Is Deep Reinforcement Learning ready for applications?

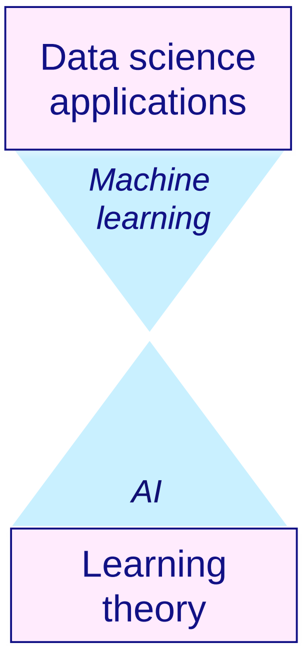
Elena Mocanu

Mechanical Enginering,
Control System Technology group
Dynamics and Control group

Technische Universiteit **Eindhoven** University of Technology

Where innovation starts

Research background



Classification	Event detection, pattern recognition and dimensionality reduction
Prediction	Supervised, Unsupervised (Reinforcement learning, Transfer learning)
Optimization and control	Deep Reinforcement Learning High-order restricted Boltzmann machine
Scalability	Restricted Boltzmann Machine Artificial Neural Networks
Evolutive networks	Restricted Boltzmann Machine Deep learning

Research background

Data science applications

Machine learning

Machine learning control

Autonomous agents

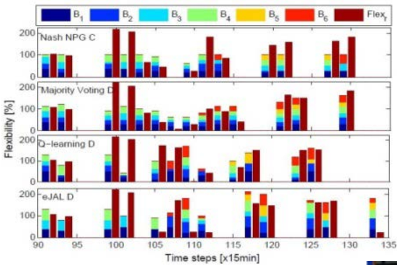
AI

Learning theory

Classification	Event detection, pattern recognition and dimensionality reduction
Prediction	Supervised, Unsupervised (Reinforcement learning, Transfer learning)
Optimization and control	Deep Reinforcement Learning High-order restricted Boltzmann machine
Scalability	Restricted Boltzmann Machine ANN
Evolutive networks	Restricted Boltzmann Machine Deep learning
New learning theories	e.g. Quantum machine learning

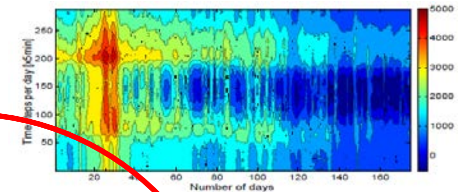
Data science applications

Resource allocation (MAS strategies)



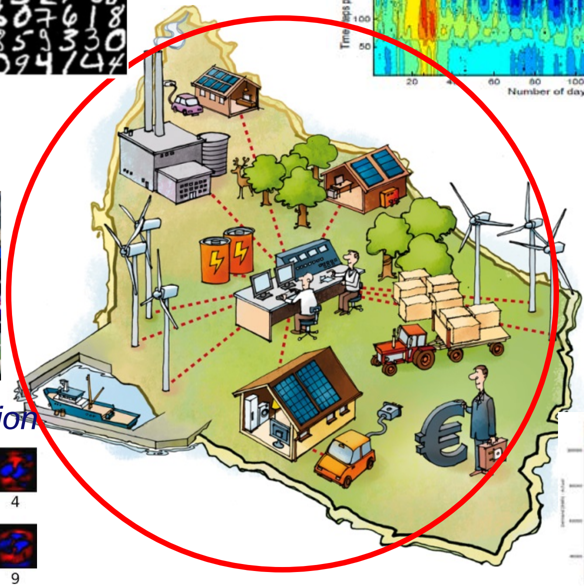
Handwritten Digit Recognition

Energy prediction



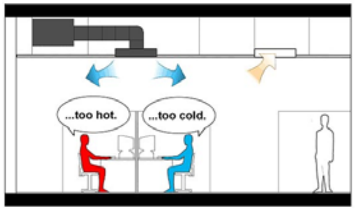
Machine learning control of respiratory system

Robotics



Learning theory

Thermal comfort



People detection and tracking

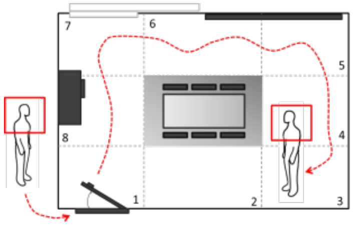
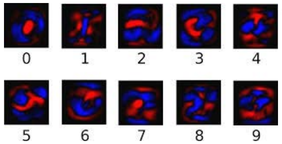
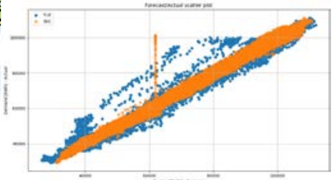


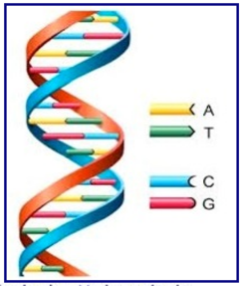
Image classification



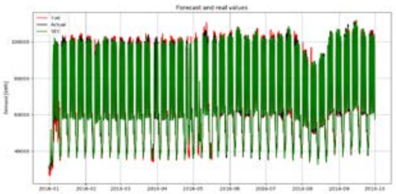
Pattern recognition



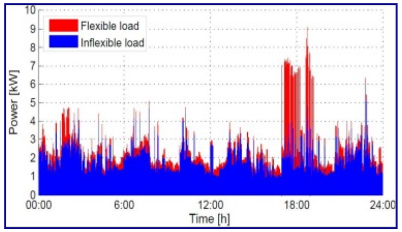
MicroRNA identification



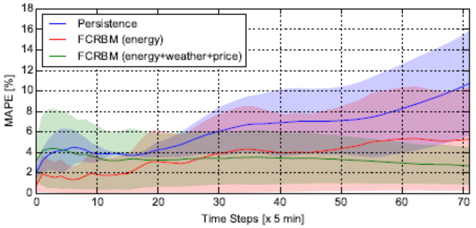
Price forecasting



Energy disaggregation



Demand forecast in a price-responsive context

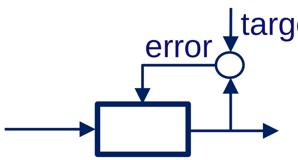


Technische Universiteit **Eindhoven** University of Technology

Why Deep Reinforcement Learning?

- ML in a Nutshell

Supervised learning



Methods
- Support Vector Machine
- Artificial Neural Networks
- Decision Trees,...

Specific tasks:
- Classification
- Regression (prediction)

Type of data:
- **Training data includes desired outputs.**

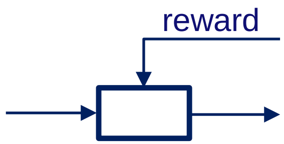
Unsupervised learning

Methods
- k-mean clustering
- Restricted Boltzmann Machine

Specific tasks:
- Discover clusters
- Discover factors/structures

Type of data:
- **Training data does not include desired outputs.**
- 75% data are unlabeled

Reinforcement learning



Methods
- Q-learning
- SARSA

Specific tasks:
- Decision making process
- Control
- Strategic optimization
- On-line learning

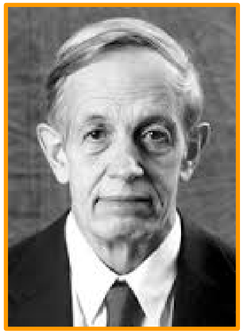
Rewards from sequence of actions.

Reinforcement Learning

- **Fundamental concepts - Dynamic Game Theory**

From 2-player Zero-Sum Games...

(Neumann,1928)



... to a solution for n-player Games.

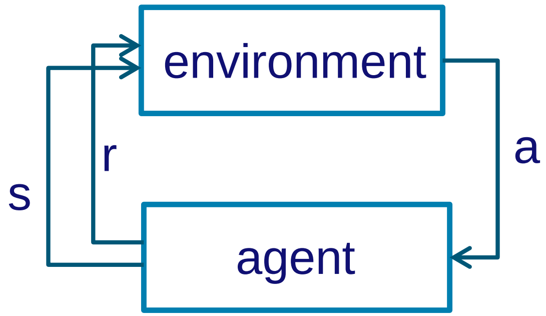
(John Nash,1951)

MIT Technology Review

10 Breakthrough Technologies 2017

1. Reversing Paralysis
2. Self-Driving Trucks
3. Paying with Your Face
4. Practical Quantum Computers
5. The 360-Degree Selfie
6. Hot Solar Cells
7. Gene Therapy 2.0
8. The Cell Atlas
9. Botnets of Things
10. Reinforcement Learning

Reinforcement Learning



An RL agent may include one or more of these components:

- Policy (π): agent's behaviour function (s →a)
- Value function (Q): how good is each state and/or action
- Model (M): agent's representation of the environment

- Q-value function gives expected total reward:

$$Q^{\pi}(s,a) = \mathbb{E}\left[r_{t+1} + \gamma r_{t+2} + \gamma^2 r_{t+3} + \ldots \mid s,a\right]$$

- Value functions decompose into a Bellman equation

$$Q^{\pi}(s,a) = \mathbb{E}_{s',a'}\left[r + \gamma Q^{\pi}(s',a') \mid s,a\right]$$

Optimal Value Functions

$$Q^{*}(s,a) = \max_{\pi} Q^{\pi}(s,a)$$

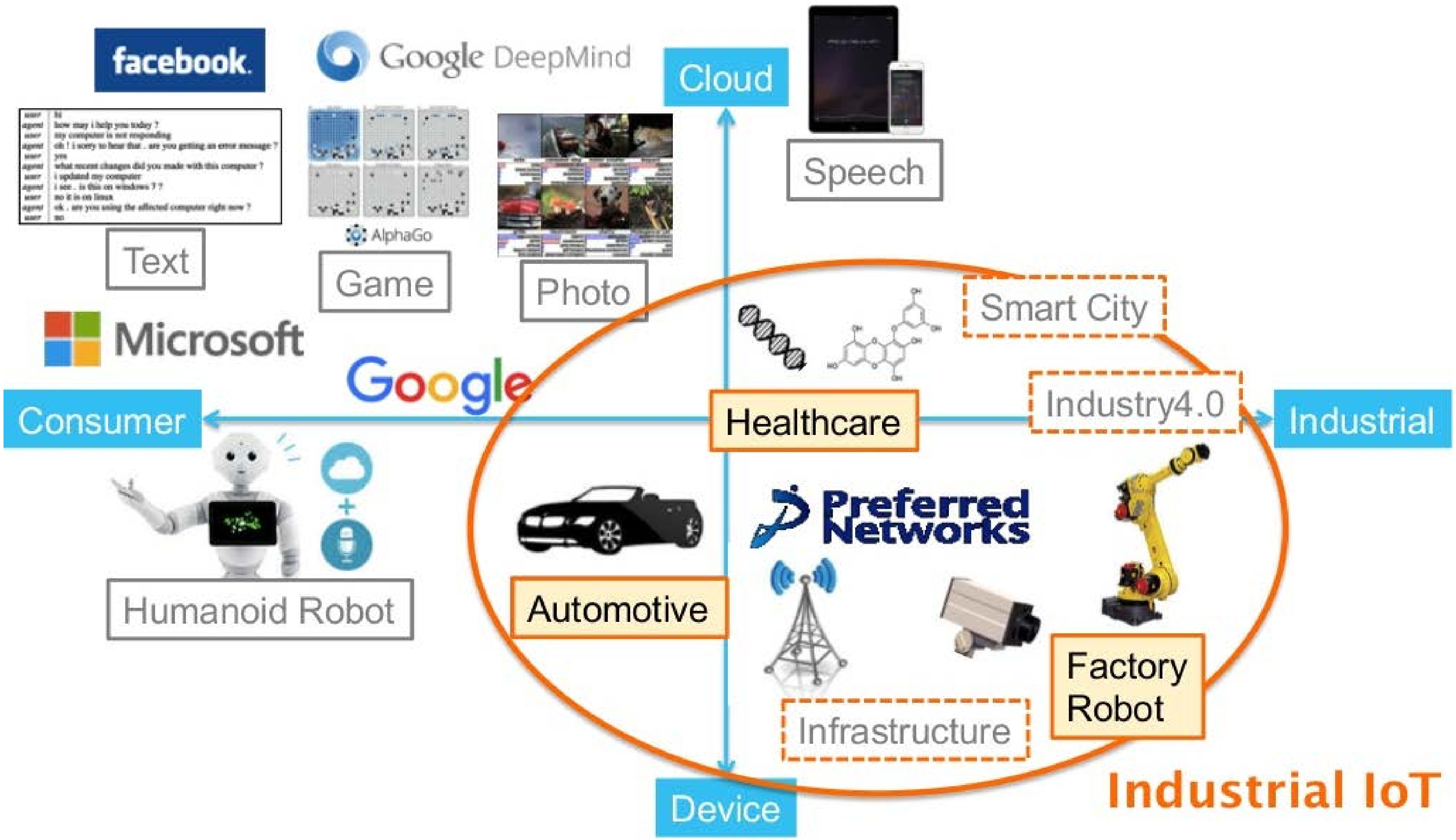
[2015] Human-level control through deep reinforcement learning, Minh et al., Nature

"This work bridges the divide between high-dimensional sensory inputs and actions, resulting in the first artificial agent that is capable of learning to excel at a diverse array of challenging tasks."

[2016] AlphaGO versus Lee Sedol - professional Go player

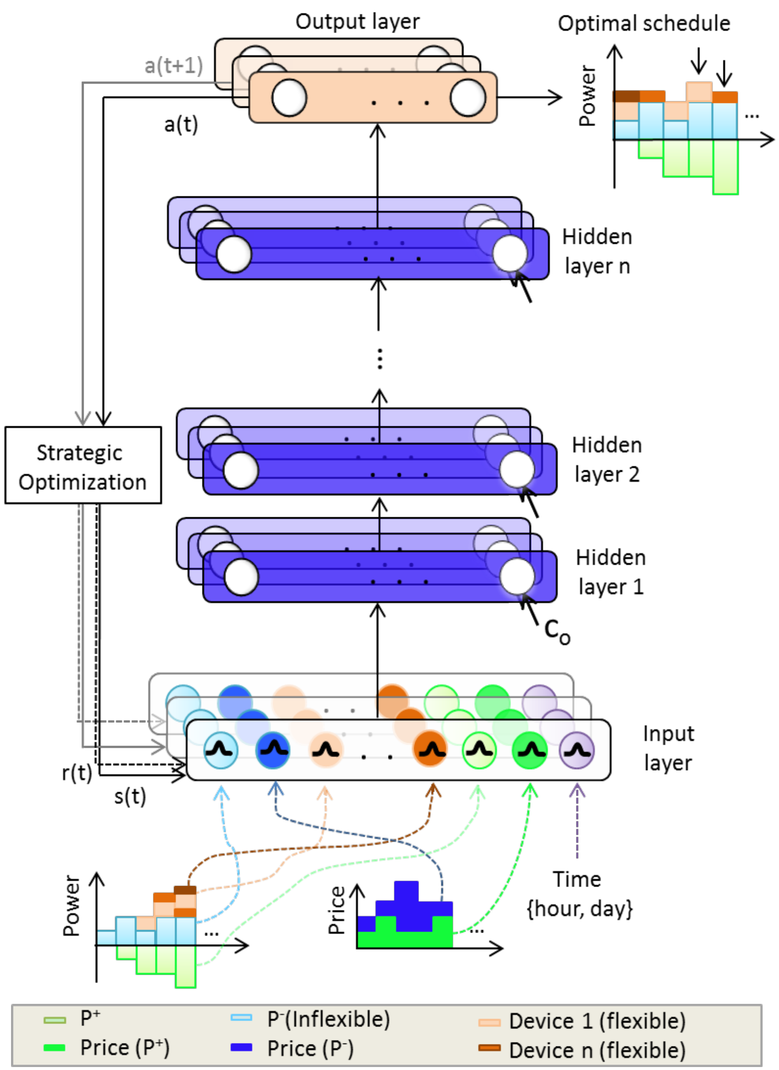
~10^{170} compared to ~10^{50} for chess (Kasparov, 1997)

Deep Learning



Credit: Yahoo Japan, 20 March 2016

On-line building energy optimization



Reinforcement learning

$$\{s,a\} \rightarrow Q(s,a)$$

Deep learning

$$\xrightarrow[Data]{\text{In}} DNN_{(k)} \xrightarrow[Data\ estimation]{\text{Out}}$$

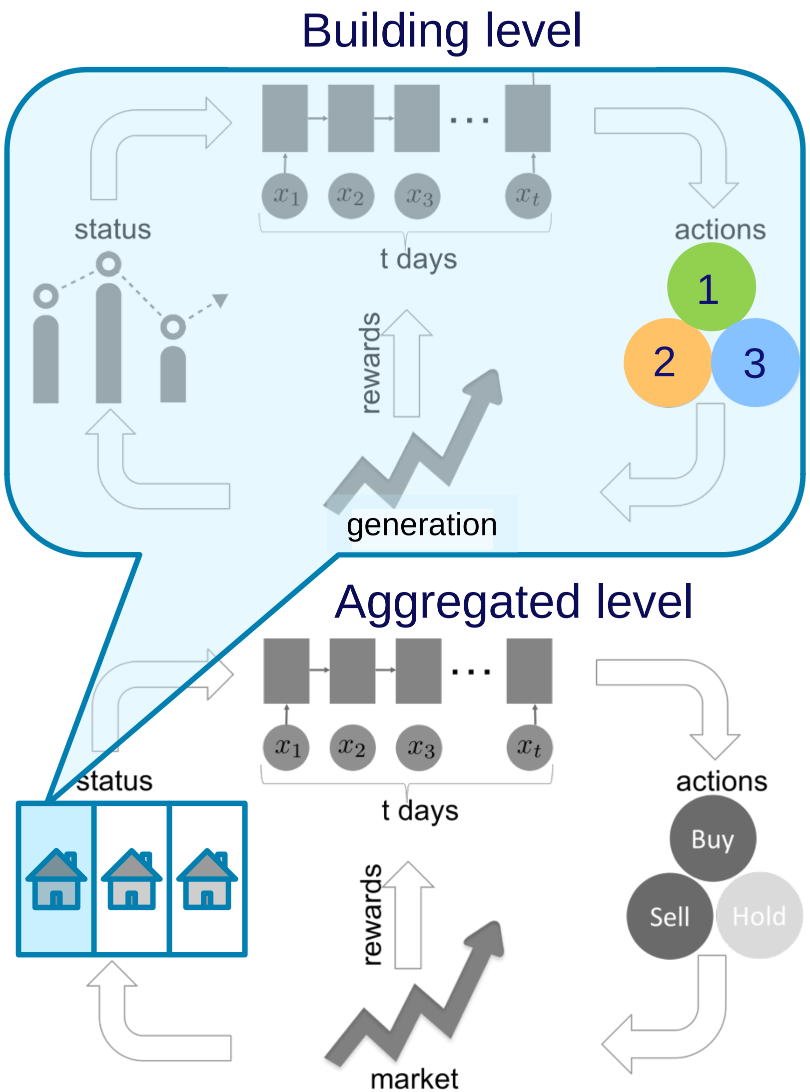
Deep reinforcement learning

$$\xrightarrow[States]{\text{In}} DNN_{(k)} \xrightarrow[Q(s,a)]{\text{Out}}$$ Deep Q-learning

$$\xrightarrow[States]{\text{In}} DNN_{(k)} \xrightarrow[p(a|s)]{\text{Out}}$$ Deep Policy Gradient

E. Mocanu, D.C. Mocanu, P.H. Nguyen, A. Liotta, M.E. Webber, M. Gibescu, and J.G. Slootweg, On-line Building Energy Optimization using Deep Reinforcement Learning, IEEE Transactions on Smart Grid, 2018

On-line building energy optimization



Energy minimization → Price reduction

Reinforcement learning
- Q-learning (Watkins, 1989)

  $Q^{\pi}(s,a)$

Deep reinforcement learning
- Deep Q-learning

  $Q(s,a,\theta) \approx Q^{\pi}(s,a)$
- Deep Policy Gradients

  $p(a|s_t,\theta), \forall a \in A$

Our contribution:
- multiple actions simultaneously

Experimental Results

Reinforcement learning formalism (Markov Decision Process)

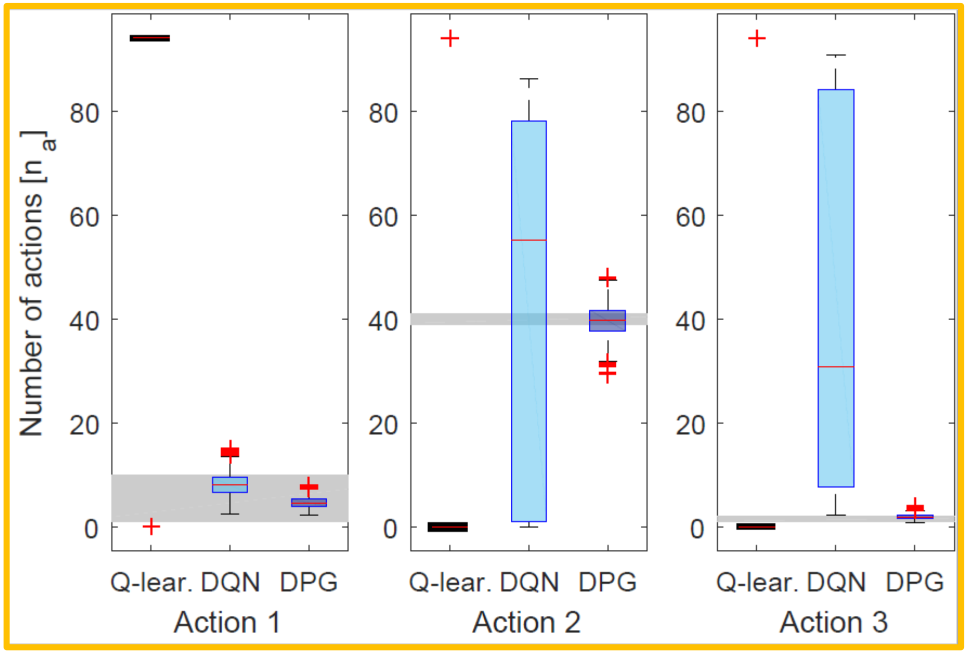
States:
DRL: continuous states
Q-learing: 11 inputs (2048 discrete states)

Actions:
1 : time-scaling electrical device
e.g. air conditioning load
2 : time-shifting electrical device
e.g. dishwasher
3 : time-scaling and time-shifting
e.g. electric vehicle

Reward: joint reward function

Scale: up to 50 buildings
Pecan Street datasets

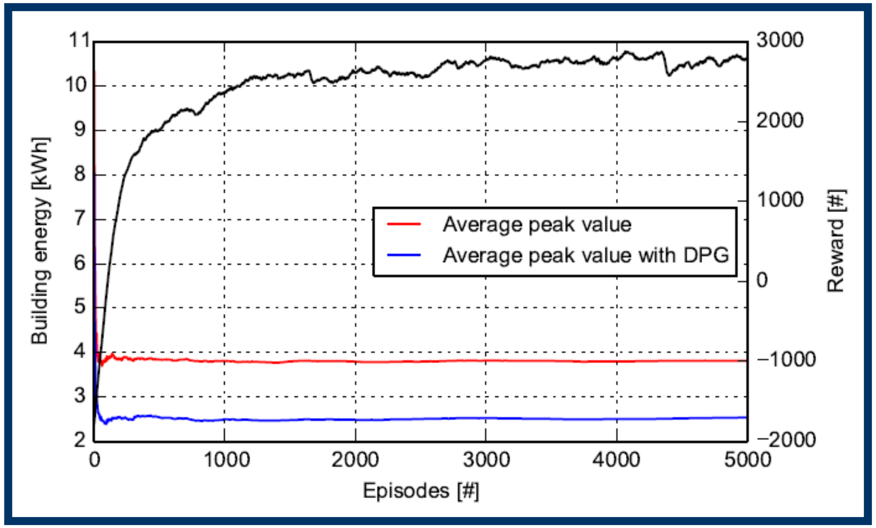
Learning to satisfy the constraints bounds on one building



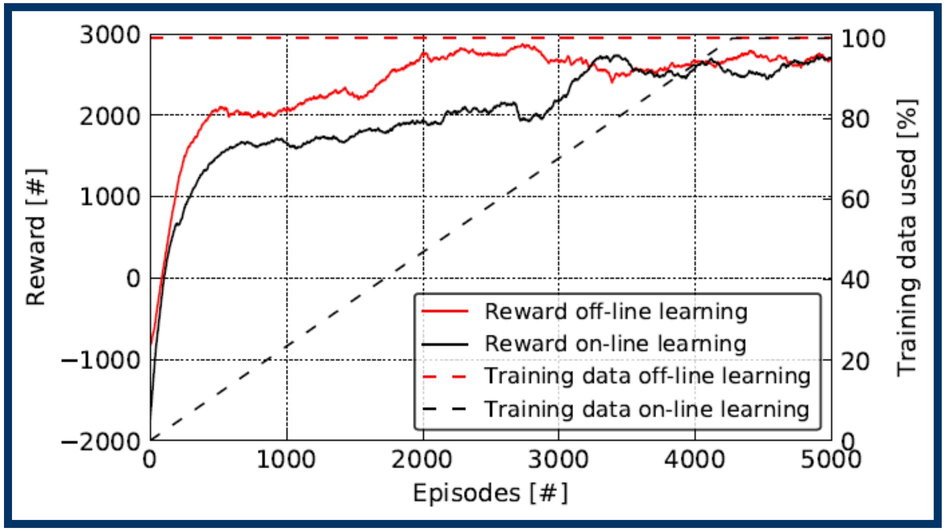
Technische Universiteit Eindhoven University of Technology

Learning capabilities of DRL

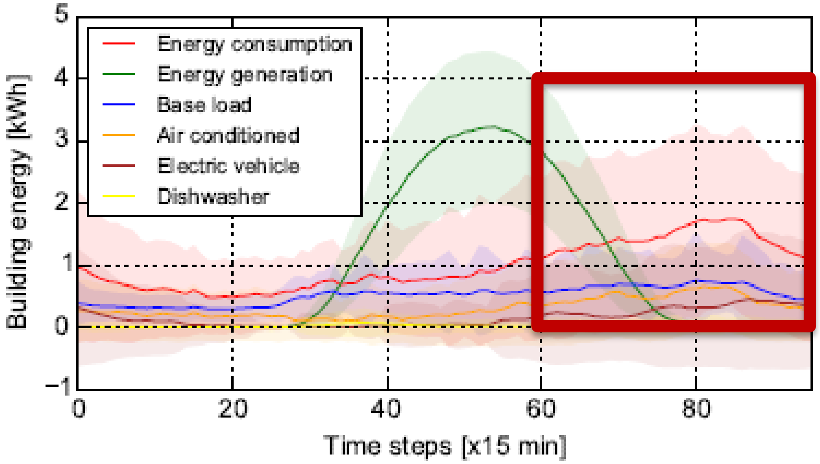
Learning capabilities of DPG method in terms of peak reduction and their corresponding reward function



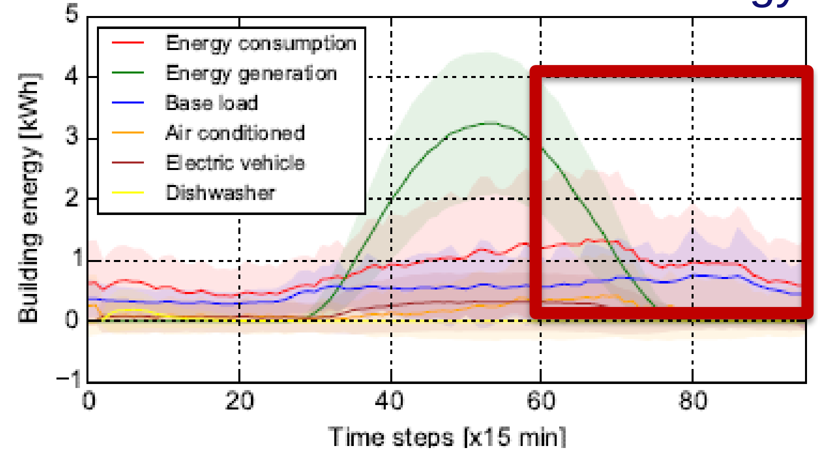
On-line versus off-line learning capabilities of DPG in terms of reward function and their corresponding training data used



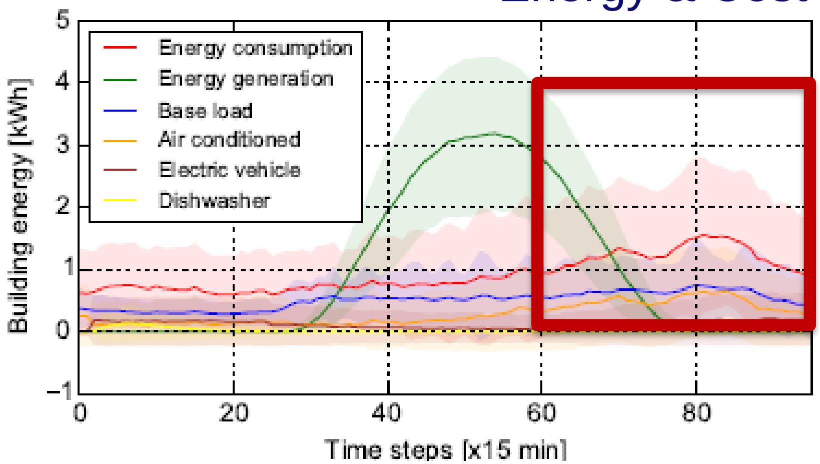
Scalability capabilities of DRL



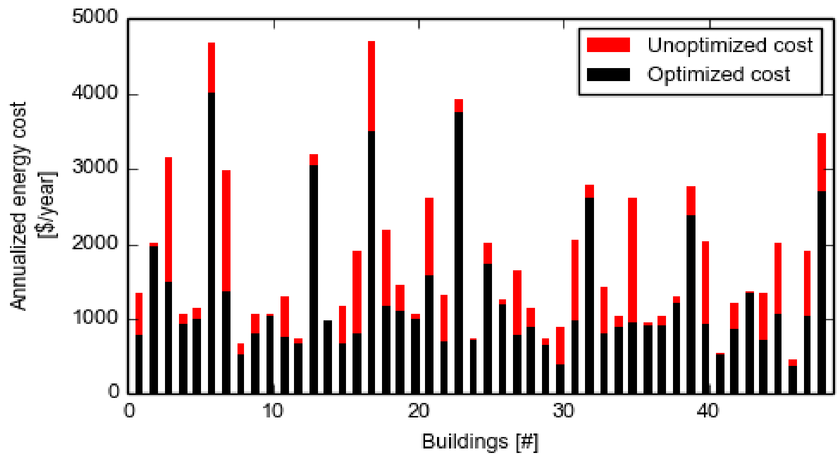
Energy



Energy & Cost



Results: 27% cost reduction



Is Deep Reinforcement Learning ready for applications?

- For this particular case and few others, YES!
- Use further for system automation

identification

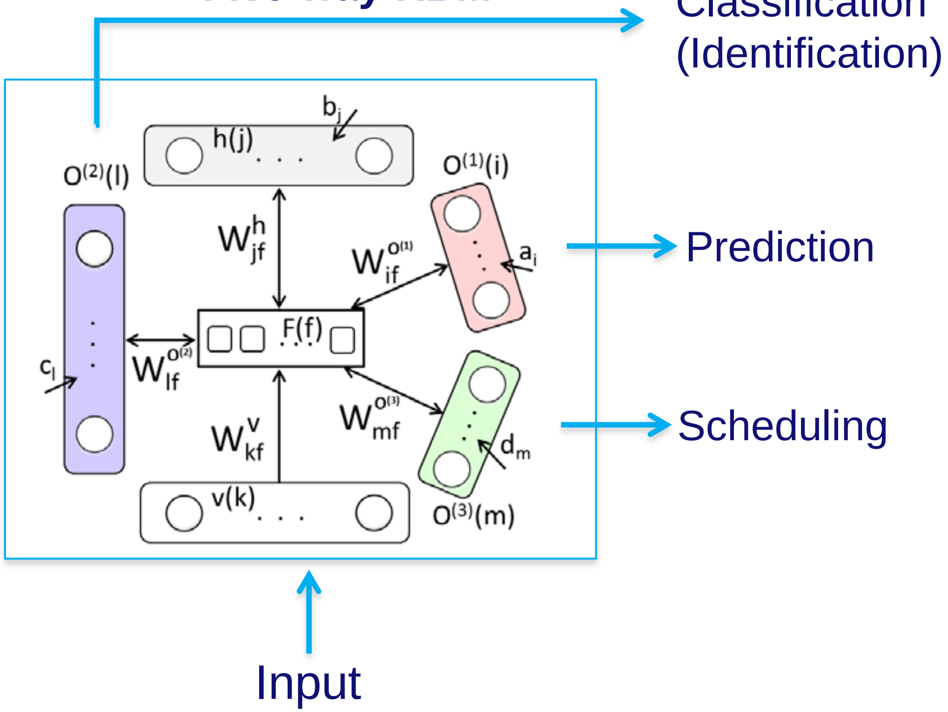
\+

prediction

\+

control

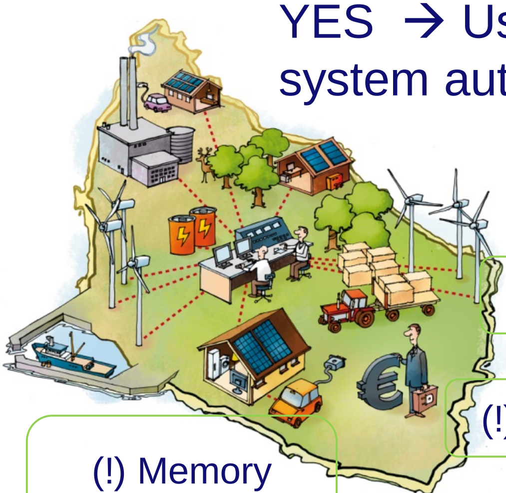
Five-way RBM



References: E. Mocanu, Machine Learning applied to Smart Grids, PhD thesis, 2017 (Chapter 5)

Technische Universiteit Eindhoven University of Technology

Is Deep Reinforcement Learning ready for applications?



YES → Use further for system automation

(!)Reward function

(!)Convergence

(!) Memory requirements

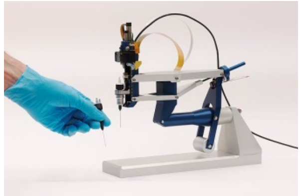
(?) small datasets

DRL and

- one-shot learning
- transfer learning
- sparse ANN
- …

(?) large datasets

- generalization to multi-tasks
- efficient training of deep learning models to apply them at scale across increasingly more complex and diverse tasks.



(?) very large datasets

Computational power:

- scalable ANN
- some AI hardware under development, such as neuromorphic chips or even quantum computing systems, could factor into the new equation for AI innovation.

References [4], [5] , [6] and [7]

Is Deep Reinforcement Learning ready for applications?

References

1. E. Mocanu, D.C. Mocanu, P.H. Nguyen, A. Liotta, M.E. Webber, M. Gibescu, and J.G. Slootweg, On-line Building Energy Optimization using Deep Reinforcement Learning, **IEEE Transactions on Smart Grid**, **2018.**

2 L. A. Hurtado Munoz, E. Mocanu, H.P. Nguyen, M. Gibescu & I.G. Kamphuis, Enabling cooperative behavior for building demand response based on extended joint action learning, **IEEE Transactions on Industrial Informatics**, **2018**.

3. E. Mocanu, Machine Learning applied to Smart Grids, **PhD thesis**, **2017** (Ch 5)

AI for stochastic decision making

4. D.C. Mocanu, E. Mocanu, P.H. Nguyen, M. Gibescu, and A. Liotta, Evolutionary Training of Sparse Artificial Neural Networks: A Network Science Perspective, **Nature Communications, 2018**.

5. D.C. Mocanu, E. Mocanu, P.H. Nguyen, M. Gibescu, and A. Liotta, A topological insight into restricted Boltzmann machines, **Machine Learning**, **2017**.

6. D.C. Mocanu and E. Mocanu. One-shot learning using mixture of variational autoencoders: a generalization learning approach. **AAMAS 2018**, Sweden.

7. E. Mocanu, P.H. Nguyen, M. Gibescu, and W. Kling Unsupervised energy prediction in a smart grid context using reinforcement cross-buildings transfer learning, **Energy and Buildings, 2016**.

Learning Theory

One-shot learning

Transfer learning

Technische Universiteit Eindhoven University of Technology