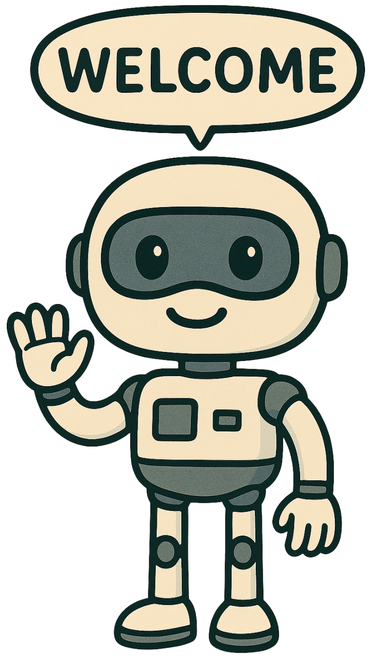
L1: Course Introduction

EECE 571N | Sequential Decision Making | Fall 2025

Cyrus Neary | cyrus.neary@ubc.ca



Who Am I?

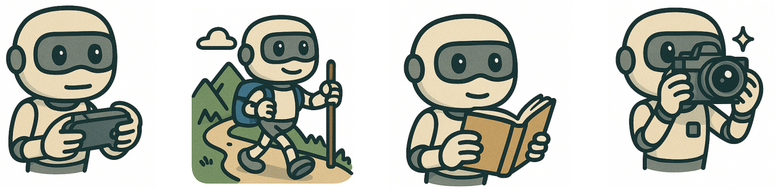
Undergrad from UBC in Engineering Physics and Mathematics.

MSc and PhD from The University of Texas at Austin in Computational Science, Engineering, and Mathematics.

Postdoctoral research at Mila - The Québec AI Institute.

Assistant professor in Electrical and Computer Engineering.
Research: Theory, algorithms, and applications of AI-driven autonomous systems in robotics and engineering.

Hobbies: Reading (sci-fi/fantasy), photography, hiking, gaming.



Contact: cyrus.neary@ubc.ca. Include [Fall2025SDM] in subject line.

Who are all of you?

What is Sequential Decision-Making?

Sequential decision making

Add languages

Article Talk | Read Edit View history Tools

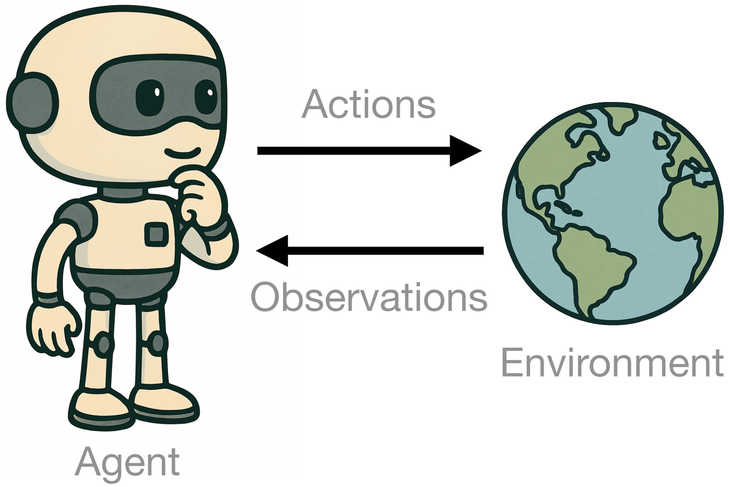
From Wikipedia, the free encyclopedia

Sequential decision making is a concept in control theory and operations research, which involves making a series of decisions over time to optimize an objective function, such as maximizing cumulative rewards or minimizing costs. In this framework, each decision influences subsequent choices and system outcomes, taking into account the current state, available actions, and the probabilistic nature of state transitions.[1] This process is used for modeling and regulation of dynamic systems, especially under uncertainty, and is commonly addressed using methods like Markov decision processes (MDPs) and dynamic programming.[2]

Mathematical modeling
Algorithms and solution techniques
Learning from data
Verifying generated behaviors

Approximate Dynamic Programming

Optimal Control



Game Theory

Reinforcement Learning

AI Planning

Robotics & Autonomous Systems

Economics & Decision Theory

Multiagent Systems

Why Study Sequential Decision-Making?

Continuous Control

Flight control

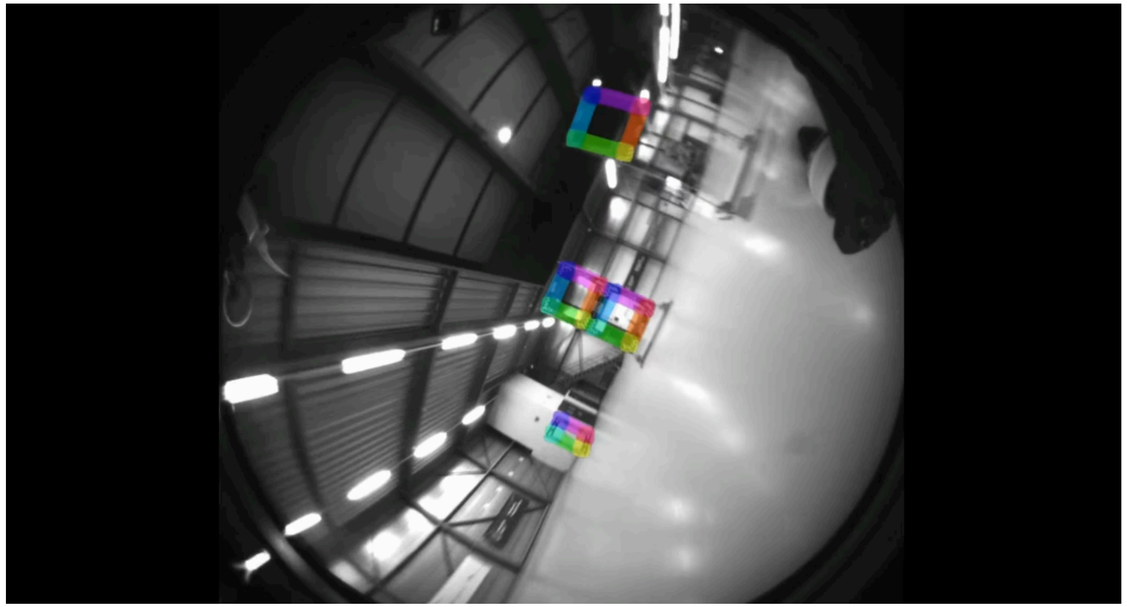
Article

Champion-level drone racing using deep reinforcement learning

https://doi.org/10.1038/s41586-023-06419-4
Received: 5 January 2023
Accepted: 10 July 2023
Published online: 30 August 2023
Open access

Elia Kaufmann[1✉], Leonard Bauersfeld[1], Antonio Loquercio[1], Matthias Müller[2], Vladlen Koltun[3] & Davide Scaramuzza[1]

First-person view (FPV) drone racing is a televised sport in which professional competitors pilot high-speed aircraft through a 3D circuit. Each pilot sees the



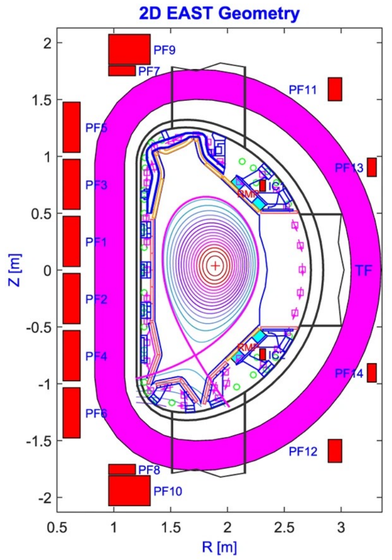
Controlling fusion reactors

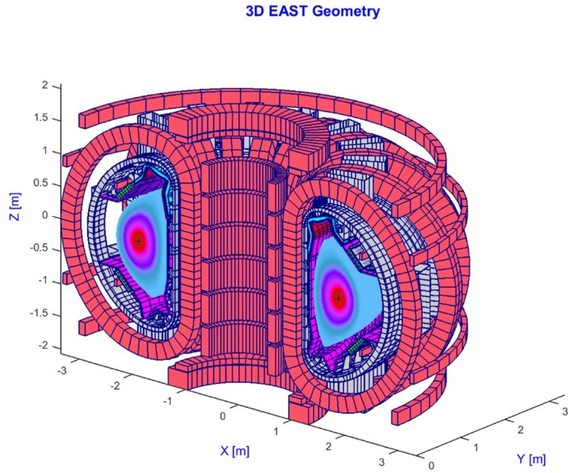
Article

Magnetic control of tokamak plasmas through deep reinforcement learning

https://doi.org/10.1038/s41586-021-04301-9
Received: 14 July 2021
Accepted: 1 December 2021
Published online: 16 February 2022
Open access

Jonas Degrave[1,3], Federico Felici[2,3✉], Jonas Buchli[1,3✉], Michael Neunert[1,3], Brendan Tracey[1,3✉], Francesco Carpanese[1,2,3], Timo Ewalds[1,3], Roland Hafner[1,3], Abbas Abdolmaleki[1], Diego de las Casas[1], Craig Donner[1], Leslie Fritz[1], Cristian Galperti[2], Andrea Huber[1], James Keeling[1], Maria Tsimpoukelli[1], Jackie Kay[1], Antoine Merle[2], Jean-Marc Moret[2], Seb Noury[1], Federico Pesamosca[2], David Pfau[1], Olivier Sauter[2], Cristian Sommariva[2], Stefano Coda[2], Basil Duval[2], Ambrogio Fasoli[2], Pushmeet Kohli[1], Koray Kavukcuoglu[1], Demis Hassabis[1] & Martin Riedmiller[1,3]



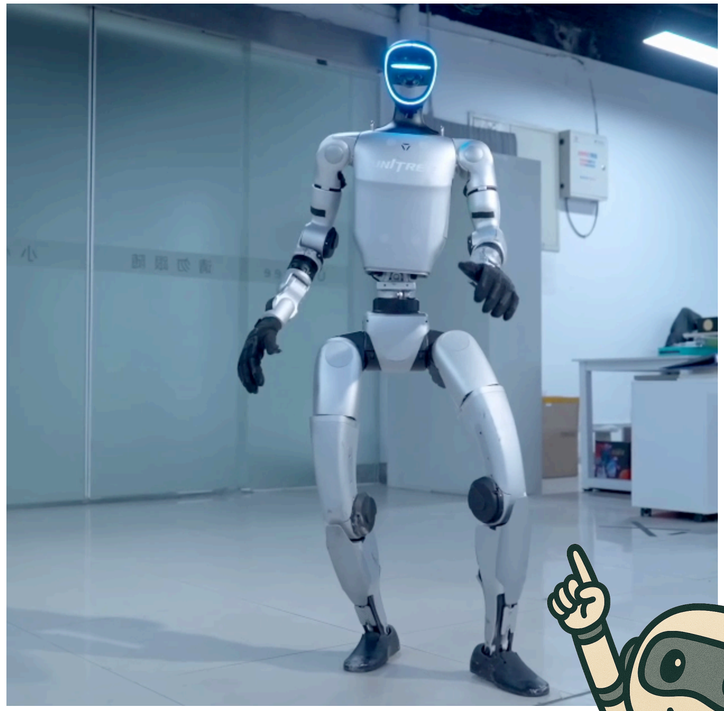


Why Study Sequential Decision-Making?

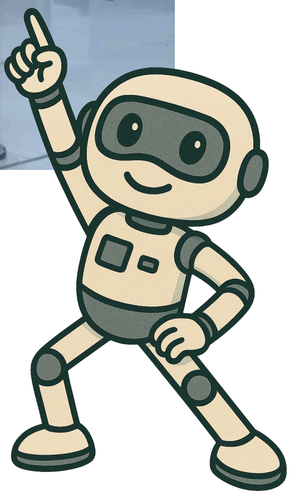
Robots…

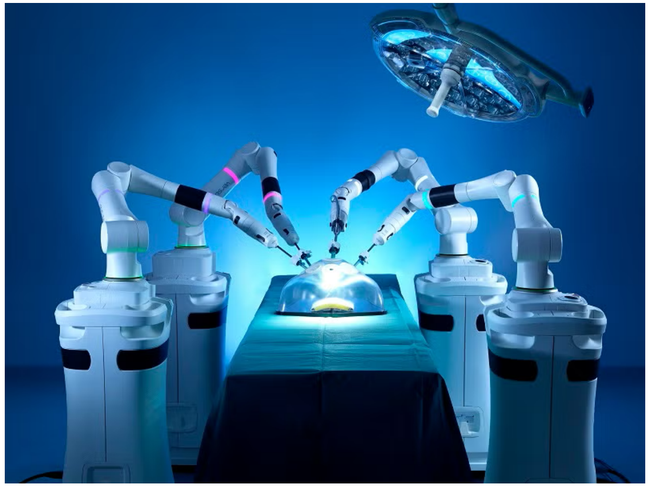
That optimize supply chains…

that explore the unknown…



and that dance!

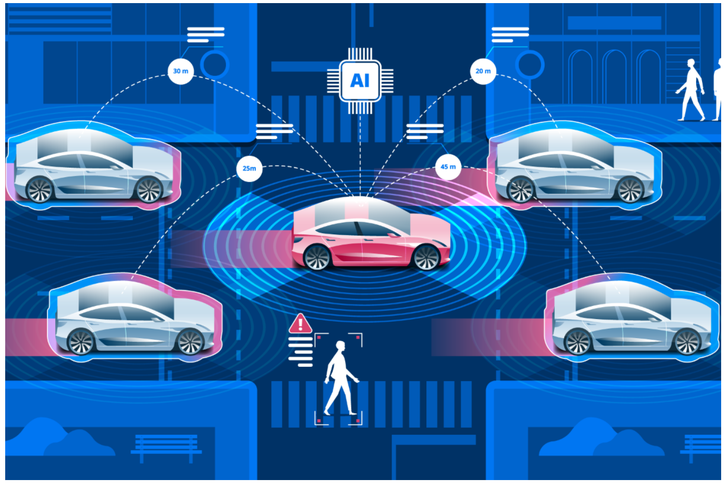




that save lives…

Why Study Sequential Decision-Making?

Autonomous Decision-Making

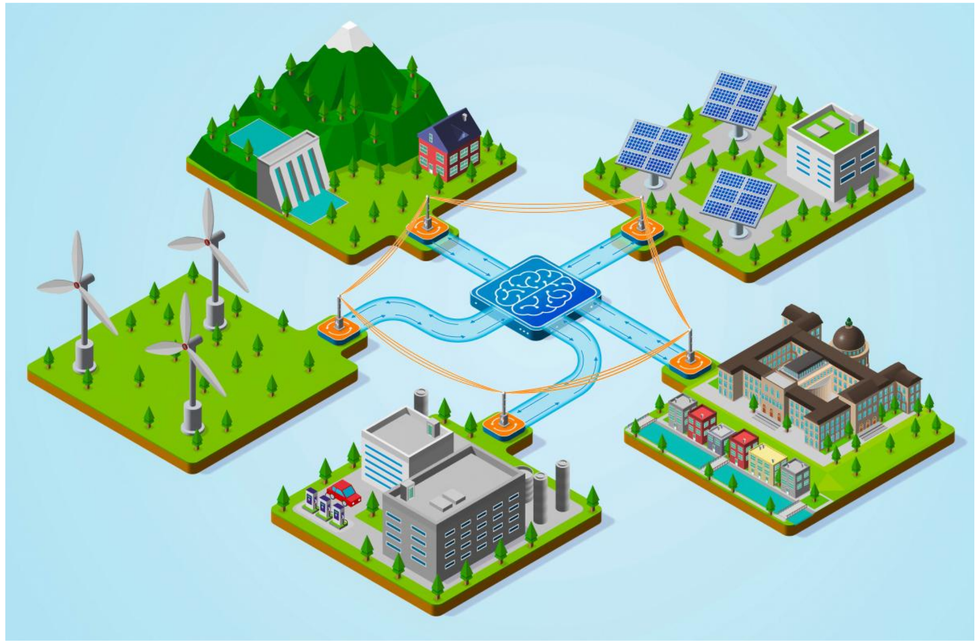


Autonomous vehicles

Automated trading

Why Study Sequential Decision-Making?

Controlling Networked Systems



Smart power grids

City-wide traffic control

Why Study Sequential Decision-Making?

Game Playing

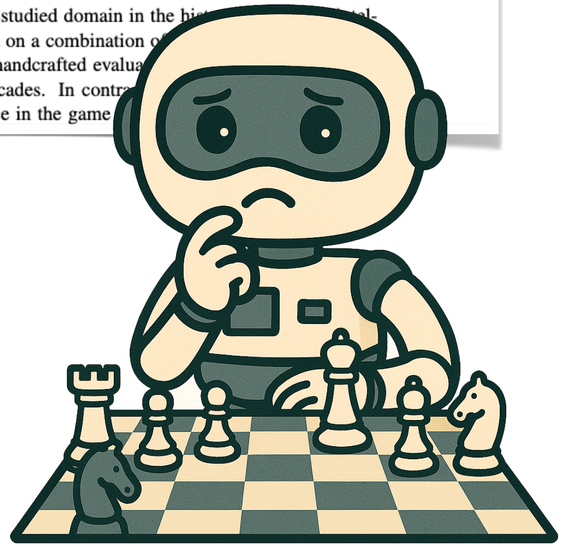
Mastering Chess and Shogi by Self-Play with a General Reinforcement Learning Algorithm

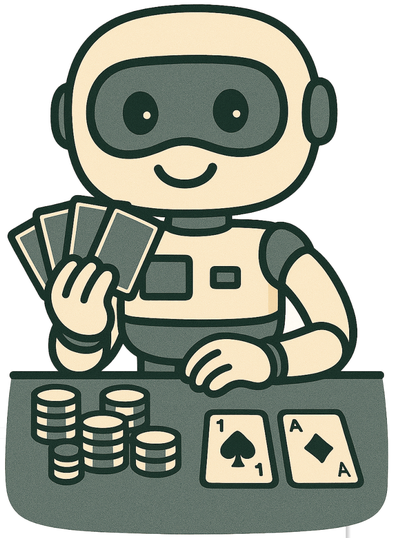
David Silver,[1*] Thomas Hubert,[1*] Julian Schrittwieser,[1*]
Ioannis Antonoglou,[1] Matthew Lai,[1] Arthur Guez,[1] Marc Lanctot,[1]
Laurent Sifre,[1] Dharshan Kumaran,[1] Thore Graepel,[1]
Timothy Lillicrap,[1] Karen Simonyan,[1] Demis Hassabis[1]

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

Abstract

The game of chess is the most widely-studied domain in the hi... ligence. The strongest programs are based on a combination o... niques, domain-specific adaptations, and handcrafted evalua... refined by human experts over several decades. In contra... recently achieved superhuman performance in the game...





RESEARCH ARTICLE

COMPUTER SCIENCE

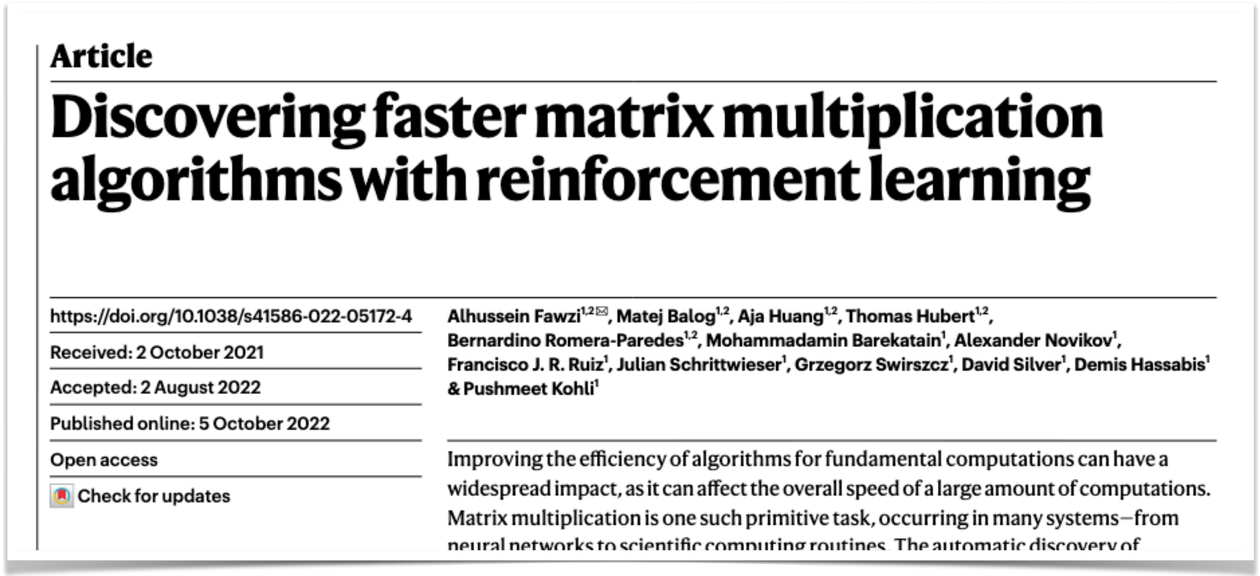
Superhuman AI for multiplayer poker

Noam Brown[1,2*] and Tuomas Sandholm[1,3,4,5*]

In recent years there have been great strides in artificial intelligence (AI), with games often serving as challenge problems, benchmarks, and milestones for progress. Poker has served for decades as such a challenge problem. Past successes in such benchmarks, including poker, have been limited to two-player games. However, poker in particular is traditionally played with more than two players. Multiplayer games present fundamental additional issues beyond those in two-player games, and multiplayer poker is a recognized AI milestone. In this paper we present Pluribus, an AI that we show is stronger than top human professionals in six-player no-limit Texas hold'em poker, the most popular form of poker played by humans.

Why Study Sequential Decision-Making?

Automating Design and Discovery



Article

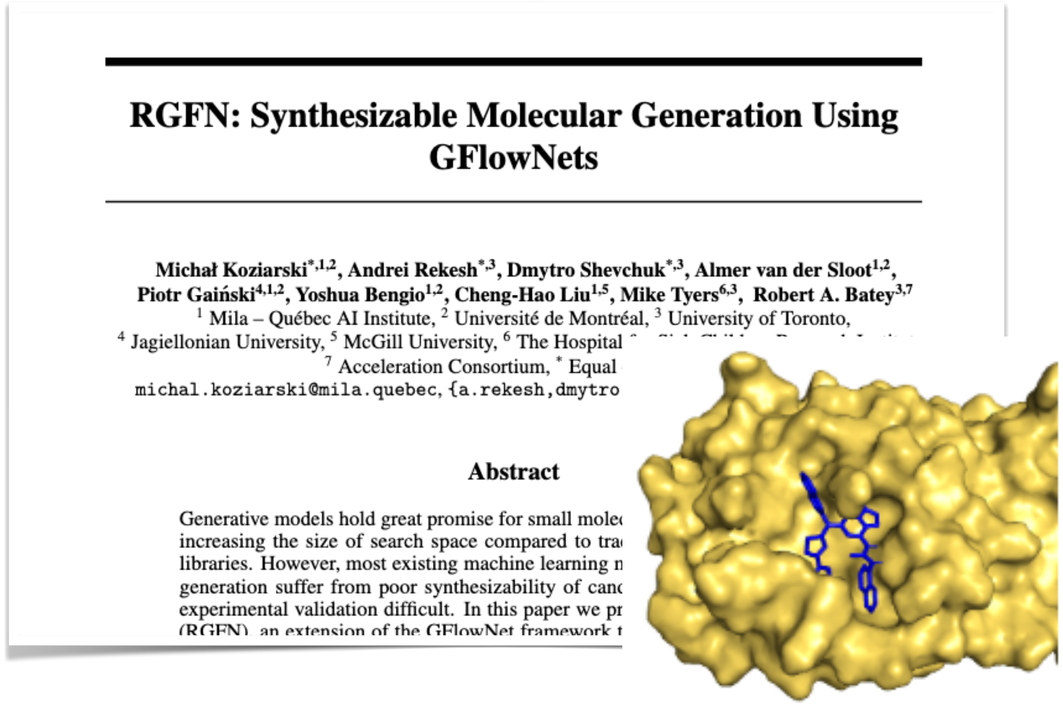
Discovering faster matrix multiplication algorithms with reinforcement learning

https://doi.org/10.1038/s41586-022-05172-4
Received: 2 October 2021
Accepted: 2 August 2022
Published online: 5 October 2022
Open access
Check for updates

Alhussein Fawzi[1,2], Matej Balog[1,2], Aja Huang[1,2], Thomas Hubert[1,2], Bernardino Romera-Paredes[1,2], Mohammadamin Barekatain[1], Alexander Novikov[1], Francisco J. R. Ruiz[1], Julian Schrittwieser[1], Grzegorz Swirszcz[1], David Silver[1], Demis Hassabis[1] & Pushmeet Kohli[1]

Improving the efficiency of algorithms for fundamental computations can have a widespread impact, as it can affect the overall speed of a large amount of computations. Matrix multiplication is one such primitive task, occurring in many systems—from neural networks to scientific computing routines. The automatic discovery of

Searching for efficient numerical algorithms



RGFN: Synthesizable Molecular Generation Using GFlowNets

Michał Koziarski[*,1,2], Andrei Rekesh[*,3], Dmytro Shevchuk[*,3], Almer van der Sloot[1,2], Piotr Gaiński[4,1,2], Yoshua Bengio[1,2], Cheng-Hao Liu[1,5], Mike Tyers[6,3], Robert A. Batey[3,7]
[1] Mila – Québec AI Institute, [2] Université de Montréal, [3] University of Toronto,
[4] Jagiellonian University, [5] McGill University, [6] The Hospital
[7] Acceleration Consortium, [*] Equal
michal.koziarski@mila.quebec, {a.rekesh,dmytro

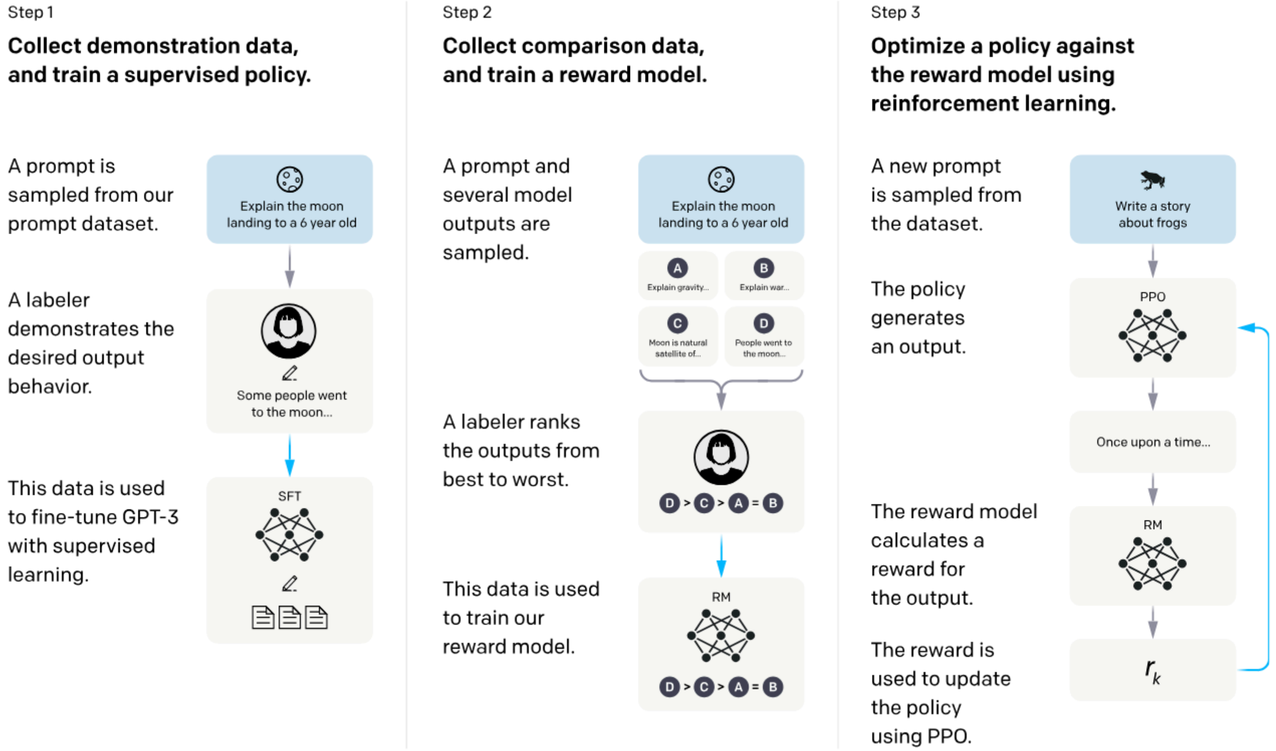
Abstract

Generative models hold great promise for small molec
increasing the size of search space compared to tra
libraries. However, most existing machine learning n
generation suffer from poor synthesizability of canc
experimental validation difficult. In this paper we pr
(RGFN), an extension of the GFlowNet framework t

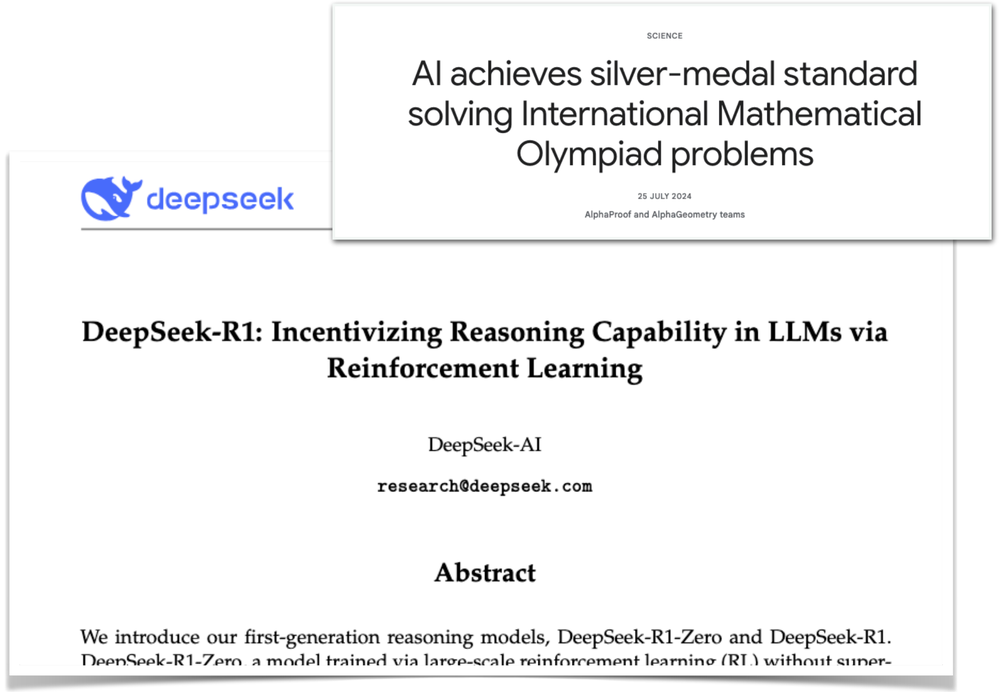
Generating candidate structures for materials and drug design.

Why Study Sequential Decision-Making?

Language Models and Reasoning



Finetuning language models via RL from Human Feedback.

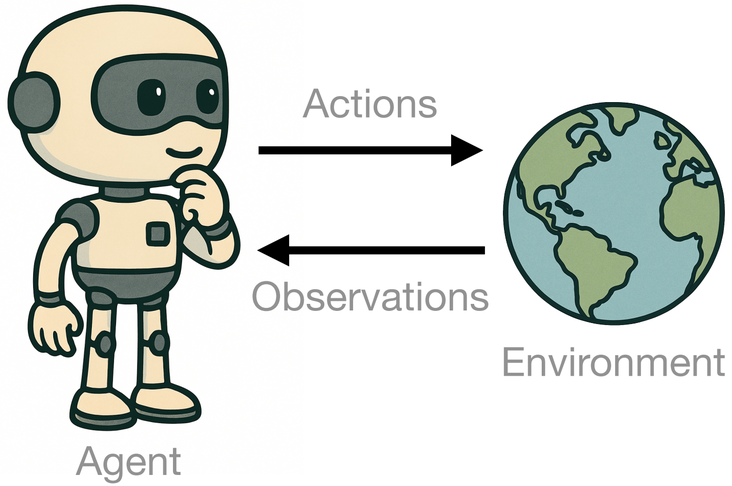


Reasoning as search via RL in the token space of language models.

Why Study Sequential Decision-Making?

Approximate Dynamic Programming

Optimal Control



Game Theory

Reinforcement Learning

AI Planning

Robotics & Autonomous Systems

Economics & Decision Theory

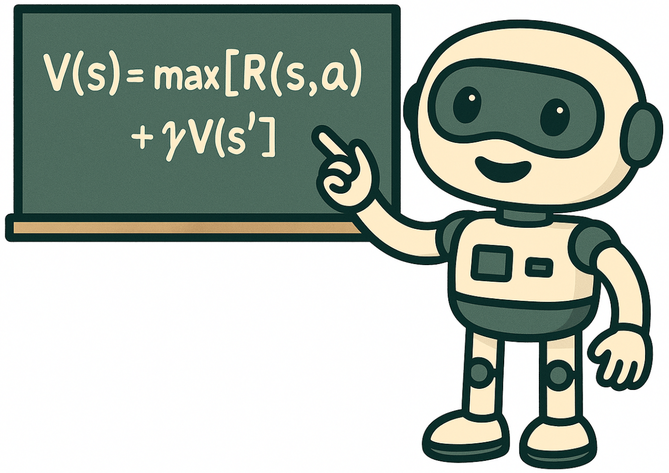
Multiagent Systems

Course Objectives

Formulate real-world decision-making problems mathematically.

Implement and apply dynamic programming and RL algorithms to solve meaningful real-world problems in robotics, control, and autonomous systems.

Critically analyze current research in the field.



Course Outline

Part 1

Modeling and solving sequential decision-making problems via dynamic programming.

- Markov Decision Processes
- Value and Policy Iteration
- LP solutions
- Connections with optimal control
- POMDPs

Part 2

The RL setting and fundamental RL algorithms.

- Exploration vs. Exploitation: Bandits algorithms
- Temporal differencing, Q-learning
- Model-based RL and sample complexity

Part 3

Survey of deep RL algorithms.

- Function approximation
- Deep Q learning
- Challenges with deep RL
- Policy gradient algorithms
- Landmark deep RL algorithms

Part 4

Advanced topics and current research.

- Intro to hierarchical RL
- Intro to multiagent decision-making
- Sim-to-real methods
- Behavioral cloning
- Generalist robotic agents
- Safe sequential decision-making

Course Project

The course project will allow students to explore an application or extension of course concepts in depth, and aims to prepare students to complete related independent research projects. The project will be completed in teams of up to 2–3 students and include a short proposal (graded and required for approval by the course instructor), a midterm progress update, and a final written report and in-class presentation.

Project proposals

- Due September 24th

Project updates

- Due October 29th

Project presentations

- December 1st and 3rd

Project reports

- December 5th

Weekly project surveys

- Every Wednesday at the end of lecture.

Grading

- 5% - In class participation.
- 30% - Take-home homework and programming assignments. The lowest assignment score will be dropped.
- 65% - Course project.
  - 5% - Project proposal.
  - 5% - Progress update report.
  - 10% - Final project presentation.
  - 30% - Final project report.
  - 5% - Weekly project update survey (every Wednesday at the end of lecture).
  - 10% - Individual contribution to project. Assessed through responses to the weekly survey.

Grading will allow for generative AI models to be used as a research tool. However, all submitted work must be produced by the students and reflective of their own ideas. Submitted assignments or course project materials that are produced primarily by generative AI will receive a score of zero.

Let’s get started!