6th International Conference on Control, Decision and Information Technologies

Universal Notice Networks: Transferable Knowledge Among Agents

Mehdi Mounsif[1], Sébastien Lengagne[1], Benoit Thuilot[1], Lounis Adouane[1]

[1] **Institut Pascal**: UMR CNRS 6602, Clermont-Ferrand, France

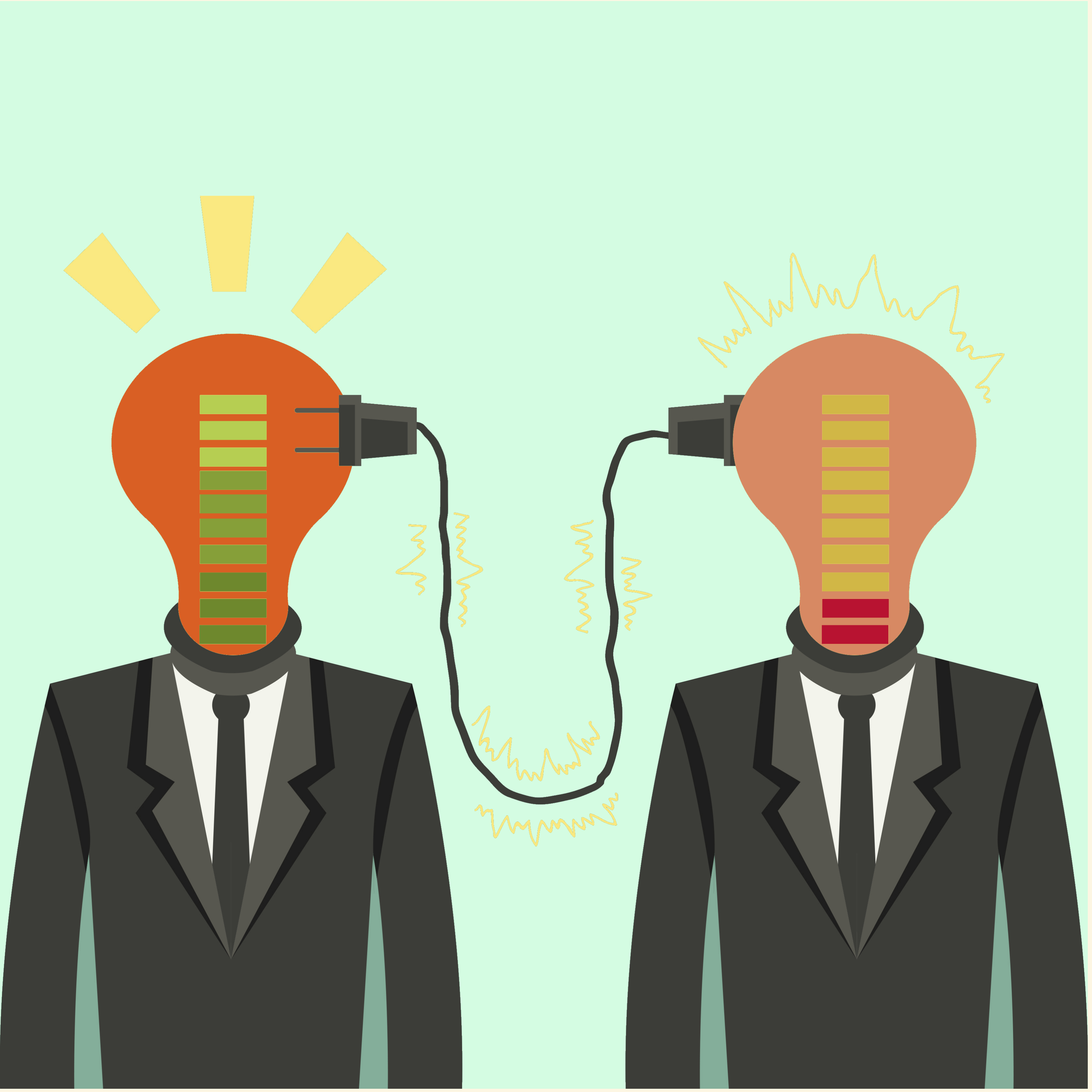
A highly intelligent invertebrae

- *9 brains*
- *500 billons neurons, mostly located in the ams*
- *Brain/body ratio comparable to mammal's*
- *Use of tools*
- *Learns through observation and play*

- Octopus and squids are one of the most intelligent species on the planet
- They also are very ancient
- **How come we do not have a sub-marine neighbour civilization ?**

Potential causes

1. It is difficult for them to discover fire
2. They have a much shorter lifespan
3. **Once their eggs are created, the mother usually goes away, leaving them figure out how the reinvent the wheel**



Humans use several mediums to share their knowledge:

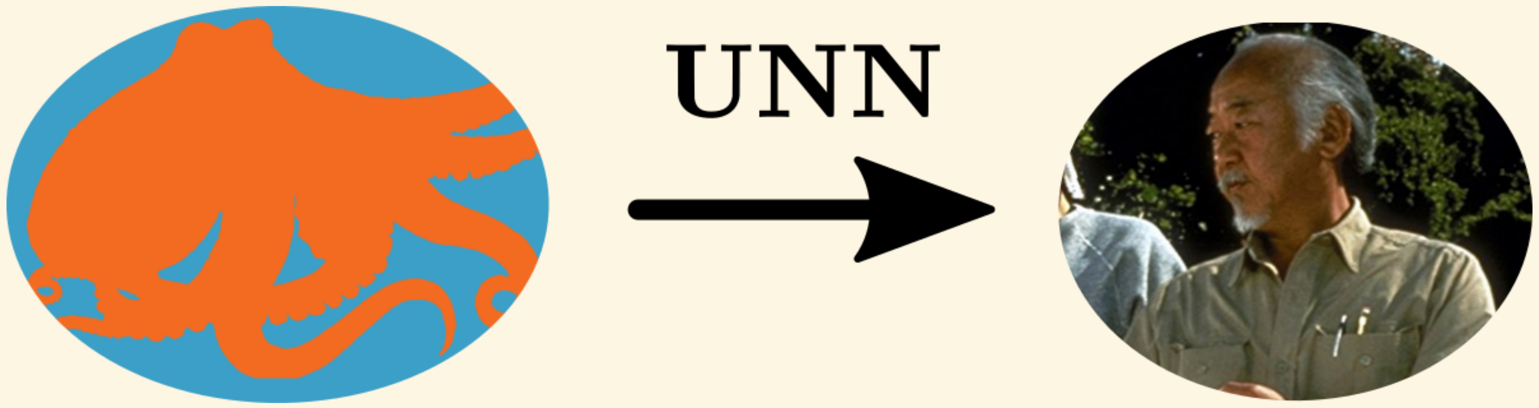
- Speeches / Talks
- Demonstrations
- Written instructions

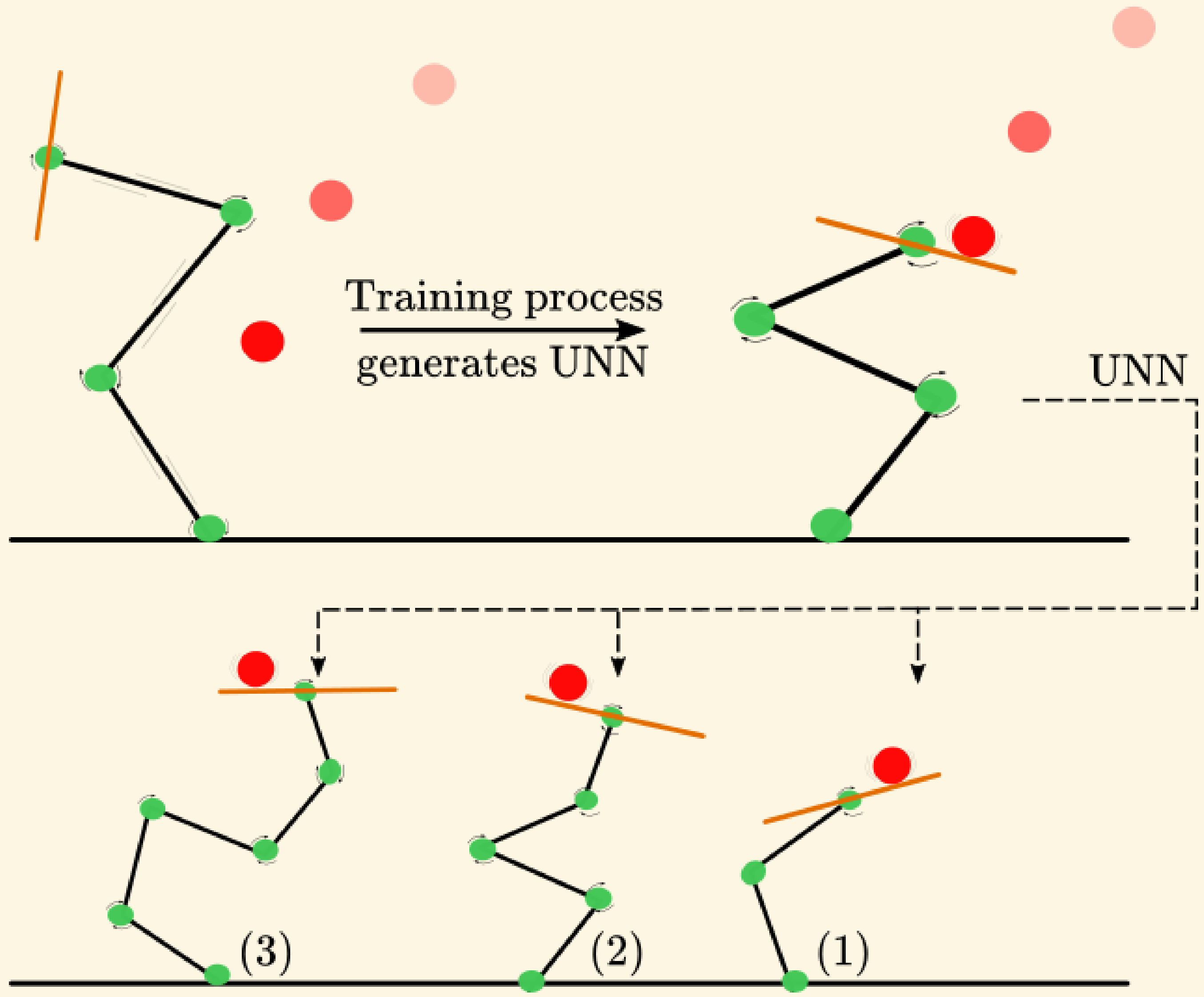
However, most of these do not convey informations in a perfect way. Which implies that the information exists on a medium but usually goes through several filters before reaching the target

As opposed to octopuses, humans spend most of their lives receiving knowledge from their peers.

Intuition

In order to tackle some of these issues, we propose the **Universal Notice Network** (UNN). The UNN is a pipeline architecture that enable the transmissions of skills from one agent to another which:





- **Separates clearly the task knowledge** from the other control modules, which in turn:
- Allows **transmission** to other agents
- Improves **reusability**
- Provides means for **life-long adaptation**

SotA: Model-Based

An important part of robotic research relies on model-based approaches. Indeed, these powerful methods provide:

- Theoretical garantees
- Throrough analysis

For instance, the Stack of Tasks [1] is a solid theoretical framework that stacks constraints with a task priority order and finds solutions complying with these constraints.

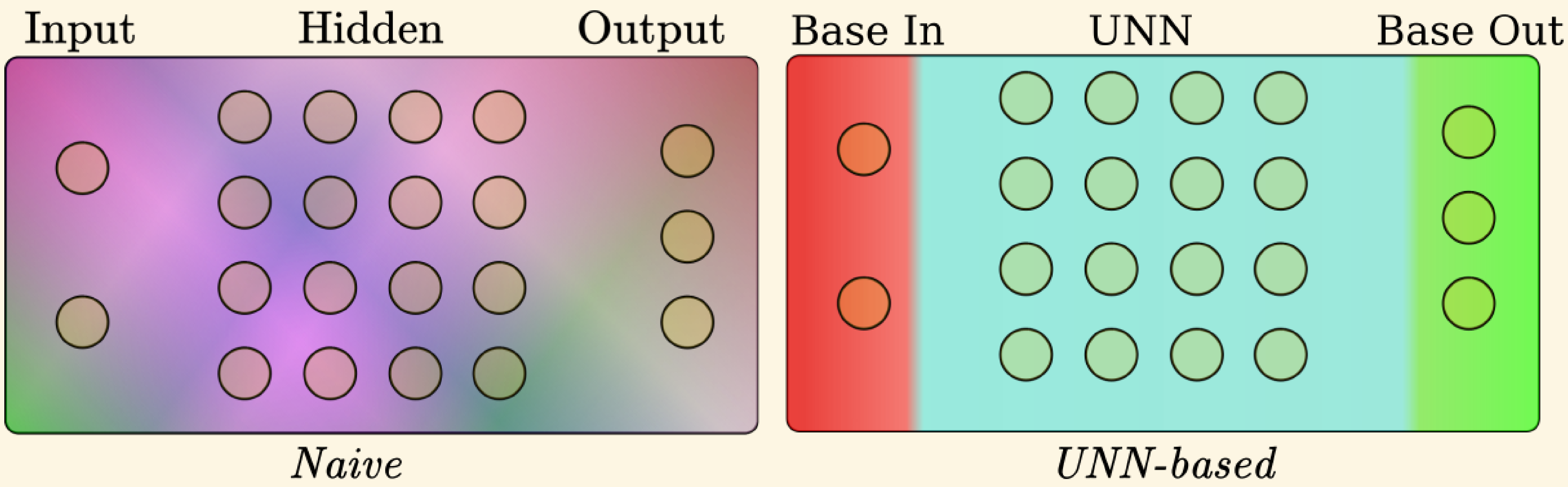
- However these methods rely entirely on the **mathematical problem formulation** that can be hard, if not impossible to obtain.

[1] N. Mansard & *al*. **A Versatile Generalized Inverted Kinematics Implementationfor Collaborative Working Humanoid Robots:The Stack of Tasks**. 2009 International Conference on Advanced Robotics

The UNN Method

Goal: Create a module for a task that is model-agnostic

Problem: Naive training results in unstructured knowledge inside weight matrices.



The UNN strengh is to **enforce, by construction, a segmentation** between the task model and the other modules, called **bases**, making the task **model-agnostic**. Hence, knowledge to solve the task can be passed from one agent to another, ideally seamlessly.

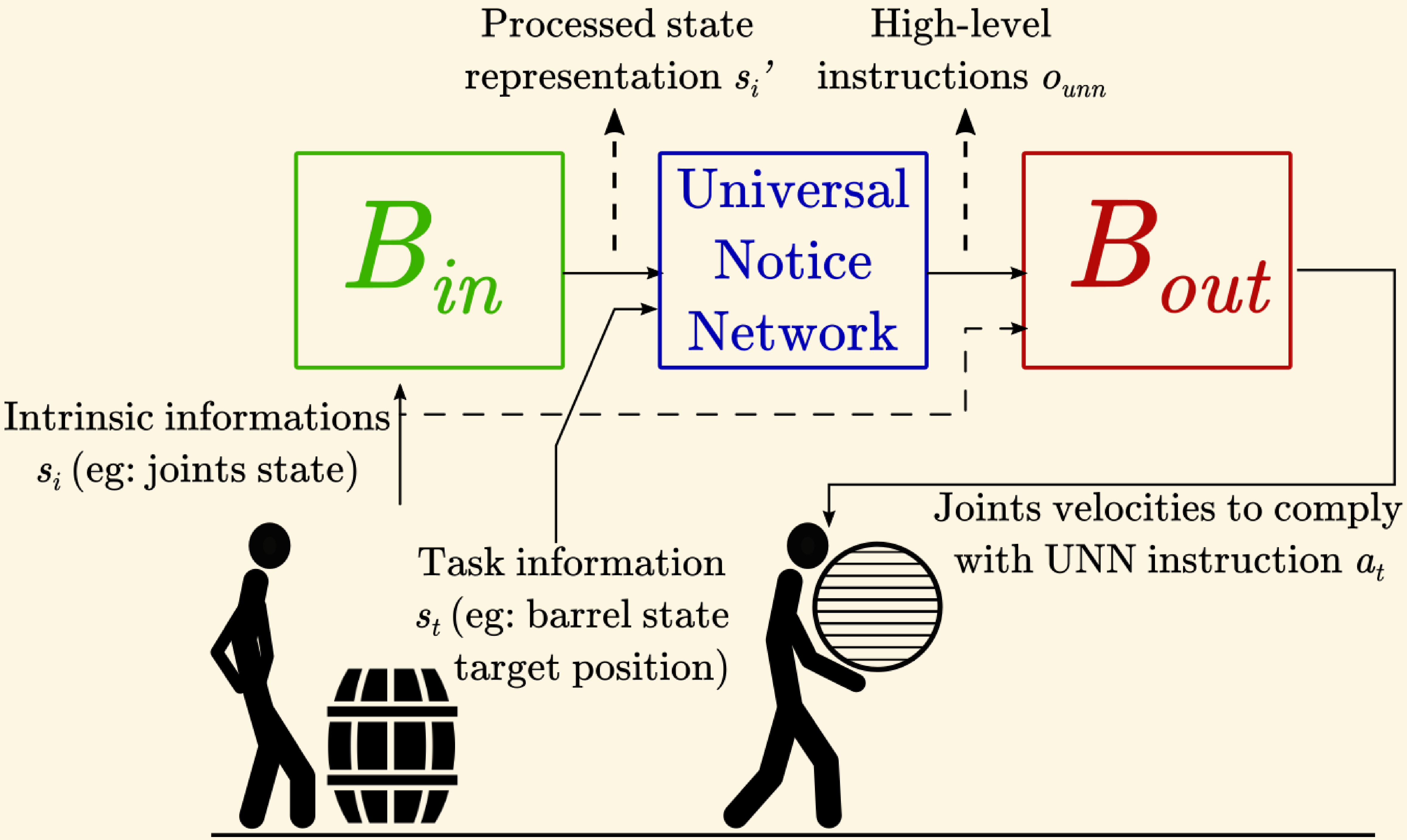
SotA: Model-Free

- **Broaden agent skills**
  - DIAYN [2]: Entropy-based loss to incentize learning new skills
  - Curiosity based approaches [3]: Intrinsic reward for more exploration
- **Transfer learning from one environment to another**
  - Impala, PopArt [4], [5]: Navigation tasks mostly

The focus is rarely given to transmitting knowledge to another entity.

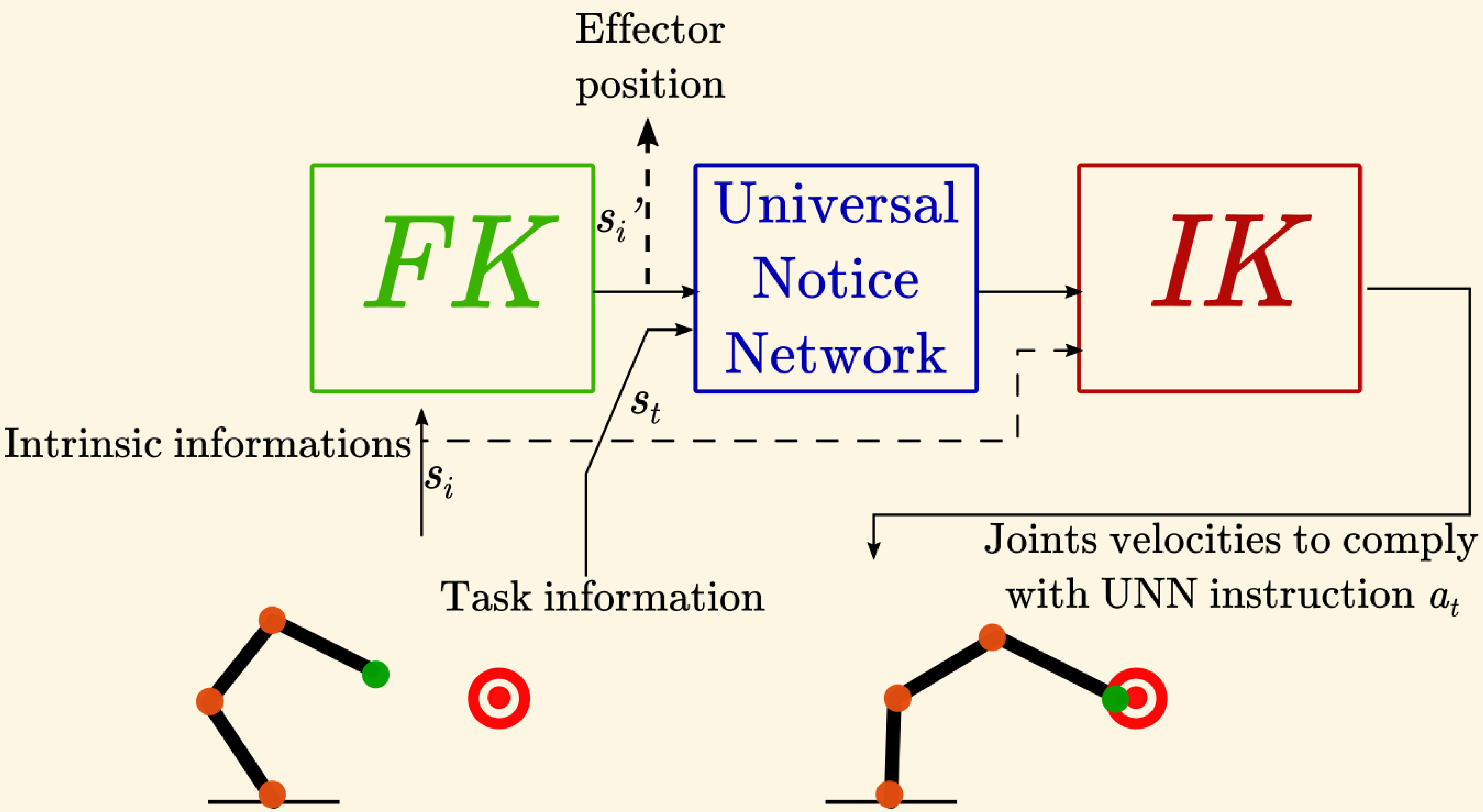
[2] B Eysenbach & *al* **Diversity is all you need: Learning skills without a reward function** arXiv preprint arXiv:1802.06070
[3] Y. Burda & *al* **Exploration by random network distillation** arXiv preprint arXiv:1810.12894
[4] E. Lasse & *al* **IMPALA: Scalable Distributed Deep-RL with Importance Weighted Actor-Learner Architectures** arXiv preprint arXiv:1802.01561
[5] H. Matteo & *al* **Multi-task Deep Reinforcement Learning with PopArt** arXiv preprint arXiv:1809.04474

The UNN Architecture



The UNN Proof of Concept

As a proof of concept, we consider the simplest case



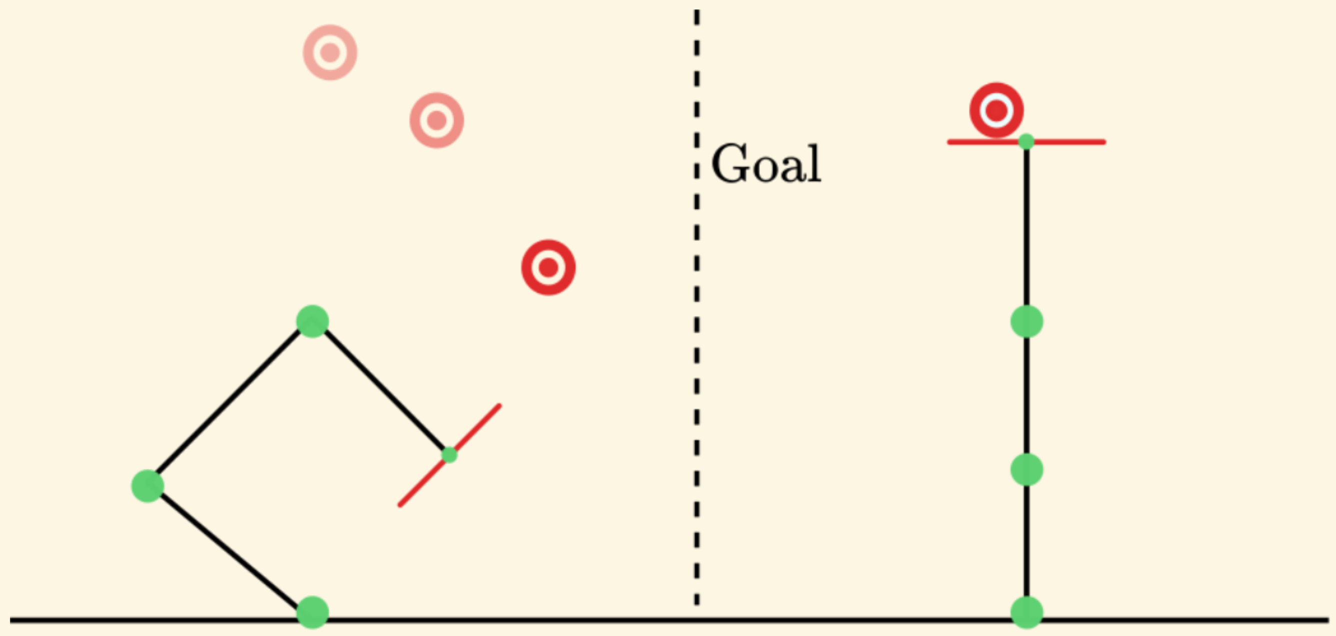
This simple setup has two main advantages:

1. Works for any serial manipulator configuration and is easy to transfer
2. Absolute segmentation: Ensure that the task knowledge is held in the central UNN

The Catcher Task

Goal: Catch and raise the ball

1. Observe world state
   - Intrinsic State: Joints angles
   - Task Informations: Ball position & speed
2. FK: Outputs Effector position based on intrinsic informations
3. UNN uses FK output and task informations $\rightarrow$ Effector direction + bar orientation
4. IK uses intrinsic state and UNN output $\rightarrow$ Joints speeds

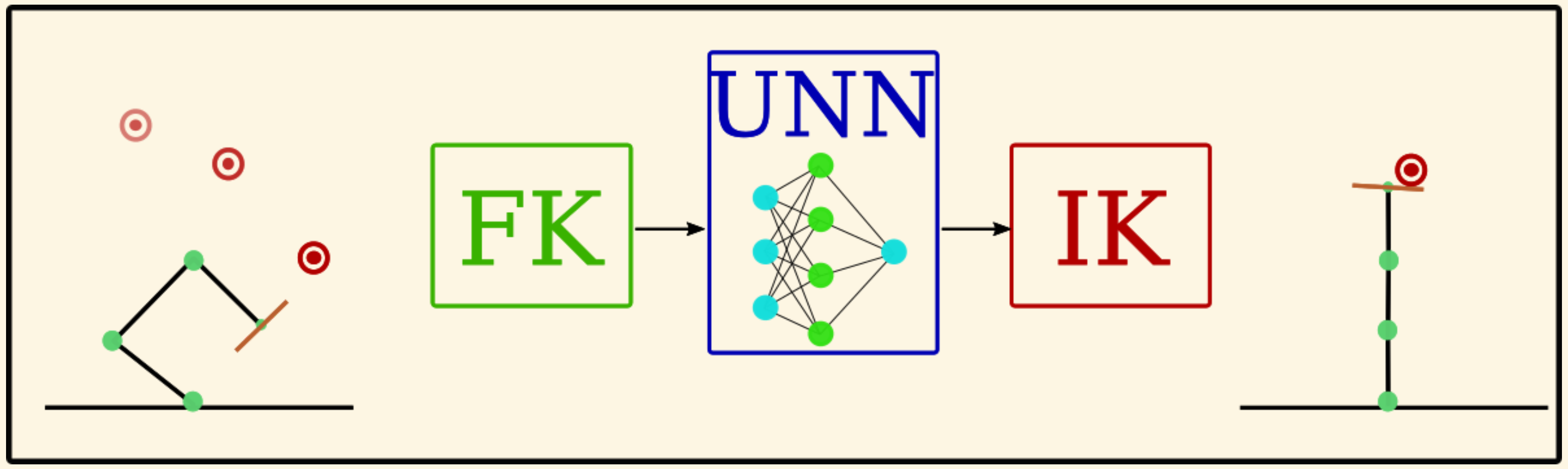


Process

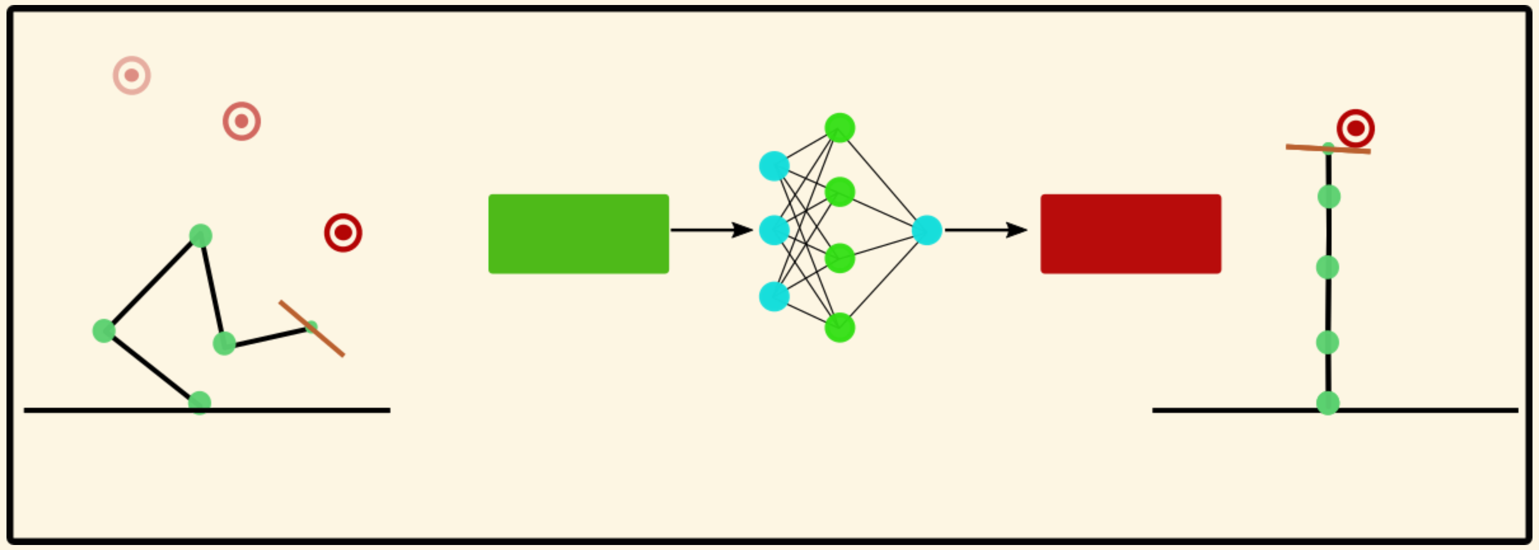
1. Train one agent configuration *via* Reinforcement Learning
2. Transmission of learned UNN to various different configurations

Results on Catcher

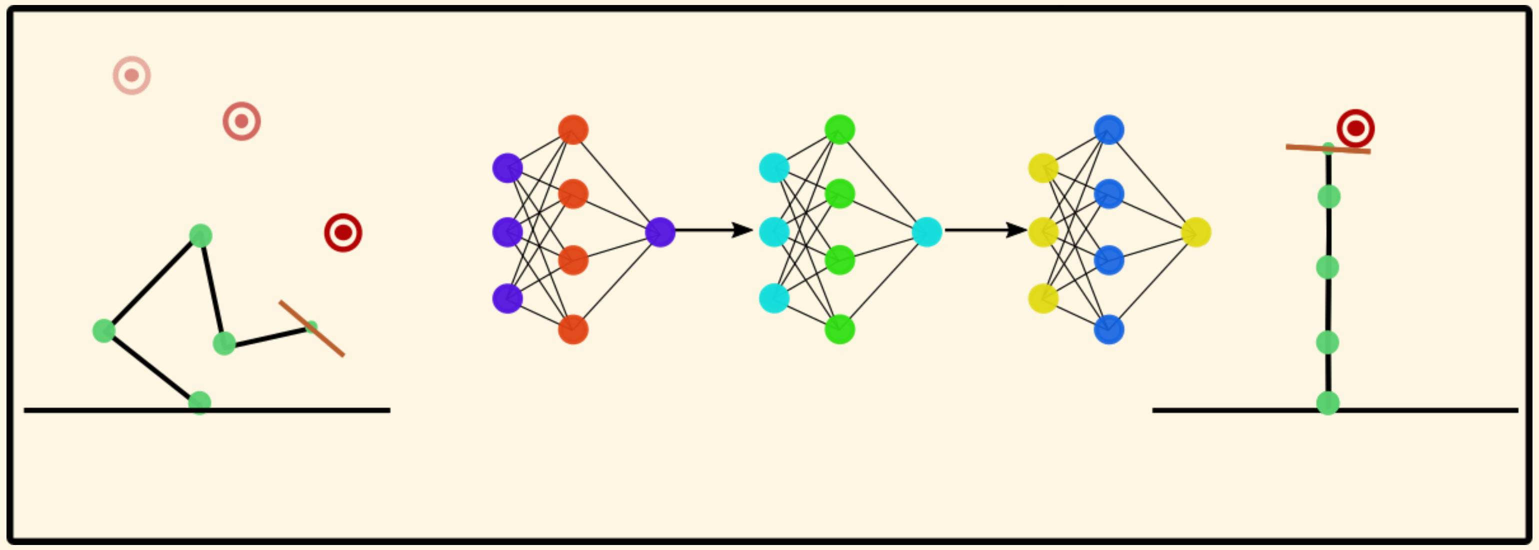
I. Training the UNN



II.1 UNN Transmission



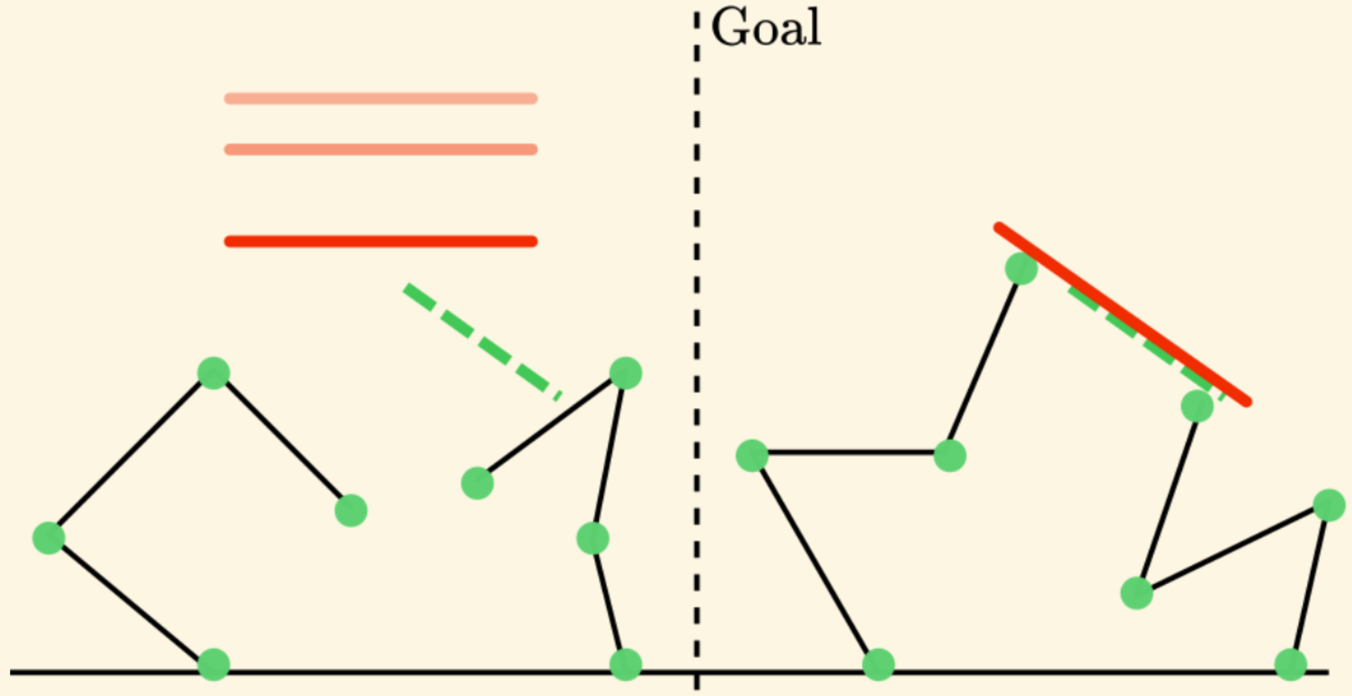
II.2 Base Recovery

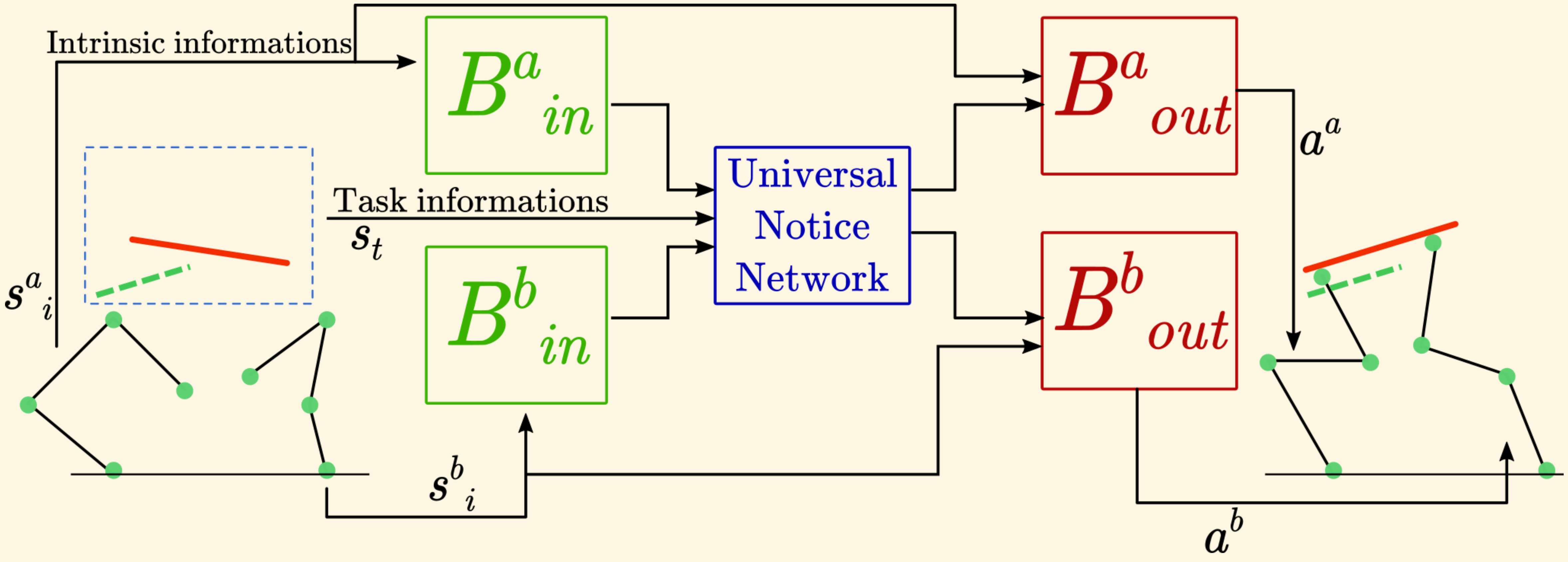


Results on Double Catcher

Goal: Catch and raise the bar

- State observation:
  - Intrinsic State: Joints angles for each robot
  - Task Informations: Bar position, angular position, speed & angular speed





Conclusions

We introduced the UNN:

- A method for transfer learning between agents of different configurations
- Compatible with domain-knowledge and heuristics
- Compatible with multi-agents/hierarchical settings
- Displaying interesting capabilities for real-world transfers

Perspectives

- Scale up to 3D environments with less physics glitches and more complex tasks
- Formalize the use of heuristics as the Physical Curriculum learning
- Abstract from the physics-based state representation

THANK YOU FOR YOUR ATTENTION !