Merging the Worlds of Simulation and Machine Learning:

What's the Art of the Possible with Deep Reinforcement Learning

anylogic®

dse CONSULTING

AnyLogic Partner for over 10 years

skymind | DEEP LEARNING FOR ENTERPRISE

Bespoke & open training events

Mentoring

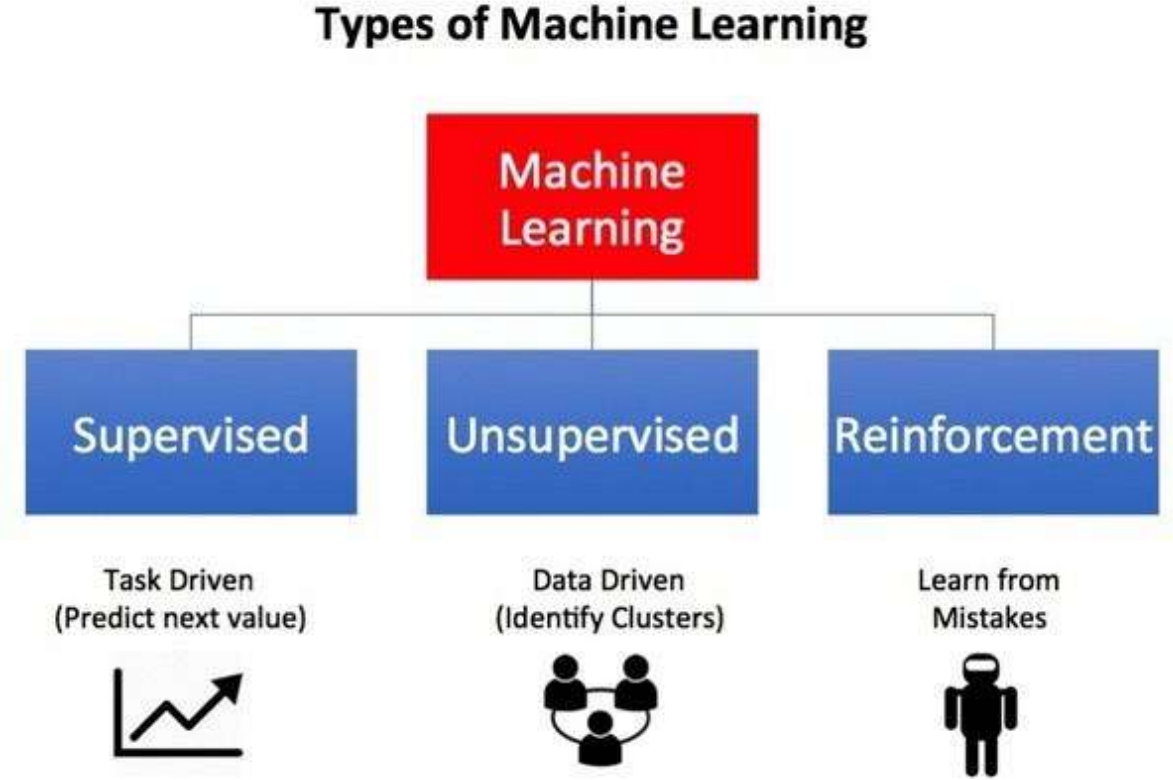
Rapid prototyping

How I came to simulation modelling & AI

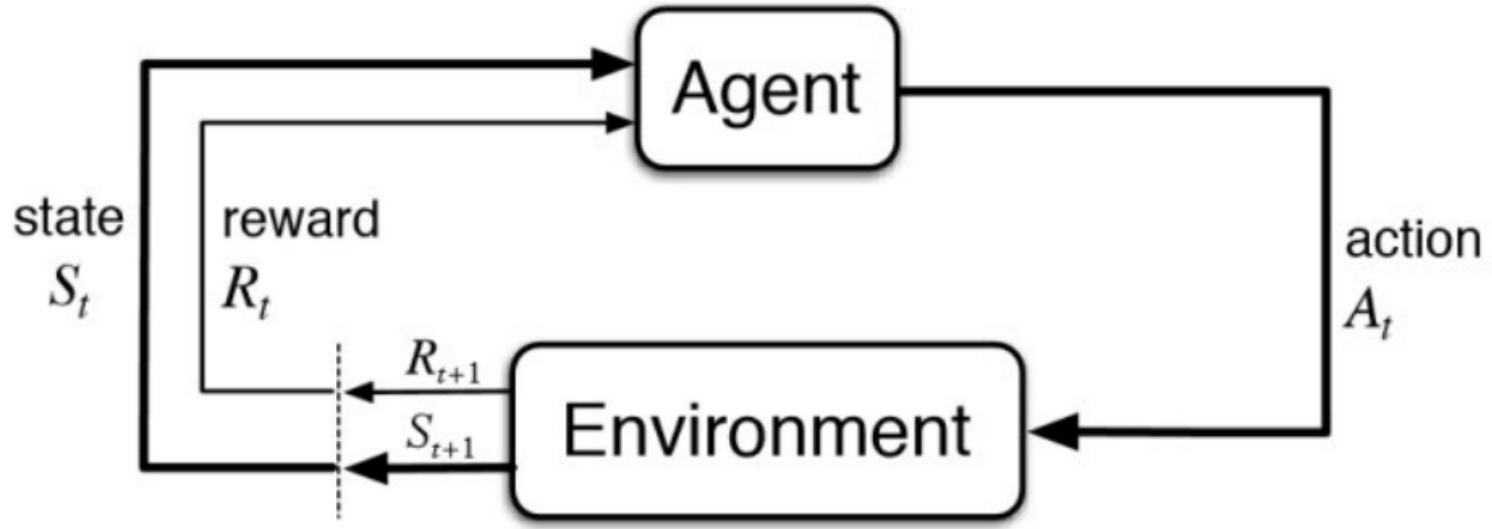
My life in numbers -

21	years since i discovered simulation when writing my a geography dissertation
15	years since i first started using AnyLogic
11	years since i become a freelance AnyLogic consultant
9	years since DSE became the UK distributor of AnyLogic
7	years ago dL launches as a decision support model development company
5	years ago dL became a Data Science technology company
1	year ago dL started doing commercial R&D projects in Reinforcement Learning

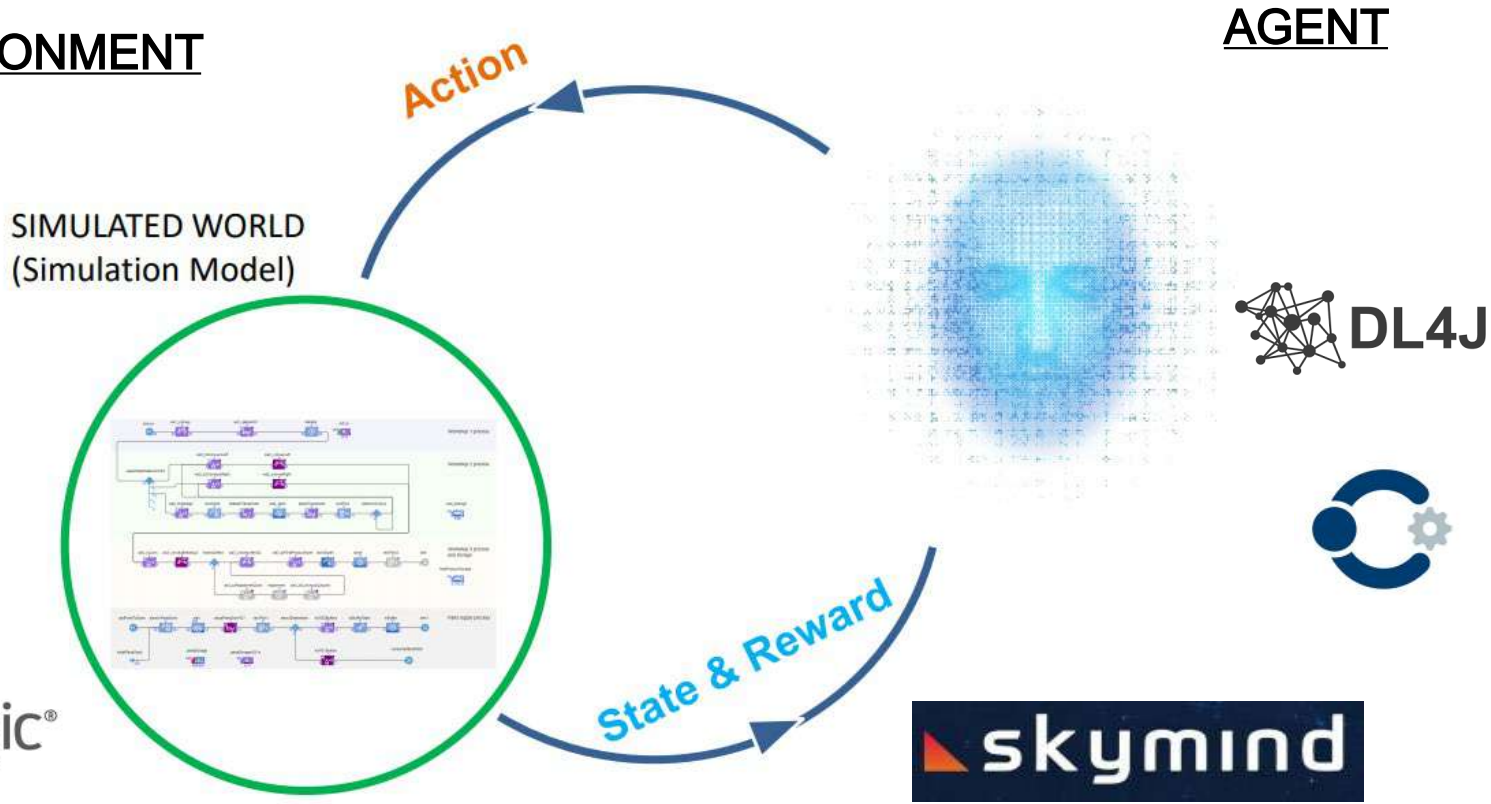
We don’t learn to walk by following rules. We learn by doing and by falling over *(Richard Branson)*..



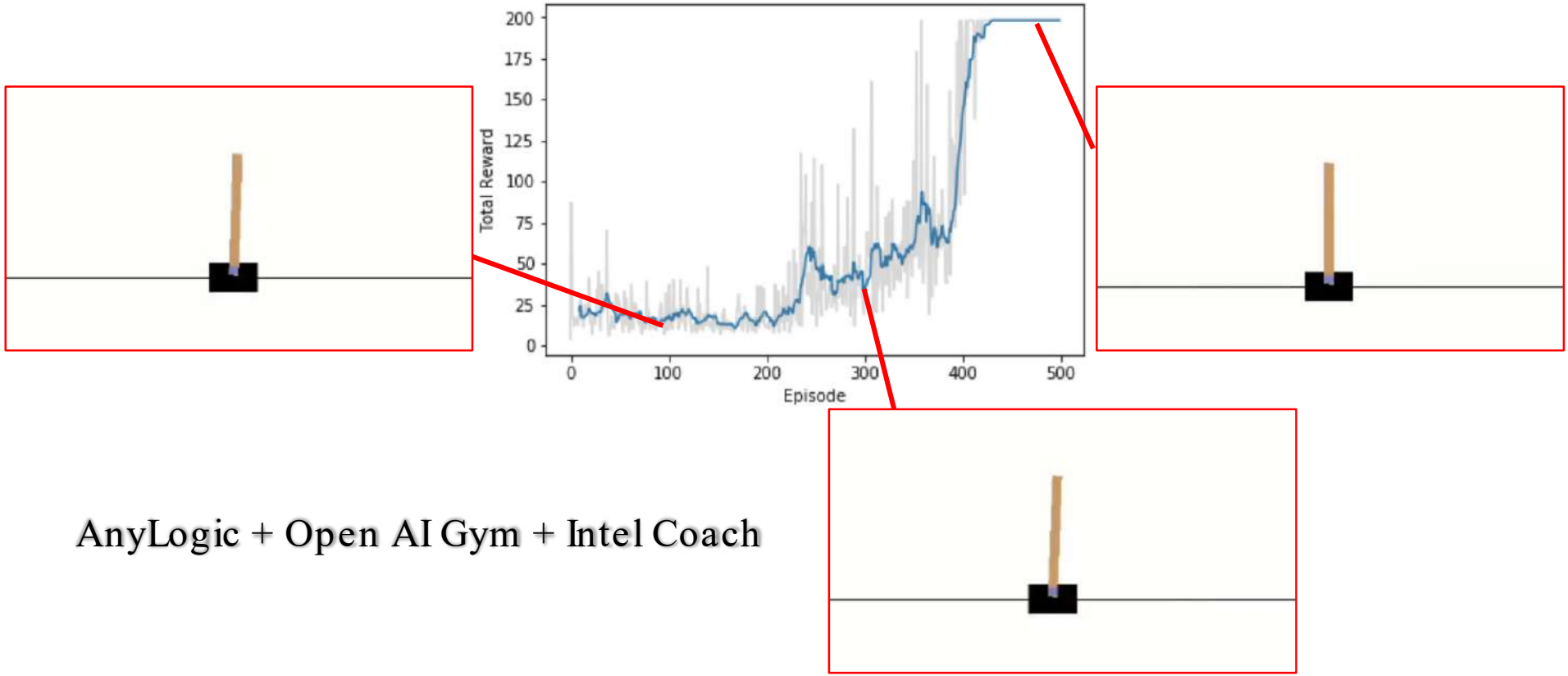
Text book reinforcement learning architecture



And in practice, this looks like...



And then you get to do things like this...



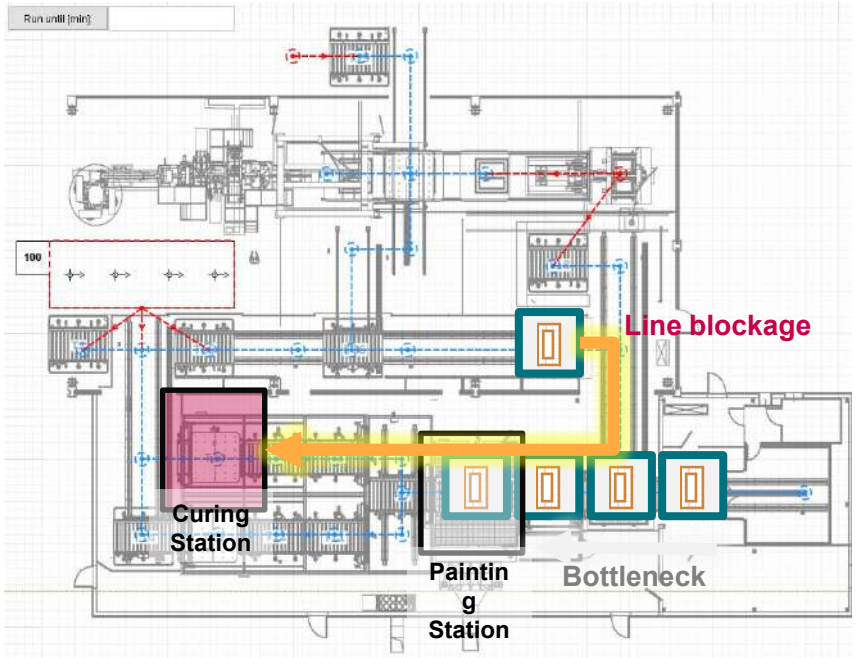
AnyLogic + Open AI Gym + Intel Coach

Cartpole is interesting... but it's only a game

- Very quick look at a use case - Manufacturing Routing Problem
    - https://www.anylogic.com/upload/conference/2019/presentations/fair-dynamics-moving-towards-autonomous-digital-twin.pdf

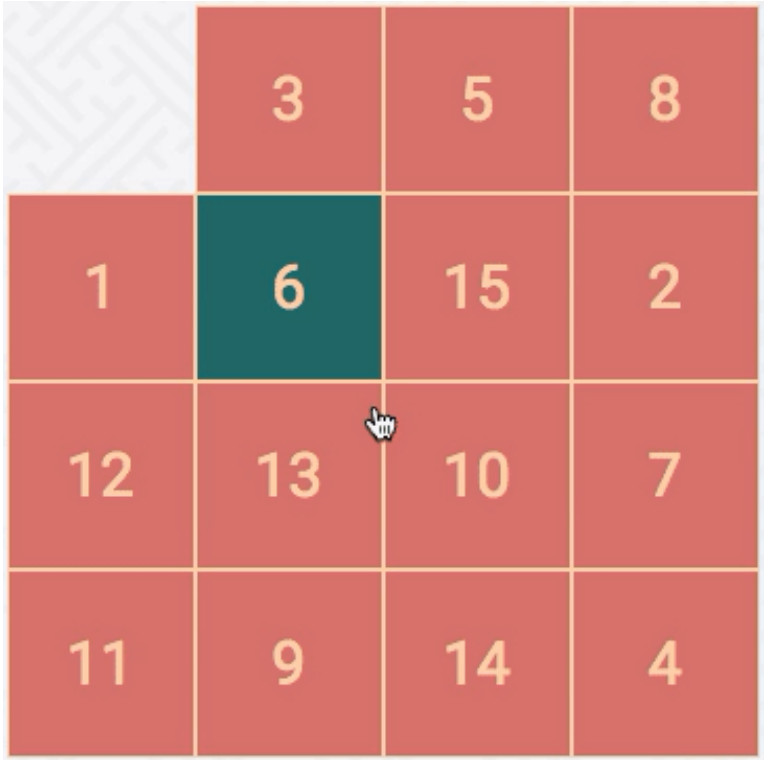
Use case 1 – Lagor Power Transformer

- PT Cores are essentially built from many layers of coils which can weigh up to **8 tons** and require **different production cycles** depending on their size and client-specific customizations.
- Since the very first operation, the cores are loaded on steel pallets which are moved across the different work stations using roller or shuttle conveyors.
- At any given time, **several different cores** can be found on the line.
- Due to the above features, sequencing the operations in the production line can be critical to avoid **bottlenecks** and **line blockage.**

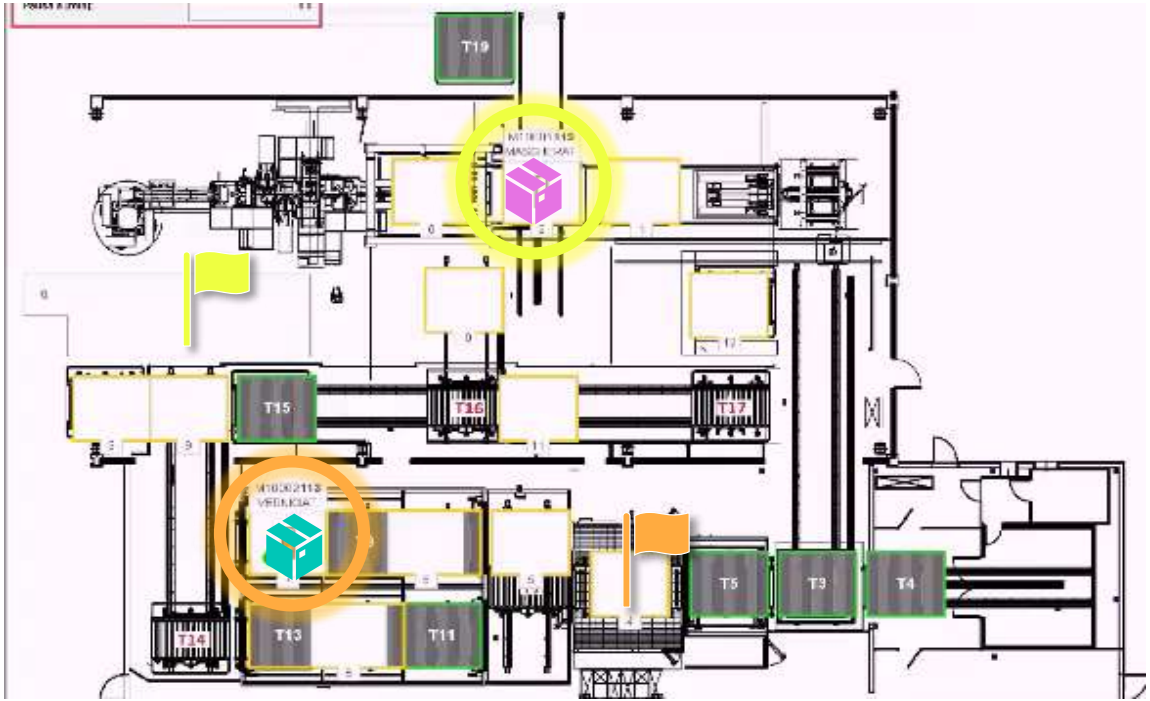


The job of the Line Manager

Evaluate the “best” route for each situation in order to avoid unnecessary movements, anticipate possible criticalities, resolve conflicts (both for Cores and Steel Pallets), and to ultimately respect delivery dates.



The actual problem



Implementation approach

Reinforcement algorithm

DL4J

PyTorch

Apache Spark

Apache kafka
A distributed streaming platform

Apache hadoop

Ramin Mirzazadeh

Cristina Morariu

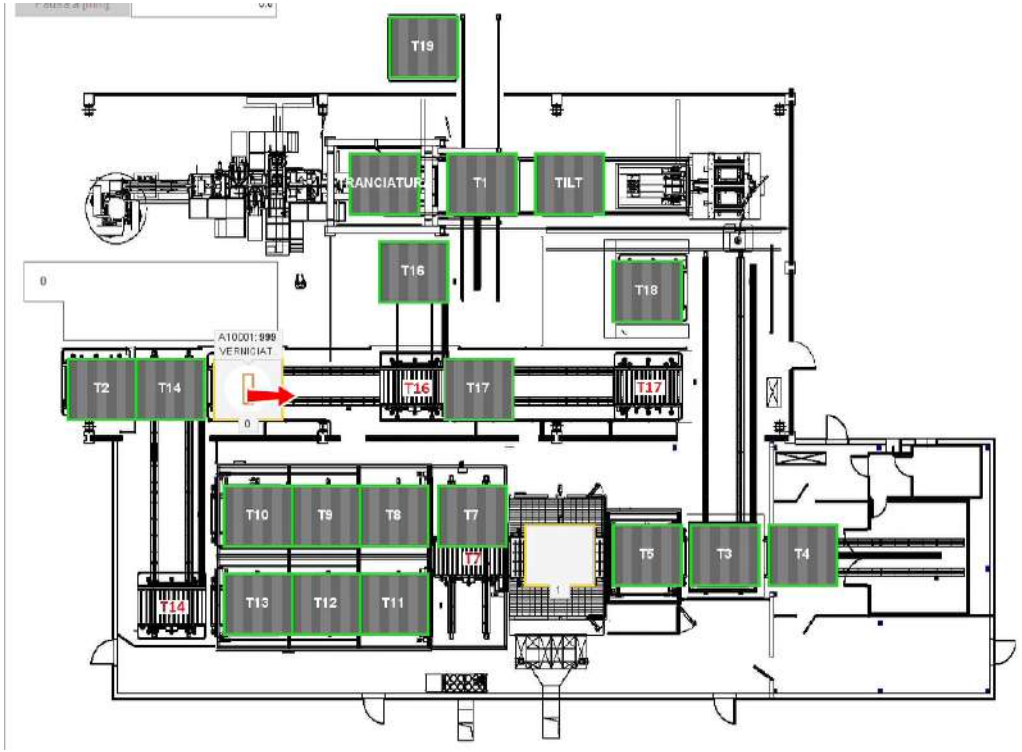
Environment

anylogic®

Leading simulation software for business applications

Roberto Grugni

The learnt policy



The model (Learnt Policy) can take the core to the target position starting from any given initial layout!

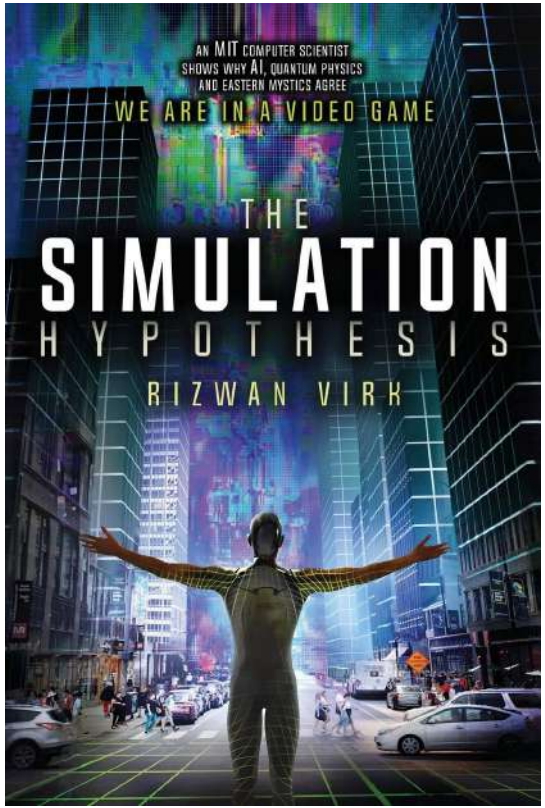
So...current challenges

- Sim + Learning = BIG problems…
  - ML is hard (consistent available of data, quality of data, aggregation of data, hyperparameter tuning, over fitting, ...)
  - RL adds new complexity (Shaping the reward and penalty system, communication between tools that are good for RL)
  - Simulation modelling can be hard (conceptualisation, choosing the appropriate level of detail, verification, validation, accurate representation of the problem)
- The learning process is opaque and the resulting algorithm is black box
  - So the reasons for algorithm performance can be difficult be explain / trust
  - Need to be able to take stakeholders on the journey!
- RL is not an off-the-shelf-product
  - Either technologically or conceptually
  - Most development is in Python, but the IT business system community still prefer Java
  - Which algorithm
    - QLearning Vs PPO Vs *Clipped PPO*

If it's so hard...why then would we bother to do this?

- Enabler of a true Digital Twin!!
- You can do far more than traditional simulation based optimisation
- The **self discovery** of alternative to the heuristics currently used in industrial control theory offer huge opportunity to optimise **sequencing / scheduling** problems
- Simulation becomes a practical **tactical decision support tool**
  - Sim = strategic planning models
  - Sim + optimisation = best possible course of action given a known set of conditions
  - Sim + learning = new better decision logic, outperformance of human decision making, **best course of action regardless of the starting conditions**

Final thoughts...



Is the simulation hypothesis now becoming a reality?

https://www.anylogic.com/blog/machine-learning-and-simulation-example-and-downloads/

Questions & Comments...