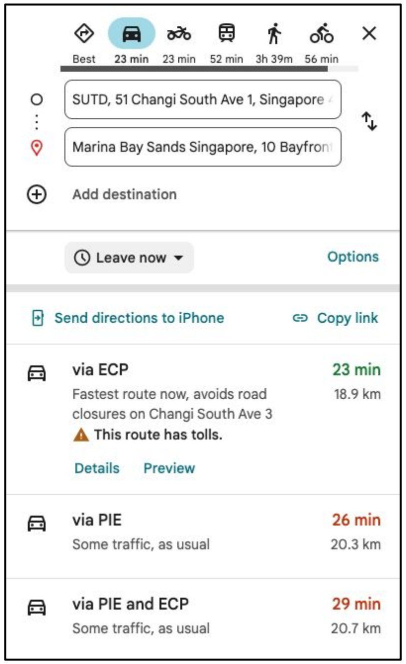


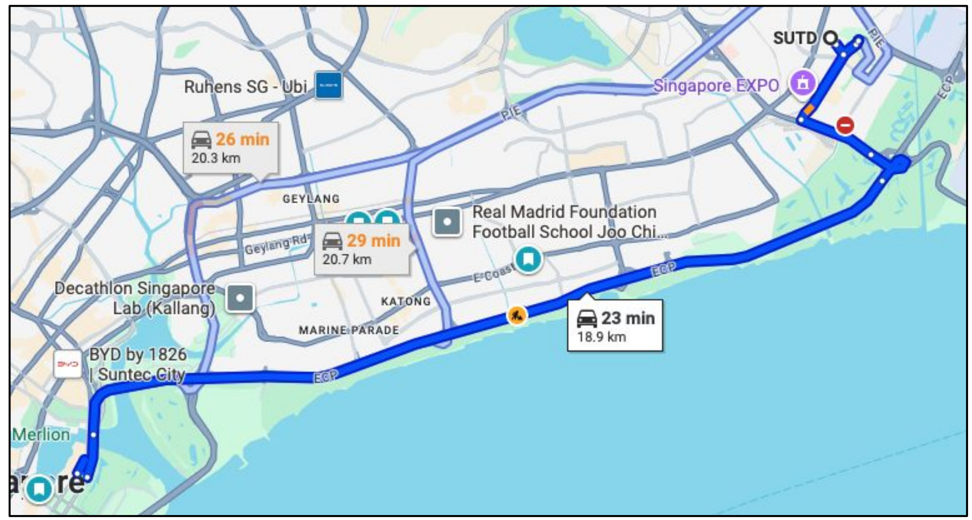
Setting: investment manager must allocate money across a set of stocks (e.g., 20% in Apple, 10% in Google, etc.)

After each day, can observe a company's change in share price and update portfolio weightings.

Online Learning

Example 3: Online Shortest Paths





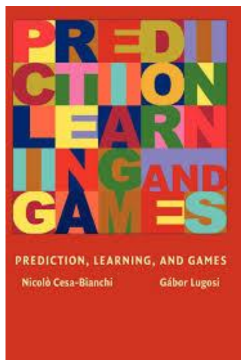
Setting: recommend route to user requesting directions.

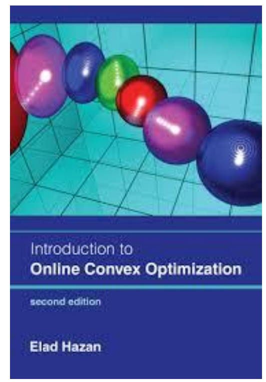
Can observe true travel time for a route after a user completes a trip.

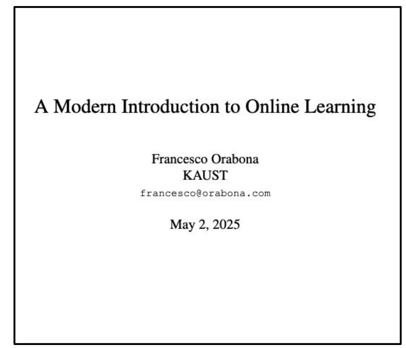
Online Learning

Modern paradigm: iterative, gradient-based algorithms.

- Online learning / online convex optimization: active area of research since 1990s!







- Today: standardized models and performance metrics (***regret***).
- Powerful toolkit of (near-)optimal algorithms and analysis techniques (*e.g.*, *regularization*, *adaptive learning rates*, *convex analysis tools*, *etc.*).

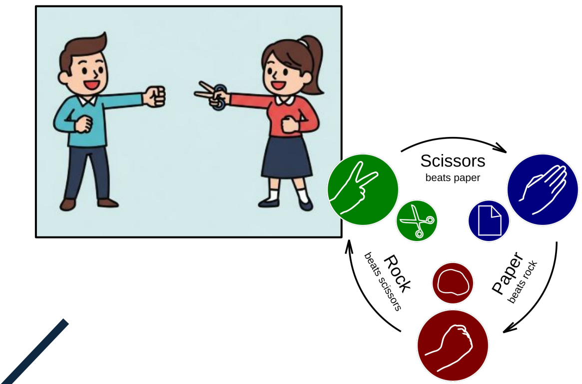
Online learning/optimization: provides the algorithmic/theoretical foundation for problems in statistical learning, RL, economics, … .

Learning in Games

Learning in Games

Online Learning

Game Theory



Learning in Games

Game Theory

Framework for studying strategic interactions and incentives.

Players/Agents: choose actions and receive a utility/payoff.

Agents' utility functions depend jointly on actions of all agents!

Example: *Rock-Paper-Scissors*

Alice: chooses action x from {R, P, S}
Bob: chooses action y from {R, P, S}

Given (x, y):
- Alice gains **A**(x, y)
- Bob loses **A**(x,y)

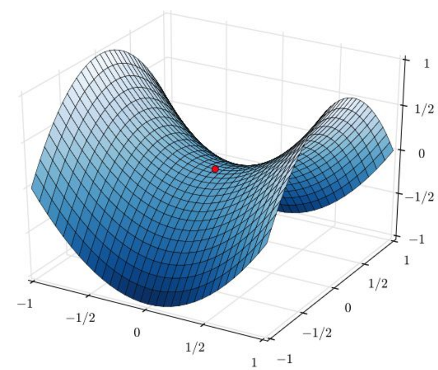
A =

	R	P	S
R	0	-1	1
P	1	0	-1
S	-1	1	0

Key object of study: **equilibria** of games

Informally: actions that capture *stability:* agents have no incentive to change their actions.

Mathematically: solutions to certain (constrained) optimization problems



Learning in Games

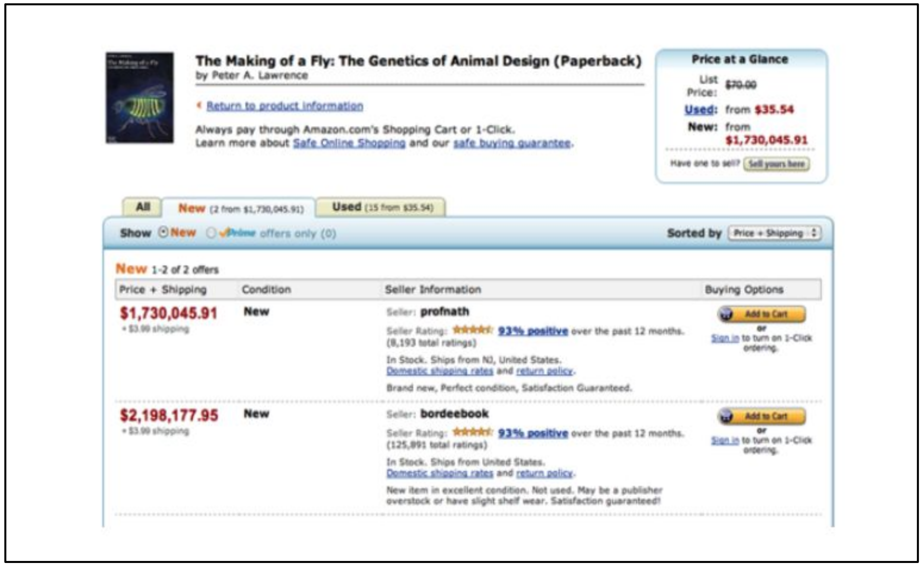
- **Main view:** decentralized learning of equilibria via repeated play.
- ***Motivation*:** study collective behavior of individual *learning agents* in multi-agent settings. *What outcomes emerge from simultaneous use of online learning algorithms?*

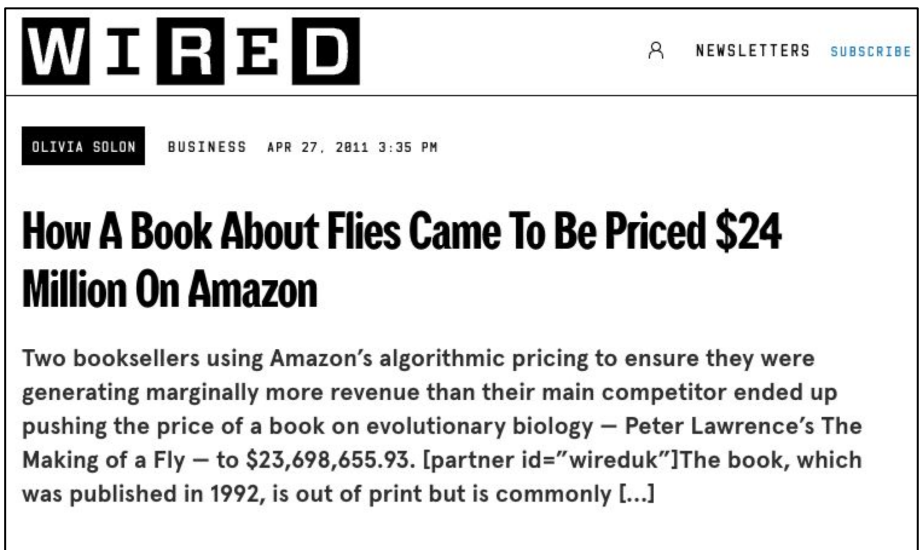
Fundamental Theorem of Learning in Games: in many games settings, performance guarantees for *single-agent online learning* correspond to global convergence guarantees to an underlying equilibrium.

- **In the past decade:** many advances in the theory and practice of learning in games (training GANs, self-play in RL, solving large games).
- **Today:**
  - Complex, multi-agent systems increasingly prevalent.
  - Heightened need for understanding (global) behavior of learning algorithms in these systems.

Goal of this class: give foundational tools for addressing these problems.

Example 1: Algorithmic Pricing in Online Marketplaces





WIRED

NEWSLETTERS SUBSCRIBE

OLIVIA SOLON BUSINESS APR 27, 2011 3:35 PM

How A Book About Flies Came To Be Priced $24 Million On Amazon

Two booksellers using Amazon's algorithmic pricing to ensure they were generating marginally more revenue than their main competitor ended up pushing the price of a book on evolutionary biology – Peter Lawrence's The Making of a Fly – to $23,698,655.93. [partner id="wireduk"]The book, which was published in 1992, is out of print but is commonly [...]

- Sellers use algorithmic pricing methods based on other observable prices in market.
- Simultaneous deployment of such learning algorithms can lead to undesirable outcomes.

Example 2: Online Advertising Systems

THE WALL STREET JOURNAL.

English Edition ▾ | Print Edition | Video | Audio | Latest Headlines | More ▾

World Business U.S. Politics Economy Tech Markets & Finance Opinion Arts Lifestyle Real Estate Perso

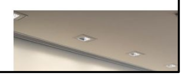
CMO TODAY

AI Will Soon Dominate Ad Buying, Whether Marketers Like It or Not

Tech giants want marketers to let go of some control and trust their algorithms

By Patrick Coffee Follow

March 6, 2025 at 6:00 am ET



- **Multi-billion dollar industry:** complex incentive structure between platform, advertisers, and users.

TECHNOLOGY

Facebook Algorithm Shows Gender Bias in Job Ads, Study Finds

Researchers found the platform's algorithms promoted roles to certain users; company pledges to continue work in removing bias from recommendations

CNN BUSINESS Audio Log In

Facebook's algorithm shows gender bias when promoting ads in several countries, rights group finds

Data collected in France, the Netherlands, India, South Africa, Ireland and the UK suggests that Facebook's ad platform shows open job postings to users based on historical gender stereotypes*, rights group Global Witness says.

- **Undesirable outcomes may emerge –** broader questions on *social welfare.*

Example 3: Rideshare Surge Pricing

Second study finds Uber used opaque algorithm to dramatically boost profits

US academics say computer code systematically raised fares at expense of drivers and passengers

Simon Goodley
Wed 25 Jun 2025 10.45 BST
Share

In 2024, Uber reported it had generated $6.9bn of cash durin[…] Hart/Alamy

The Guardian

UK Uber drivers' earnings cut after changes to secretive algorithm

Introduction of 'dynamic pricing' algorithm also coincided with company raising trip prices, study finds

Home 〉 News 〉 New Oxford research reveals Uber's algorithmic pricing leaves drivers and passengers worse off

New Oxford research reveals Uber's algorithmic pricing leaves drivers and passengers worse off

PUBLISHED
23 JUN 2025

SHARE THIS

MATHEMATICAL, PHYSICAL AND LIFE SCIENCES | RESEARCH | SOCIETY

A new study from researchers in the University of Oxford's Department of Computer Science has found that Uber's use of dynamic pricing has led to higher fares for passengers and lower earnings for drivers, whilst increasing Uber's share of revenue.

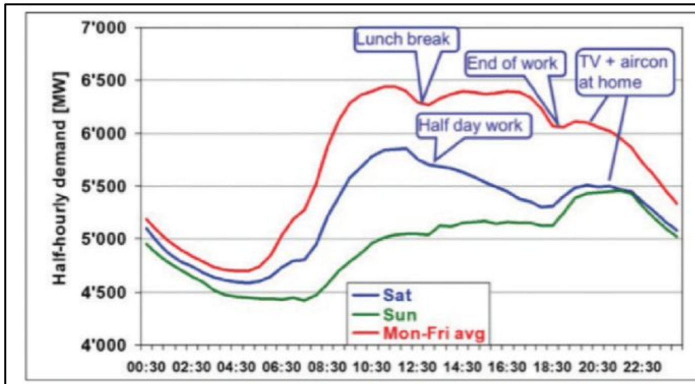
- **Setting:** rideshare platforms use *surge pricing* to adjust to demand/supply in real time.
- **Non-equilibrium outcomes:** resulting from pricing algorithms and strategic behavior of drivers

Example 4: Energy Markets

The AI Energy Crisis: How Big Tech's Power Demands Are Reshaping Global Infrastructure Markets

TrendPulse Finance • Thursday, Aug 14, 2025 7:18 pm ET

- **New paradigms for energy markets:** by 2025, AI data centers will consume 22% of the electricity used by all U.S. households.
- **More agents in system:** individual households increasingly sell excess energy back to power grids.

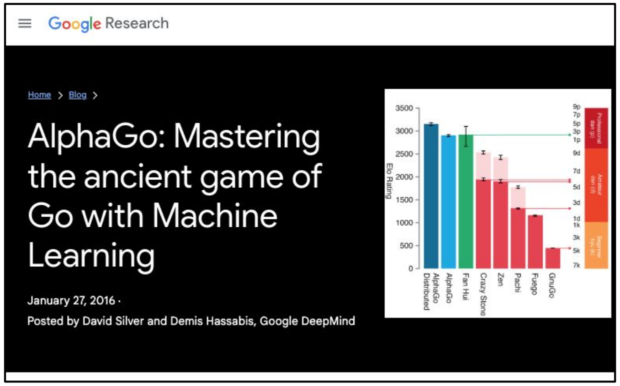


Source: Image based on data from EMA—half-hourly demand from 23–29 Sep 2013.

- **Non-equilibrium behavior** can have negative environmental impacts.

Image sources: altergra.com, Hickem (2020),

Example 5: Superhuman Game Play and RL



Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm

David Silver,[1*] Thomas Hubert,[1*] Julian Schrittwieser,[1*] Ioannis Antonoglou,[1] Matthew Lai,[1] Arthur Guez,[1] Marc Lanctot,[1] Laurent Sifre,[1] Dharshan Kumaran,[1] Thore Graepel,[1] Timothy Lillicrap,[1] Karen Simonyan,[1] Demis Hassabis[1]

[1]DeepMind, 6 Pancras Square, London N1C 4AG.
*These authors contributed equally to this work.

- **Learning in games paradigm:** used to achieve *superhuman levels* in Go via *self-play.*

COMPUTER SCIENCE

Human-level play in the game of *Diplomacy* by combining language models with strategic reasoning

Meta Fundamental AI Research Diplomacy Team (FAIR)†, Anton Bakhtin[1]‡, Noam Brown[1*]‡, Emily Dinan[1*]‡, Gabriele Farina[1], Colin Flaherty[1]‡, Daniel Fried[1,2], Andrew Goff[1], Jonathan Gray[1]‡, Hengyuan Hu[1,3]‡, Athul Paul Jacob[1,4]‡, Mojtaba Komeili[1], Karthik Konath[1], Minae Kwon[1,3], Adam Lerer[1*]‡, Mike Lewis[1*]‡, Alexander H. Miller[1]‡, Sasha Mitts[1], Adithya Renduchintala[1]‡, Stephen Roller[1], Dirk Rowe[1], Weiyan Shi[1,5]‡, Joe Spisak[1], Alexander Wei[1,6], David Wu[1]‡, Hugh Zhang[1,7]‡, Markus Zijlstra[1]

- **More recently:** extend approach to more complex games requiring cooperation, like *Diplomacy*.

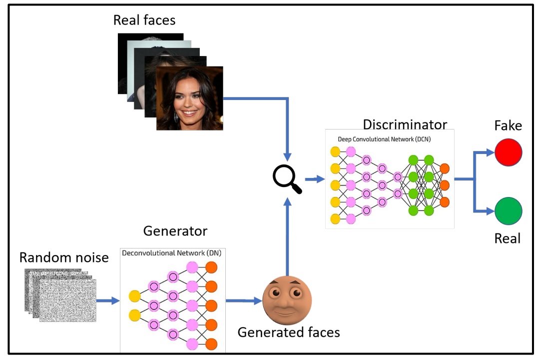
- **Also:** learning in games approach for LLM alignment via RLHF.

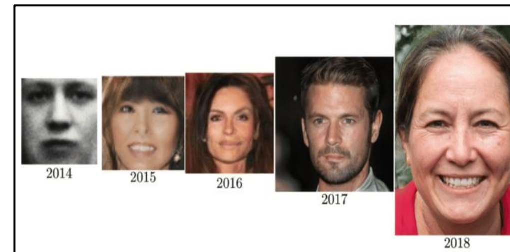
COMAL: A **Co**nvergent **M**eta-**A**lgorithm for **Al**igning LLMs with General Preferences

Example 6: Adversarial Training and Robust ML

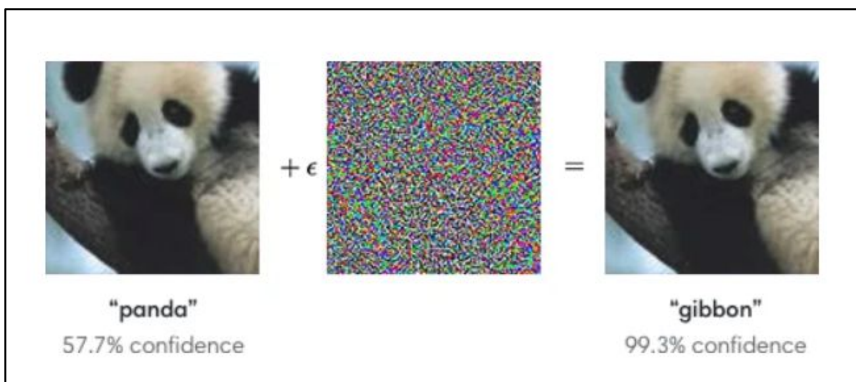
- **Generative Adversarial Networks (GANS):** training objective as minmax optimization problem.

TRAINING GANS WITH OPTIMISM





Progress of human face generation with GAN



- **Adversarial training:** make models robust to "attacks".

$$\underset{\theta}{\text{minimize}} \frac{1}{|S|} \sum_{x,y \in S} \max_{\|\delta\| \le \epsilon} \ell(h_\theta(x+\delta), y).$$

Image sources: *Cheng+2020, ResearchGate; Spindox.it, OpenAI, adversarial-ml-tutorial.org;*

Goals of this Course

Goal 1: Teach fundamental techniques and concepts, in order to understand latest research in these and adjacent fields.

Goal 2: Prepare students to pursue research in these areas.

Learning outcomes – understand the online learning paradigm for learning in games, including:

- Fundamental algorithms
- Key results and the landscape of open questions
- Core analysis techniques for proving rigorous guarantees.

Course Structure

- **Lectures:** Tuesdays/Thursdays, 10am–12pm

  Four parts: ***Part I: Introduction to Online Learning*** (John)
  Part II: Learning in Normal Form and Stochastic Games (Anas)
  Part III: Learning in Extensive Form and Continuous Games (Joseph)
  Part IV: Special Topics (all 3 Instructors)

  Notes/slides posted on eDimension/course website.

- **Evaluation:** 2 problem sets and Final Project
  - Problem sets – 40% of grade (20% each)
  - Final project – 60% of grade

Overview of Topics – Part I

Part I: Online Learning – *Lectures given by John*

- (L01) – Introduction to Online Learning
  Prediction with expert advice, online convex optimization, regret, Multiplicative Weights Update and Online Gradient Descent.
- (L02) – Follow-the-Regularized-Leader: No-regret via Regularization
  Family of leader-based algorithms, analysis of Follow-the-Regularized-Leader (FTRL) via coupling with Be-the-Leader/Follow-the-Leader, Multiplicative Weights Update as FTRL, lower bounds for online learning.
- (L03) – Follow-the-Perturbed-Leader and Online Mirror Descent: No-regret via Perturbation and Penalty
  Follow-the-Perturbed-Leader (FTPL) analysis, equivalence between FTPL and FTRL, Online Mirror Descent analysis.
- (L04) – Online Learning with Bandit Feedback
  Bandit feedback model, expected regret and pseudo-regret, EXP3 algorithm for adversarial bandits, Explore-then-Commit and UCB algorithms for stochastic bandits.
- (L05) – Φ-Regret Minimization
  Beyond external regret: swap-regret, internal-regret, and Φ-regret framework. Blum-Mansour and Stoltz-Lugosi algorithms.
- (L06) – Blackwell Approachability and Regret Matching
  Blackwell's Approachability theorem, Regret Matching (RM) and Regret Matching+ (RM+) algorithms.

Overview of Topics – Part II

Part II: Learning in Normal-Form and Stochastic Games – *Lectures given by Anas*

- (L07) – Introduction to Normal-Form Games and Nash Equilibria
  Normal-form games, Nash equilibria (NE), game classes (potential, zero-sum, decomposition).
- (L08) – No-Regret Learning in Games and Learning NEs in Zero-Sum and Potential Games
  Hindsight rationality, proof of minimax theorem via online learning, learning NE in potential games.
- (L09) – Learning (Coarse)-Correlated Equilibria in General-Sum Games
  (Coarse)-correlated equilibria, time-average convergence via no-ϕ-regret learning, average vs. last-iterate convergence.
- (L10) – Optimistic Learning and Social Welfare of No-Regret Dynamics
  Optimistic FTRL algorithms, RVU bounds, individual vs. sum of regrets, fast convergence of social welfare.
- (L11) – Introduction to Stochastic Games and Multi-Agent Reinforcement Learning
  Introduction to Markov Decision Processes (MDPs) and Reinforcement Learning, definition of stochastic games, Shapley's minimax theorem, existence of Nash equilibria.
- (L12) – Learning Equilibria in Stochastic Games
  Independent and decentralized learning, zero-sum Markov games and Markov potential games, policy gradient methods.

Overview of Topics – Part III

Part III: Learning in Extensive-Form and Continuous Games – *Lectures given by Joseph*

- (L13) – Introduction to Extensive-Form Games
  Game trees, imperfect information, perfect recall, strategy representations, Kuhn's theorem.
- (L14) – Learning Equilibria in Extensive-Form Games
  Counterfactual Regret Minimization (CFR) algorithm and speedups.
- (L15) – Introduction to Continuous Games
  Concave games, Rosen's theorem, variational inequalities, monotone games, zero-sum games and Gradient Descent Ascent (GDA), divergence of GDA in bilinear case.
- (L16) – Learning Equilibria in Continuous Games
  Proximal point method, Optimistic GDA and Extragradient algorithms for zero-sum games, learning equilibria in potential games, general concave games.
- (L17) – Price of Anarchy and Equilibrium Selection
  Braess's paradox, Pigou's network, smooth games, introduction to Price of Anarchy (PoA) bounds.

Overview of Topics – Part IV

Part IV: Special Topics

The final six lectures will cover more advanced topics based on results in the field over the past five years:

- (L18) – Online Learning in Time-Varying Games (Anas)
- (L19) – (Multi-Agent) Online Nonstochastic Control (Anas)
- (L20) – Bridging Continuous-time and Discrete-time Learning in Games (John)
- (L21) – Unregularized Learning in Zero-Sum Games (John)
- (L22) – Sum-of-Squares Optimization in Games (Joseph)
- (L23) – Hidden Games (Joseph)

Assignments

- **2 Problem Sets (40% of total grade)**
  - Problem Set A: released Friday, Sep 18; due Friday, Oct 10.
  - Problem Set B: released Friday, Oct 17; due Friday, November 14.
  - Each problem set will consist of roughly 4–5 exercises related to algorithms and analysis techniques covered in lectures.
  - Can collaborate with classmates, but each student must submit their own assignment.

Final Project

- **Final project containing 3 components (60% of total grade).**
  - Project is based on reading, synthesizing, and presenting on several related research papers based on topics covered in class.
  - Topics and research papers will be based on recent works published in top ML venues (ICML, NeurIPS, COLT, ICLR) etc.

- **Structure of project / important dates**
  - Start of Week 2 (Sep 23): list of project topics announced by instructors
  - End of Week 3 (Oct 02): students rank topics; matched with project.
  - **Midterm Presentation** (10% of grade) – in class on Tues, November 4.
    - 5-10 min presentations per student (including questions)
    - Goal: give initial overview of paper topics; receive some feedback.
  - **Final Presentation** (20% of grade) – in class on Dec 16 and Dec 18
  - **Final Report** (30% of grade) – due Friday, Dec 19.

Final Project

- **Purpose of project**: gain methodological research experience
  - After matching with topic (based on an active area of research in online learning and learning in games), instructors will provide 2-3 recent papers on the topic.
  - **Your goal**: understand the main contributions, techniques, and connections between the papers. Identify possible directions for future work.
  - **Why?** Gain experience in quickly (but deeply) learning about and understanding a new line of work.

    Midterm presentation (5min), final presentation (15min), and final report will involve demonstrating this understanding.
  - For motivated students, your project can serve as a springboard for a full research project supervised by the instructors.
- **More details on rubric/structure/examples in upcoming lectures**

Resources

- Course material is designed from scratch, but further reading on topics can be found in following texts:
  - **Cesa-Bianchi and Lugosi, 2006**. *Prediction, Learning, Games.*
  - **Hazan**, **2016**. *Introduction to Online Convex Optimization.*
  - **Nisan, Roughgarden, Tardos, and Vazirani**, **2007**. *Algorithmic Game Theory.*
  - **Orabona**, **2019**. *A Modern Introduction to Online Learning.*

- Additional references/pointers will be given throughout the lectures.

- All lecture material will be posted on eDimension and course website: learning-in-games.github.io

- **Prerequisites:** Prior courses/background in calculus, linear algebra, and basic notions of probability and analysis. A prior course in optimization is helpful (*See instructors after class if you have questions*).

- **Office Hours:** 1 hour per week (to be announced); or by appointment – email for all course staff: sutd.glo.course@gmail.com