



Agent-centric learning

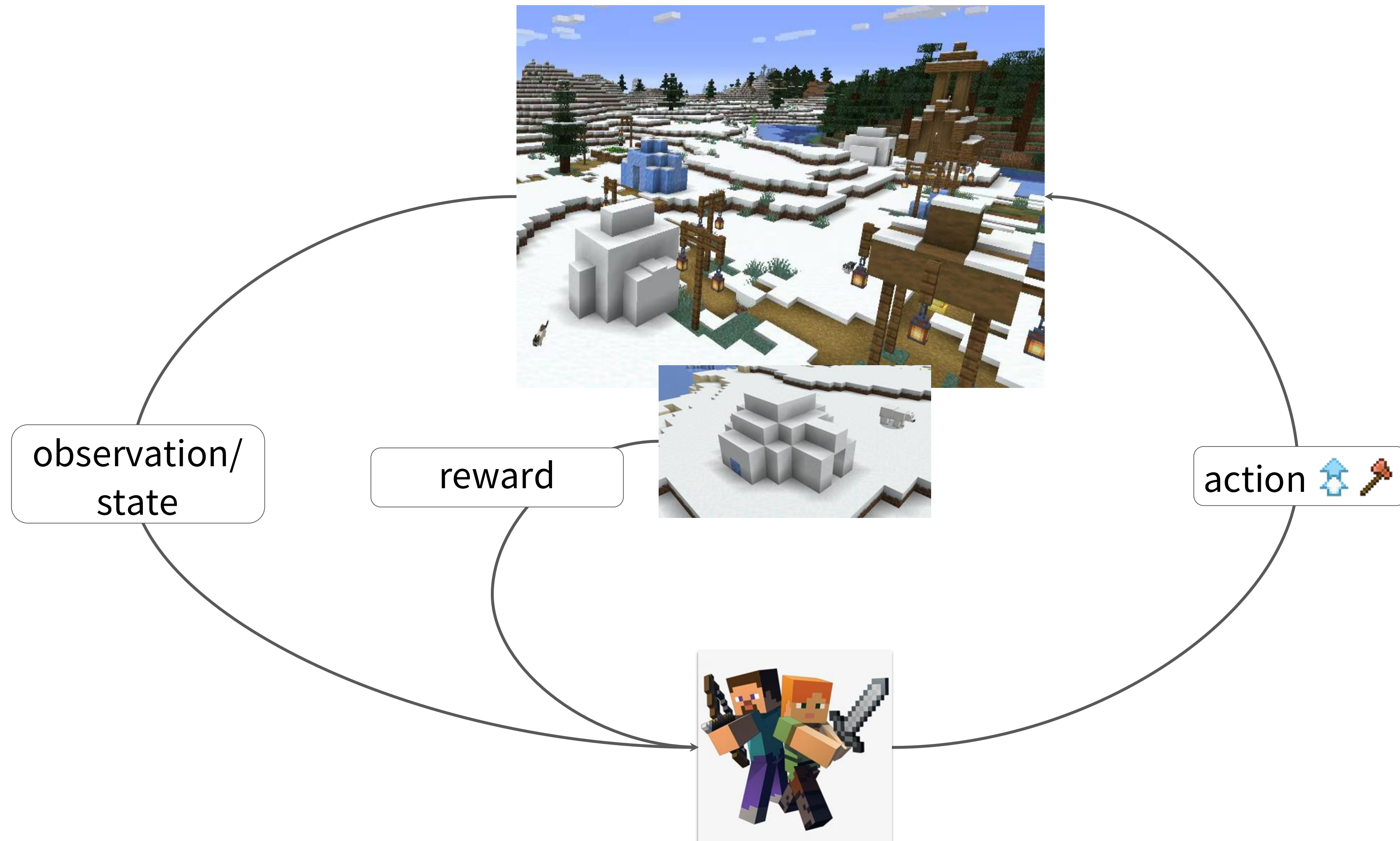
from external reward maximization to internal knowledge curation

Hanqi Zhou, Fryderyk Mantiuk, David Nagy, Charley Wu

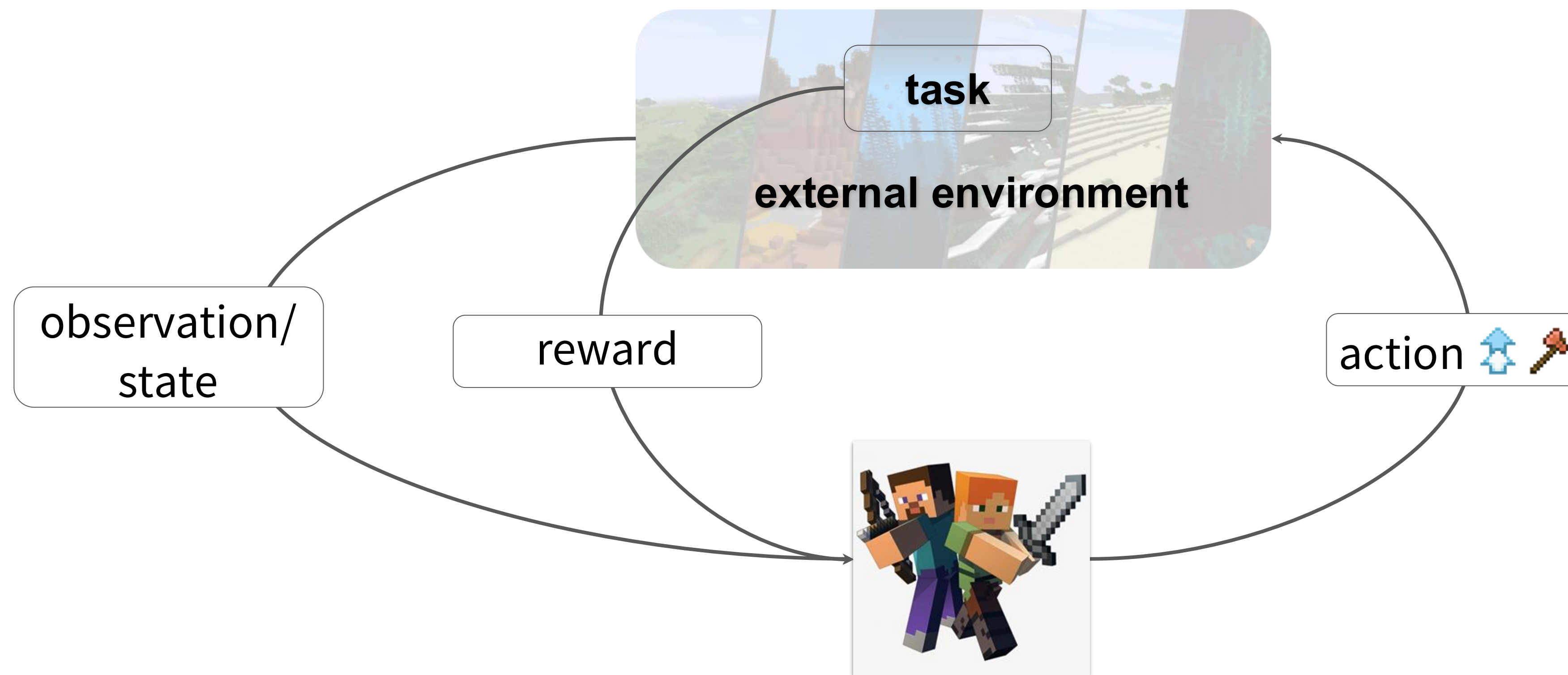
hanqi.zhou@uni-tuebingen.de

University of Tübingen

Expert for one task

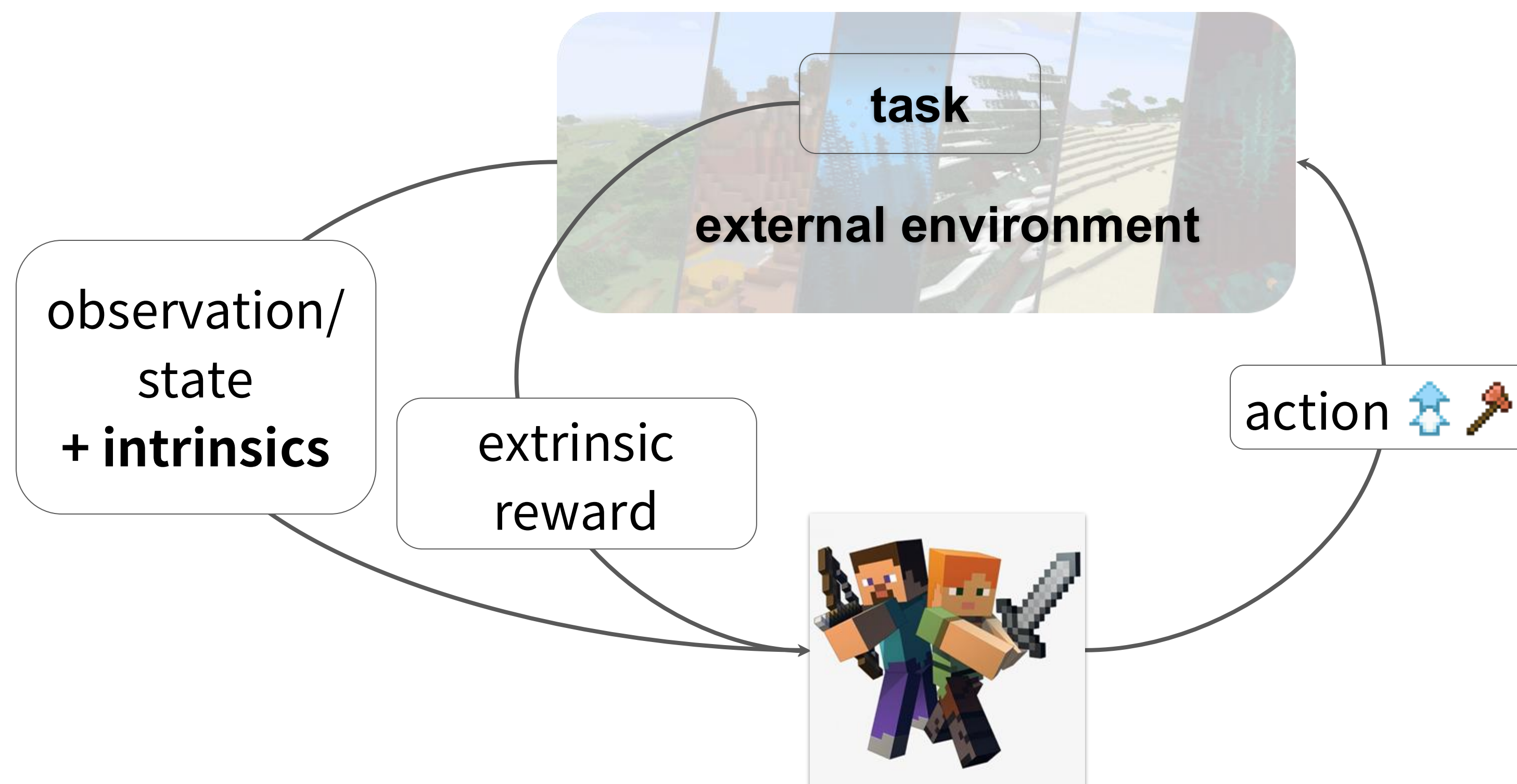


Narrow tasks + fixed rewards → brittle agents that know what to do only when we tell them.





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





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

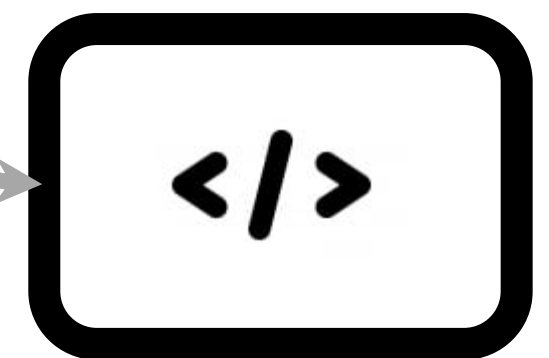
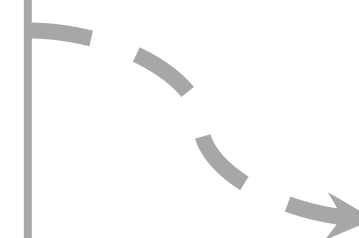
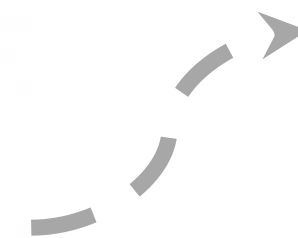


Abel, D., Ho, M. K., & Harutyunyan, A. Three Dogmas of Reinforcement Learning. In Reinforcement Learning Conference.

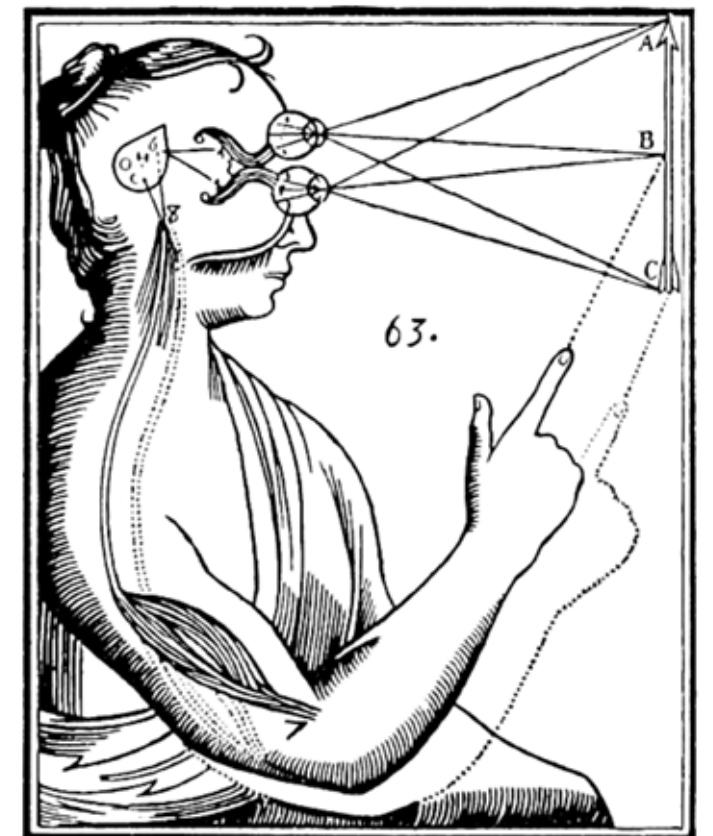
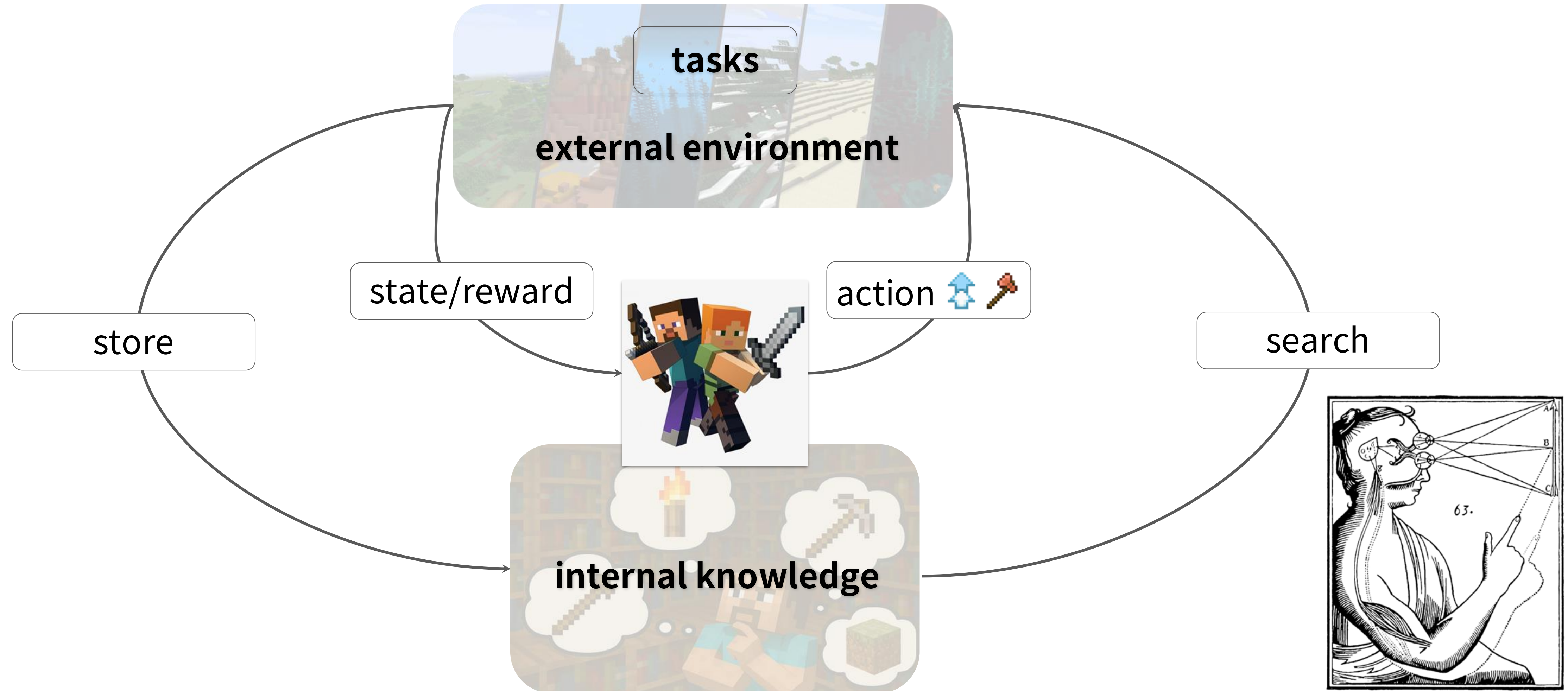
How to define & prepare an agent for unknown tasks?



⋮

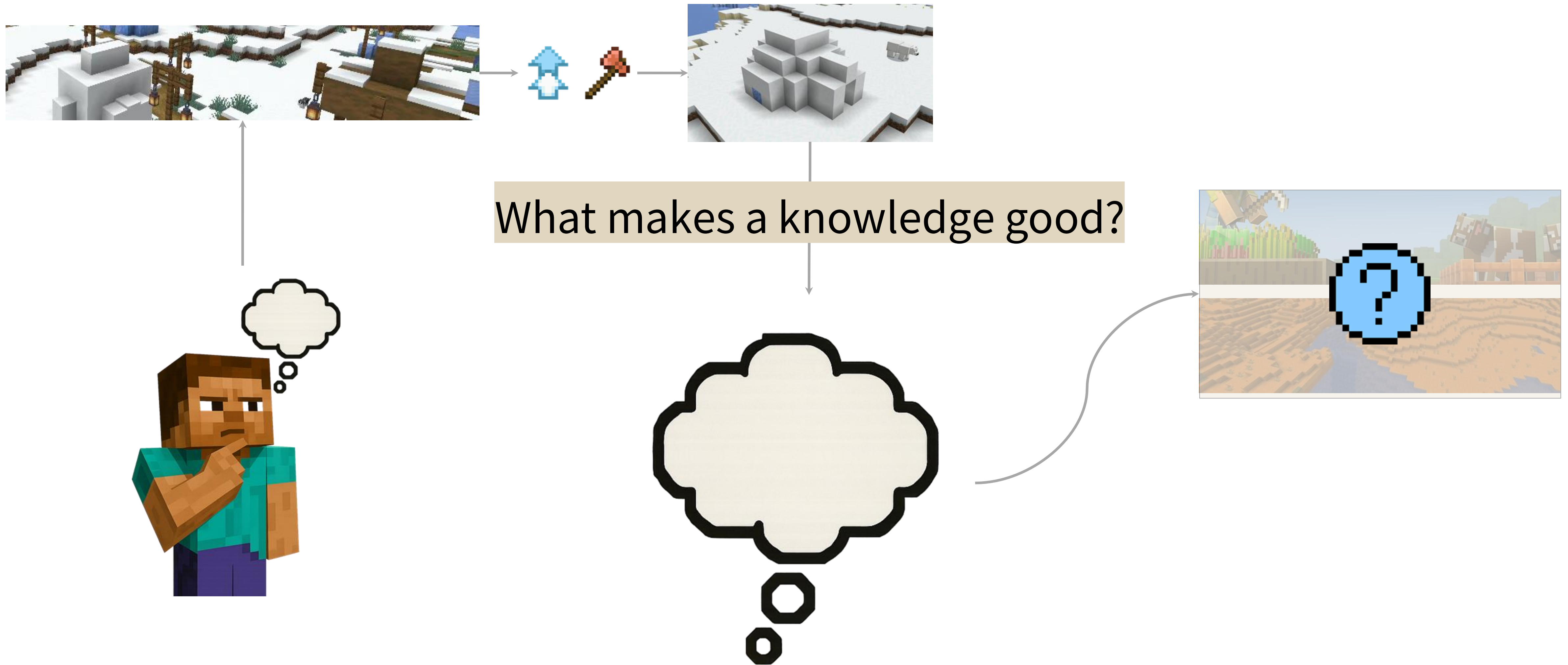


How to define & prepare an agent for unknown tasks?



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



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>

Empowerment

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

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

$H(s' \mid s)$ diversity

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



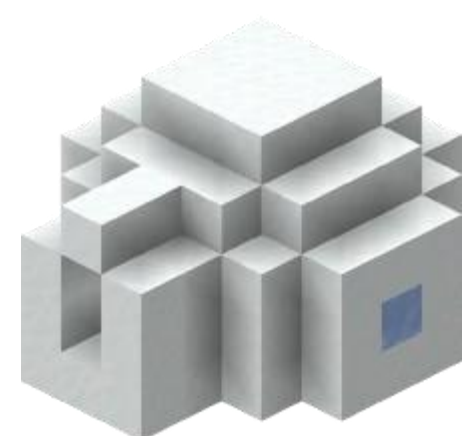


$$\text{RepEmp}(Z_k) = \max_{\omega_k \in \Omega^T} I(Z'_k; \omega_k \mid Z_k)$$

$H(Z'_k \mid Z_k)$ diverse candidates
as starting point

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

What makes a knowledge good?

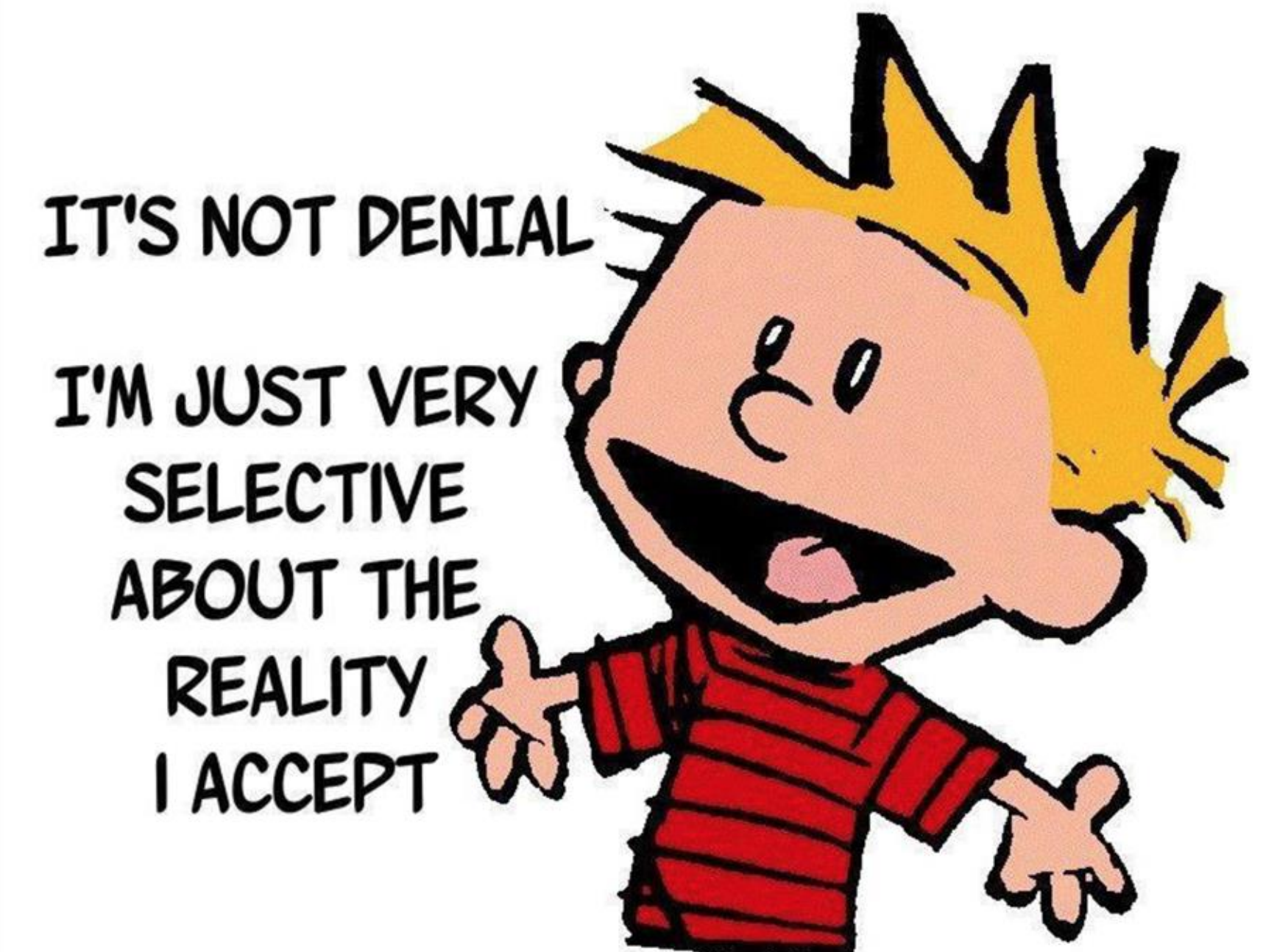


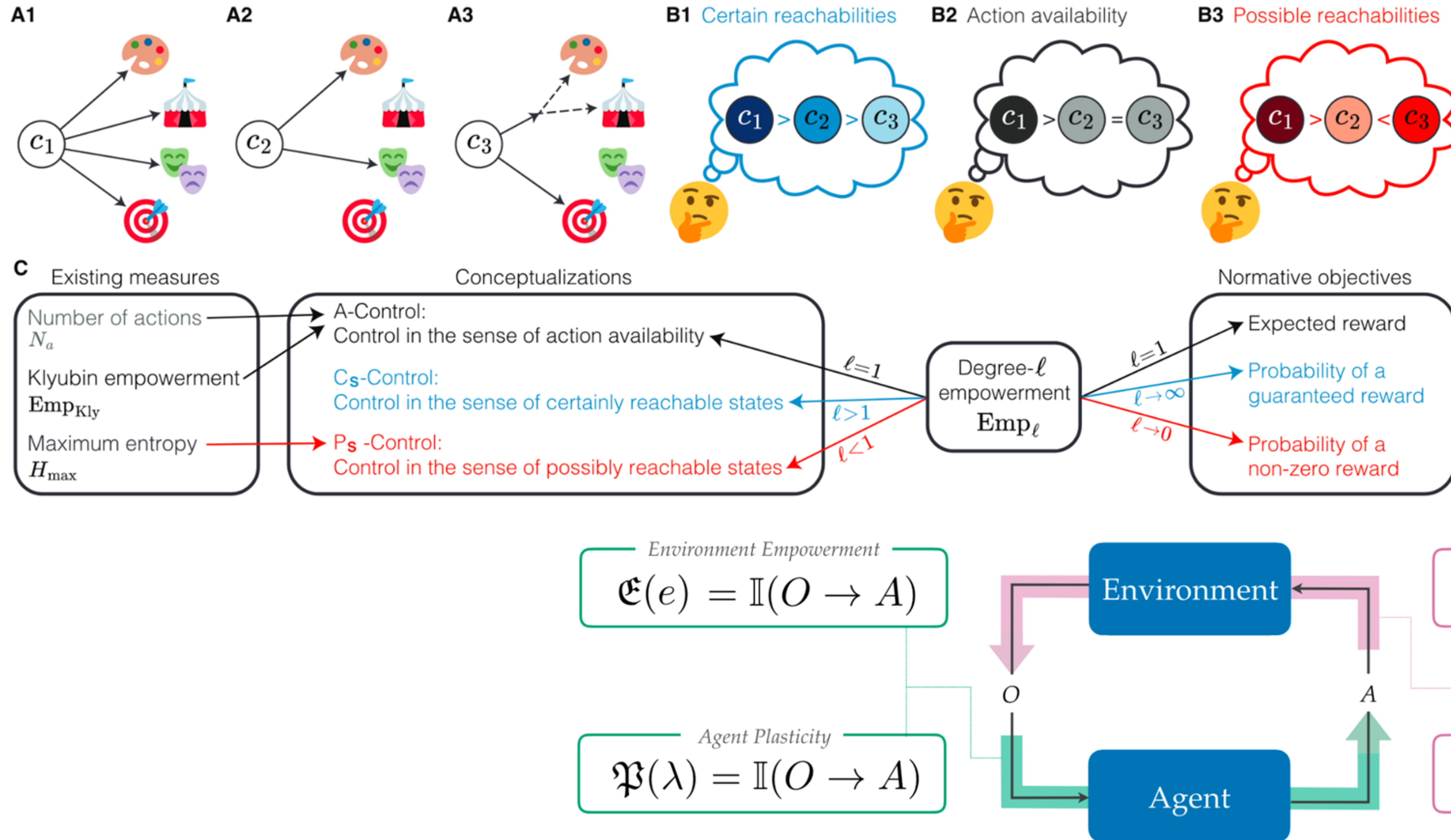
Diversity: Can I combine what I know in many new ways to produce a rich variety of useful ideas?

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



Why agent-centric & empowerment





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.