Autonomous Robots Realising their Abilities

Adaptive & Intelligent Robotics Lab

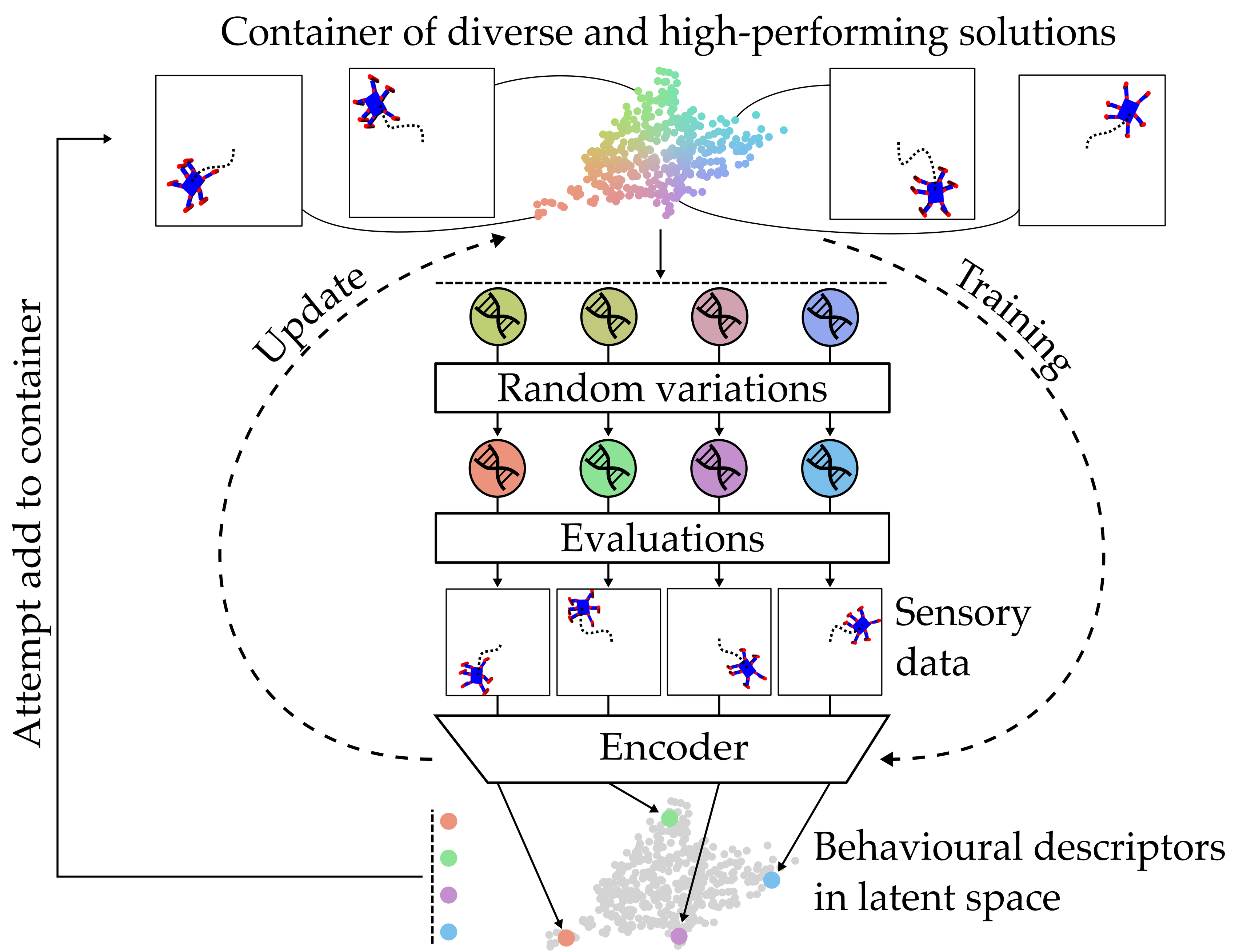
Imperial College London

Student: Luca Grillotti *First Supervisor:* Antoine Cully

Problem statement

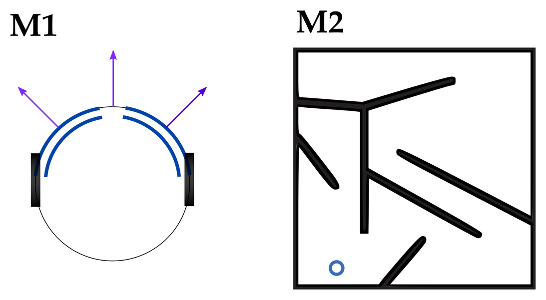
- Robots are expected to exhibit versatility and resilience if they fall into an unforeseen situation.
- From an engineering point of view, it is technically challenging to anticipate every possible situation a robot could have to face.
- Quality-Diversity (QD) algorithms [1] can find a collection of diverse and high-performing behaviours to improve the resilience of robotic systems. However, such algorithms require a hand-coded definition of behavioural descriptors, which requires prior expert knowledge about the robotic system.

We introduce **AUtonomous RObots that Realise their Abilities (AURORA)**, a QD algorithm that:
- returns a collection of diverse high-performing robotic behaviours.
- learns how to define behavioural descriptors in an unsupervised manner.

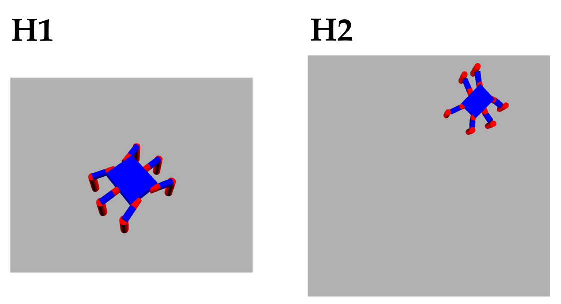


AURORA alternates between two phases:
- A **Quality-Diversity phase** (plain arrows), which improves the diversity and performance of the collection
  1. Parent behaviours are selected from the collection
  2. Those parent behaviours are copied and undergo random variations
  3. We evaluate their new performance and the sensory data they collect
  4. An encoder then maps that sensory data into low-dimensional behavioural descriptors
  5. If the novelty of their behavioural descriptor and their performance are high enough, they can be added to the collection.
- An **Encoder update phase** (dashed arrows), which learns a behavioural descriptor definition that characterises the sensory data collected by the individuals in the container.
  1. The encoder is trained to learn a low-dimensional embedding of the sensory data collected by the individuals in the container.
  2. Each time the encoder is updated, it is used to re-compute the behavioural descriptors for all the individuals.in the container

Experimental tasks

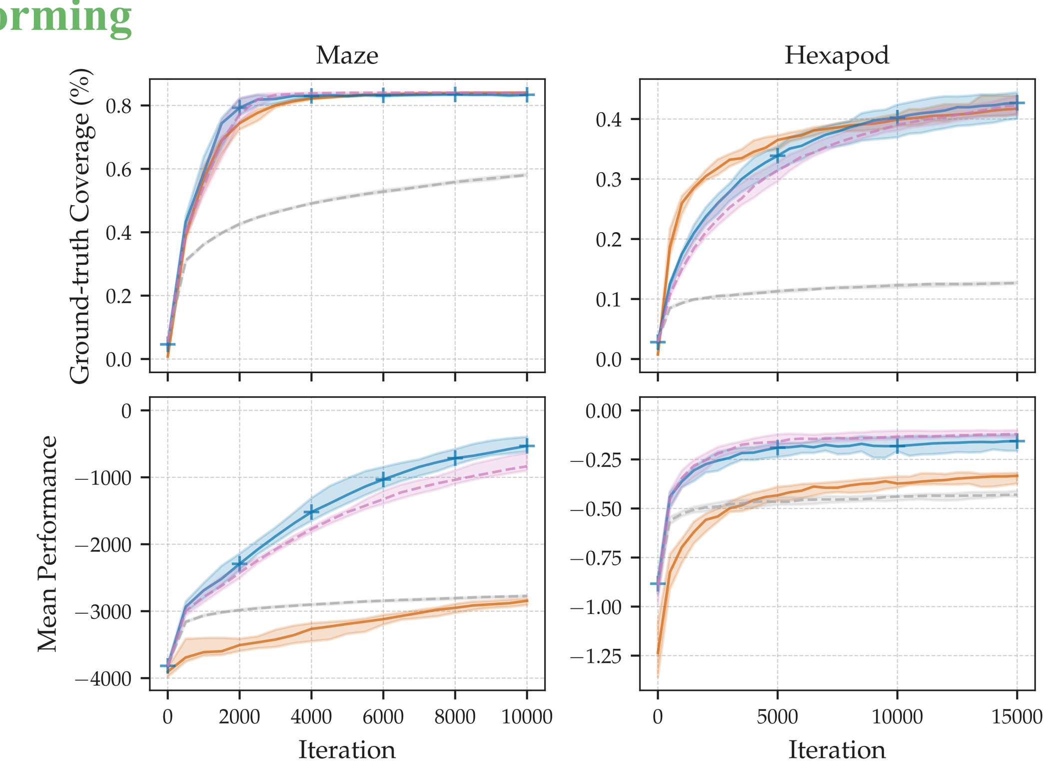


Maze task:
- *environment:* a 2D circular robot (M1) moves in a maze (M2).
- *task:* find a collection of controllers ending at diverse final positions in the maze M2.
- *controllers:* multi-layer perceptron outputting wheel velocities.
- *sensory data:* image of the maze.



Hexapod task:
- *environment:* a simulated hexapod robot moves in a 3D DART environment (H1).
- *task:* find a collection of controllers ending at diverse final positions.
- *controllers:* parameterisation of sinewave-like functions controlling the 18 robot actuators (three per leg).
- *sensory data:* top-down picture from the vertical of the origin (see H2).

AURORA manages to find collections of behaviours, which:
- achieve diverse final positions
- are high-performing



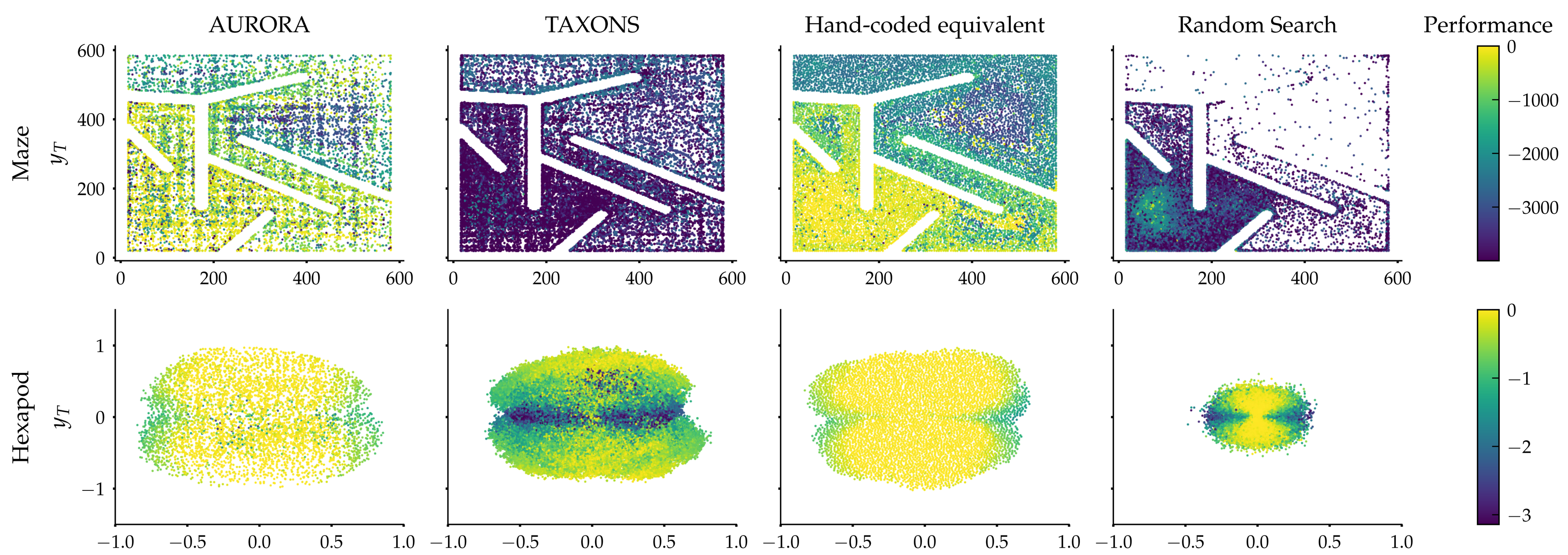
AURORA Proposed algorithm.

TAXONS [2] Prior approach similar to AURORA, but based on another framework: Novelty Search [3].

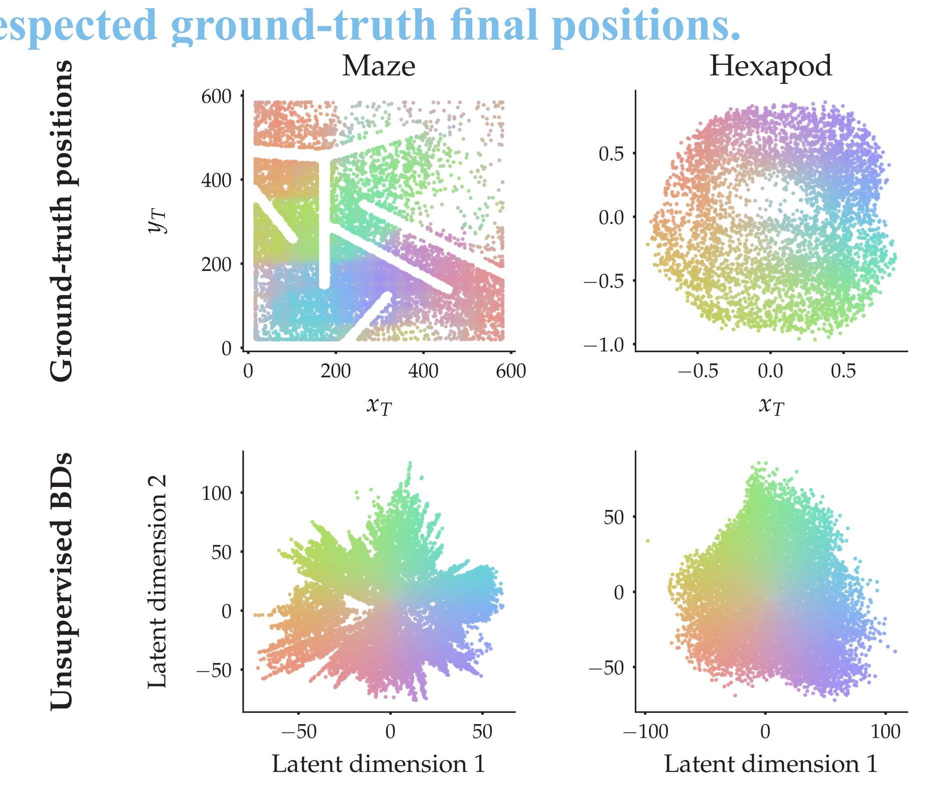
Hand-coded equivalent Same as AURORA, but using directly the ground-truth position as an hand-coded behavioural descriptor.

Random Search Generating random behaviours.

Collections of behaviours obtained by those algorithms:
(each dot represent the ground-truth final position obtained by one behaviour in the collection)

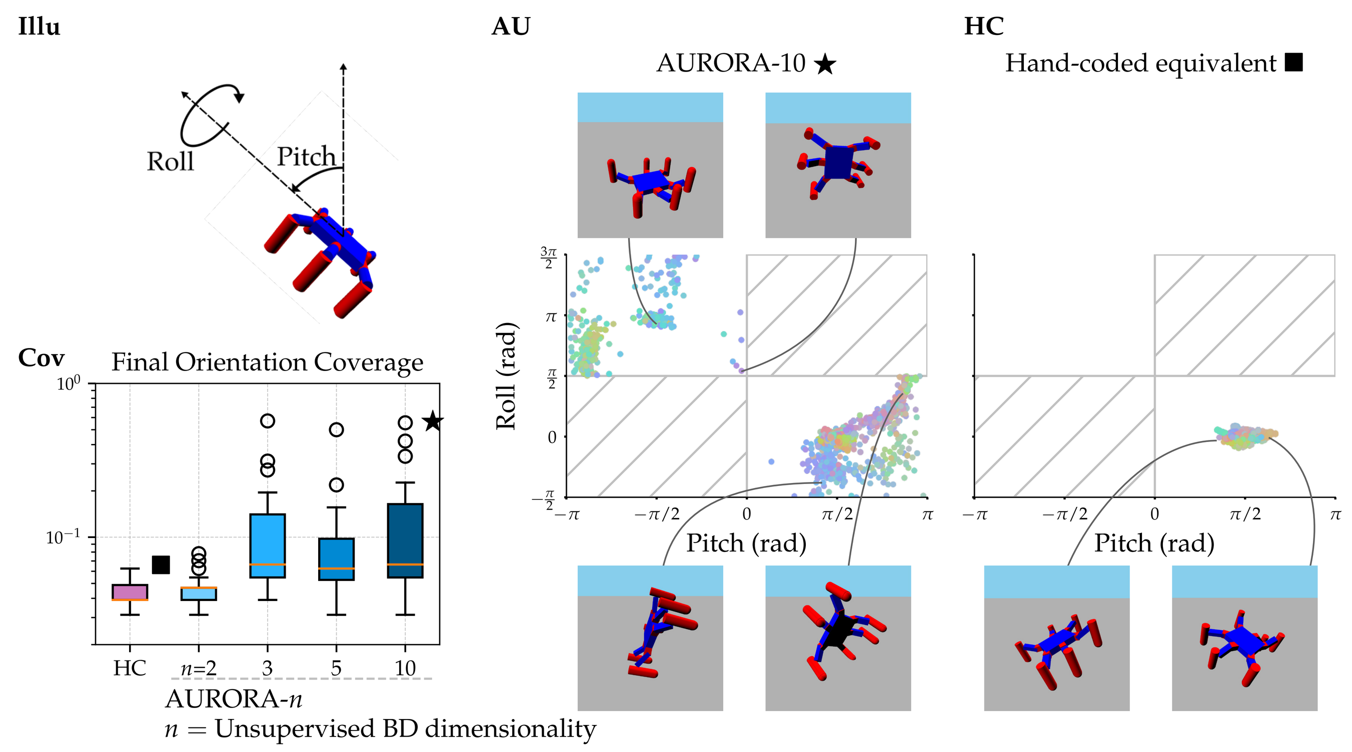


The unsupervised behavioural descriptors learnt by AURORA are meaningful with respect to their respected ground-truth final positions.



AURORA can improve the container diversity with respect to more features than the hand-coded behavioural descripts.

On the *Hexapod task*, AURORA learns behavioural descriptors based on full hexapod images. Those images contain some information regarding the orientation of the robot. The collections returned by AURORA present some diversity in terms of pitch-roll final orientations.



Future Works:

- Use AURORA to **learn diverse Hexapod trajectories; use those trajectories for real-world planning** using the Reset-free Trial and Error algorithm [5].
- So far, the sensory data always contains full information about the ground-truth position of the robot. We plan to evaluate **what happens if the sensory data provides incomplete information** (e.g. if the sensory data comes from embedded camera).
- All the presented results have been obtained using a fully connected auto-encoder as Encoder. **We intend to also study the effect of non-reconstruction-based techniques** such as UMAP [6].

Submitted Papers:

- Grillotti, L., & Cully, A. (2021). Unsupervised Behaviour Discovery with Quality-Diversity Optimisation. *arXiv preprint arXiv:2106.05648*.
- Lim, B., Grillotti, L., Bernasconi, L., Cully, A (2021). Dynamics-Aware Quality-Diversity for Efficient Learning of Skill Repertoire. Submitted to *Conference on Robot Learning*.

Bibliography:

[1] Pugh, J. K., Soros, L. B., & Stanley, K. O. (2016). Quality diversity: A new frontier for evolutionary computation. Frontiers in Robotics and AI, 3, 40.
[2] Paolo, G., Laflaquiere, A., Coninx, A., & Doncieux, S. (2020, May). Unsupervised learning and exploration of reachable outcome space. In 2020 IEEE International Conference on Robotics and Automation (ICRA) (pp. 2379-2385). IEEE.
[3] Lehman, J., & Stanley, K. O. (2011). Abandoning objectives: Evolution through the search for novelty alone. Evolutionary computation, 19(2), 189-223.
[4] Cully, A. (2019, July). Autonomous skill discovery with quality-diversity and unsupervised descriptors. In Proceedings of the Genetic and Evolutionary Computation Conference (pp. 81-89).
[5] Chatzilygeroudis, K., Vassiliades, V., & Mouret, J. B. (2018). Reset-free trial-and-error learning for robot damage recovery. Robotics and Autonomous Systems, 100, 236-250.
[6] McInnes, L., Healy, J., & Melville, J. (2020). UMAP: uniform manifold approximation and projection for dimension reduction.