

Collective Cognition Emerging from the Free Energy Principle

INAUGURAL WORKSHOP ON
ACTIVE INFERENCE IN ITALY
NAPLES, 2026 JUN 23-24



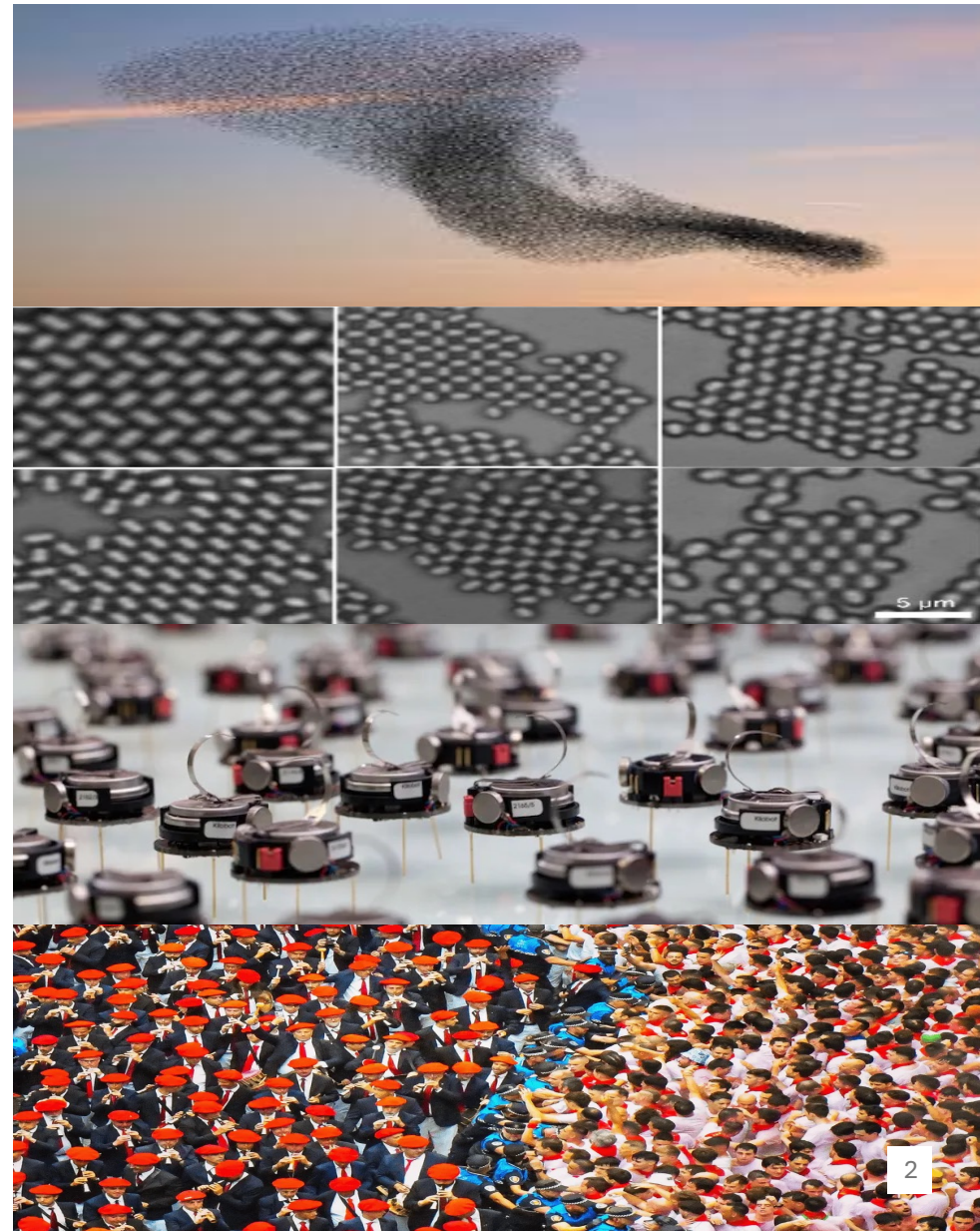
DOMENICO MAISTO

Collective Cognition

The ability of a group of individuals to process information and act as a unified system

Communication, knowledge aggregation, and synchronisation as organisation mechanisms

Top-down and bottom-up as organisation forms

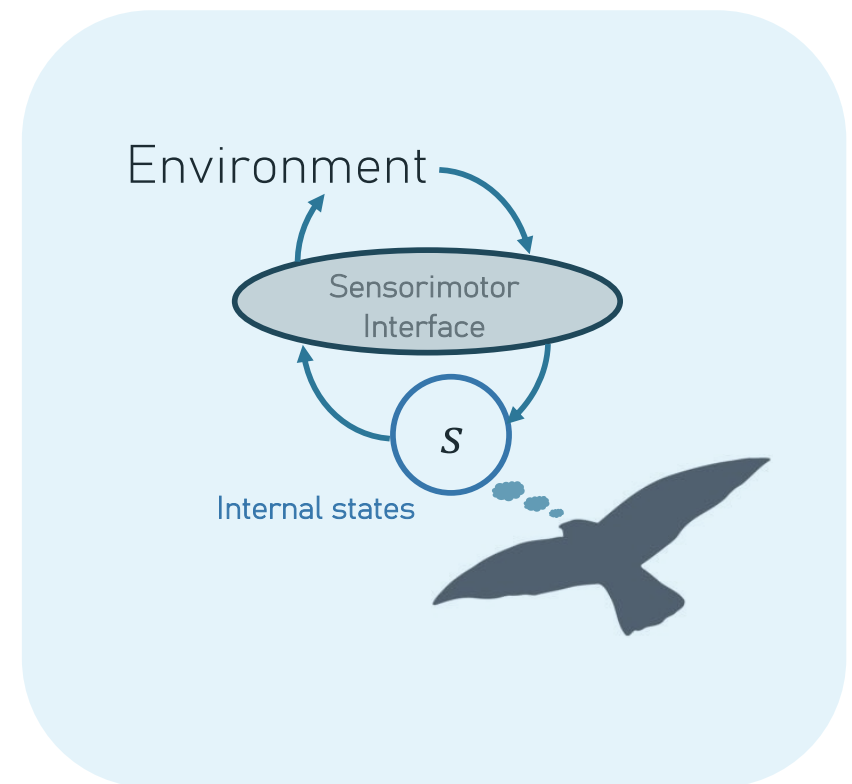
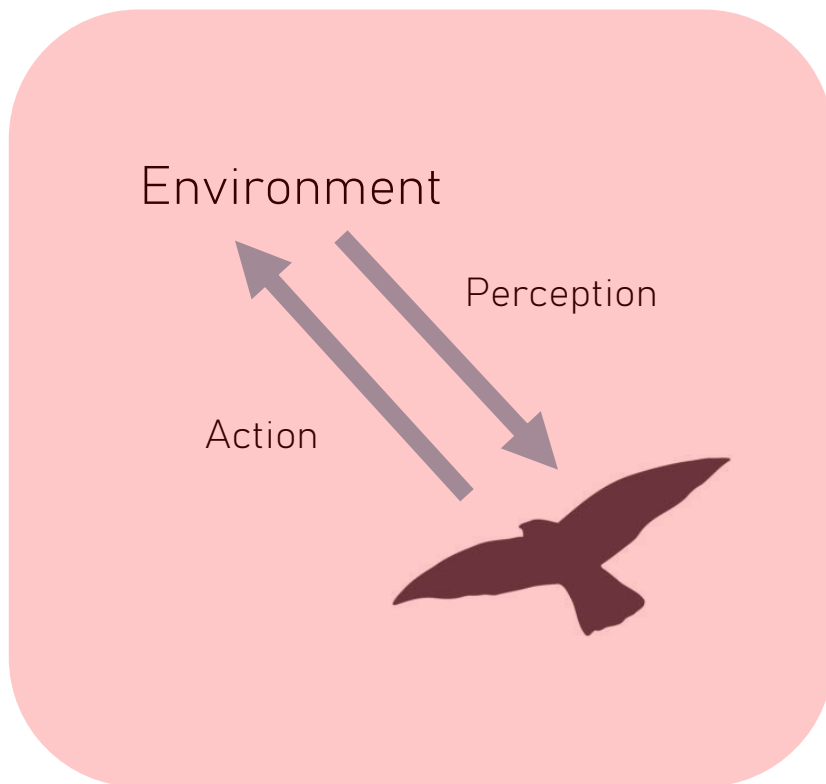


An exemplificative case

(somewhere in time...)



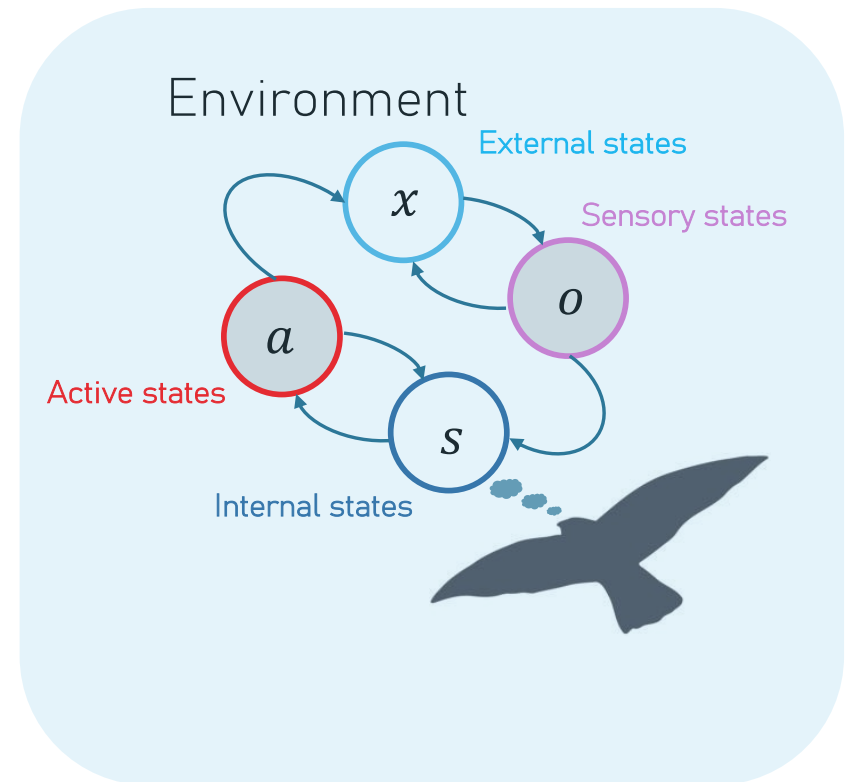
Classical vs Bayesian approach



Active Inference in Nutshell (1/2)

■ MARKOV BLANKET ASSUMPTION

s and x are statistically independent because of o and a



Active Inference in Nutshell (1/2)

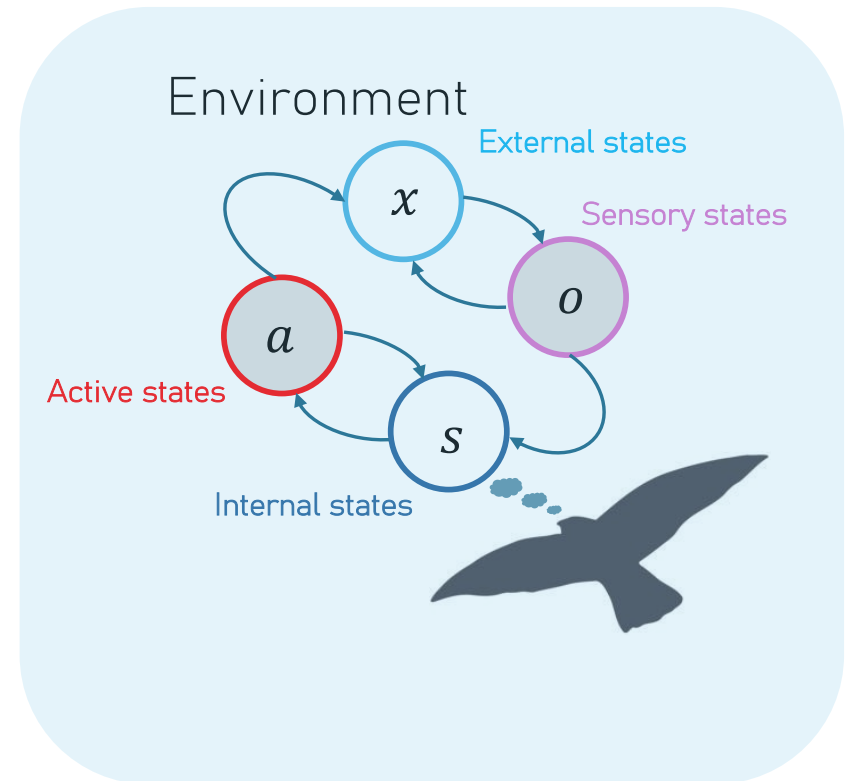
■ MARKOV BLANKET ASSUMPTION

s and x are statistically independent because of o and a

■ GENERATIVE MODEL

a.k.a. an internal model of environment:

$$P(s, o, a) = P(o|s)P(s|a)P(a)$$



Active Inference in Nutshell (1/2)

- MARKOV BLANKET ASSUMPTION

s and x are statistically independent because of o and a

- GENERATIVE MODEL

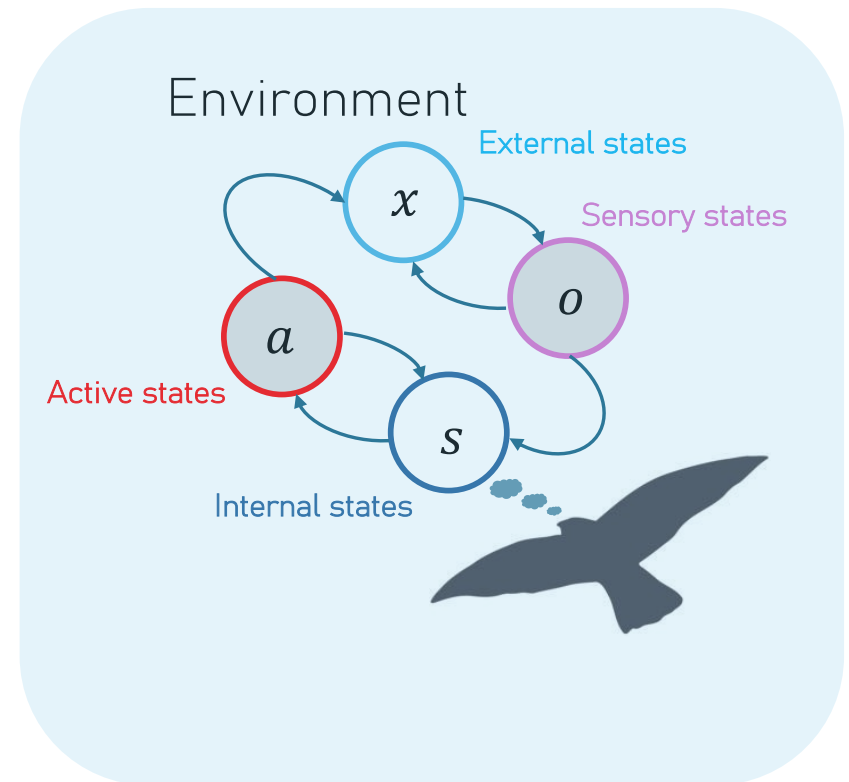
a.k.a. an internal model of environment:

$$P(s, o, a) = P(o|s)P(s|a)P(a)$$

- APPROXIMATE POSTERIOR

a feasible way to invert the Generative Model:

$$Q(s, a) \approx P(s, a|o)$$



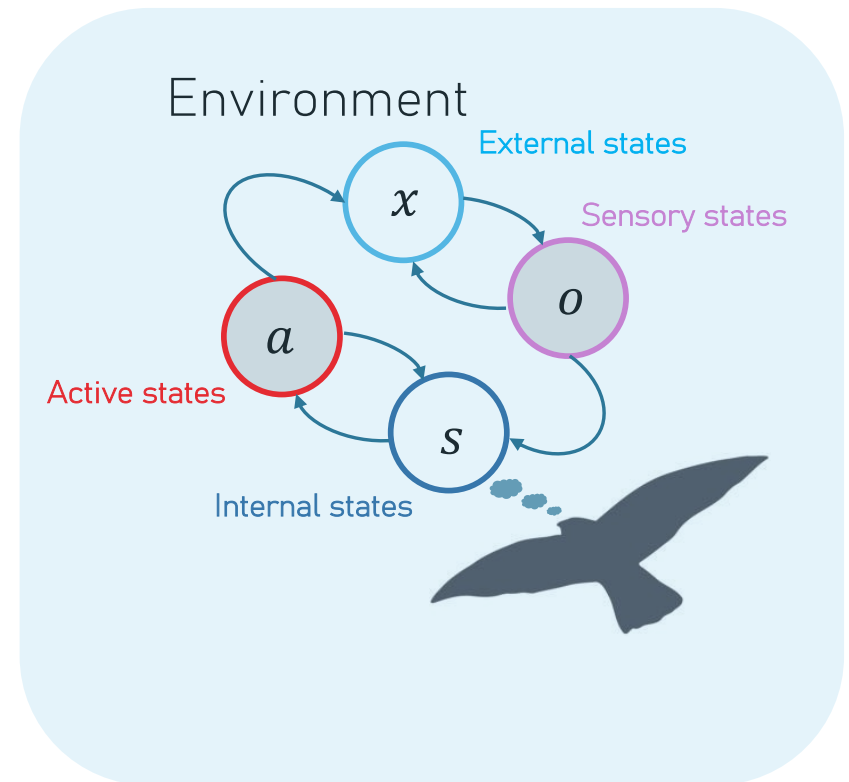
Active Inference in Nutshell (2/2)

■ THE FREE ENERGY PRINCIPLE

- Perception: to approximate the true posterior we can minimise the *variational free energy*:

$$0 \leq D_{KL}[Q(s, a) \parallel P(s, a | o)] = \log P(o) + F[Q(s, a)]$$

$$\Rightarrow F[Q(s, a)] \geq -\log P(o)$$



Active Inference in Nutshell (2/2)

■ THE FREE ENERGY PRINCIPLE

- **Perception**: to approximate the true posterior we can minimise the *variational free energy*:

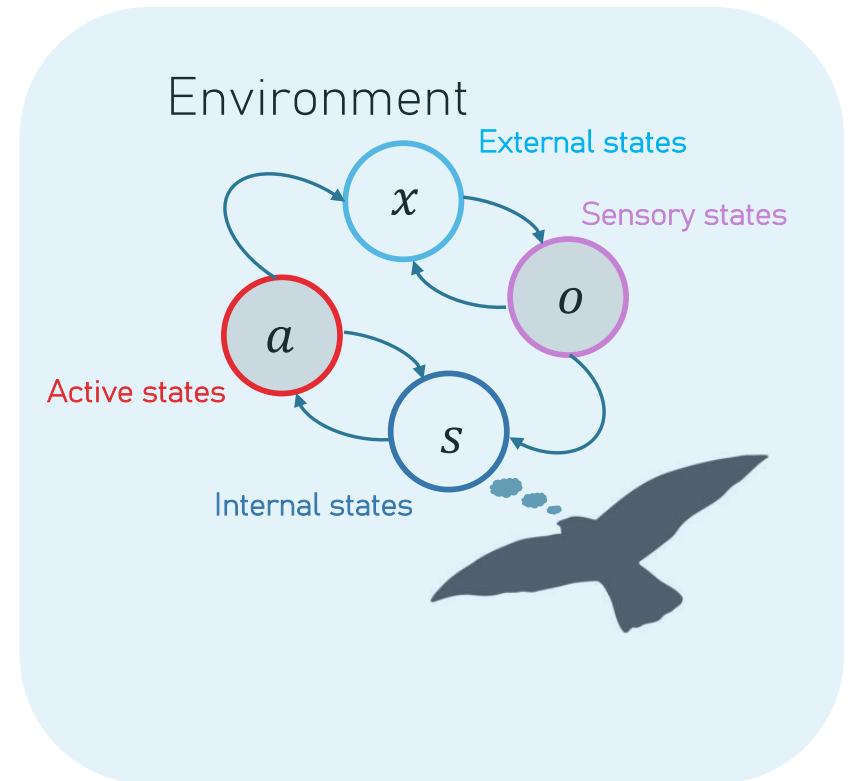
$$0 \leq D_{KL}[Q(s, a) \parallel P(s, a | o)] = F[Q(s, a)] + \log P(o)$$

$$\Rightarrow F[Q(s, a)] \geq -\log P(o)$$

- **Action**: the most likely course of active states minimises the *expected free energy*.

$$G(a) \cong \underbrace{D_{KL}[Q(o|a) \parallel P(o)]}_{\text{risk}} + \underbrace{\mathbb{E}_{Q(s|a)}[H[P(o|s)]]}_{\text{expected ambiguity}}$$

preferred observation



Active inference agents in flock

A generative model

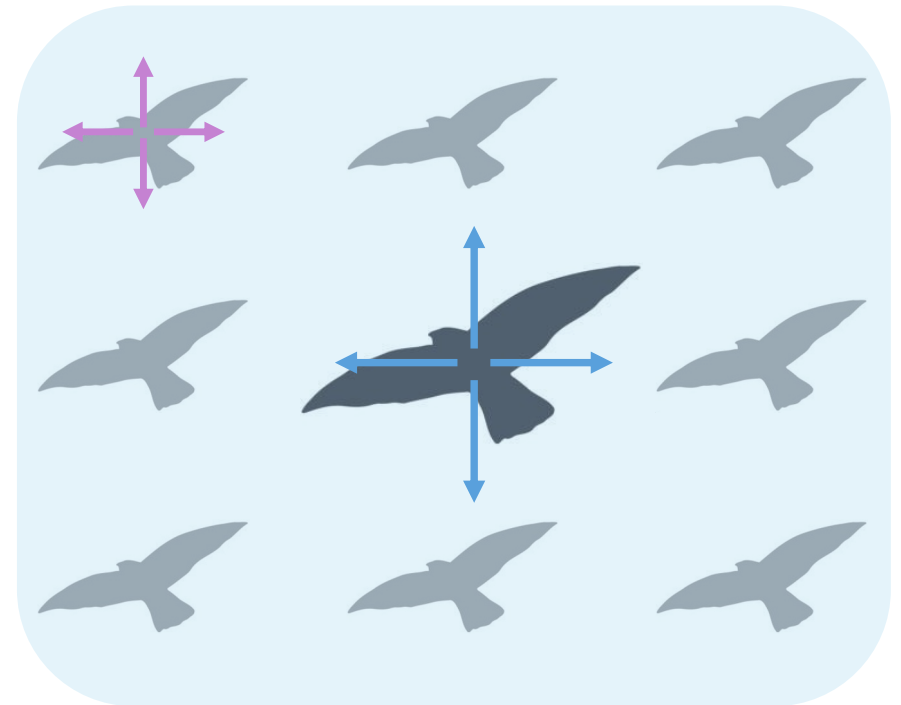
$$P(\mathbf{s}_t, \mathbf{o}_t, \mathbf{a}_t) = P(\mathbf{o}_t | \mathbf{s}_t) P(\mathbf{s}_t | \mathbf{a}_t) P(\mathbf{a}_t)$$

with a hybrid definition combining:

- lattice spin glass methods (Potts model):

$$\mathbf{s}_t \in \{\uparrow, \rightarrow, \downarrow, \leftarrow\},$$

$$\mathbf{o}_t = \bigcup_{j=1}^8 \sigma_t^j, \text{ with } \sigma_t^j \in \{\uparrow, \rightarrow, \downarrow, \leftarrow\}$$



Active inference agents in flock

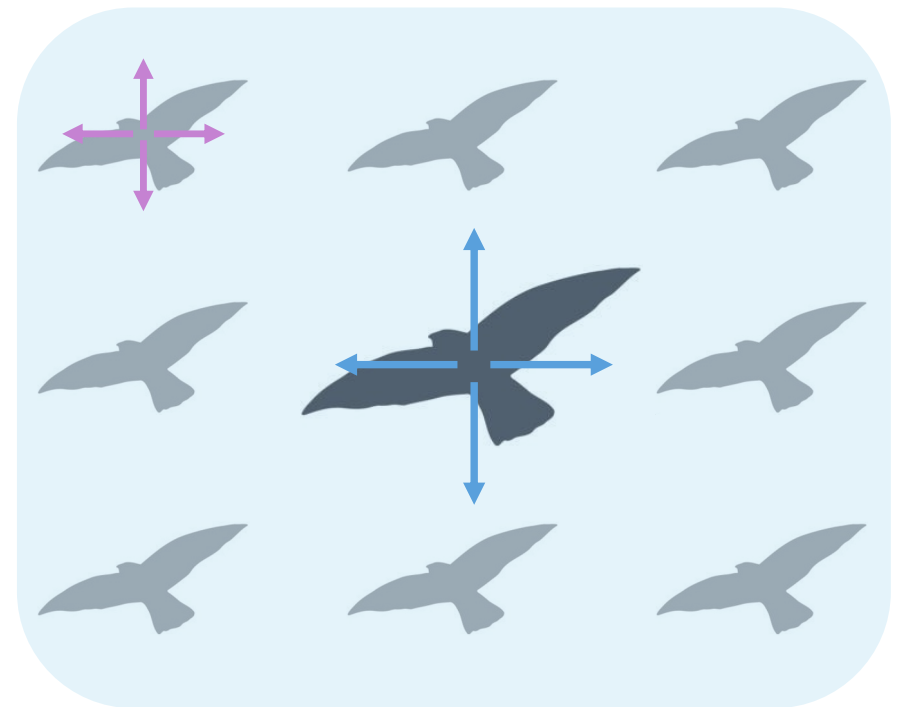
A generative model

$$P(\mathbf{s}_t, \mathbf{o}_t, \mathbf{a}_t) = P(\mathbf{o}_t | \mathbf{s}_t) P(\mathbf{s}_t | \mathbf{a}_t) P(\mathbf{a}_t)$$

with a hybrid definition combining:

- lattice spin glass methods (Potts model);
- the free energy principle for transition:

$$P(\mathbf{a}_t) = \underset{\text{precision}}{\text{softmax}}(-\gamma G(\mathbf{a}_t))$$



Active inference agents in flock

A generative model

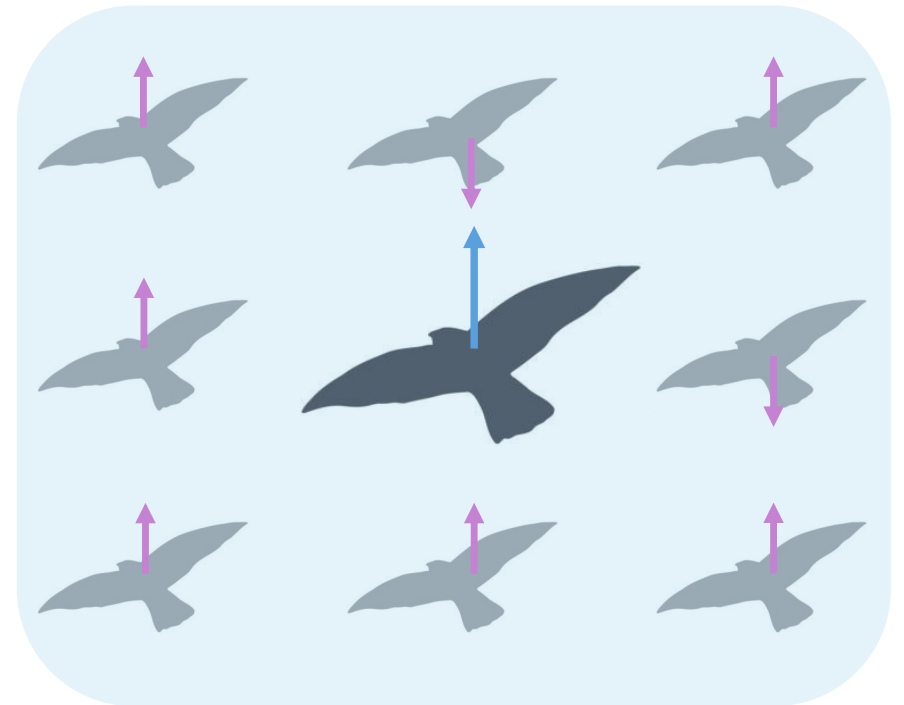
$$P(\mathbf{s}_t, \mathbf{o}_t, \mathbf{a}_t) = P(\mathbf{o}_t | \mathbf{s}_t) P(\mathbf{s}_t | \mathbf{a}_t) P(\mathbf{a}_t)$$

with a hybrid definition combining:

- lattice spin glass methods (Potts model);
- the free energy principle for transition;
- Reynolds' rules (computer graphics) for likelihood mappings:

$$P(\mathbf{o}_t | \mathbf{s}_t) = \prod_{j=1}^8 P(\sigma_t^j | \mathbf{s}_t) \text{ and } P(\sigma_t^j | \mathbf{s}_t) \propto \exp\{-R(j)\}$$

$$\text{with } R(j) = \begin{cases} vm, & \text{velocity matching} \\ ca, & \text{collision avoidance} \\ fc, & \text{flock centering} \end{cases}$$



Active inference agents in flock

A generative model

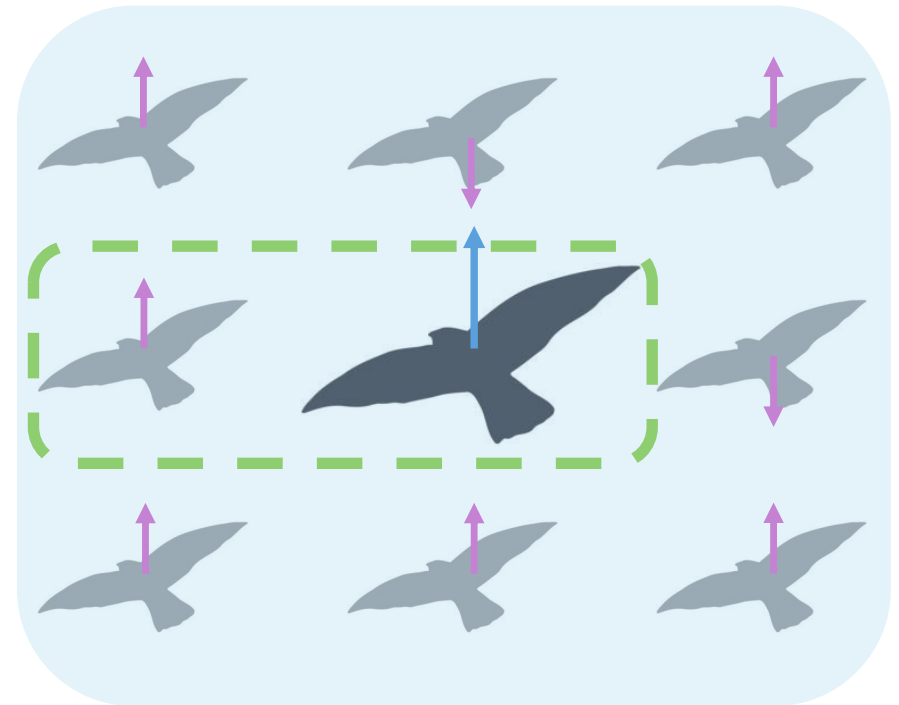
$$P(\mathbf{s}_t, \mathbf{o}_t, \mathbf{a}_t) = P(\mathbf{o}_t | \mathbf{s}_t) P(\mathbf{s}_t | \mathbf{a}_t) P(\mathbf{a}_t)$$

with a hybrid definition combining:

- lattice spin glass methods (Potts model);
- the free energy principle for transition;
- Reynolds' rules (computer graphics) for likelihood mappings:

$$P(\mathbf{o}_t | \mathbf{s}_t) = \prod_{j=1}^8 P(\sigma_t^j | \mathbf{s}_t) \text{ and } P(\sigma_t^j | \mathbf{s}_t) \propto \exp\{-R(j)\}$$

$$\text{with } R(j) = \begin{cases} vm, & \text{velocity matching} \\ ca, & \text{collision avoidance} \\ fc, & \text{flock centering} \end{cases}$$



Active inference agents in flock

A generative model

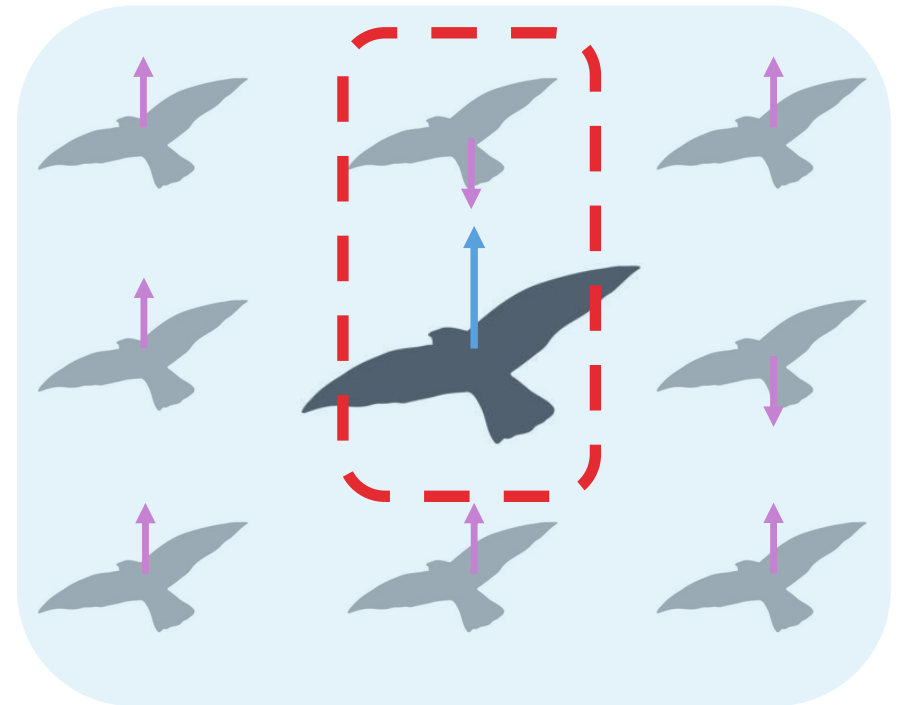
$$P(\mathbf{s}_t, \mathbf{o}_t, \mathbf{a}_t) = P(\mathbf{o}_t | \mathbf{s}_t) P(\mathbf{s}_t | \mathbf{a}_t) P(\mathbf{a}_t)$$

with a hybrid definition combining:

- lattice spin glass methods (Potts model);
- the free energy principle for transition;
- Reynolds' rules (computer graphics) for likelihood mappings:

$$P(\mathbf{o}_t | \mathbf{s}_t) = \prod_{j=1}^8 P(\sigma_t^j | \mathbf{s}_t) \text{ and } P(\sigma_t^j | \mathbf{s}_t) \propto \exp\{-R(j)\}$$

$$\text{with } R(j) = \begin{cases} vm, & \text{velocity matching} \\ ca, & \text{collision avoidance} \\ fc, & \text{flock centering} \end{cases}$$



Active inference agents in flock

A generative model

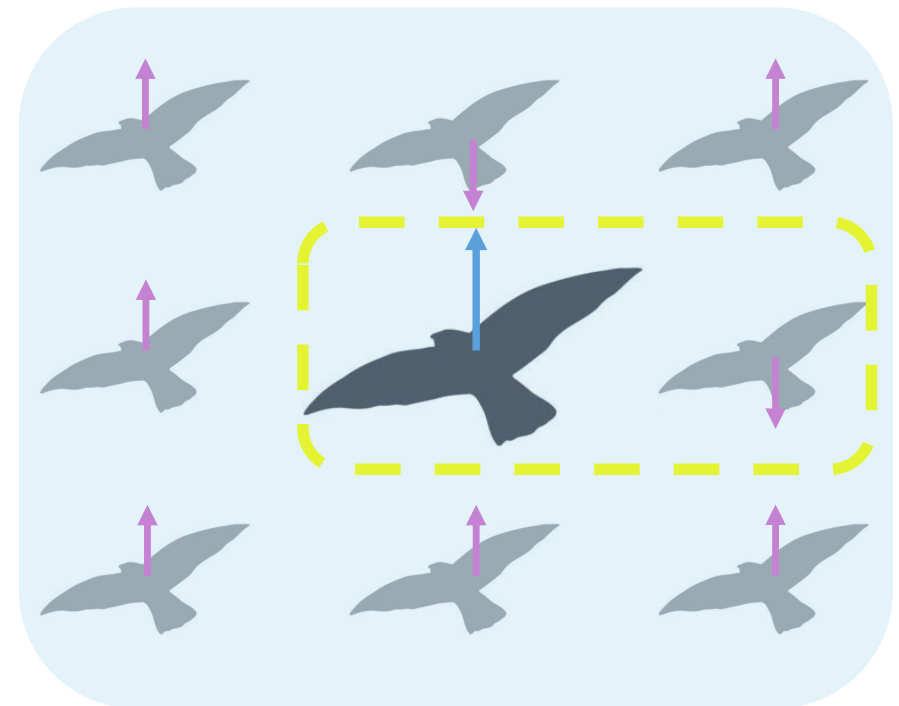
$$P(\mathbf{s}_t, \mathbf{o}_t, \mathbf{a}_t) = P(\mathbf{o}_t | \mathbf{s}_t) P(\mathbf{s}_t | \mathbf{a}_t) P(\mathbf{a}_t)$$

with a hybrid definition combining:

- lattice spin glass methods (Potts model);
- the free energy principle for transition;
- Reynolds' rules (computer graphics) for likelihood mappings:

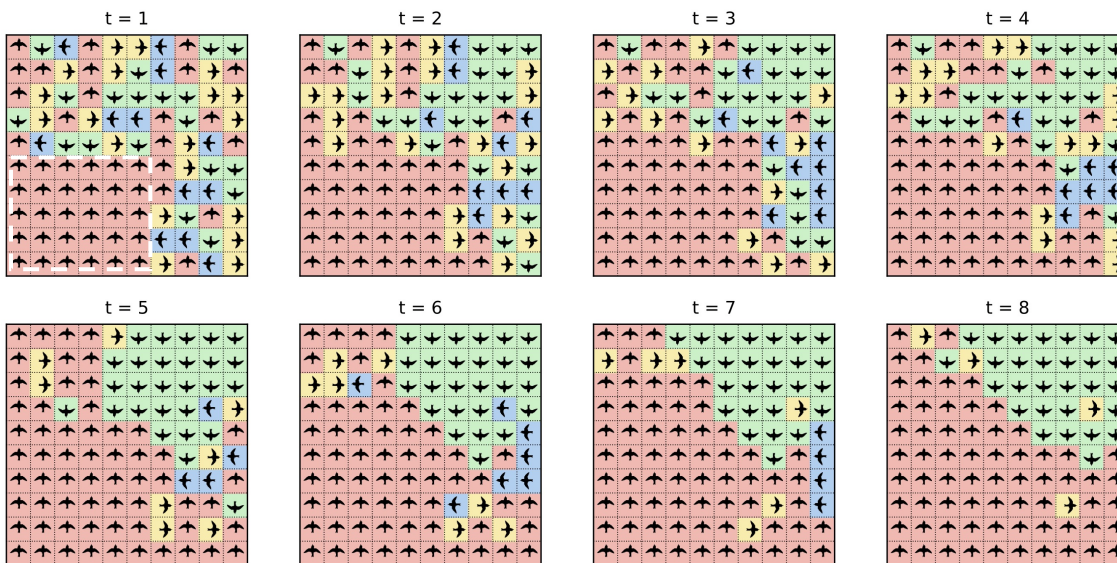
$$P(\mathbf{o}_t | \mathbf{s}_t) = \prod_{j=1}^8 P(\sigma_t^j | \mathbf{s}_t) \text{ and } P(\sigma_t^j | \mathbf{s}_t) \propto \exp\{-R(j)\}$$

$$\text{with } R(j) = \begin{cases} v_m, & \text{velocity matching} \\ c_a, & \text{collision avoidance} \\ f_c, & \text{flock centering} \end{cases}$$



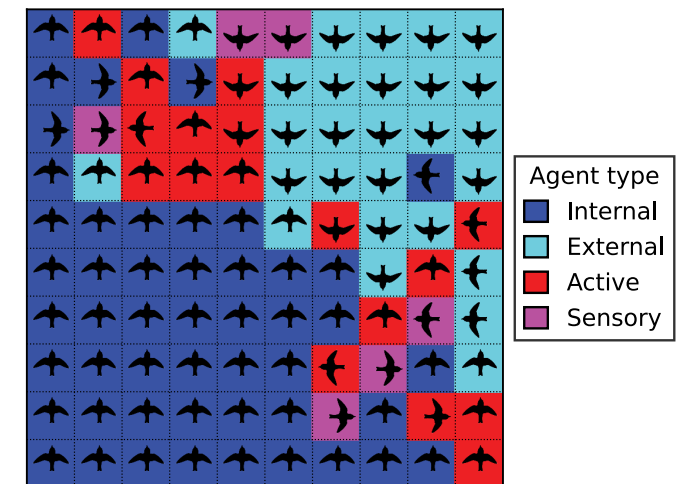
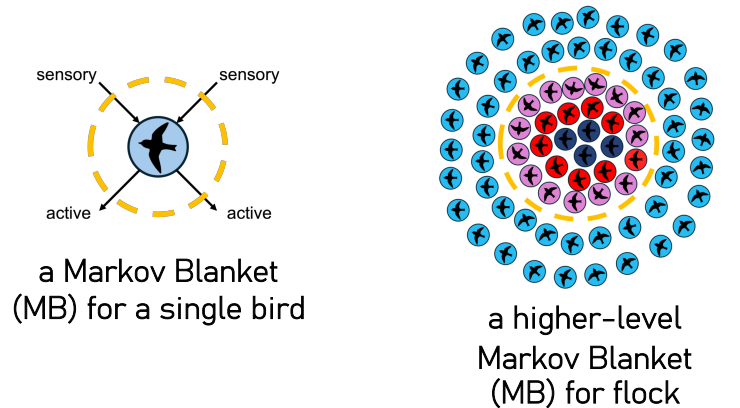
In our simulations, $v_m = 4$, $c_a = 2$, and $f_c = 1$.

Flocking as Joint Agency



gradual alignment of 100 birds over 8 steps

- Lattice size 10x10; one active inference agent (bird) per cell.
- Random 70 out of 100 birds updated at every time step
- Markov Blankets (MB) detected by using spectral techniques



Emergent MB in the time window with $t \in \{1, \dots, 6\}$

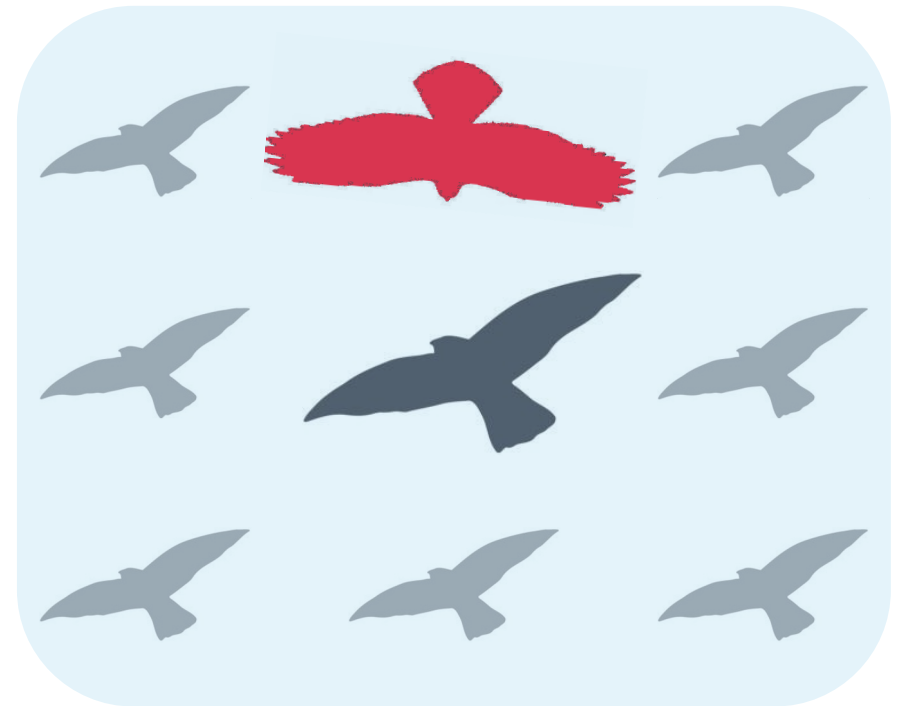
Extending the GM to escape from predators

Let's add a 'stress state' s_t^* to the model:

$$s_t' = s_t \otimes s_t^* \quad \text{with} \quad s_t^* \in \{0, 1\}$$

and the related stress observations σ_t^{j*}

$$\sigma_t^{j'} = \sigma_t^j \otimes \sigma_t^{j*} \quad \text{with} \quad \sigma_t^{j*} \in \{0, 1\}$$



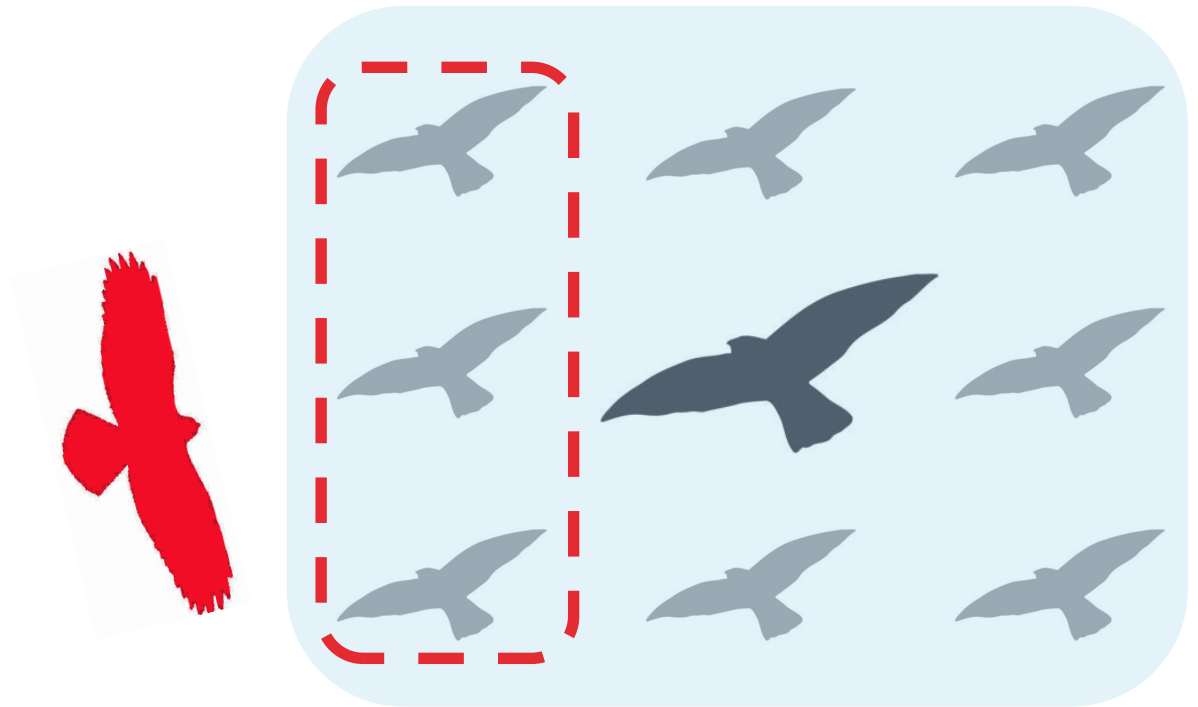
Extending the GM to escape from predators

Let's add a 'stress state' s_t^* to the model:

$$s'_t = s_t \otimes s_t^* \quad \text{with} \quad s_t^* \in \{0, 1\}$$

and the related stress observations σ_t^{j*}

$$\sigma_t^{j'} = \sigma_t^j \otimes \sigma_t^{j*} \quad \text{with} \quad \sigma_t^{j*} \in \{0, 1\}$$



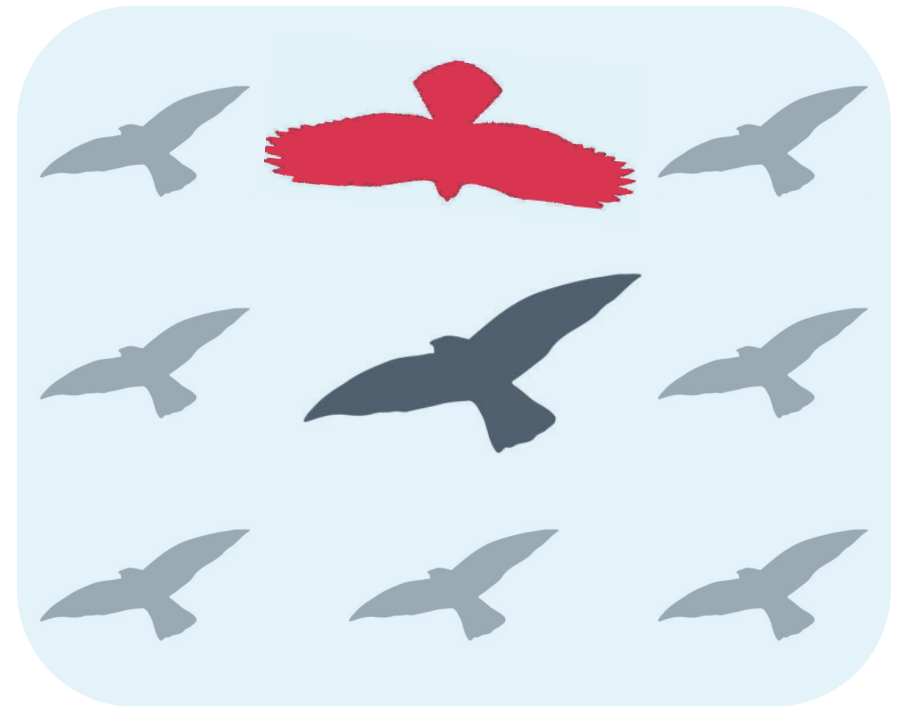
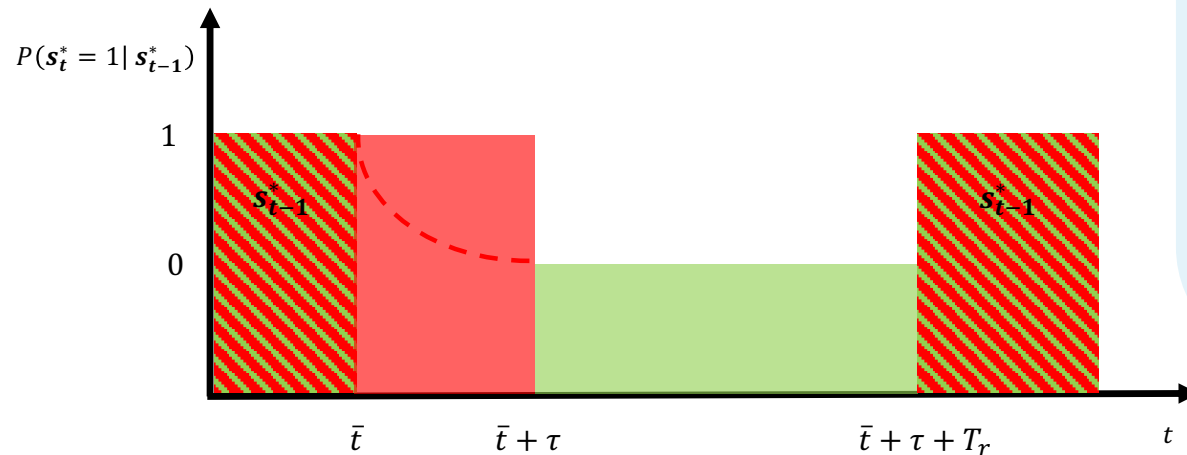
Extending the GM to escape from predators

Let's add a 'stress state' s_t^* to the model:

$$s'_t = s_t \otimes s_t^* \quad \text{with} \quad s_t^* \in \{0, 1\}$$

whose activation decays with a constant τ and a refractory period T_r after a 'danger' occurs in its neighbourhood at the time $\bar{t}$:

In our simulations, $\tau = 2$ and $T_r = 10$.



Extending the GM to escape from predators

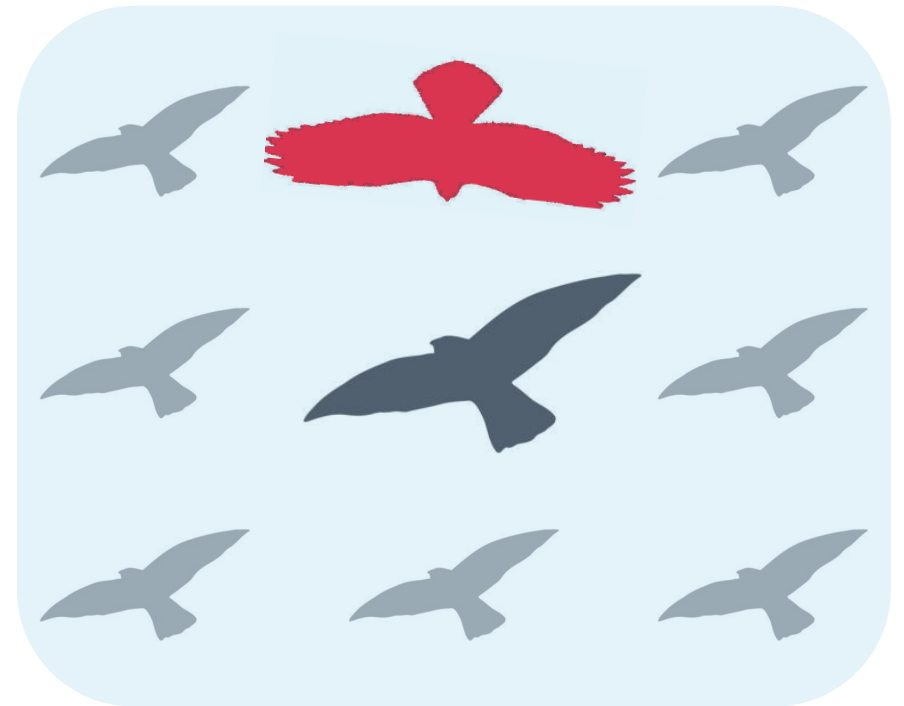
Let's add a 'stress state' s_t^* to the model:

$$s_t' = s_t \otimes s_t^* \quad \text{with} \quad s_t^* \in \{0, 1\}$$

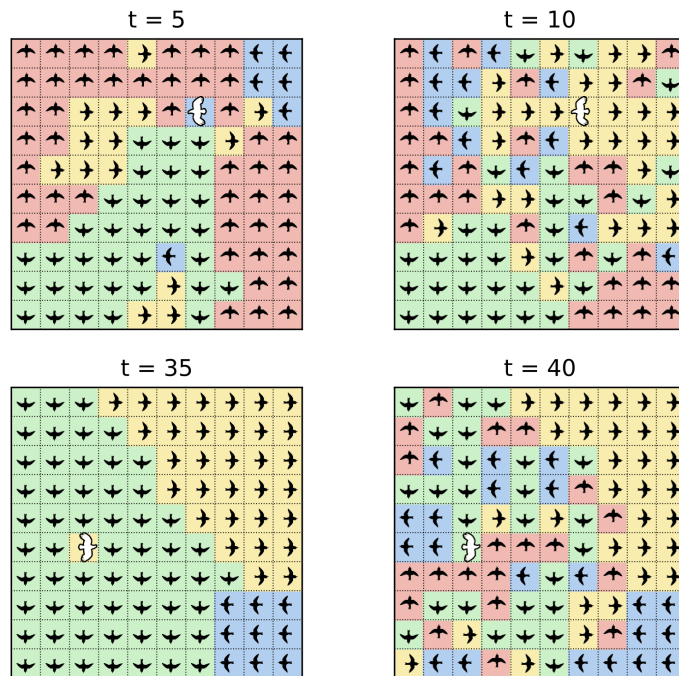
whose activation decays with a constant τ and a refractory period T_r after a 'danger' occurs in its neighbourhood at the the time $\bar{t}$:

In our simulations, $\tau = 2$ and $T_r = 10$.

$$P(\sigma_t^j | s_t, s_t^*) = \begin{cases} P(\sigma_t^j | s_t, 0) = P(\sigma_t^j | s_t) \\ P(\sigma_t^j | s_t, 1) = 1 - P(\sigma_t^j | s_t) \end{cases}$$

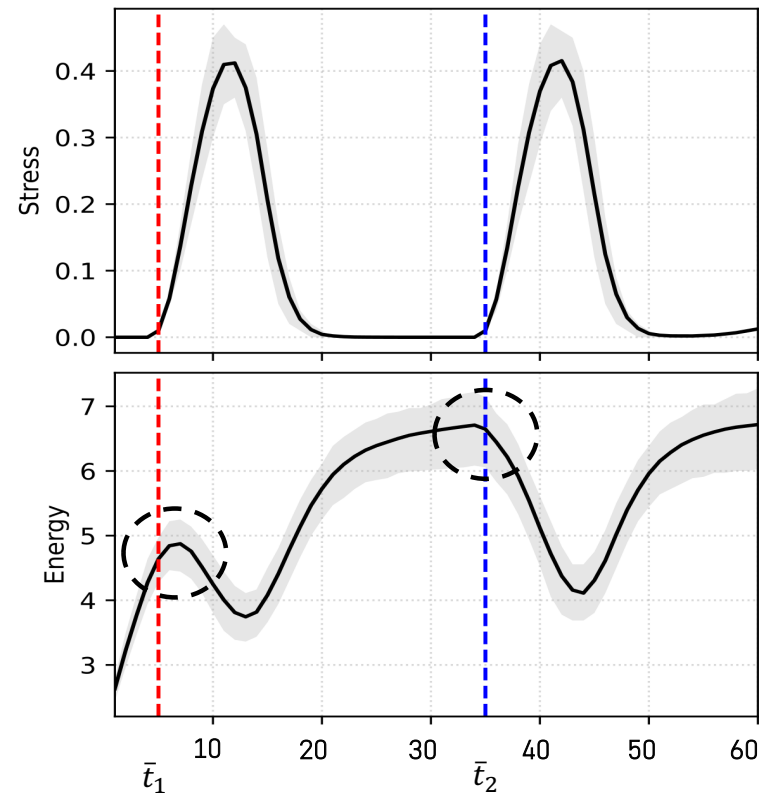


Sensing and escaping predators in the flock

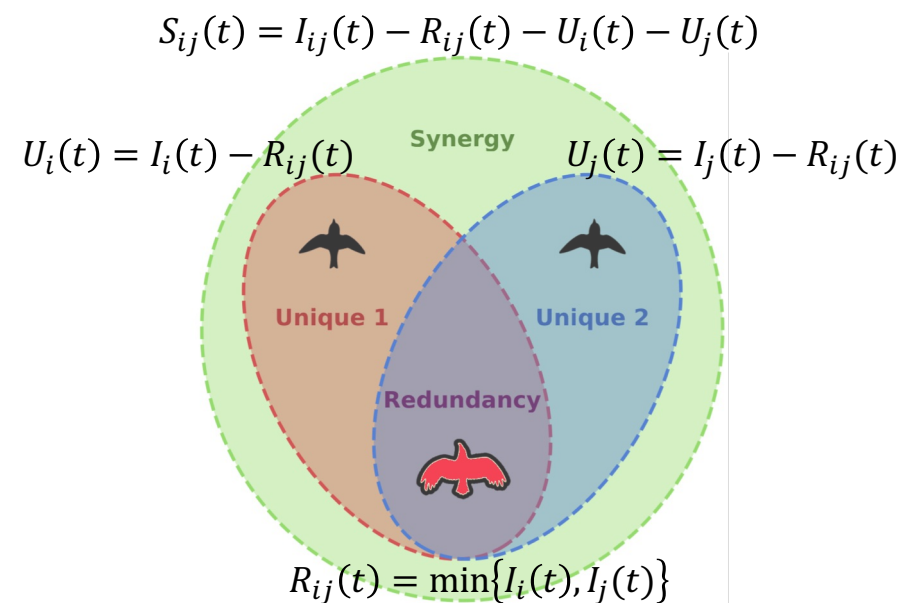
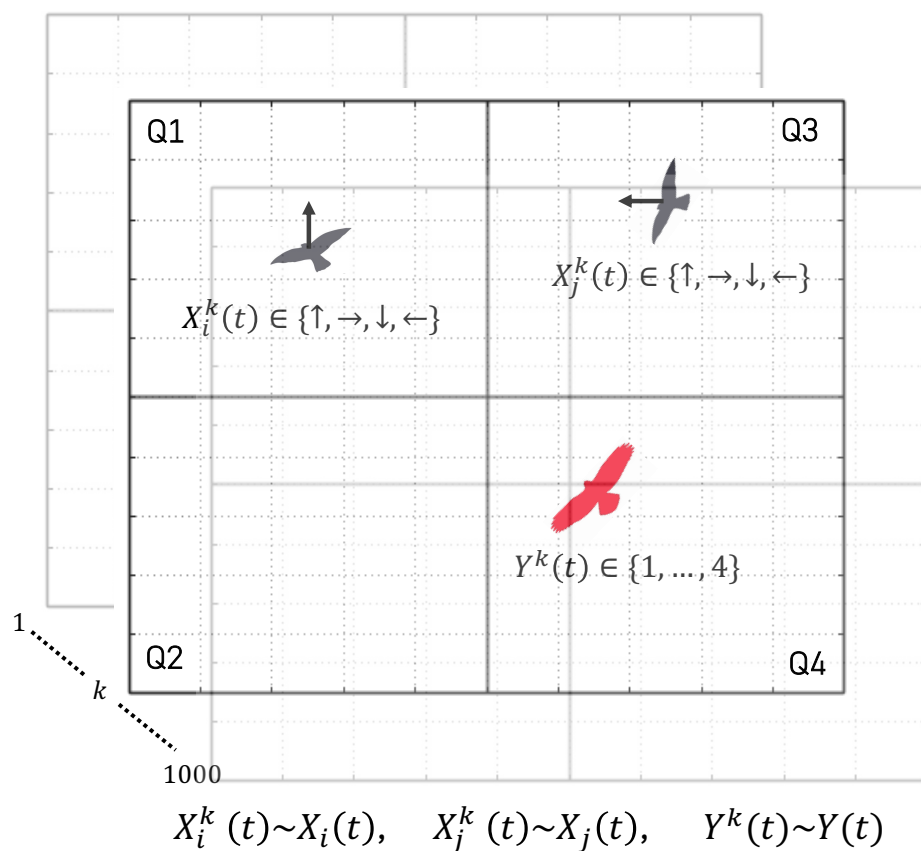


A predator (white bird) arrives at $\bar{t}_1 = 5$ and $\bar{t}_2 = 35$, at two distinct, random locations.

results of 500 simulations lasting 60 time steps



Partial Information Decomposition

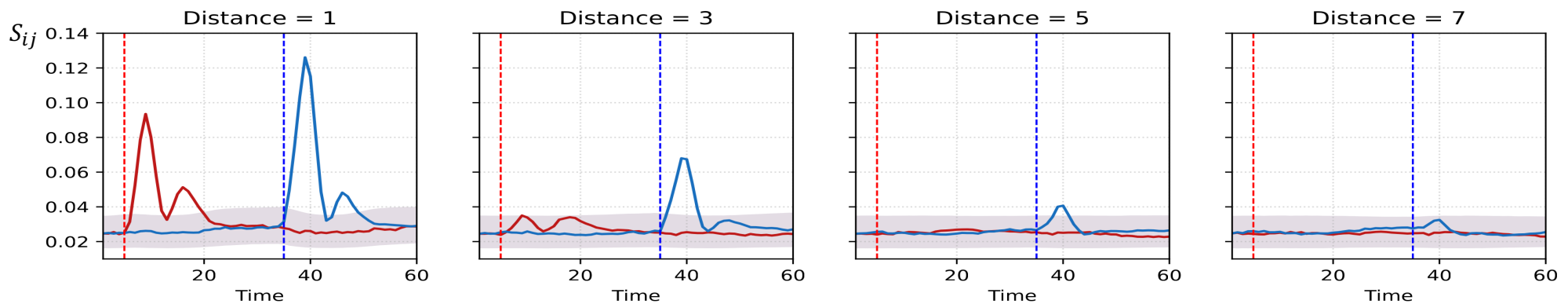
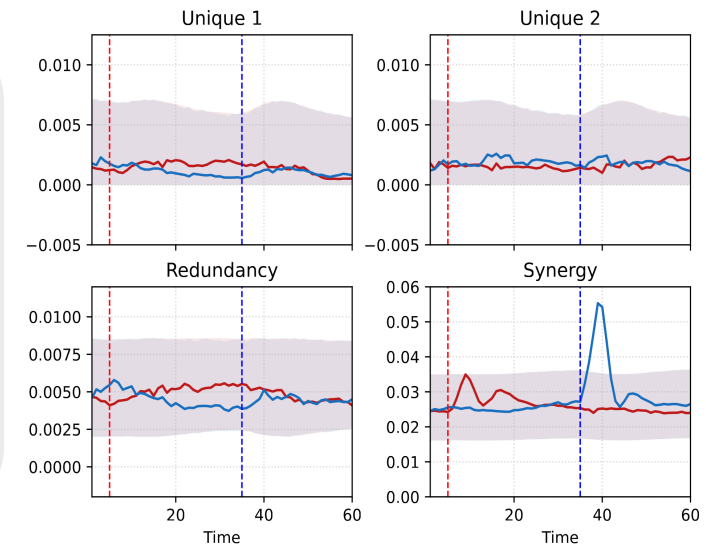


$$I_{ij}(t) = I(X_i(t), X_j(t); Y(t))$$

$$I_i(t) = I(X_i(t); Y(t)) \quad I_j(t) = I(X_j(t); Y(t))$$

Information decomposition about predator location

- PID atoms remain zero, except for synergy
- Only the synergy with Predator 2 appears significant
- Synergies decrease as spatial separation increases



Conclusions & Takeaways

- Active Inference as a deeply integrative, principled, and bio-inspired framework, also for collective behaviour.
- The flock exhibits collective knowledge beyond individual birds
 - The flock emerges as a collective agent (joint agency) with its own statistical boundary (nested Markov blanket)
 - Flocks respond more rapidly to threats when organized. Larger flocks react faster and more effectively
- Implicit collective cognition arises without explicit models
 - flocking require no shared intentions and beliefs
 - Suggests a multi-scale joint agency in more complex systems

Acknowledgements & References

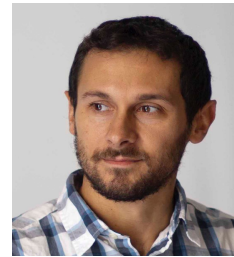
Thanks to the other co-authors:



Francesco Donnarumma



Davide Nuzzi



Giovanni Pezzulo

My contact:

domenico.maisto@cnr.it

Some of our references on Collective Cognition and Joint Agency:

- What the flock knows that the birds do not: exploring the emergence of joint agency in multi-agent active inference, *Philosophical Transactions of the Royal Society B* (2026)
- Collective field-based self-organisation of Active Particles through the Free Energy Principle, (in preparation)
- A predictive processing framework for joint action and communication, *Physics of Life Reviews* (2026)
- Interactive Inference: A Multi-Agent Model of Cooperative Joint Actions, *IEEE Transactions on Systems, Man, and Cybernetics: Systems* (2023), pp. 1–12,

POMDP generative model

