Human-like Artificial Agency Needs Curiosity, Compression, and Resource Constraints

Richard Csaky Barcelona Computational Foundation | Foresight Institute
richard.csaky@gmail.com | richardcsaky.notion.site/main | arxiv.org/abs/2602.24100

Thesis

Agents are **embedded systems** in an environment: they choose observations, spend compute, and act through limited channels.

observe act think communicate

Reward **predictive compression gain**, penalize **resource cost**.

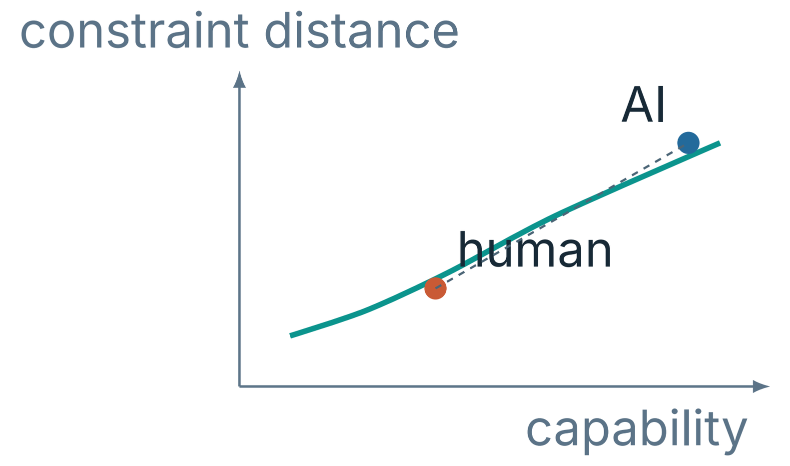
Agent loop

$$O_t \sim p_O(\cdot \mid X_t;\, c_t^O),$$
$$S_t \sim p_S(\cdot \mid S_{t-1}, O_t, U_{t-1}, V_{t-1};\, c_t^C),$$
$$(U_t, V_t) \sim \pi(\cdot \mid S_t;\, c_t^A),$$
$$X_{t+1} \sim p_X(\cdot \mid X_t, U_t),$$

- p_S updates recurrent state, memory, and optional online learning.
- The policy selects **external actions** and **internal computation**.
- Budgets attach to every channel: observation, action, compute, memory, and communication.

Human frames

Measure both competence and **distance from human constraints**: sensing, actuation, memory, time, compute, tools, and social coupling. This distance characterizes the **alignment gap**.



capability $\neq$ human-likeness

Resource objective

$$J = \mathbb{E}\sum_{t=1}^{T} \gamma^{t-1}\Big(r_t - \lambda_O C_O - \lambda_E C_E - \lambda_C C_C - \lambda_M C_M\Big)$$

C_O observe C_E act C_C compute C_M memory

The comparison target is **return per unit cost**. Price sensor bandwidth, action authority, private tokens, updates, and memory writes. Compare agents at matched budgets.

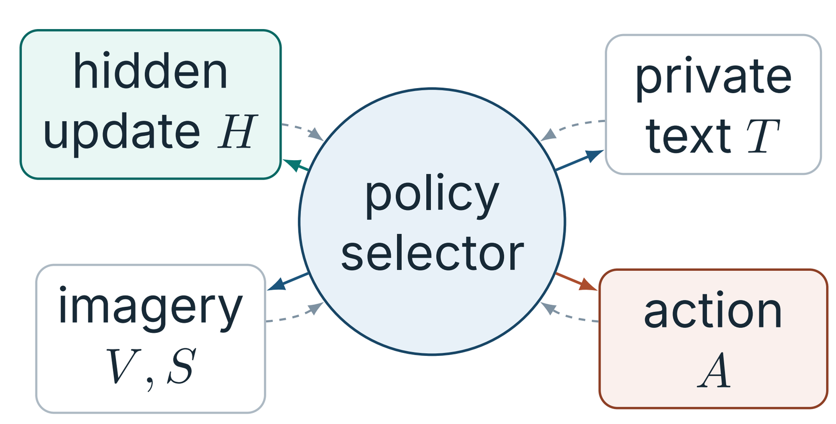
Curiosity as learning progress

$$r_t = \mathcal{L}_{\text{pred}}(\theta_t^-;\, \mathcal{D}_{t:t+K}) - \mathcal{L}_{\text{pred}}(\theta_t^+;\, \mathcal{D}_{t:t+K})$$

- Reward prediction improvement on **future observations** not used for the update.
- Avoids both random noise and already-trivial patterns.

Flexible deliberation

The agent can pause before, during, or after output and choose the **cheapest useful reasoning operation**.



silent hidden updates optional self-talk
interleaved action tokens multimodal imagery

Interface quality

$$\mathcal{U}_t = \sum_{q\in\{O,A,L\}} w_q \left(1 - \frac{\mathcal{R}_q}{\mathcal{R}_q^{\max}}\right)$$

- **Observation:** losslessness, latency, noise, bandwidth.
- **Action:** authority, precision, reversibility, energy.
- **Language:** public and private channels as priced actions.

Hypotheses

H1	Predictive-compression progress tracks useful control under flexible bottlenecks.
H2	Interface upgrades continue only while marginal gains exceed marginal costs.
H3	Costly interaction favors compact predictive states and selective control.
H4	Adaptive compute beats fixed schedules at matched budget.
H5	Private communication helps long-horizon tasks beyond latent recurrence.

Takeaway

Intelligence should be evaluated as embedded agency: an agent must observe and act through limited channels, communicate across bottlenecks, and allocate finite resources over finite time.

For human-AI systems, the key problem is coupling: building machines that expand collective agency.

The proposed framework treats prediction, control, communication, memory, and cost as one coupled object of study.