# DELTA: Dynamically Evolving Long-Term Autonomy

## Partners

**UPF**: Universitat Pompeu Fabra, Spain (coordinator)
**INRIA**: INRIA Lille, France
**MUL**: Montanuniversität Leoben, Austria
**ULg**: Université de Liège, Belgium

## Motivation

**Reinforcement learning (RL)** offers a powerful framework for acquiring adaptive behaviour in autonomous systems, but makes strong assumptions:

- The environment model is stationary.
- The objective is fixed.
- The task ends once the objective is met.

In **lifelong learning**, autonomous systems are operative during long time periods, and the above stationarity assumptions typically do not hold.

## Project Objectives

Develop several novel algorithms for **lifelong RL** that relax common stationarity assumptions, with the following desired properties:

- Robust to environmental changes, both in terms of observations and actions.
- Able to operate over long time periods while achieving different objectives.

The proposed algorithms will address three key problems in lifelong RL: **planning**, **exploration** and **task decomposition**.

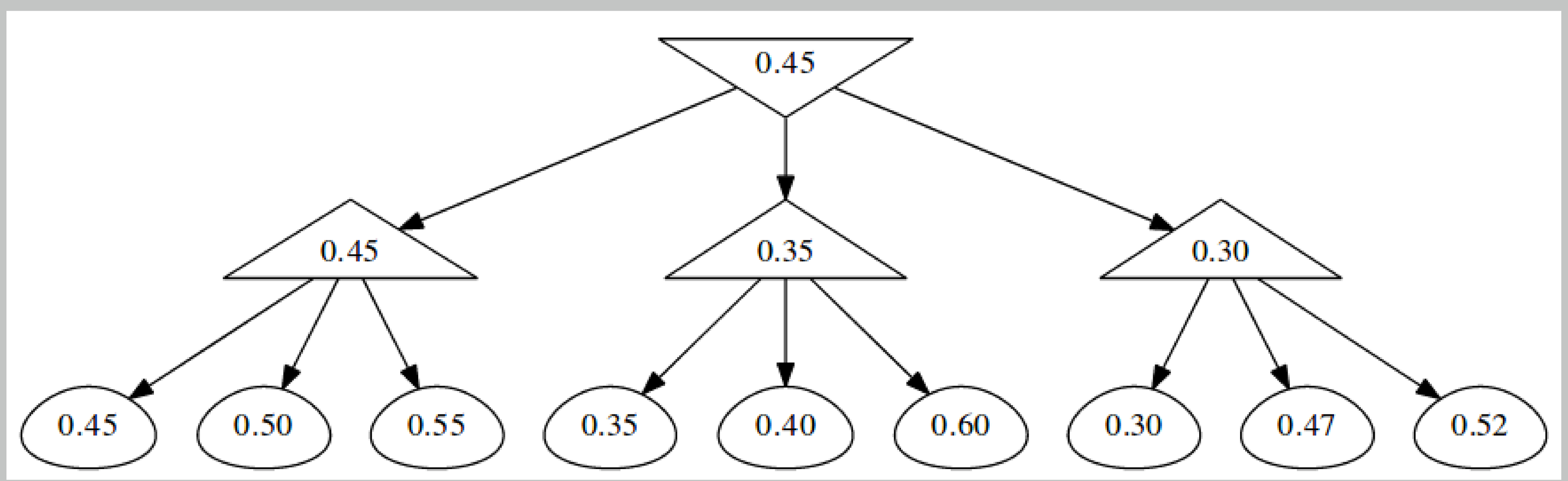
## Workpackages

- **WP1**: Management and dissemination
- **WP2**: Planning in complex changing environments
- **WP3**: Exploration
- **WP4**: Task decomposition
- **WP5**: Integration

## WP2: Planning in complex changing environments

**Responsible partner: INRIA**

**Objective**: Develop a novel planning algorithm that efficiently achieves a new, previously unknown objective given the current environment model of the system. The planning algorithm should account for the fact that the environment model may change over time.

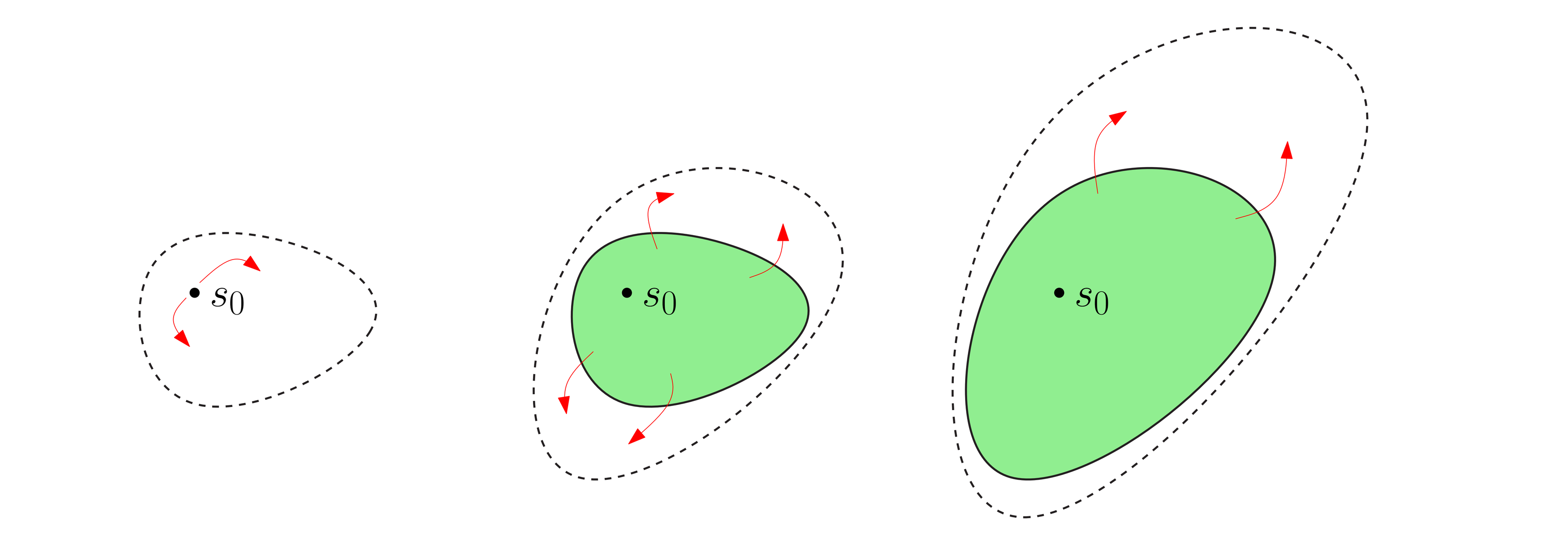



**Proposed approach**: extend Monte-Carlo tree search (MCTS) to deal with new nodes, new states, and a new model. Rather than using upper confidence bounds on the rewards, adapt algorithms for best arm identification in bandit models to achieve improved sample complexity.

## WP3: Exploration

**Responsible partner: MUL**

**Objective**: Develop a novel exploration strategy for RL that automatically and efficiently updates the environmental model, by selecting actions that explore parts of the environment that the system is not yet familiar with.

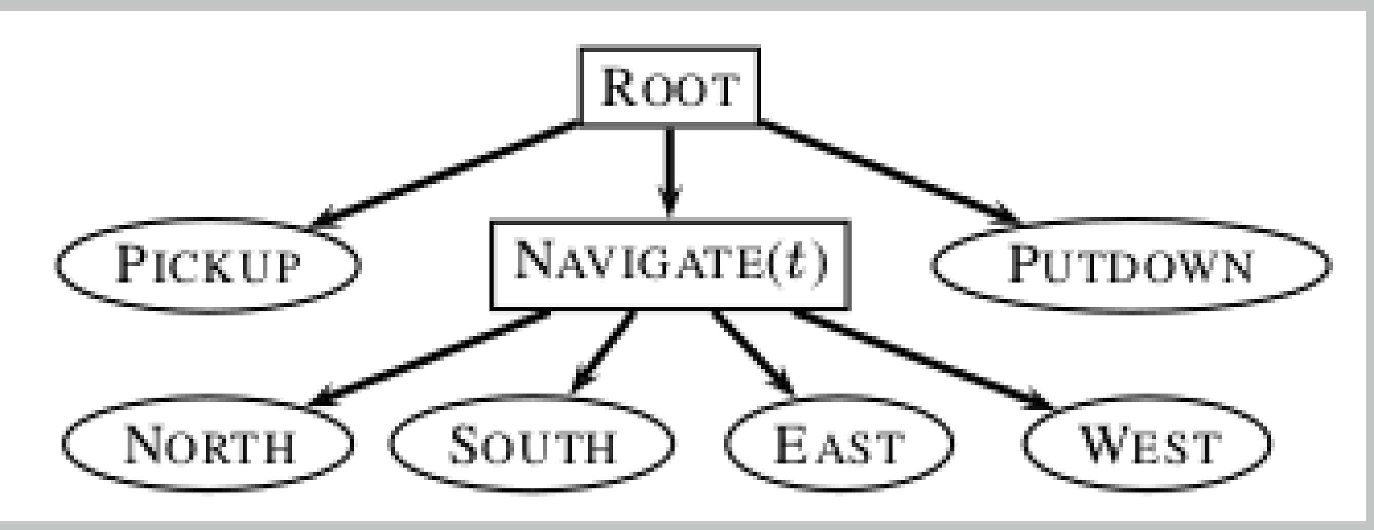



**Proposed approach**: Modify the UCRL algorithm and its analysis to cope with gradually changing environments. UCRL can already tolerate disruptive changes using restarts. Starting with a (nearly) optimal policy for a small part of the environment, this small known part can be enlarged incrementally.

## WP4: Task Decomposition

**Responsible partner: UPF**

**Objective**: Develop a novel framework for task decomposition that automatically creates and evaluates tasks, discarding tasks that are not deemed useful. Each task has its own associated decision strategy.




**Proposed approach**: Develop novel methods for subtask identification which assume a factored state representation in which subtasks correspond to certain assignments to state variables. Develop novel metrics for evaluating the utility of subtasks based on the notion of regret.

## WP5: Integration

**Responsible partner: ULg**

**Objective**: Evaluate the novel planning and RL algorithms in two realistic scenarios: **active network management** for electrical distribution networks, and **microgrid management**. Apart from using these scenarios for evaluation, the project also aims at improving on the state-of-the-art in these two applications.