Using self-organised behaviours generated by Differential Extrinsic Plasticity (DEP) to guide the exploratory phase of Reinforcement Learning

Simon Birrell
Bio-Inspired Robotics Laboratory (BIRL), University of Cambridge
25th March 2021
Standard RL control models very far from reality

Reality is much more complex:
- Muscles
- Tendons
- Skin
- Reflexes, sensorimotor loops etc.

Can the body’s complexity simplify the problem?
Differential Extrinsic Plasticity (DEP)

“Interesting” natural behaviours generated by DEP plasticity over a wide variety of agents and situations

The Playful Machine – Der, Martius (2011)
Novel plasticity rule can explain the development of sensorimotor intelligence - Der, Martius (2015)
How it works

- Simple plasticity rule spontaneously generates a variety of goalless, environmentally-aware, edge-of-chaotic behaviours

Motor Activation

\[ y_i = g \left( \sum_{j=1}^{n} C_{ij} x_j + h_i \right) \]

\[ g(z) = \tanh(z) \]

Inverse Model

\[ F(x') = M \dot{x}' \]

\[ \ddot{y}_i = \sum_j M_{ij} \dot{x}_j \]

Plasticity

\[ \tau \dot{C}_{ij} = \ddot{y}_i \dot{x}_j - C_{ij} \]

\[ \ddot{y} = \dot{y} + \delta \dot{y} = F(x') \]

Global normalization

the entire weight matrix is normalized

\[ C \leftarrow \kappa C / (\|C\| + \rho) \]

Main Hyperparameters

\( \tau \) – timescale plasticity
\( \kappa \) – feedback strength

Novel plasticity rule can explain the development of sensorimotor intelligence - Der, Martius (2015)
DEP driving OpenAI Gym Reacher-v2
Can we use complex behaviours generated by DEP to enhance learning?

- Enhance
  - The skill level (reward obtained)
  - The sample efficiency (speed of learning)

- Different approaches:
  - Learn to switch between behaviours generated by DEP-enabled sensorimotor loops (hierarchical RL)
  - Use DEP-generated behaviour to improve exploration for off-policy Reinforcement Learning algorithms
Leveraging pre-existing complexity

• *Creatio ex nihlo* in videogames
  • Elite on the BBC Micro (1984): universe created from Fibonacci series
  • Procedural Generation now an established technique
  • Creating data from nothing

• Physical Reservoir Computing
  • Exploit the complex dynamics of physical systems as information processing devices

Nakajima (2020), Physical reservoir computing—an introductory perspective