EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

machine learning
new perspectives
for science

Agent-centric learning

from external reward maximization to internal knowledge curation

Hanqi Zhou, Fryderyk Mantiuk, David Nagy, Charley Wu

hanqi.zhou@uni-tuebingen.de

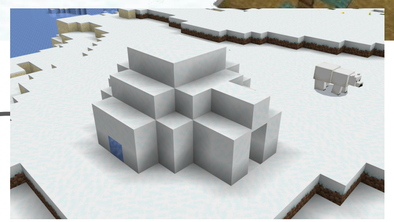
University of Tübingen

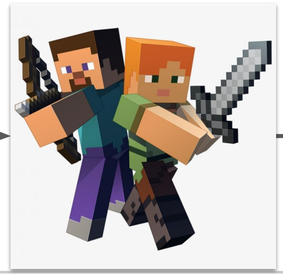
Expert for one task

observation/
state

reward

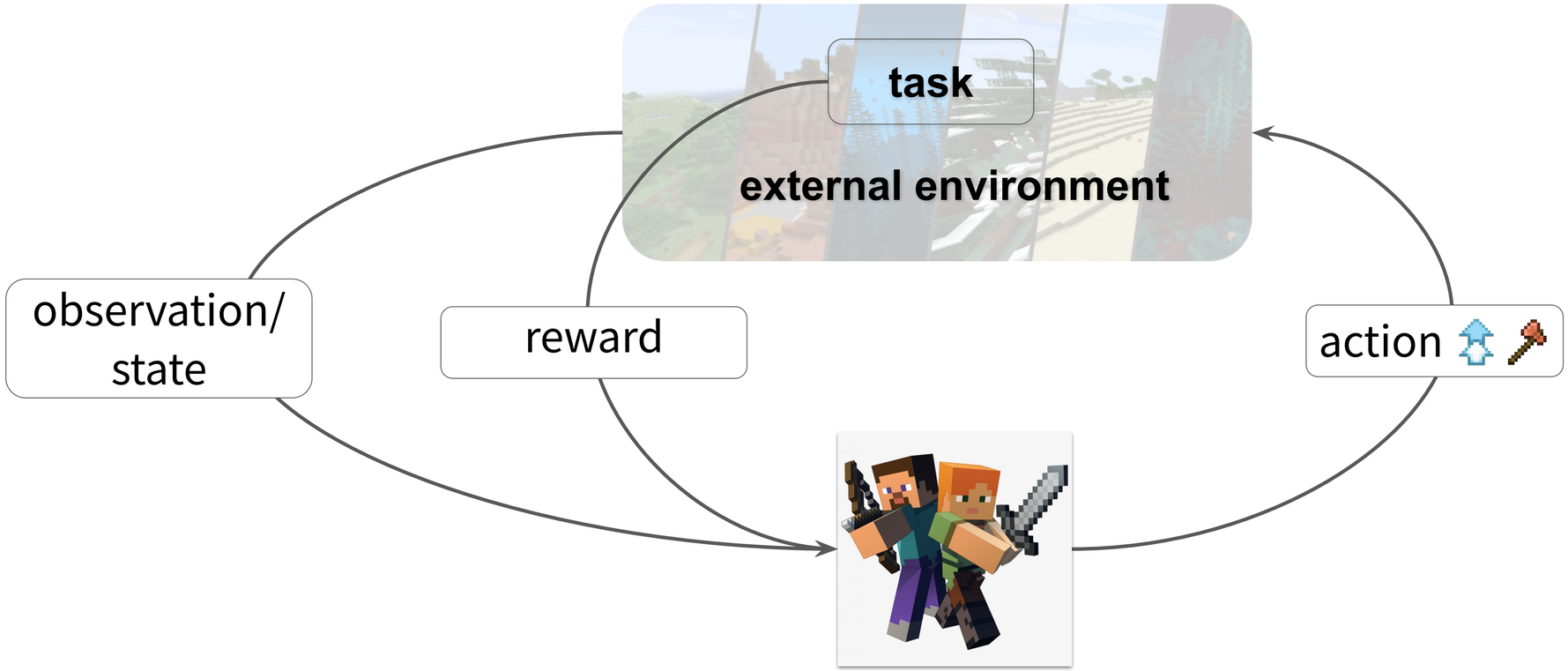
action





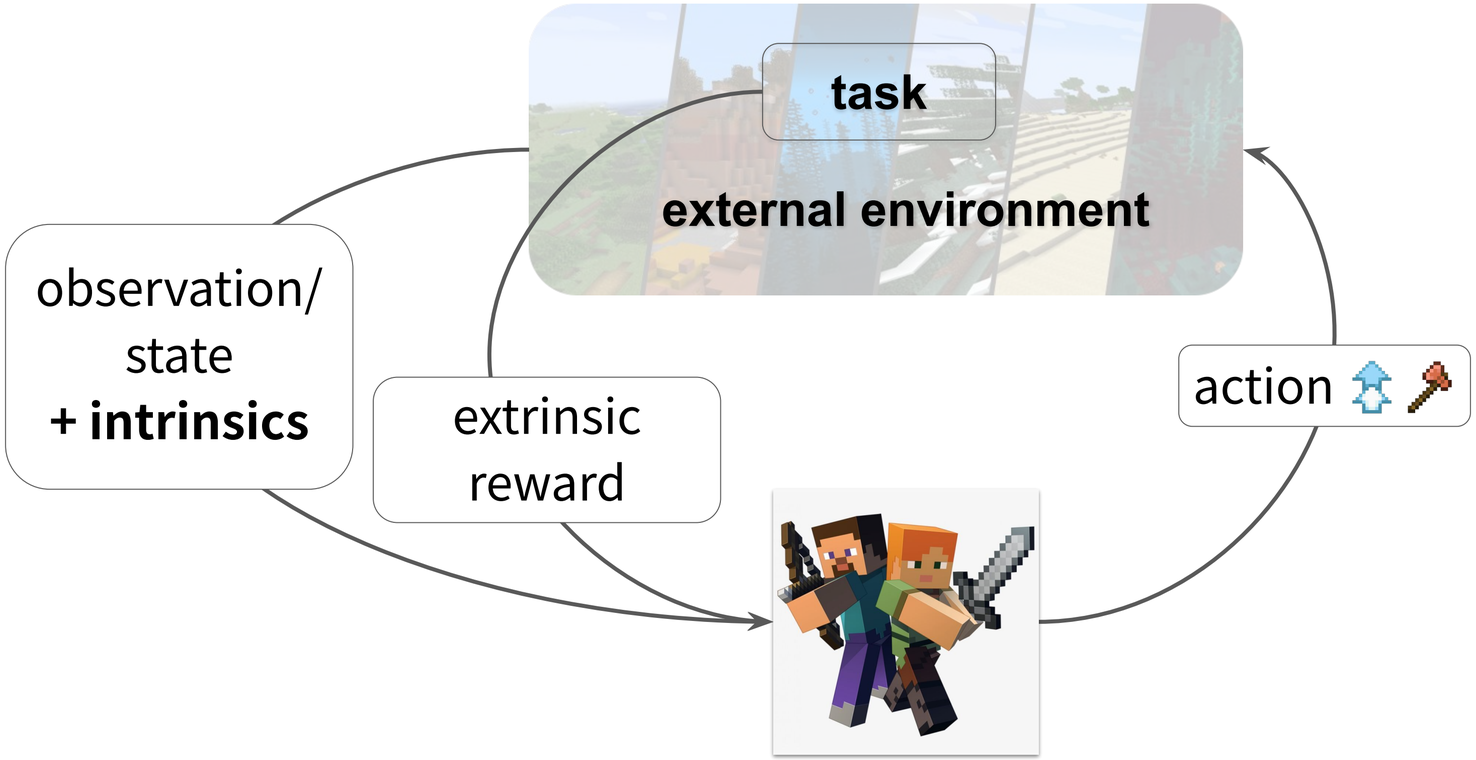
Expert for one task

Narrow tasks + fixed rewards → brittle agents that know what to do only when we tell them.



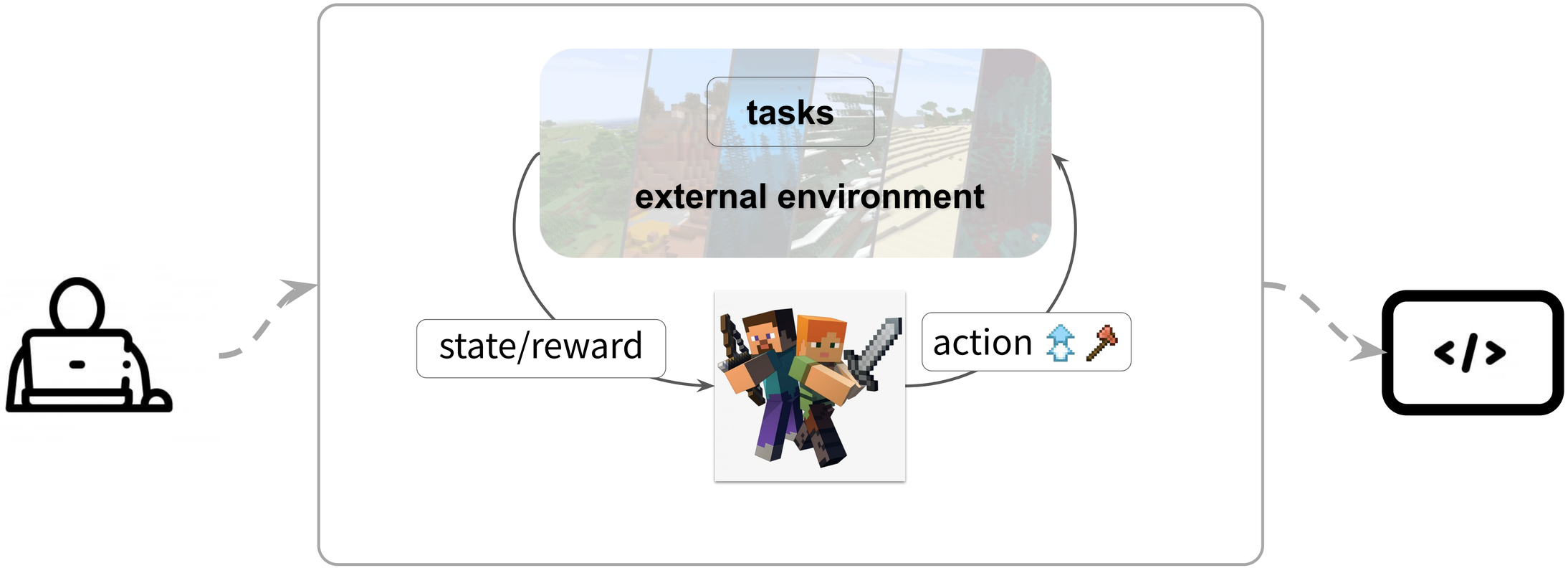
Even for intrinsic rewards

Curiosity (novelty, information gain, etc.) still ties learning to this environment, so agents overfit the world they were born in.



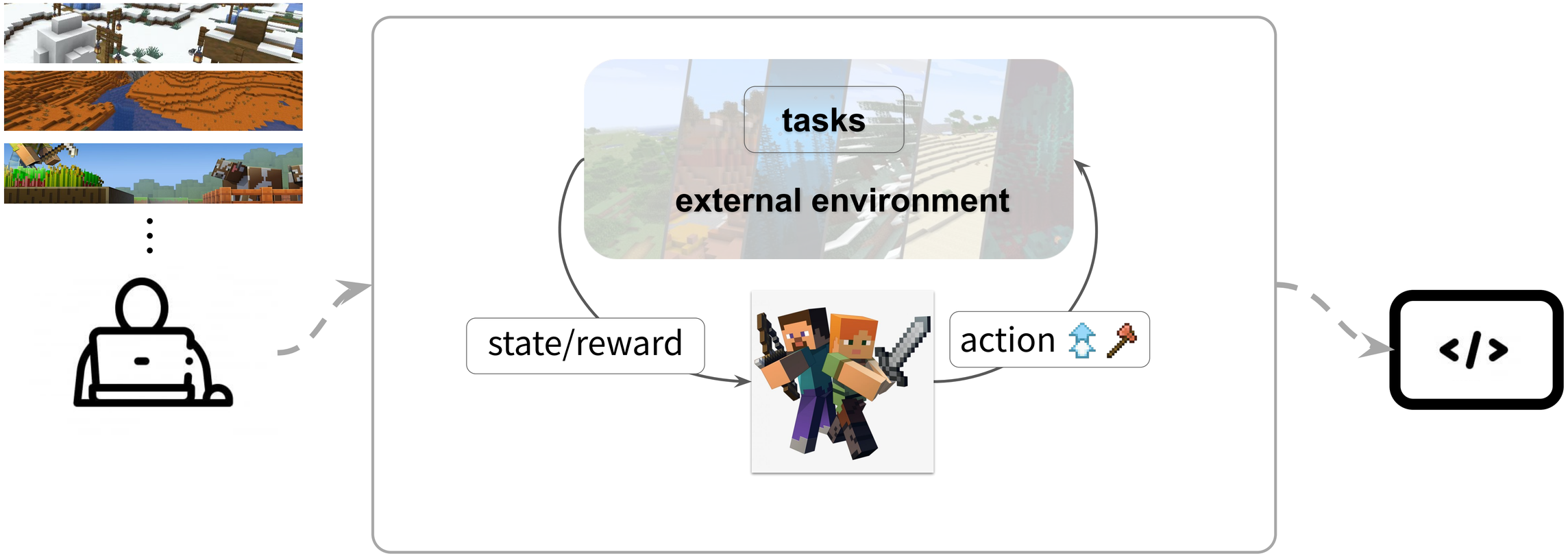
Spotlights on tasks & environments

The outward-facing view: we define a narrow task/env distribution and the agent over-specializes.

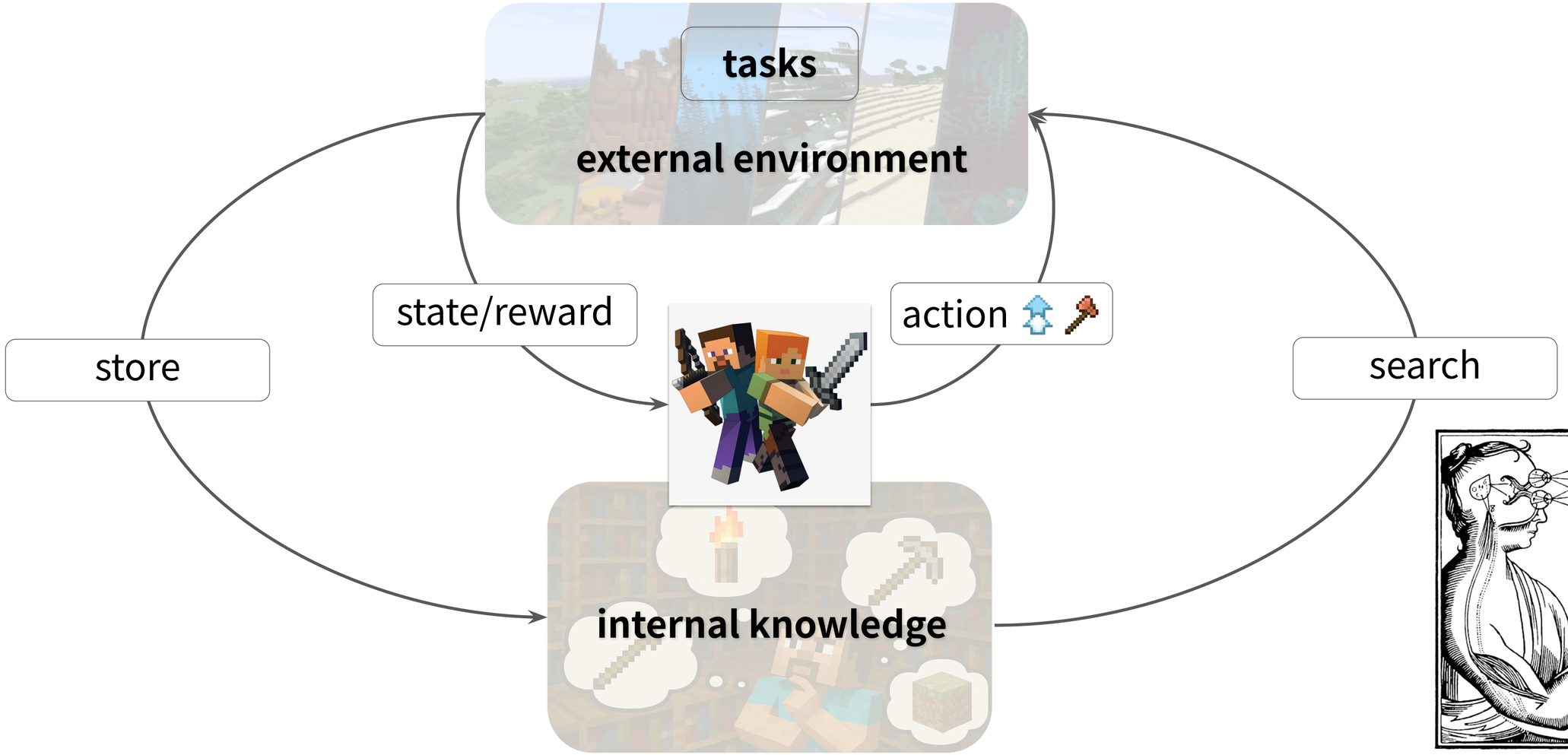


Abel, D., Ho, M. K., & Harutyunyan, A. Three Dogmas of Reinforcement Learning. In Reinforcement Learning Conference.

How to define & prepare an agent for unknown tasks?

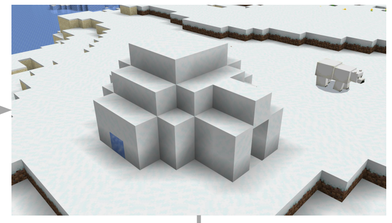


How to define & prepare an agent for unknown tasks?

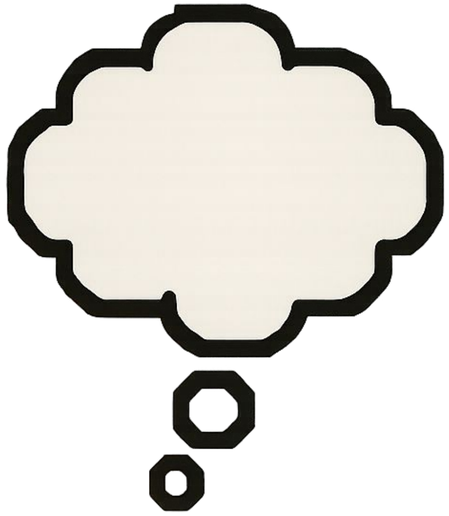


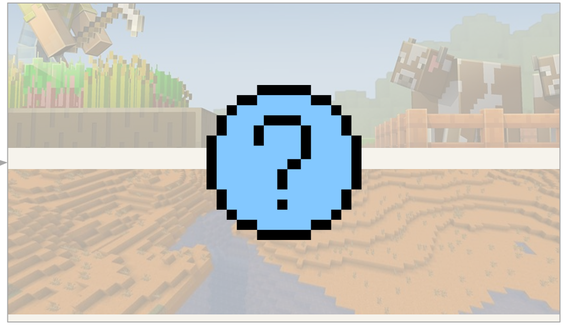
Harutyunyan, A. (2020). What is an agent. Preprint.
Descartes's illustration of dualism. Inputs are passed on by the sensory organs to the brain and from there to the immaterial spirit.

An agent's power comes from its internal world



What makes a knowledge good?





What should the agent learn?

Instrumental convergence: some goals are useful no matter what you want to do.

"All computation and physical action requires the **physical resources of space, time, matter, and free energy.** Almost any goal can be better accomplished by having more of these resources."

Omohundro, S. M. (2018). The basic AI drives. In Artificial intelligence safety and security (pp. 47-55). Chapman and Hall/CRC.
Turner, A. M., Smith, L., Shah, R., Critch, A., & Tadepalli, P. (2019). Optimal policies tend to seek power. arXiv preprint arXiv:1912.01683.
https://www.alignmentforum.org/w/instrumental-convergence

What should the agent learn?

Empowerment

- Measures an agent's control over its future
- It’s about keeping your options open in the world.

$$\mathrm{EnvEmp}(s) = \max_{\pi(a_{1:T})} I\left(s'; a_{1:T} \mid s\right)$$

$H(s' \mid s)$ diversity

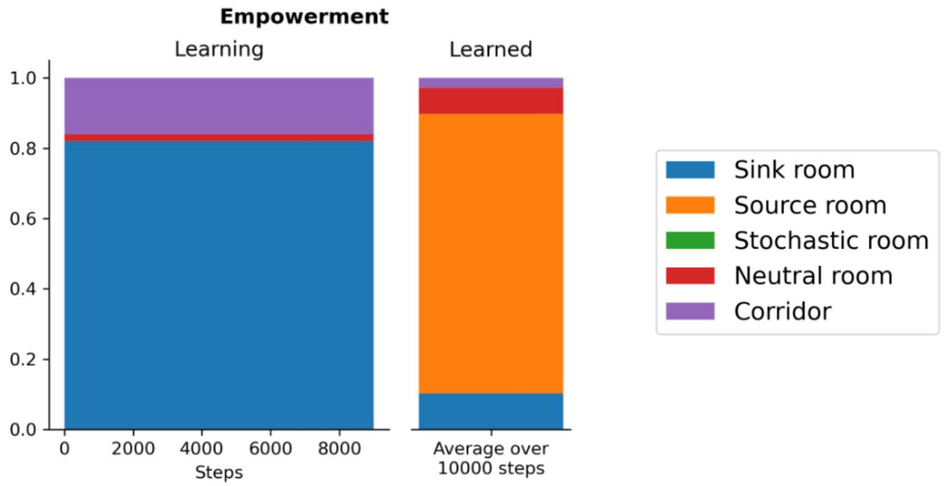
$-\,H(s' \mid s, a_{1:T})$ controllability

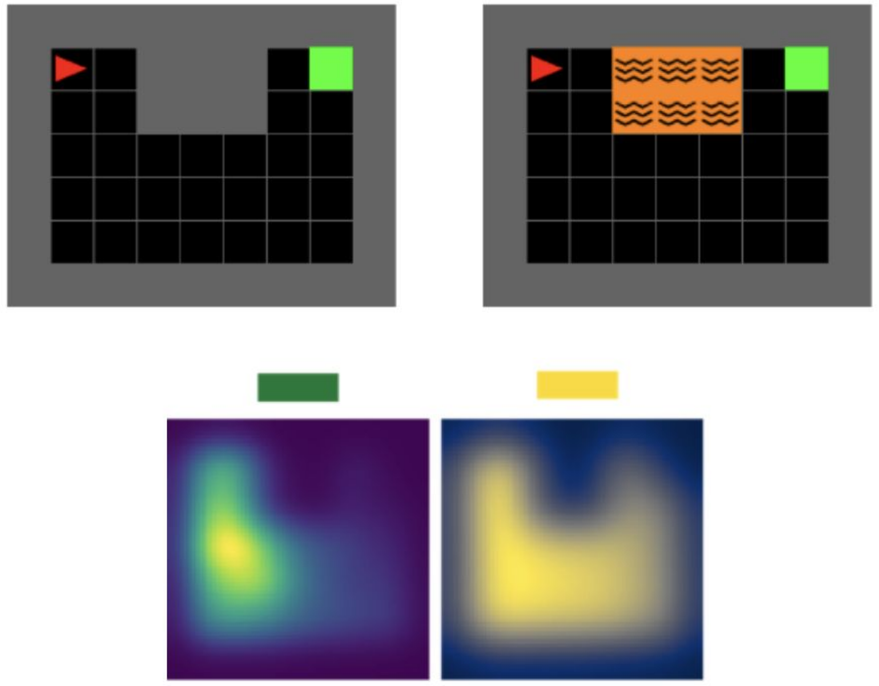
Salge, C., Glackin, C., & Polani, D. (2014). Empowerment–an introduction. In Guided Self-Organization: Inception (pp. 67-114). Berlin, Heidelberg: Springer Berlin Heidelberg.

What should the agent learn?

Empowerment

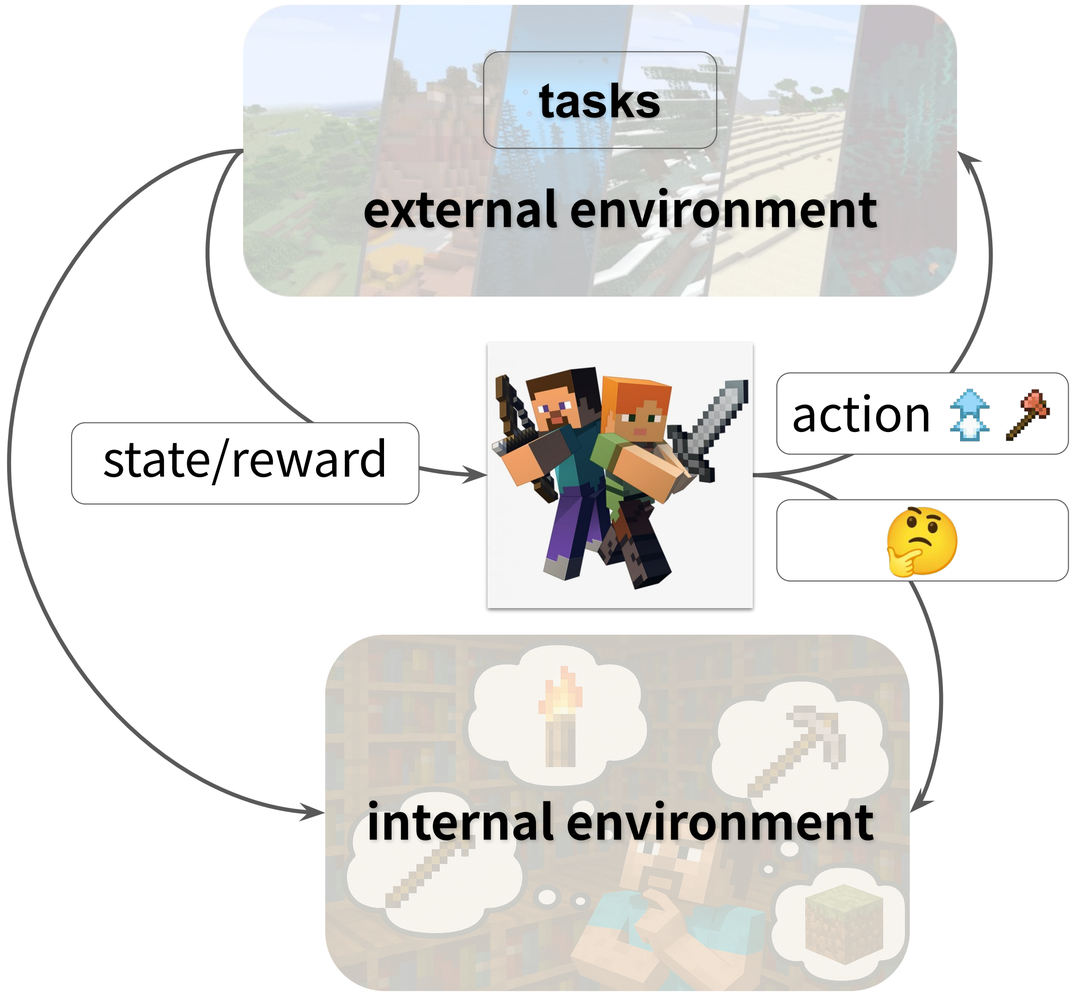
- Measures an agent's control over its future
- It’s about keeping your options open in the world.
- **What should be controlled over?**





Gruaz, L., Modirshanechi, A., Becker, S., & Brea, J. (2024). Merits of curiosity: a simulation study. PsyArXiv (to appear in Open Mind).
Mantiuk, F., Zhou, H., & Wu, C. M. (2025). From Curiosity to Competence: How World Models Interact with the Dynamics of Exploration.

Library curation for maximal empowerment

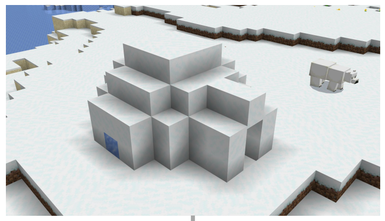


$$\mathrm{RepEmp}\left(Z_k\right) = \max_{\boldsymbol{\omega}_k \in \Omega^T} I\left(Z_k'; \boldsymbol{\omega}_k \mid Z_k\right)$$

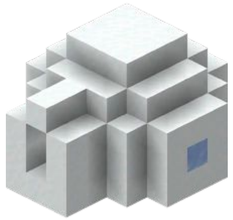
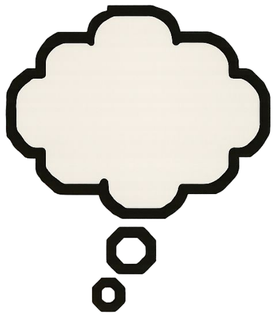
$H\left(Z_k' \mid Z_k\right)$ diverse candidates as starting point

$-H\left(Z_k' \mid Z_k, \omega_k\right)$ regularization on overly flexible knowledge

Library curation for maximal empowerment



What makes a knowledge good?


$$\mathrm{RepEmp}\left(Z_k\right) = \max_{\boldsymbol{\omega}_k \in \Omega^T} I\left(Z_k'; \boldsymbol{\omega}_k \mid Z_k\right)$$

$H\left(Z_k' \mid Z_k\right)$ diverse candidates as starting point

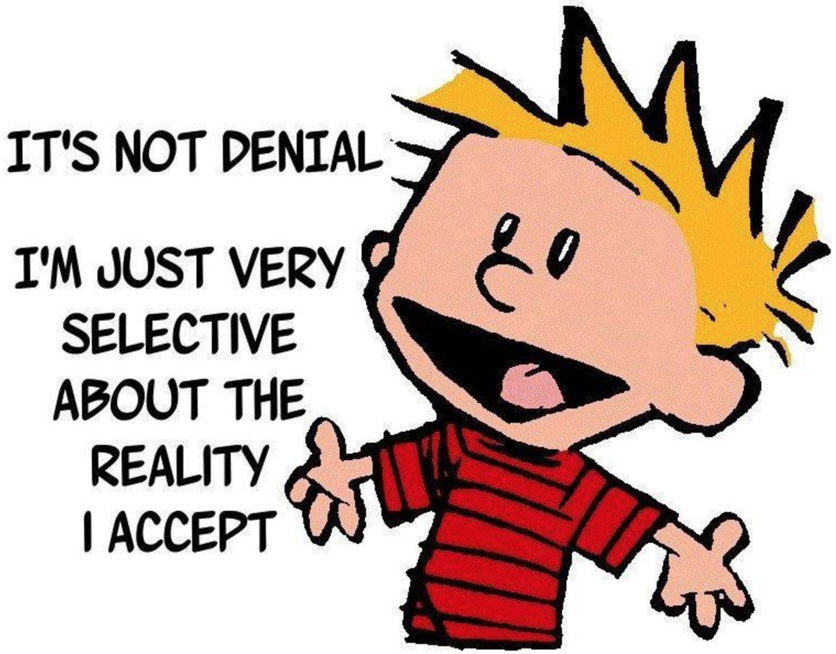
$-H\left(Z_k' \mid Z_k, \omega_k\right)$ regularization on overly flexible knowledge

Diversity: Can I combine what I know in many new ways to produce a rich variety of useful ideas?

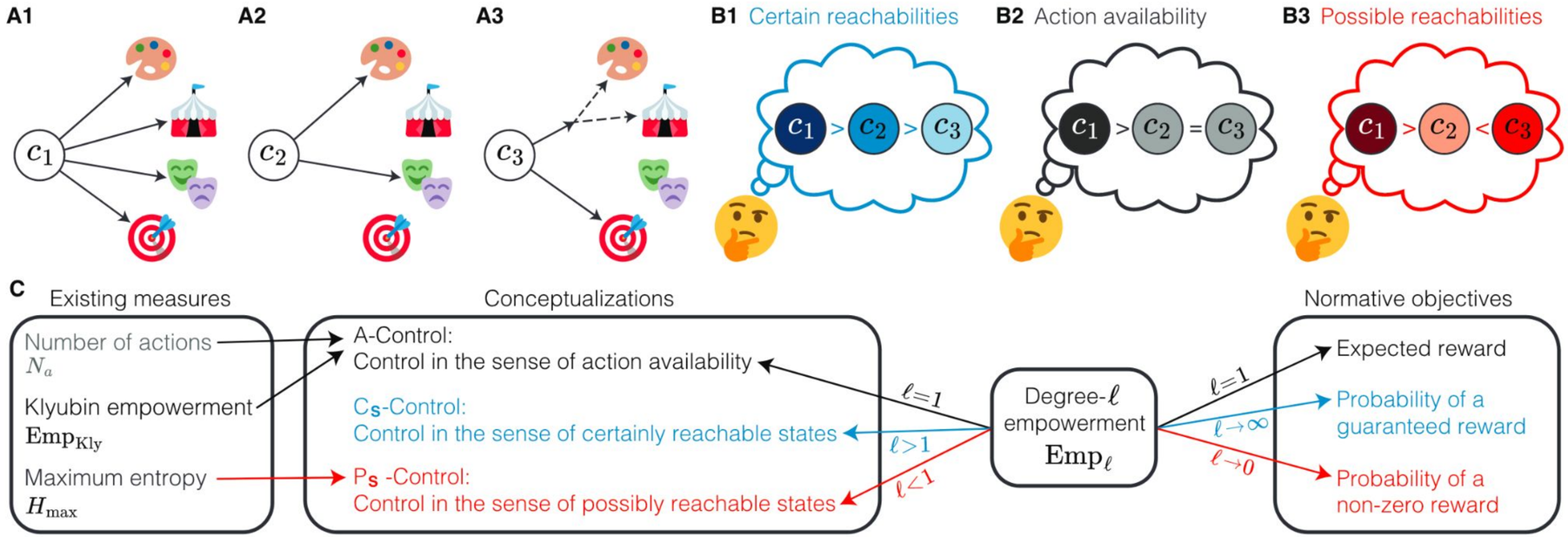
Controllability: Can I reliably shape my knowledge to create a specific new idea when I need it?

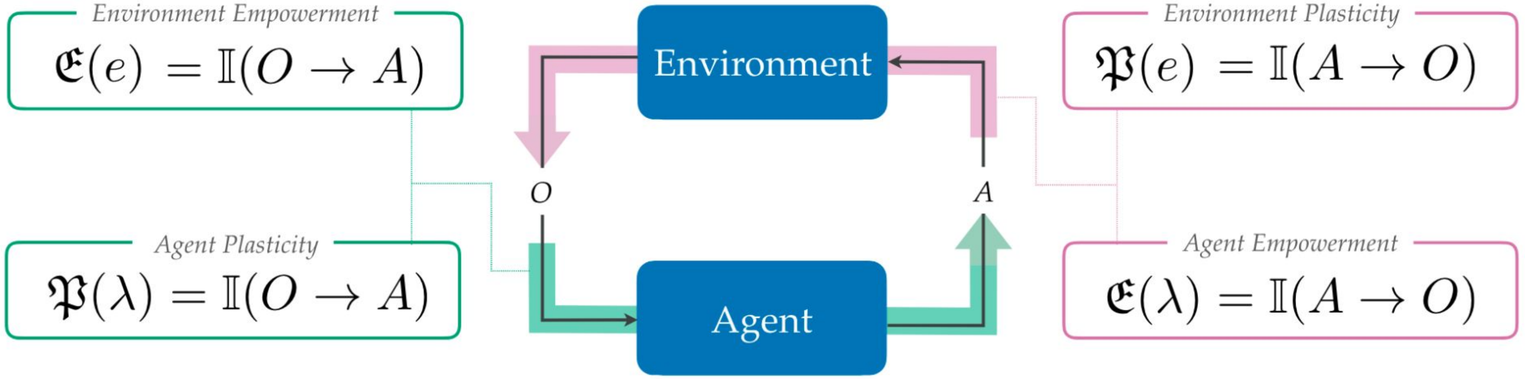
Final remarks

Why agent-centric & empowerment



Final remarks





Abel, D., Bowling, M., Barreto, A., Dabney, W., Dong, S., Hansen, S., ... & Singh, S. (2025). Plasticity as the Mirror of Empowerment. arXiv preprint arXiv:2505.10361.
Modirshanechi, A., Dayan, P., & Schulz, E. (2025). An integrative framework for the human sense of control.