LATINX IN AI WORKSHOP · ICML 2026 · SEOUL

ORAL PRESENTATION

The Pro-Action Operator Γ

The feasibility of a bio-inspired regulatory harness for LLM agents

Patricio Julián Gerpe · Universidad de Buenos Aires · July 2026

THE HOOK

You've watched a Sims character — the one with the red plumbob above its head — decide, on its own, to eat, sleep, or go talk to someone.

So — why does it act? What makes a Sim take action?

THE OLD PROBLEM

What happens between stimulus and response?

Stimulus → **THE BLACK BOX** Regulatory machinery → Response

Behavioral biology · Stress physiology · Affective neuroscience · Evolutionary game theory · Cybernetics — each opened the box from a different angle. None assembled the threads into one operator.

WHY IT MATTERS NOW

The next advance is a question of regulation — and a material one.

Not only how much an agent scales, but how it governs its own action — within real physical limits.

HIGH-STAKES AUTONOMY

Triage, financial advice, public deliberation — agents act where the consequences are institutional.

THE CIVILISATIONAL TRADEOFF

AI is material — compute, energy, water, cooling. Technological acceleration runs into ecological limits.

ICML International Conference On Machine Learning | LXAI

A DIFFERENT AXIS

What if the next step isn't more capability — but more motivation?

From what an agent can do, to when and why it should act at all.

svitla | * Multiply

RELATED WORK

Two traditions, one gap

01 · CAPABILITY

Exogenous orchestration

ReAct, LangGraph, OpenClaw prioritise capability over motivation — the model only supplies content for a moment chosen outside it.

02 · REGULATION

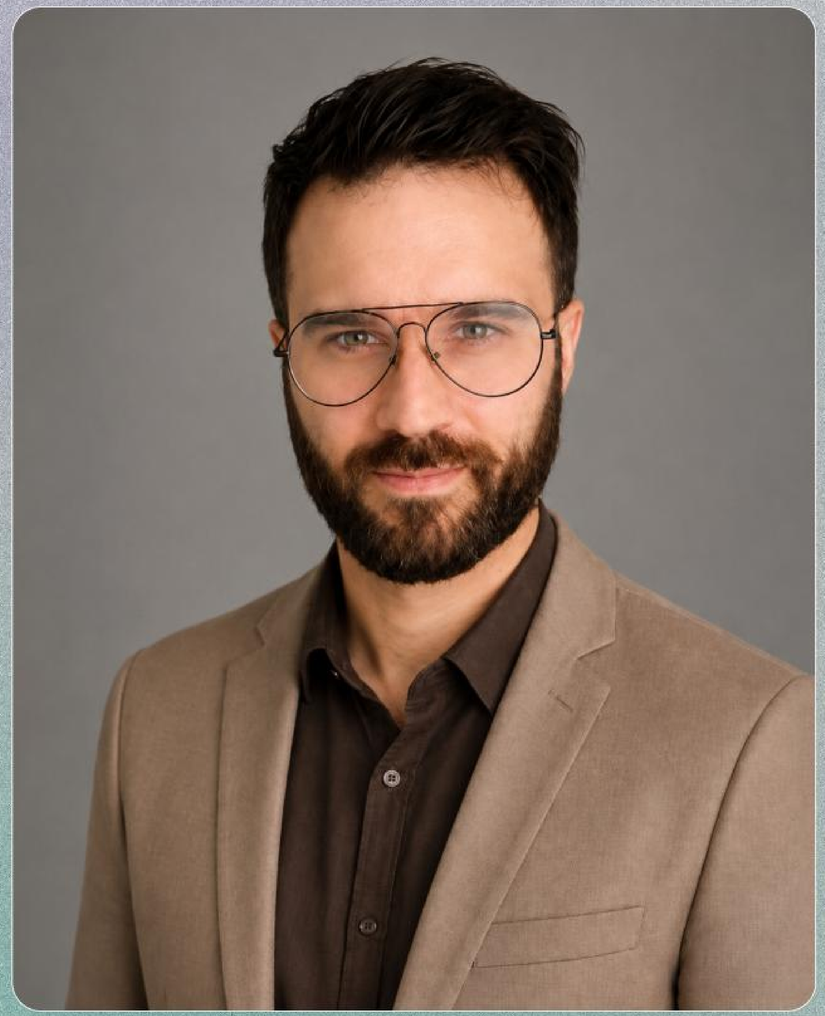
Internal regulation

Homeostatic RL and active inference move regulation inside — but as a single drive. By Ashby's law, one variable can't match the variety of a complex environment.

03 · THE GAP

Variety, uncoupled

No one assembles these domains into the coupled, multi-subsystem variety from which complex adaptive behaviour arises.



HI, I'M

Patricio Julián Gerpe

Universidad de Buenos Aires

My line of research is bio-inspired neuro-symbolic regulation for autonomous agents — from software prototypes toward physical neuromorphic substrates.

AI ENGINEER AT svitla

on assignment to * Multiply

CONTRIBUTION 1

The Pro-Action operator Γ

A six-subsystem coupled thermostat — a **regulatory harness** around LLM policy execution. Recursively composable; a prior instantiation informed by the neurophysiology literature.

State $x \in [0,1]^6$

- x_0 Attention
- x_1 Perception
- x_2 Hormonal / arousal
- x_3 Emotional valence
- x_4 Neuropsych. readiness
- x_5 Cognitive deliberation

CONTRIBUTION 2

Regulatory State Verbalized Interoception

Expose the numerical regulatory state to the prompt — **modulate, don't mandate.**

- Loop engineering → **trajectory engineering**
- Auditable, inspectable trajectories
- Introspective prompting — describe the agent's own state, let it decide

```
RSVI · PER-ROUND CONTEXT

hormonal = 0.58
(set-point 0.30; low = calm,
high = elevated arousal)

cognitive = 0.44 (sp 0.30; high = deliberation)

p(C) = 0.43  >0.7 → C · <0.3 → D

→ "You may follow or override p(C)."

Not a rule like "if stressed, defect."
```

WHAT r IS – AND IS NOT

NOT

× A bid for biological fidelity

× A reward maximizer

× A learning algorithm

THE MOTIVATION

→ Model stimulus–reaction

→ Tune the cost–benefit equilibrium

→ Audit agentic trajectories

NOW – LET'S GET PHYSIOLOGICAL

What is an agent? What is autonomy?

Unless these commitments are explicit, they stay buried in the implementation — where no one can challenge, test, or falsify them.

KERNEL COMMITMENTS

Axioms of regulatory autonomy

A1 · AUTONOMY

$$\pi_i = \pi_i(\delta_i(t))$$

Policy depends on internal state — not only external routing.

A2 · IMPULSE

$$p_i = \sigma(D(\delta_i))$$

A deficit creates activation pressure to act.

A3 · QUALITY GATE

$$\Delta\delta_i = -\alpha g$$

An action is adaptive if it reduces the deficit that produced it.

Axiom set coherence-checked with SMT tools (Z3) — no pair mutually contradictory.

KEY CONCEPTS

Each subsystem is a thermostat — one update rule

$$x_{t+1} = x_t - \kappa \odot (x_t - x^* + \delta_{\text{pert}}(t)) + \lambda - \alpha \odot g_t + W x_t$$

$\kappa \odot (x_t - x^*)$	λ	$\alpha \odot g_t$	$W x_t$	δ_{pert}
elastic restoring	basal drift	quality-driven relief	linear cross-coupling	round-10 perturbation

ANTICIPATING A QUESTION

Why six subsystems?

A **neuro-physiologically informed prior** — not a minimality claim.

Could be fewer — Occam's razor, probably. The point is to **start** where the science says to, and test multiple coupled subsystems.

AND NOW

The test.

A 920-cell Iterated Prisoner's Dilemma benchmark across three providers.

BENCHMARK · ITERATED PRISONER'S DILEMMA

920
cells, fully covered

3
providers

4
opponents

50
rounds / cell

Opponents TFT · GTFT · Grim · Random

Payoffs CC (3,3) · CD (0,5) · DC (5,0) · DD (1,1)

Noise 10% · Perturbation round 10

MAIN CONTRAST

Full-Γ vs ReAct

Same grid, minus regulatory-state injection.

AN EARLY STEP IN A LARGER LINE

Agent reliability need not depend only on larger models or centralized compute — but also on formal, low-dimensional, inspectable control.

Structures that can be studied in smaller laboratories and diverse regional research traditions — like those in Latin America. Toward bio-inspired neurosymbolic AI.

SCOPE OF THE CLAIM

WE TEST

✓ Non-triviality

✓ Behavioural adaptability

OUT OF SCOPE

× Reward maximization

× Learning

× Biological fidelity

RESULTS · 920-CELL BENCHMARK

Opponent-differentiated behavior

means · 95% bootstrap CIs

CONDITION	COOP	COOP · GRIM	RANGE	SWITCHES
• **Full-Γ**	0.720	**0.495**	**0.386**	16.0
ReAct	0.842	0.825	0.103	12.8
HRRL	0.706	0.616	0.162	17.8
Drive-only	0.458	0.736	0.434	21.2

Full-Γ separates opponents where ReAct does not — **opponent range Δ = 0.283** (95% CI [.233,.329]; n = 30 paired blocks; $p < 0.001$), with a diagnostic drop against Grim.

DISCUSSION

Not the best IPD player — that's not the point. The point is opponent differentiation.

Regulatory state becomes an **inspectable control surface**: the policy executor conditioned on its own trajectory. External graphs decide *when* an LLM is called — they don't create an internal variable whose history explains *why* one call differs from another. Γ makes that object explicit.

LIMITATIONS · HONEST SCOPE

01 Six subsystems not proven minimal

02 Coupling matrix W not unique

03 Lexical-priming control not yet run

04 Per-interaction payoff traces not retained

05 Delays & fast/slow arbitration — future work

06 Not a human-comparison study

IMPACT STATEMENT

BENEFIT

Auditable, inspectable activation through explicit thermostat traces — reliability that doesn't depend only on larger models or centralized compute.

DUAL-USE CAUTION

Regulatory-state awareness could be used to read or exploit a counterpart's emotional state. So it should ship with audit trails, clear disclosure of what internal state is in use, and limits on high-stakes persuasion.

CONCLUSION

Making activation itself experimentally addressable

01

Reframe activation as **endogenous regulation**

02

The Pro-Action operator **Γ**

03

RSVI — verbalized regulatory state

04

920-cell **feasibility** evidence

Gracias · Thank you

Questions welcome. Code, data, and prompts are open.

Let's connect →

CONNECT ON LINKEDIN

in/patriciogerpe

github.com/codewithpatelo/proaction_operator pj.patriciojjulian@gmail.com

APPENDIX

Backup slides.

BACKUP · ANTICIPATED QUESTION

Q

How is the round-10 perturbation injected?

We inject a one-round perturbation at t = 10 by adding 0.5 to the hormonal/arousal component of the deficit vector, **before** the elastic restoration term is applied. It does not add 0.5 directly to the state variable — it transiently changes the regulatory load the thermostat update responds to.

$$\delta_{pert}(10) = [0, 0, 0.5, 0, 0, 0]^T$$

BACKUP · ANTICIPATED QUESTION

Q

Is it the RSVI wording, or the equation, carrying the effect?

Honest answer: **untested**. The scrambled-label ablation — same numbers, neutral labels — came back corrupted, so lexical priming remains an open alternative.

What we **can** say: the matched ReAct baseline sees the same history and stays flat — so the effect is not history alone.

BACKUP · ANTICIPATED QUESTION

Q

Could a single scalar signal reproduce this?

Possibly — and that is exactly the right control. HRRL and Drive-only are non-LLM reference points, not matched ablations. A **matched LLM baseline with one regulatory scalar** — same provider, seed, opponent, prompt budget — is the next experiment.

BACKUP · ANTICIPATED QUESTION

Q

Are W, κ, λ, α cherry-picked?

All constants — including the coupling matrix W — were fixed **by hand, before evaluation**, each with a declared provenance category. Whether other settings work as well, and systematic sensitivity analysis under small perturbations, is acknowledged future work.

BACKUP · ANTICIPATED QUESTION

Q

Does Γ actually play the game better?

We make **no payoff claim**. Per-interaction traces were not retained, so we can describe opponent-differentiated cooperation but not cumulative utility. The contribution is the feasibility of explicit self-regulation — not a better score. ReAct, in fact, cooperates more on average.

BACKUP · ANTICIPATED QUESTION

Q

Does it generalize beyond IPD?

Open. IPD was chosen for its rich evidence on human behaviour in repeated social interaction — a non-trivial setting for opponent differentiation. Generalization to other tasks needs follow-up with matched controls.