Factorised Active Inference for Strategic Multi-Agent Interactions

AAMAS 2025

Jaime Ruiz-Serra, Patrick Sweeney, Michael Harré

Centre for Complex Systems,
School of Computer Science,
The University of Sydney

May 2025

1 Introduction and Background
2 Technical Approach
3 Experiments
4 Conclusions

Research questions

- How do agents learn and adapt when interacting with others, especially under uncertainty and changing conditions (non-stationarity)?
- How does this manifest at the collective level?

Bridging two frameworks [DHK24]:

- **Game Theory**: Formalises strategic interactions between agents with potentially competing goals. Classically assumes perfect rationality. Difficulties with bounded rationality and learning dynamics.
- **Active Inference (AIF)**: A process theory from neuroscience describing perception, action, and learning under uncertainty via generative models. Agents act to minimize free energy (surprise/uncertainty). Lacks a clear framework for strategic multi-agent interactions.

Background concepts

- **Iterated Normal-Form Games**: Standard framework for studying repeated interactions (as determined by payoffs defined by a game matrix, g, e.g., *Prisoner's Dilemma*, *Chicken*, *Stag Hunt*). Agents learn, build reputations, and use reciprocity.
- **Active Inference**:
  - Agents use an internal generative model (beliefs about world dynamics, POMDP)
  - Perception: infer hidden states (s) from observations (o) by minimising *Variational Free Energy* (VFE) (approximating Bayesian inference)
  - Action: select actions (u) to minimise *Expected Free Energy* (EFE). EFE balances:
    - Pragmatic value: achieving preferred outcomes ($p^*(o)$, related to utility/payoffs)
    - Epistemic value: reducing uncertainty about hidden states (salience) and model parameters (novelty)

Core Ideas

Factorised Generative Model

We propose a **factorisation** of the AIF generative model, enabling agents to maintain **individual beliefs about others** while **planning strategically in a joint context**.

Non-stationary Environment

We employ **game transitions** to induce non-stationarity in the environment and study **how AIF agents adapt** their planning and behaviour.

1 Introduction and Background
2 Technical Approach
3 Experiments
4 Conclusions

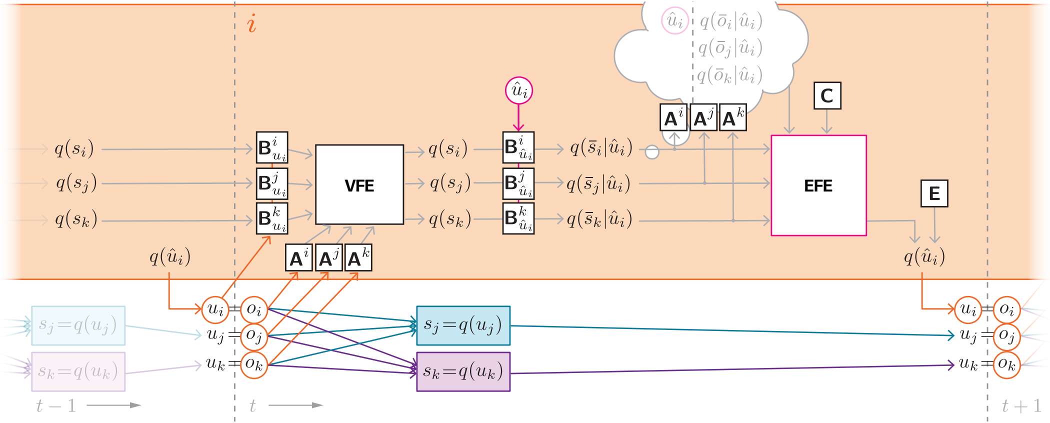
The Factorised AIF Model

Core idea: Ego agent (i) models itself and each alter agent ($j, k, \ldots$) separately.

- Based on the idea that given an agent's Markov Blanket (past actions of all agents), its internal state is conditionally independent of other agents' internal states.
- Factorised Beliefs: Ego maintains individual beliefs $q(s_i), q(s_j), q(s_k)$ about the hidden state (e.g., propensity to cooperate, 'type') of each alter agent (and itself).
- Inference (VFE): Beliefs $q(s_n)$ for each factor $n \in \{i, j, k\}$ are updated based on observed actions $o = (o_i, o_j, o_k)$ by minimising VFE

$$F[q, o] = \mathrm{D}_{KL}\left[q(s) \,\|\, p(s|o)\right] - \log p(o)$$

Generative Model Structure



Planning and Action Selection

Planning (EFE minimisation): ego evaluates counterfactual actions ($\hat{u}_i$) by calculating their EFE

$$G[\hat{u}_i] = -\underbrace{\rho[\hat{u}_i]}_{\text{Pragmatic Value}} - \underbrace{\varsigma[\hat{u}_i]}_{\text{Salience (Epistemic)}} - \underbrace{\eta[\hat{u}_i]}_{\text{Novelty (Learning)}}$$

Planning and Action Selection

$$G[\hat{u}_i] = -\underbrace{\rho[\hat{u}_i]}_{\text{Pragmatic Value}} - \underbrace{\varsigma[\hat{u}_i]}_{\text{Salience (Epistemic)}} - \underbrace{\eta[\hat{u}_i]}_{\text{Novelty (Learning)}}$$

Pragmatic value (joint interaction context): Preferences are derived from the game payoff matrix [DHK24]: $p^*(o_i, o_j, o_k) = \sigma\big(\mathrm{g}(o_i, o_j, o_k)\big)$

- Requires predicting joint outcomes: $q(\bar{o}_i, \bar{o}_j, \bar{o}_k|\hat{u}_i)$
- Conditional independence allows factorisation: $q(\bar{o}_i, \bar{o}_j, \bar{o}_k|\hat{u}_i) = \prod_n q(\bar{o}_n|\hat{u}_i)$
- Pragmatic value becomes expected utility under predicted opponent actions:

$$\rho[\hat{u}_i] = \mathbb{E}_{q(\bar{o}_j)q(\bar{o}_k)}\big[\log p^*(\hat{u}_i, \bar{o}_j, \bar{o}_k)\big]$$

Planning and Action Selection

$$G[\hat{u}_i] = -\underbrace{\rho[\hat{u}_i]}_{\text{Pragmatic Value}} - \underbrace{\varsigma[\hat{u}_i]}_{\text{Salience (Epistemic)}} - \underbrace{\eta[\hat{u}_i]}_{\text{Novelty (Learning)}}$$

Salience: Uncertainty about predicted outcomes

$$\begin{aligned}\varsigma[\hat{u}] &= \mathbb{E}_{q(\bar{o}|\hat{u})}\left[\mathrm{D}_{KL}\left[q(\bar{s}|\hat{u},\bar{o}) \,\|\, q(\bar{s}|\hat{u}) \right]\right] \\ &= H\big(q(\bar{o}|\hat{u})\big) - \mathbb{E}_{q(\bar{s}|\hat{u})}\left[H\big(p(\bar{o}|\bar{s})\big)\right]\end{aligned}$$

- Due to identity observation model **A**, the second term is 0.
- Each factor's uncertainty contributes additively $\varsigma[\hat{u}_i] = \sum_n H(q(\bar{o}_n|\hat{u}_i))$.

Planning and Action Selection

$$G[\hat{u}_i] = -\underbrace{\rho[\hat{u}_i]}_{\text{Pragmatic Value}} - \underbrace{\varsigma[\hat{u}_i]}_{\text{Salience (Epistemic)}} - \underbrace{\eta[\hat{u}_i]}_{\text{Novelty (Learning)}}$$

Novelty: Uncertainty about model parameters

$$\eta[\hat{u}_i] = \sum_n \mathrm{D}_{KL}\left[\bar{\mathbf{B}}_{\hat{u}_i,n} \middle\| \mathbf{B}_{\hat{u}_i,n}\right]$$

Learning: Transition model parameters (**B**) updated periodically based on observed transitions; Bayesian model reduction used to combat overfitting.

Action Selection: Softmax policy over habits $\mathbf{E}$, and EFE values, $\mathbf{G}$, modulated by precision (γ):

$$u_i \sim q(\hat{u}_i) = \sigma(\log \mathbf{E} - \gamma \mathbf{G})$$

Precision is dynamically-updated:

$$\gamma = \frac{\beta_1}{\beta_0 - \langle \mathbf{G} \rangle}$$

where $\langle \mathbf{G} \rangle = \mathbb{E}_{q(\hat{u}_i)}\left[G[\hat{u}_i]\right]$

1 Introduction and Background
2 Technical Approach
3 Experiments
4 Conclusions

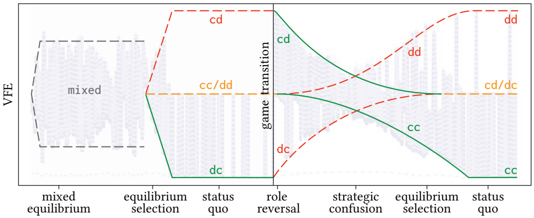
Experiments: Game Transitions

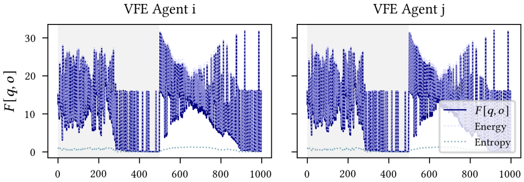
Setup: two- and three-player iterated games (Chicken, Stag Hunt).

Non-stationarity: Introduced *game transitions*: payoff matrices linearly interpolated over time (e.g., Chicken → Stag Hunt). Models changing social contexts beyond just reciprocal adaptation.

Analysis: Tracked key AIF quantities over time: VFE, EFE components, policies. Examined dynamics during transitions.

VFE and Strategic Uncertainty





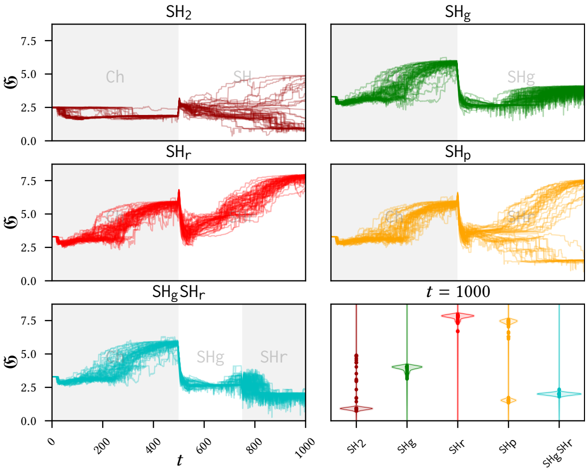
VFE

- Tracks ‘surprise’ given current beliefs.
- Fluctuates high during mixed strategy phases or strategic confusion.
- Drops when agents settle into predictable equilibria (status quo).
- Spikes when unexpected actions occur (deviations from equilibrium or during transitions).

Example (Fig 2): Ch $\longrightarrow$ SH (2 players)

- Initial mixed phase in Ch, then convergence to asymmetric Nash Equilibrium (e.g., dc).
- Transition induces role reversal (dc $\longrightarrow$ cd), strategic confusion (high VFE/policy entropy), then re-convergence in SH.

Ensemble-level Expected EFE



Ensemble-level expected EFE (cf. [Hyl+24])

$$\mathfrak{G} = \sum_i \langle \mathbf{G} \rangle^{(i)} = \sum_i \mathbb{E}_{q(\hat{u}_i)} \left[G[\hat{u}_i] \right]$$

Stag Hunt game variants

$$\mathrm{SH_g} = \left[\begin{bmatrix} R & R \\ R & S \end{bmatrix}, \begin{bmatrix} T & P \\ P & P \end{bmatrix} \right]$$

$$\mathrm{SH_r} = \left[\begin{bmatrix} R & S \\ S & S \end{bmatrix}, \begin{bmatrix} T & P \\ P & P \end{bmatrix} \right]$$

$$\mathrm{SH_p} = \left[\begin{bmatrix} R & S \\ S & S \end{bmatrix}, \begin{bmatrix} P & P \\ P & P \end{bmatrix} \right]$$

Findings

1. $\mathfrak{G}$ dynamics characterise equilibria and attractor basins (Fig 4). Lower $\mathfrak{G}$ generally indicates 'better' collective outcomes.
2. Nash Equilibria aren't always socially optimal (minimal $\mathfrak{G}$).
3. Bifurcations in $\mathfrak{G}$ show convergence to different equilibria (e.g., payoff-dominant vs risk-dominant in SH) across trials. Shows relative basin size.
4. Game structure significantly impacts equilibrium selection (e.g., $\mathsf{SH_g}$ vs $\mathsf{SH_r}$ vs $\mathsf{SH_p}$). Paradoxical results observed (requiring more cooperation sometimes led to less).
5. Strategic intervention possibility: Transitioning through a trust-building game ($\mathsf{SH_g}$) can steer the collective to a better equilibrium (PDE) more effectively than penalizing defection ($\mathsf{SH_p}$).

1 Introduction and Background
2 Technical Approach
3 Experiments
4 Conclusions

Summary

We introduced a factorised generative model for AIF agents enabling strategic planning in joint game contexts by maintaining individual beliefs about others. We tested the model in non-stationary environments with changing preferences (game payoffs).

Key Insights for MAS

- Provides a principled way (via AIF) to model boundedly rational agents learning and adapting in non-stationary strategic settings.
- VFE and EFE dynamics offer analytical tools to understand strategic uncertainty, equilibrium selection, and collective behaviour.
- Demonstrates how interaction history (priors/pre-equilibrium) and game structure shape collective outcomes.

Broader Implications: Connects cognitive process models (AIF) with economic/MAS models (Game Theory). Potential for understanding collective intelligence and designing interventions.

Future Work: Learning opponent rationality (via observation model **A**); more complex transition models **B**; inferring opponent preferences; exploring different EFE formulations; application to networked games.

Code: `github.com/RuizSerra/factorised-MA-AIF`

References I

[DHK24] D. Demekas, C. Heins, and B. Klein. "An Analytical Model of Active Inference in the Iterated Prisoner's Dilemma". In: *Active Inference*. Ed. by C. L. Buckley et al. Cham: Springer Nature Switzerland, 2024, pp. 145–172.

[Hyl+24] D. Hyland et al. "Free-Energy Equilibria: Toward a Theory of Interactions Between Boundedly-Rational Agents". In: ICML 2024 Workshop on Models of Human Feedback for AI Alignment. July 2, 2024.

Factorised Active Inference for Strategic Multi-Agent Interactions

Jaime Ruiz-Serra, Patrick Sweeney, Michael Harré
Centre for Complex Systems, University of Sydney

Thank you!
Questions?

Jaime.RuizSerra@sydney.edu.au

github.com/RuizSerra/factorised-MA-AIF