UNIVERSITAT POLITÈCNICA
DE CATALUNYA
BARCELONATECH
UPC

Institut de Robòtica
i Informàtica Industrial

Barcelona
Supercomputing
Center
Centro Nacional de Supercomputación
BSC

PiCoEDL: Discovery and Learning of Minecraft Navigation Goals from Pixels and Coordinates

Juan José Nieto

Roger Creus Castanyer

Xavier Giro-i-Nieto

CVPR
VIRTUAL JUNE 19-25
CVPR 2021
EMBODIED AI

Task description

Discover navigation goals in Minecraft by using an intrinsic motivation **(empowerment)**. Then, **learn** some **conditioned policies** that drive the agent towards the discovered goals.

(x_1,y_1,z_1) (x_2,y_2,z_2)

3D world. Input consists in RGB observations + relative position (x,y,z) to a fixed initial point.

(x_3,y_3,z_3) (x_4,y_4,z_4)
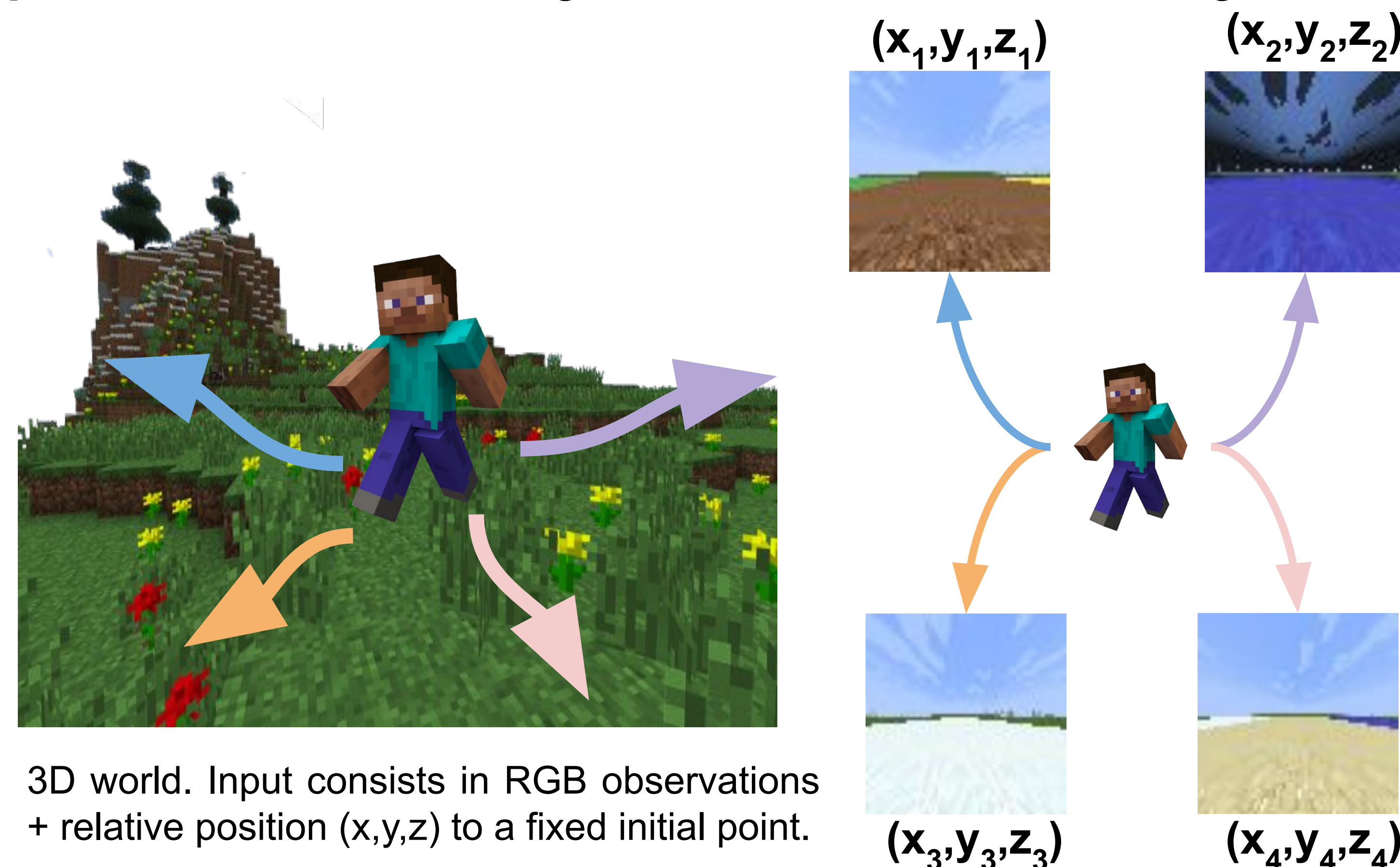
Motivation

EMPOWERMENT

Maximize influence over the environment.

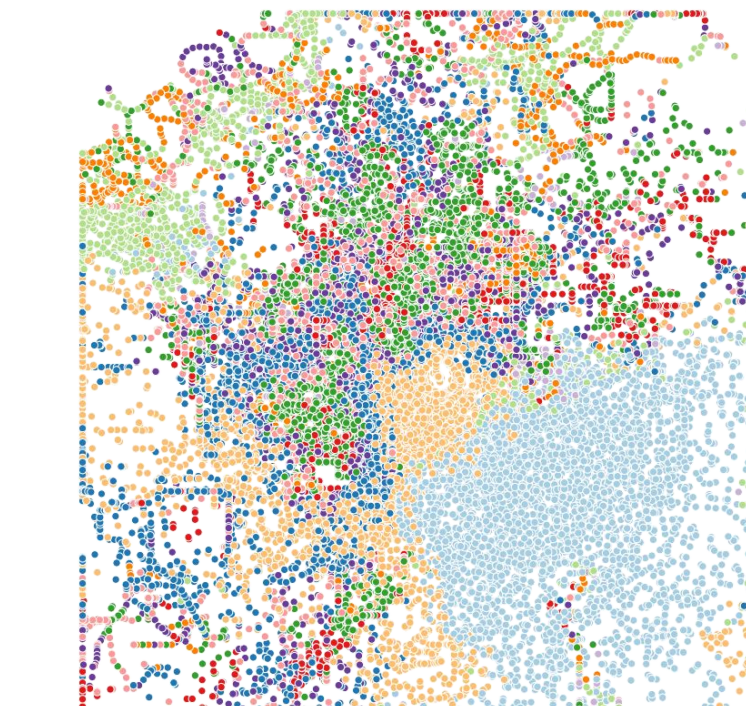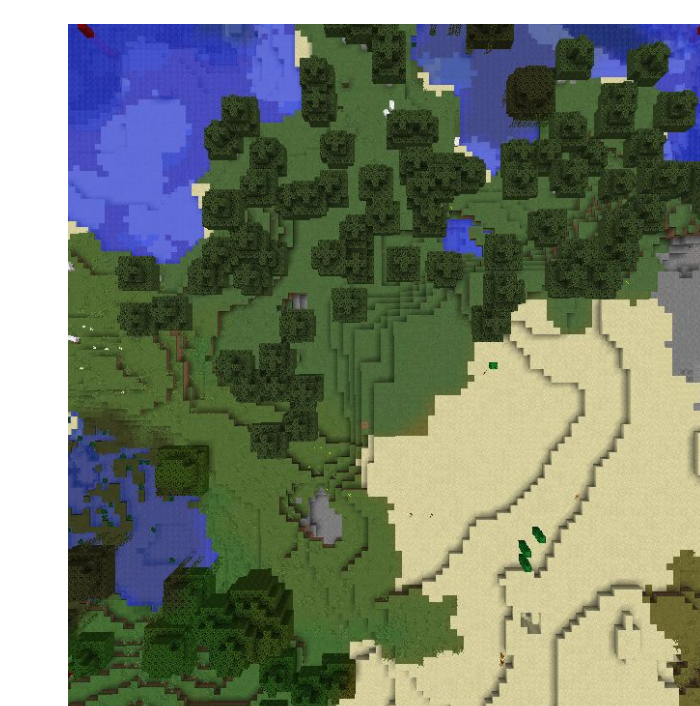**SKILL DISCOVERY**

Abstract a set of temporal actions into skills.

INFORMATION THEORETIC OBJECTIVE

Maximize skill discovery $\longrightarrow$ Maximize mutual information between states and some latents

$$I(S;Z) = H(S) - H(S|Z)$$
$$= H(Z) - H(Z|S)$$

Challenges

1. **Intrinsic reward** derived from the **forward form** of the Mutual Information **does not match a distance** in the environment.

2. Discover disjoint navigation goals in simple maps by taking into account **high-dimensional inputs** like pixel observations**.**

3. Discovering skills from pixels does not scale to realistic maps.

Navigation goals are too spread out over the map and **some latent** skills **dominate** over the others.

Our approach

Based on **EDL [1]** (Explore, Discover and Learn).
- Use the **reverse** form of the Mutual Information.
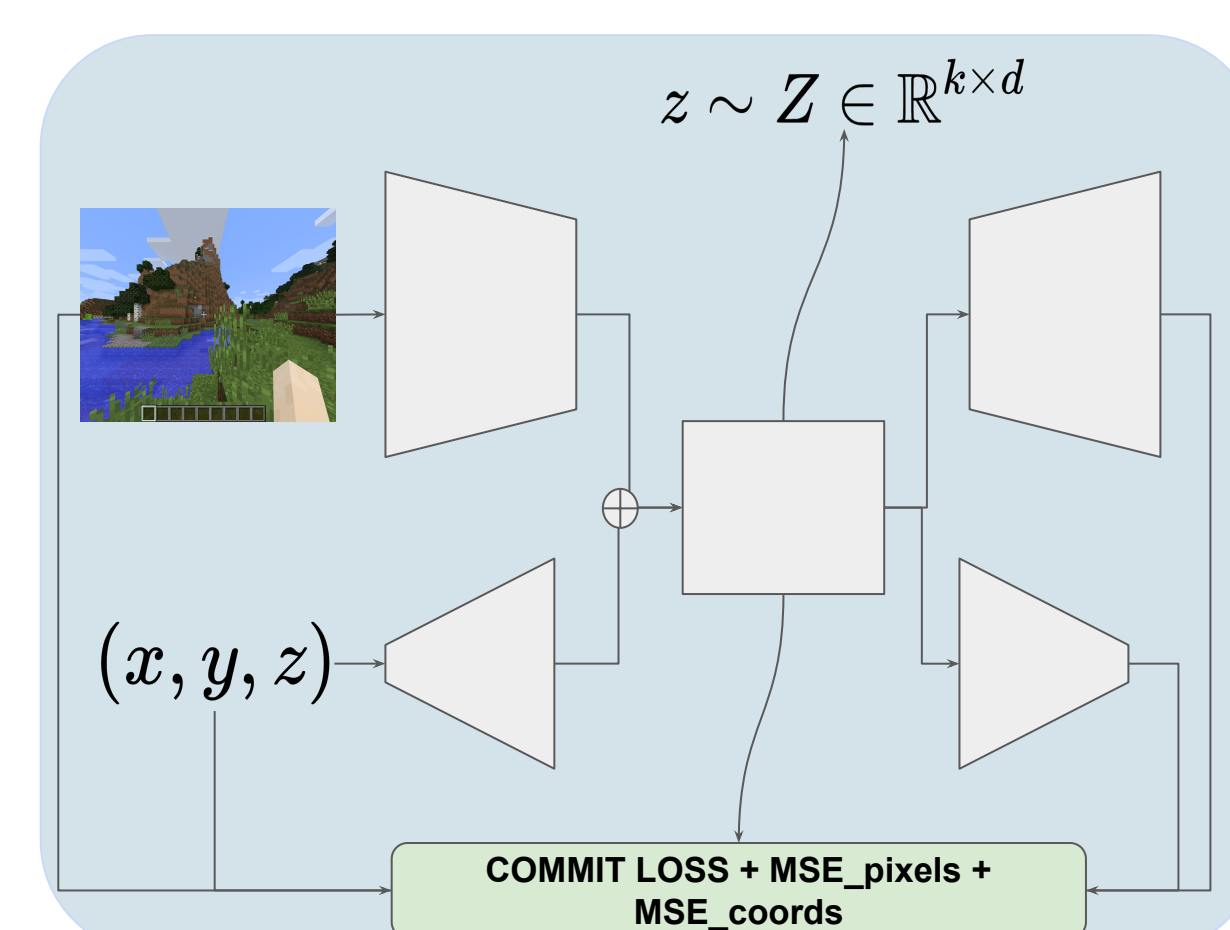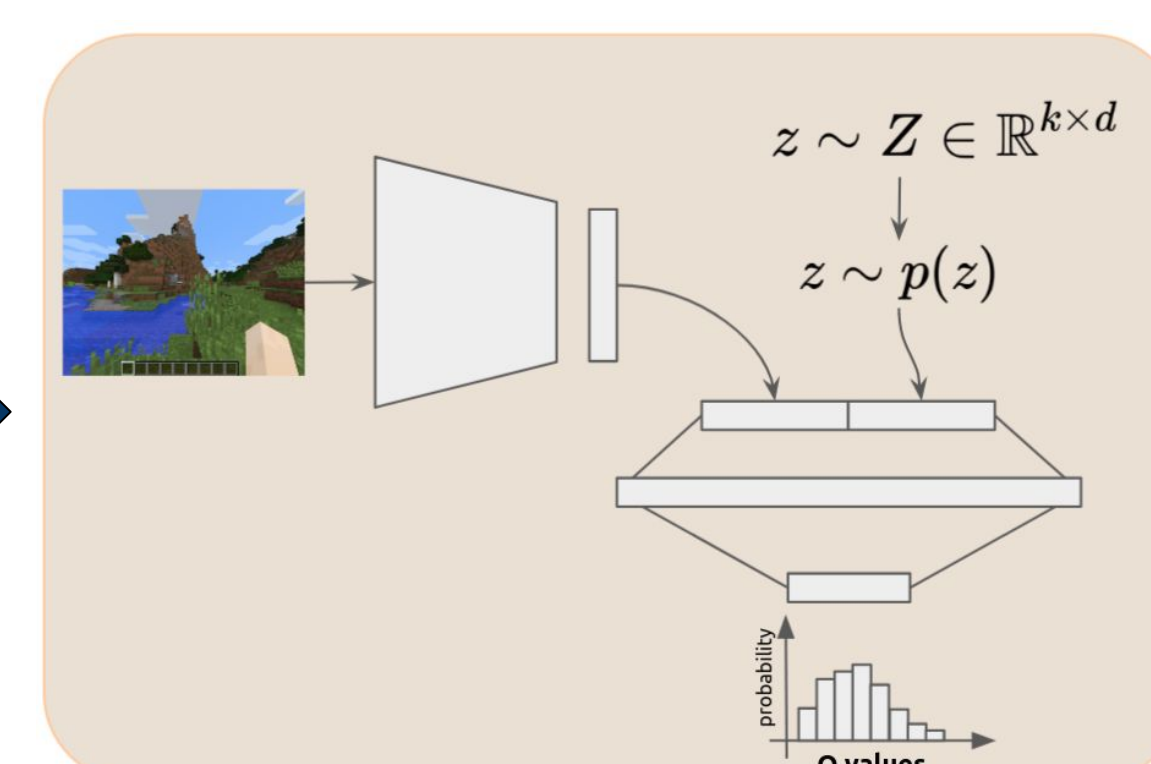- Add agent's **relative position** respect to a initial point.

EXPLORE   **DISCOVER**   **LEARN**

VARIATIONAL

GENERATE RANDOM TRAJECTORIES

$z \sim Z \in \mathbb{R}^{k \times d}$

(x, y, z)

COMMIT LOSS + MSE_pixels + MSE_coords

$z \sim Z \in \mathbb{R}^{k \times d}$

$z \sim p(z)$

probability

Q values

Results

MAP	EXPERT PLAYS PIXELS	RANDOM PLAYS PIXELS	RANDOM PLAYS COORDINATES	RANDOM PLAYS PIXELS & COORDINATES

SKILL DISCOVERY

SKILL LEARNING

(one skill example)

$r(s, z = z_3)$

$t_x(z = z_3)$

$r(z = z_3)$

[1] Campos, Víctor, et al. "Explore, discover and learn: Unsupervised discovery of state-covering skills." International Conference on Machine Learning. PMLR, 2020.

https://imatge-upc.github.io/PiCoEDL/