

# What would a curious robot do?

## Exploration and learning in complex environments

**The Curious Robot (TCR)** is about the challenge of gaining controllability over external DoFs in complex environments.

The challenge: huge continuous state and action space.

Two modes of exploration:

- **exploration for discovering DoFs**

Where are the DoFs?

- **exploration for modeling DoFs**

How to best explore a DoF to maximize learning?

Long term plan: goal-directed manipulation of external DoFs

### The Method

We approximate the **belief** with a physics simulation.

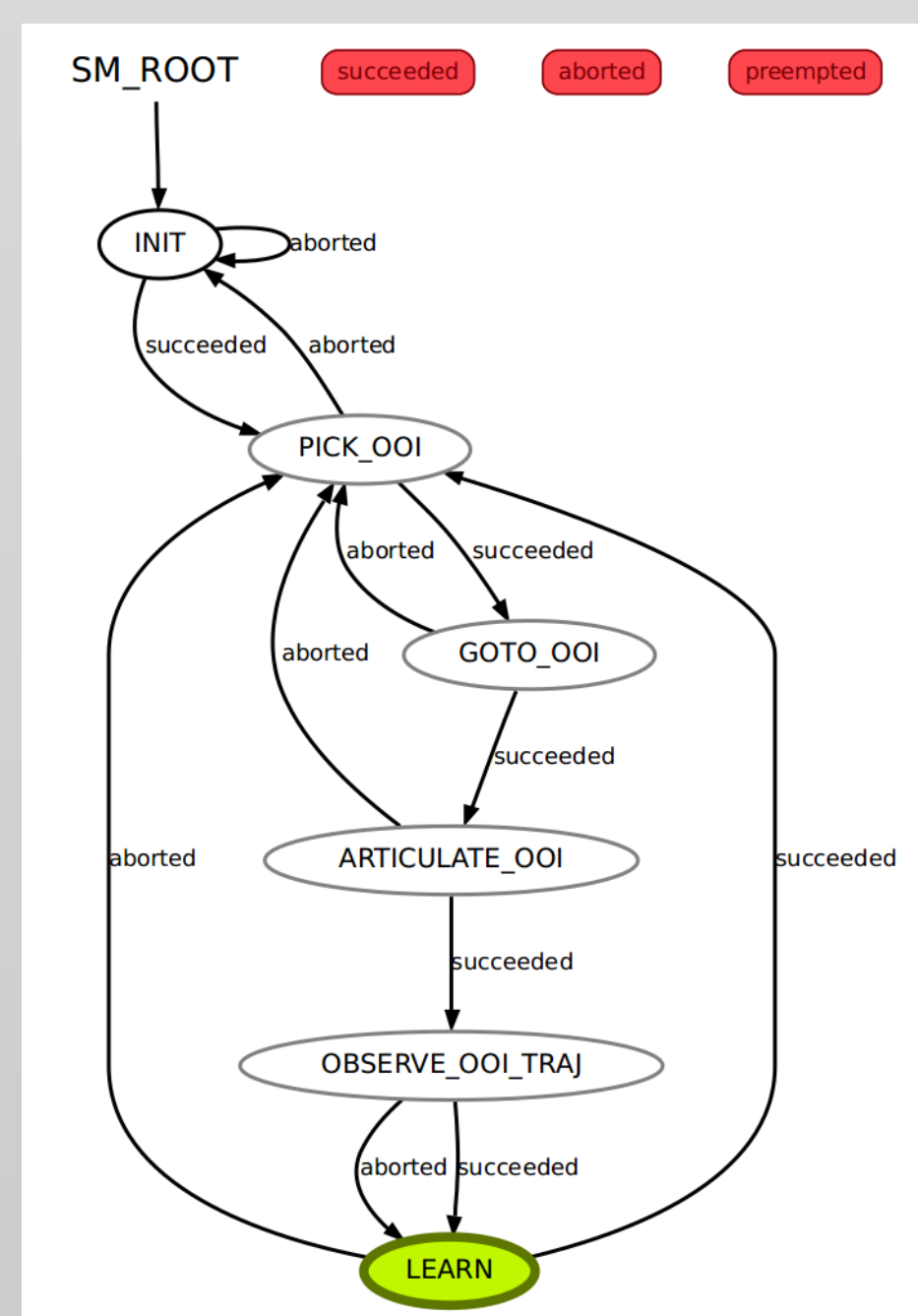
We focus on learning the geometry and kinematic structure of the environment, i.e., we learn the **object type** (static, free, articulated) and its **properties** (joint type, pose, limits, etc.).

Probability distributions are used for the types and properties of the objects.

The exploration is guided by **active learning** and maximizing the **information gain**.

### The approach

- in the beginning the robot does not know anything about the world
- the robot starts exploring the world
- it picks an object of interest for discovering DoFs
- it interacts with an object of interest for modeling the DoFs
- it updates the belief every time it learns something



### The Exploration

#### Modes and strategies of exploration

The robot must learn how to find DoFs and how to manipulate them. The strategies can be different depending on the object type.

#### Exploration for discovering

- random
- uncertainty
- heuristic uncertainty
- learned

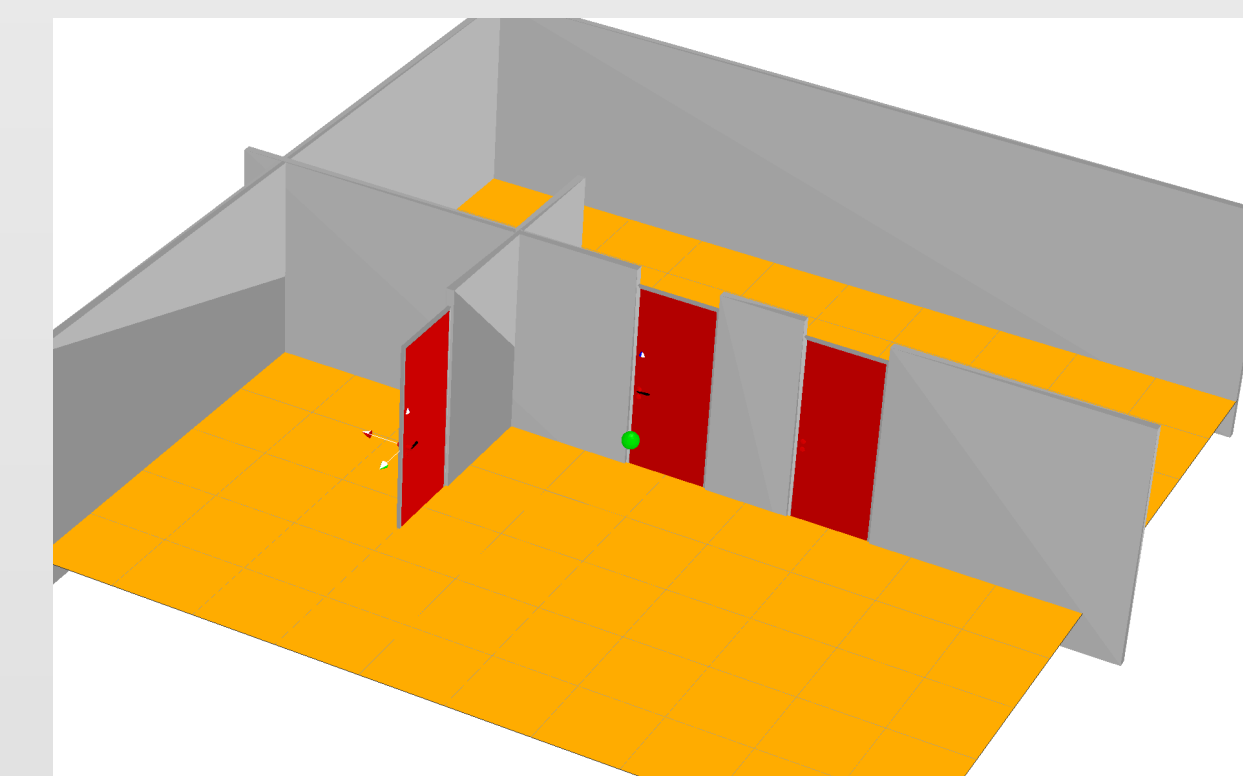
#### Exploration for modeling

- random
- heuristic
- learned

### Setup & Experiments

#### Simulation

- the world is a physics simulation of a room (walls, doors, handles, etc.)
- the robot is a flying ball / just an end-effector

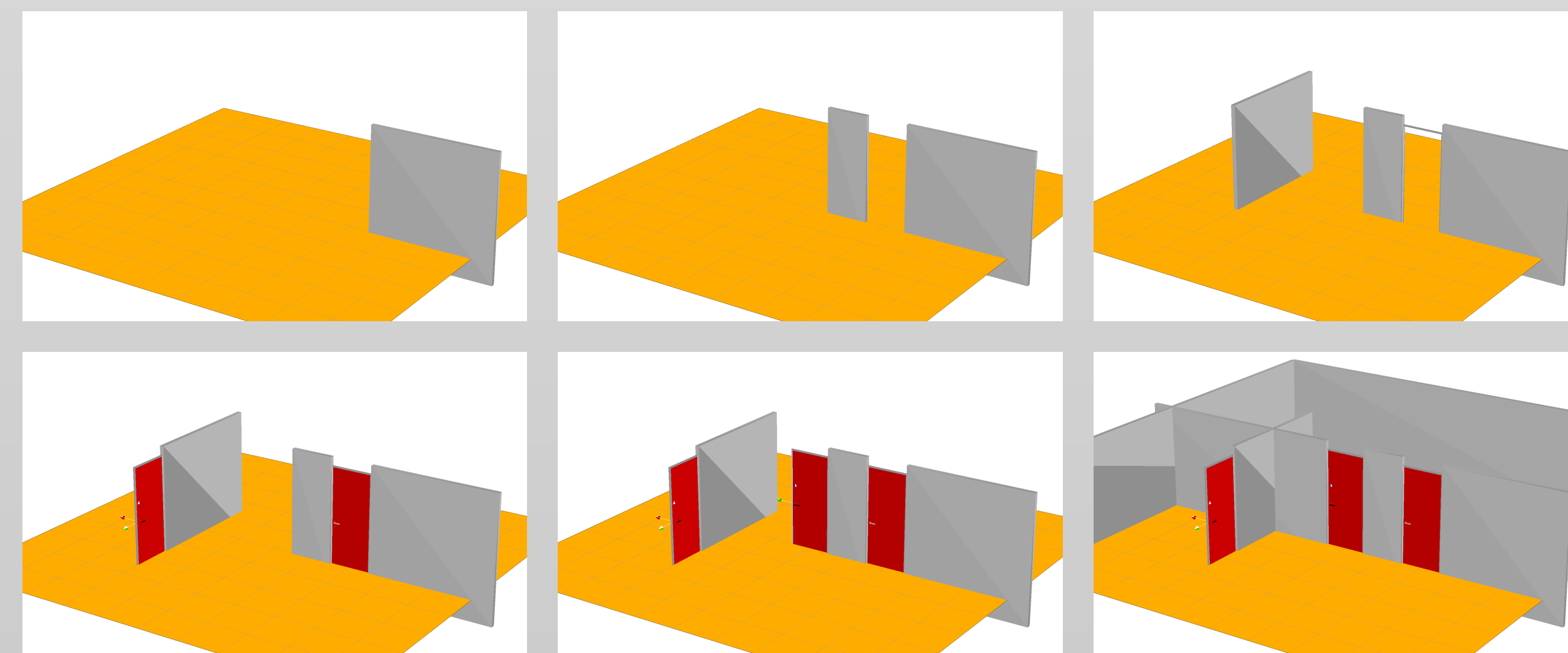


#### Real world

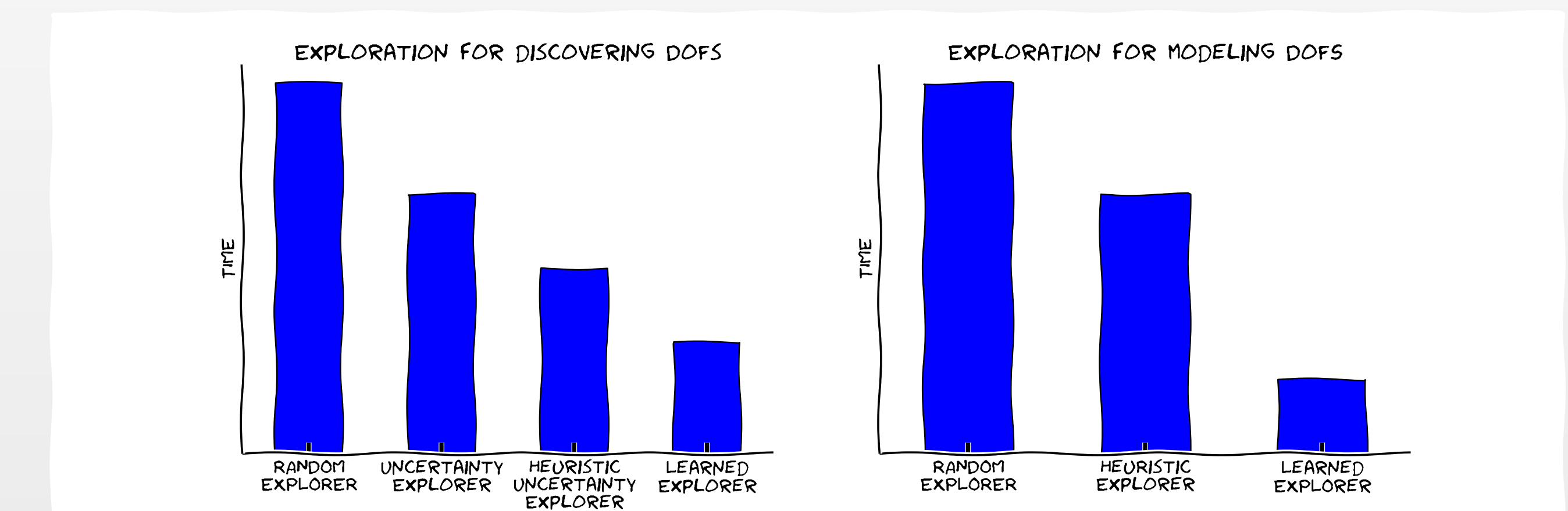
- the world is a living room (doors, drawers, handles, toys, blocks, (light) switches, pens, etc.)
- the robot is our PR2 *BigBird*



### Example of the belief update



### Dummy Results



### Outlook

#### Next steps

- use data to learn exploration strategies (probably infeasible to use the belief directly)
- more complex world
- transfer everything onto our PR2
- add symbol learning
- benchmark for exploration

Also learn from **human-robot interaction**

- learning by demonstration
- kinesthetic teaching
- directing attention

### Authors



#### Stefan Otte

<http://ipvs.informatik.uni-stuttgart.de/mlr/stefan>  
Stefan.Otte@ipvs.uni-stuttgart.de



#### Johannes Kulick

<http://ipvs.informatik.uni-stuttgart.de/mlr/johannes>  
Johannes.Kulick@ipvs.uni-stuttgart.de



#### Marc Toussaint

<http://ipvs.informatik.uni-stuttgart.de/mlr/marc>  
Marc.Toussaint@ipvs.uni-stuttgart.de