In this lecture series, leading research scientists from leading AI research lab, DeepMind, will give 12 lectures on an exciting selection of topics in Deep Learning, ranging from the fundamentals of training neural networks via advanced ideas around memory, attention, and generative modelling to the important topic of responsible innovation.

Please join us for a deep dive lecture series into Deep Learning!

#UCLxDeepMind
General information

Exits:
At the back, the way you came in

Wifi:
UCL guest
Thore Graepel is a research group lead at DeepMind and holds a part-time position as Chair of Machine Learning at University College London. He studied physics at the University of Hamburg, Imperial College London, and Technical University of Berlin, where he also obtained his PhD in machine learning in 2001. After postdoctoral work at ETH Zurich and Royal Holloway College, University of London, Thore joined Microsoft Research in Cambridge in 2003. At DeepMind since 2015, Thore leads the multi-agent research team and contributed to AlphaGo, the first computer program to defeat a human professional player in the full-sized game of Go.
In this lecture Thore will explain DeepMind's machine learning based approach towards AI. He will give examples of how deep learning and reinforcement learning can be combined to build intelligent systems, including AlphaGo, Capture-The-Flag, and AlphaFold. This will be followed by a short introduction to the different topics and speakers coming up in the subsequent lectures.
Introduction to Machine Learning and AI

Prof Thore Graepel
Chair of Machine Learning, UCL
Research Scientist, DeepMind

03/02/2020
Thanks to all these people (organisers)

- David Barber
  UCL

- Mark Herbster
  UCL

- Pontus Stenetorp
  UCL

- Sarah Hodkinson
  DeepMind

- Thore Graepel
  DeepMind & UCL

- George Kraev
  DeepMind

- Jon Fildes
  DeepMind

- Danielle Breen
  DeepMind

- Dominic Barlow
  DeepMind

- Gaby Pearl
  DeepMind
Plan for this Lecture

01 Solving Intelligence
02 AlphaGo & AlphaZero
03 Learning to Play Capture The Flag
04 Folding Proteins with AlphaFold
05 Overview of Lecture Series
1 Solving Intelligence
A human being should be able to change a diaper, plan an invasion, butcher a hog, conn a ship, design a building, write a sonnet, balance accounts, build a wall, set a bone, comfort the dying, take orders, give orders, cooperate, act alone, solve equations, analyze a new problem, pitch manure, program a computer, cook a tasty meal, fight efficiently, die gallantly. Specialization is for insects.

Robert A Heinlein
Science Fiction Author
What is Intelligence?

Intelligence measures an agent’s ability to achieve goals in a wide range of environments.

\[ \Upsilon(\pi) := \sum_{\mu \in E} 2^{-K(\mu)} V_{\mu}^{\pi} . \]

Measure of Intelligence

Sum over environments

Complexity penalty

Value achieved

Universal Intelligence: A Definition of Machine Intelligence, Legg & Hutter 2007
Reinforcement Learning

- General Purpose Framework for AI
- Agent interacts with the environment
- Select actions to maximise long-term reward
- Encompasses supervised and unsupervised learning as special cases
- Module Reinforcement Learning (UCL COMP0089)
### Why use games to solve AI?

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><strong>Microcosms of the real world</strong>&lt;br&gt;Games are a proving ground for real-world situations</td>
</tr>
<tr>
<td>2</td>
<td><strong>Stimulate intelligence</strong>&lt;br&gt;By presenting a diverse set of challenges</td>
</tr>
<tr>
<td>3</td>
<td><strong>Good to test in simulations</strong>&lt;br&gt;Efficient, run thousands in parallel, faster than real time</td>
</tr>
<tr>
<td>4</td>
<td><strong>Measure progress and performance</strong>&lt;br&gt;Measure progress and compare against human performance</td>
</tr>
</tbody>
</table>
Reinforcement Learning in Games
Superhuman Skill at Playing Atari Games
(Mnih et al, Nature 2015)
Previous systems required feature engineering for every new problem

Deep Learning enables end-to-end learning for a given loss and architecture

Weak prior knowledge can be encoded via architecture (e.g. convolutions, recurrence)

Deep Learning made possible by:
- Greater computational power (GPUs, TPUs)
- More available data (mobile devices, online services, distributed sensors, crowdsourcing)
- Better understanding of algorithms and architectures
2

AlphaGo and AlphaZero
CASE STUDY

A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play learning

(Science, 2018)

David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dharshan Kumaran, Thore Graepel, Timothy Lillicrap, Karen Simonyan, Demis Hassabis
Deep Learning in AlphaZero: Policy Network
Deep Learning in AlphaZero: Value Network
Training AlphaGo

Human expert positions → Supervised Learning → Policy network → Reinforcement Learning → Value network
Exhaustive Search
Reducing breadth with the policy network
Reducing depth with the value network
AlphaGo vs Lee Sedol

Lee Sedol (9p): winner of 18 world titles

Match was played in Seoul, March 2016

No previous program had ever defeated a human professional player

AlphaGo won the match 4–1
A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play, Science 2018, Joint work with: David Silver, Thomas Hubert, Julian Schrittwieser, Arthur Guez, Ioannis Antonoglou, Matthew Lai, Karen Simonyan, Marc Lanctot, Timothy Lillicrap, Laurent Sifre, Dharshan Kumaran, and Demis Hassabis
Wins of AlphaZero against Stockfish

- 0.4% Losses
- 29% Wins
- 70.6% Draws
- 0.8% Losses
- 2.0% Wins
- 97.2% Draws

4 hours
AlphaZero surpasses Stockfish

Graph showing trends over thousands of steps.
Reinforcement Learning in AlphaZero

AlphaZero plays games against itself
Reinforcement Learning in AlphaZero

Move

Policy

Position

New policy network $P'$ is trained to predict AlphaGo's moves
Reinforcement Learning in AlphaZero

New value network $V'$ is trained to predict winner
Reinforcement Learning in AlphaZero

New policy/value network is used in next iteration of AlphaGo Zero
AlphaZero surpasses StockFish in 4 hours.

AlphaZero surpasses Elmo in 2 hours.

AlphaZero surpasses AlphaGo in 8 hours.
Amount of search per decision

State-of-the-Art Chess Engine: 10,000,000’s of positions

AlphaZero: 10,000’s of positions

Human Grandmaster: 100’s of positions

1000 x more than

100 x more than
Discovering Chess Opening Theory

A10: English Opening

1.e5 g3 d5 cxd5 ♘f6 ♘g2 ♘xd5 ♘f3

C00: French Defence

w 39/11/0, b 4/46/0

3.♘c3 ♘f6 e5 ♘d7 f4 c5 ♘f3 ♘e7
The immortal Zugzwang game

- **Stockfish**
- **Alphazero**
AlphaZero Conclusions

→ Deep Learning enables us to search the huge search space of complex board games

→ Self-play produces large amounts of data necessary for training the deep neural networks

→ Self-play provides an automatic curriculum, starting from simple opponents to stronger and stronger opponents.

→ System discovers new knowledge

→ New directions: Learn rules of the game, more than two players, imperfect information, larger action spaces etc.
# Additional Resources about AlphaGo and AlphaZero

<table>
<thead>
<tr>
<th><strong>AlphaGo</strong></th>
<th><strong>AlphaZero</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>DeepMind AlphaGo website</td>
<td>DeepMind AlphaZero website</td>
</tr>
<tr>
<td>Nature paper: “Mastering the Game of Go with Deep Neural Networks and Tree Search”</td>
<td>Nature paper: “Mastering the Game of Go without human knowledge”</td>
</tr>
<tr>
<td></td>
<td>Science paper: “Mastering Chess and Shogi with Self-Play Reinforcement Learning”</td>
</tr>
</tbody>
</table>
3 Learning to Play Capture the Flag
CASE STUDY

Human-level performance in 3D multiplayer games with population-based reinforcement learning

(Science, 2018)

Capture the Flag

Large-scale decentralised multi-agent learning, scalable computational architectures.

Population based training
Internal reward evolution,
Hierarchical temporal policies.

Agents exceed human-level,
as both teammates and opponents.

Rich emergent representation
and behaviour.
Rules of Capture the Flag

Multiplayer team game e.g. 2 vs 2.

Run to opponent base, and pick up flag.

Bring opponent flag back to your base.

Can only score if own flag is at base. Need to tag opponent flag carrier to return your flag to base.

Winner is team which scores most flag captures after five minutes.
Capture the Flag in Action
Capture the Flag environments

Based on DMLab (Quake III Arena).

Train agents on two style of maps, outdoor and indoor.

Maps are procedurally generated every game.
Procedural Generation: Every game a new map
Training Procedure

- Train a population of agents.
- CTF games played by bringing together agents from population for an episode.
- Individual streams of experience sent back to participating agents.
- Each agent trains with independent RL, independent actions, no global information.
Neural Network Architecture of FTW

Hierarchy of recurrent neural networks at two time scales: Slow RNN and Fast RNN

Internal rewards based on game events learned at even slower time scale
Population based Training

Population of agents serves two purposes.

Provides diverse teammates and opponents: robust multi-agent training without collapses found with naive self-play.

Provides meta-optimisation of agents: using population based training [Jaderberg '17]. Used for model selection, hyperparameter adaptation, internal reward evolution.
Population based Training

FTW agent is far stronger than baseline agents (UNREAL [Jaderberg ’16]).

Benchmarked by playing tournaments against humans.

Humans only win against agents when playing with an agent teammate.

Humans rated FTW agent as most collaborative!
Internal Representation of Neural Network Activity
Agent Behaviour: Playing similar to Humans

Home Base Defence
Opponent Base Camping
Teammate Following
Capture The Flag Conclusions

- Deep Reinforcement Learning can now learn to play complex multi-player video games at human level
- Train populations of agents to enable optimisation and generalisation.
- Use procedurally generated environments to produce robust, generalisable behaviours.
- We can now begin to understand how agents behave and why.
- Extra resources: blog post, arXiv paper
Beyond Games: AlphaFold
CASE STUDY

AlphaFold: Improved proteins structure prediction using potentials from deep learning

(Nature, 2020)

Proteins - Fundamental Building Blocks of Life

Proteins carry out all kinds of functions in living organisms:

- Catalysing reactions
- Transducing signals across the cell membrane
- Gene regulation
- Cellular transport
- Antibodies

Proteins are target of many drugs.

Proteins are a class of drugs

The shape of proteins tells us about their function

Mitochondrial ATP Synthase © MRC
I think that we shall be able to get a more thorough understanding of the nature of disease in general by investigating the molecules that make up the human body, including the abnormal molecules, and that this understanding will permit...the problem of disease to be attacked in a more straightforward manner such that new methods of therapy will be developed.

Linus Pauling, 1960
Nobel Prize Chemistry 1954
Complex 3D shapes emerge from a string of amino acids

Every protein is made up of a sequence of amino acids bonded together.

These amino acids interact locally to form shapes like helices and sheets.

These shapes fold up on larger scales to form the full three-dimensional protein structure.

Proteins can interact with other proteins, performing functions such as signalling and transcribing DNA.
Protein Structure Prediction

- Amino acid residues connected in a chain with a repeating \(-N-C-C-\) backbone
- Side chains connected to the C-alpha determine structure
- Structure can tell us about the function of a protein

Target amino acid sequence

```
MSEIITFPQQTVYPEINVKTLSQAVKNIWRLSHQQKSGIEIQEKTLRISLYSRDLDEARASVPPQLQTVLRQLPPQDYFLTLEIDTELEDPELD
DETRNTLLERSEHRLKKDVKGVRSLRKEANLMAISRIADVSNVVI
LERLESSLKEEQERKAIEIQADIAQKEKNKAKLVDRNKIIESQDVIRQ
YNLADMFKDYIPNISDLKDLLANPPKLIEIKQAIIKQVenIACKIKLGNIS
KGLKYIELADARALKDERINQINKDCDDLKIQLKGVEQRIAGIEDVHQQI
DKEERTTTLLQAALKLEQAWNFKQLQNTIDGKDQDLTKIHIKQLDF
LDDLALQYHSMLLS
```
Protein Structure Prediction - Parameters

- Goal is to predict the coordinates of every atom, particularly the backbone.
- Torsion angles ($\phi$, $\psi$) for each residue are a complete* parameterisation of backbone geometry.
- $2N$ degrees of freedom for chains of length $N$
Levinthal’s Paradox

“Many naturally-occurring proteins fold reliably and quickly to their native state despite the astronomical number of possible configurations”

Example: protein of 361 amino acids

- $3^{361} = 2 \times 10^{172}$ configurations
- Assume protein can sample $10^{13}$ new configurations per second or $3 \times 10^{20}$ per year
- It would take $10^{152}$ years to sample them all
- Similar number of configurations as there are legal Go positions!

Number of legal Go positions as ternary number on Go board

(John Tromp, 2016)
Why deep learning for protein folding?

Experimental techniques are expensive and time-consuming

- Cryo-electron microscopy,
- Nuclear magnetic resonance
- X-ray crystallography

Hard to model long and short range interactions explicitly

Data available from experiments:

- 150,000 proteins in Protein Data Bank (since 1971)
- But: much less data than for speech or image recognition

CASP assessment provides a benchmark with well-defined goals
What to predict?
Pairwise distances between residues!
AlphaFold System

- Combine sequence with data from protein database with coevolutionary information
- Predict pairwise distances and configuration angles
- Run gradient descent on resulting score function to obtain configuration estimate
- Code is available on github: https://github.com/deepmind/deepmind-research/tree/master/alphafold_casp13
Deep Dilated Convolutional Residual network

One residual block modifies a 64x64x128 representation from the previous block.

- Dilated convolutions
- Efficient long-range interaction

Repeat 220 times, cycling through dilations 1, 2, 4, 8

21 million parameters

N x N Input features

N x N Distance predictions

Residual network blocks
Accuracy of AlphaFold’s predictions

Top: Distance matrices for three proteins – the brighter the pixel the closer the pair.
- Top row: real experimentally determined distances.
- Bottom row: average of AlphaFold’s predicted distances
- Good match on both local and global scales

Bottom: Same comparison using 3D models.
- Blue: AlphaFold’s predictions
- Green: ground truth data
Optimise potential with gradient descent

- Repeated small steps always decreasing the potential
- Repeat the optimization to find multiple local minima
- Initialize from torsion predictions, later from corruptions of best results
CASPi3: Critical Assessment of Techniques for Protein Structure Prediction (est. 1994)

Biannual global competition for protein structure prediction

- Blind structure prediction of 82 chains – sequestered newly-solved structures
- May–August 2018 ~1 chain per day
- For each chain, 3 weeks to return up to 5 structure predictions
- 90+ groups from labs around the world
- Post-hoc scoring relative to ground truth with metrics chosen post-hoc based on backbone alignment metric GDT_TS

We are indebted to decades of prior work by the CASP organisers, as well as to the thousands of experimentalists whose structures enable this kind of assessment.
CASP13 Results

O43 AlphaFold

FM & TBM/FM domains (best-of-5)
Conclusions AlphaFold

Deep learning-based distance prediction...
- Gives more accurate predictions of contact between residues
- A richer source of information than contact prediction
- Constructs a smooth potential that is easy to optimise

Limitations:
- Accuracy still limited
- Method depends on finding homologous sequences
- Only predicts backbone, side chains filled by external tool (Rosetta)

Work builds on decades of experimental and computational work of other researchers

Deep learning can deliver solutions to science & biology problems
The future lectures
The lectures

<table>
<thead>
<tr>
<th>No.</th>
<th>Lecture Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>Introduction to Machine Learning and AI</td>
</tr>
<tr>
<td>02</td>
<td>Neural Networks Foundation</td>
</tr>
<tr>
<td>03</td>
<td>Convolutional Neural Networks for Image Recognition</td>
</tr>
<tr>
<td>04</td>
<td>Vision beyond Imagenet - Advanced models for Computer Vision</td>
</tr>
<tr>
<td>05</td>
<td>Optimization for Machine Learning</td>
</tr>
<tr>
<td>06</td>
<td>Sequences and Recurrent Networks</td>
</tr>
<tr>
<td>07</td>
<td>Deep Learning for Natural Language Processing</td>
</tr>
<tr>
<td>08</td>
<td>Attention and Memory in Deep Learning</td>
</tr>
<tr>
<td>09</td>
<td>Generative Latent Variable Models and Variational Inference</td>
</tr>
<tr>
<td>10</td>
<td>Frontiers in Deep Learning: Unsupervised Representation Learning</td>
</tr>
<tr>
<td>11</td>
<td>Generative Adversarial Networks</td>
</tr>
<tr>
<td>12</td>
<td>Responsible innovation</td>
</tr>
</tbody>
</table>
LECTURE 02
Neural Networks
Foundations
Wojtek Czarnecki
Neural Networks Foundations
(Wojtek Czarnecki)

- What are Neural Networks?
- What kinds of functions can they represent?
- How are they trained?
- What are their limitations?

\[ z = \text{XOR}(x, y) \]
LECTURE 03
Convolutional Neural Networks for Image Recognition

Sander Dieleman
Convolutional Neural Networks for Image Recognition
(Sander Dieleman)

- How can we build prior knowledge into our architectures?
- Convolutional Neural Networks encode spatial priors
- Revolutionised image recognition
- Now part of any vision based machine learning application

Architecture of LeNet 5 Gradient-based learning applied to document recognition (Lecun et al, 1998)

Size normalised examples from the MNIST dataset
Architecture of LeNet 5 Gradient-based learning applied to document recognition (Lecun et al, 1998)
LECTURE 04
Vision beyond Imagenet - Advanced models for Computer Vision

Viorica Patraucean
Vision beyond Imagenet – Advanced models for Computer Vision
(Viorica Patraucean)

- Object detection, semantic segmentation, estimation of optical flow
- Moving images: analysing videos for action recognition and tracking
- Self-supervised learning, also using additional modalities such as audio
- Computer vision for building intelligent agents

Two-Stream Convolutional Networks for Action Recognition in Videos, Simonyan & Zisserman, 2014
Optimization methods are the engines underlying neural networks that enable them to learn from data.


Gradient descent, momentum methods, 2nd-order methods, and stochastic methods.

Mode connectivity in loss landscape, Garipov et al, NeurIPS, 2018. Also see https://losslandscape.com/
LECTURE 06
Sequences and Recurrent Networks
Marta Garnelo
Sequences and Recurrent Networks
(Marta Garnelo)

- Almost all data is sequential: text, DNA, video, audio
- How can we process such data using machine learning
- Fundamentals of sequence modeling including Recurrent Neural Networks and Long-Short Term Memory (LSTMs)
- Mapping sequences to sequences as in machine translation

LSTM cell (Long–short term memory, Hochreiter & Schmidhuber, Neural Computation 1997)

Sequence to sequence learning with neural networks (Sutskever et. al, NIPS 2014)
LECTURE 07

Deep Learning for Natural Language Processing

Felix Hill
Deep Learning for Natural Language Processing

(Felix Hill)

- Why Deep Learning for language?
- Simple recurrent networks to Transformers.
- Unsupervised / representation learning for language. From Word2Vec to BERT.

SYSTEM PROMPT (HUMAN-WRITTEN)
In a shocking finding, scientist discovered a herd of unicorns living in a remote, previously unexplored valley, in the Andes Mountains. Even more surprising to the researchers was the fact that the unicorns spoke perfect English.

MODEL COMPLETION (MACHINE-WRITTEN, 10 TRIES)
The scientist named the population, after their distinctive horn, Ovid’s Unicorn. These four-horned, silver-white unicorns were previously unknown to science.

GPT-2 Language Model, OpenAI Blog post, Better Language Models and Their Implications, 2019
LECTURE 08
Attention and Memory in Deep Learning
Alex Graves
Attention and Memory in Deep Learning

(Alex Graves)

- Attention and memory: two vital new components of deep learning.
- Implicit attention, discrete and differentiable variants of explicit attention.
- Networks with external memory, attention for selective recall.
- Variable computation time, which can be seen as a form of 'attention by concentration'.

Illustration of the Differentiable Neural Computer

Hybrid computing using a neural network with dynamic external memory (Graves et al, Nature, 2016)
LECTURE 09

Generative Latent Variable Models and Variational Inference

Andriy Mnih
Generative Latent Variable Models and Variational Inference

(Andriy Mnih)

- Unsupervised Learning
- Latent variable modelling and the central role of inference
- Flow-based models which combine high expressive power with tractable inference
- Variational inference for efficient training of intractable models
- VAE modelling framework

LECTURE 10

Frontiers in Deep Learning: Unsupervised Representation Learning

Mihaela Rosca & Irina Higgins
Frontiers in Deep Learning: Unsupervised Representation Learning
(Mihaela Rosca & Irina Higgins)

- What is a good representation?
- Unsupervised learning has potential to address open problems like data efficiency, generalisation, robustness, fairness etc
- Different approaches such as disentangling, CPC, VQ–VAE, Bert, auxiliary losses for RL

β–VAE: Learning basic visual concepts with a constrained variational framework, (Higgins et al, ICLR 2017)
LECTURE 11
Generative Adversarial Networks
Mihaela Rosca & Jeff Donahue
Generative Adversarial Networks

(Mihaela Rosca & Jeff Donahue)

- Generative adversarial networks (GANs, Goodfellow et al. 2014) promising approach to generative modeling
- Two "competing" networks: a generator tries to fool a discriminator with synthesised data
- Variations, e.g., CycleGAN, VAE–GAN hybrids, bidirectional GAN
- Domains, such as video and speech synthesis

Variational Approaches for Auto-Encoding Generative Adversarial Networks

(Mihaela Rosca et al 2017)
LECTURE 12

Responsible Innovation

Iason Gabriel & Chongli Qin
Responsible Innovation
(Iason Gabriel & Chongli Qin)

- AI provides powerful tools that are shaping our lives and society
- With great power comes great responsibility
- How to build safe, robust, and verified AI systems that do exactly what we expect of them
- How to think about the ethical consequences of building and deploying AI systems

The Three Laws of Robotics (Asimov, 1942):

1. A robot may not injure a human being or, through inaction, allow a human being to come to harm.
2. A robot must obey the orders given it by human beings except where such orders would conflict with the First Law.
3. A robot must protect its own existence as long as such protection does not conflict with the First or Second Laws.
Thank you
Questions