14th Conference on Stochastic Models of Manufacturing and Service Operations (SMMSO 2024)

Energy-Efficiency Control of Manufacturing Systems via Active Inference

Yavar Taheri Yeganeh
Department of Mechanical Engineering, Politecnico di Milano, Milan, Lombardy, Italy, yavar.taheri@polimi.it

Mohsen A. Jafari
Department of Industrial and Systems Engineering, Rutgers University, Piscataway, New Jersey, United States, jafari@soe.rutgers.edu

Andrea Matta
Department of Mechanical Engineering, Politecnico di Milano, Milan, Lombardy, Italy, andrea.matta@polimi.it

We are exploring active inference theory to develop intelligent decision-making models for optimizing energy-efficient control in manufacturing systems. Inspired by insights from neuroscience, active inference provides a unified and probabilistic framework that integrates perception, learning, and action. It represents an emerging domain in artificial intelligence, bridging generative models with decision-making processes. Utilizing a deep active inference agent, we investigate potential control strategies in parallel and identical machine workstations, with a specific focus on promoting sustainability in manufacturing systems. We initially concentrate on stationary manufacturing environments, and subsequently extend our analysis to non-stationary cases. We leverage advancements in previously developed active inference agents, along with existing reinforcement solutions, for the systems under study. Our study compares the performance of the active inference-based method with reinforcement learning to evaluate the advancement of the proposed methodology.

Key words: Active Inference; Reinforcement Learning; Energy-Efficient Control; Manufacturing Systems

---

1. Introduction

Energy-efficient control (EEC) is vital in the manufacturing sector due to its substantial impact on global energy consumption. EEC strategies focus on minimizing energy usage by optimizing machine states, particularly during idle periods (Loffredo et al. 2023). Traditional EEC methods often require complete system knowledge, which is impractical in dynamic, real-world environments. Reinforcement learning has shown promising performance in optimizing manufacturing processes without prior system knowledge (Loffredo et al. 2023) but struggles with rapid adaptation to changing conditions. Active inference, based on the free energy principle (FEP), offers an alternative by unifying perception, learning, and decision-making under uncertainty through a Bayesian framework (Friston 2010). It has been applied successfully in complex decision-making tasks in fields such as robotics, enabling agents to navigate uncertain and dynamic scenarios (Pezzato et al. 2023). This research aims to build upon advancements in active inference-based decision-making and apply it to EEC in manufacturing systems to demonstrate its potential.

2. Active Inference Agent

The active inference framework posits that organisms actively interact with their environment by updating beliefs and actions based on sensory inputs to reduce *surprise*. Active inference agents have an internal generative model parameterized by θ that interacts with the world similarly to a

Partially Observable Markov Decision Process (POMDP). The core concept is FEP, which leads to optimizing the model by minimizing Variational Free Energy (VFE) to reduce *surprise*, quantified by $-\log P_\theta(o_t)$, as follows (Fountas et al. 2020):

$$\theta^* = \arg\min_\theta \left(\mathbb{E}_{Q_\phi(s_t,a_t)}\left[\log Q_\phi(s_t,a_t) - \log P_\theta(o_t,s_t,a_t)\right]\right). \tag{1}$$

The agent's actions aim to minimize Expected Free Energy (EFE or G), which is calculated during planning by simulating future trajectories, π, up to a horizon $\tau \geq t$ (Fountas et al. 2020):

$$G(\pi,\tau) = \mathbb{E}_{P(o_\tau|s_\tau,\theta)}\mathbb{E}_{\mathcal{Q}_\phi(s_\tau,\theta|\pi)}\left[\log Q_\phi(s_\tau,\theta|\pi) - \log P(o_\tau,s_\tau,\theta|\pi)\right]. \tag{2}$$

EFE comprises three terms: Reward-like expected *surprise*, which pertains to how close the future predictions are to the preference, state uncertainty, and model parameter uncertainty. The framework (as depicted in Fig. 1) includes first optimizing model parameters θ to fit observations based on VFE in Eq. 1 and then making decisions based on negative accumulated EFE in Eq. 2, incorporating a softmax function. The active inference agent architecture includes encoder, transition,

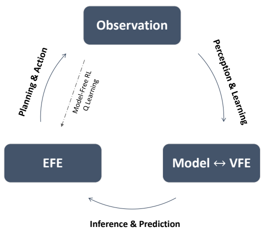

Figure 1. Active Inference Framework

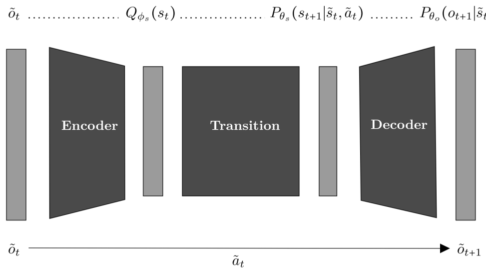

Figure 2. Agent Architecture

and decoder modules (as depicted in Fig. 2), implemented with neural networks. This architecture is probabilistic and similar to a variational autoencoder. Each module generates parameters of pre-selected distributions given a sample. Calculating EFE for all possible trajectories is infeasible, so Fountas et al. (2020) proposed using Monte Carlo Tree Search (MCTS) coupled with an inference action module (i.e., $Q_{\phi_a}(a_t)$). This module approximates the posterior distribution over actions using the prior obtained from the MCTS. Building on the proposed agent by Fountas et al. (2020), we explored various perspectives to design an agent for the EEC task in a manufacturing system. However, there are peculiar features that challenge the original agent's effectiveness. The system is highly stochastic with delays in responding to agent policy, reflected in reward functions that measure average performance (e.g., throughput) over a long horizon. This complicates learning dynamics and planning strategy. Notably, the agent architecture (Fig. 2) predicts one step ahead, often similar to the previous observation due to stochasticity. Therefore, we introduced modifications to tailor the architecture to our problem. To address the limitations of finite horizon EFE, we propose a hybrid architecture that incorporates longer horizons via deep Q-learning. This approach balances short-term EFE and long-term considerations using a hyperparameter, γ. Instead of using the previous habitual structure within Monte Carlo Tree Search (MCTS) as described by Fountas (2020), we trained $Q_{\phi_a}(a_t)$ based on deep Q-learning. Additionally, we modified the transition module to allow multiple steps, controlled by a hyperparameter (e.g., s = 90), enabling multi-step predictions. Due to the computational expense of MCTS, we replaced it with repeated actions in the transition and calculated EFE. This method assesses the impact of actions over a short period, using repeated action simulations at every decision step.

3. Results and Conclusion

Our experiments focused on controlling a real industrial workstation consisting of six parallel-identical machines with finite upstream capacity buffer, as described in (Loffredo et al. 2023). For the reward, similar to (Loffredo et al. 2023), we balanced ($\phi = 0.97$) the ratio of throughput and energy consumption against the *ALL ON* policy over the past 8 hours. We considered both 1-step transitions and multi-step transitions, taking repeated actions during planning for each of the possible policies (i.e., determining how many machines to keep *ON*) to then calculate their EFE. We tested the performance of our agents 50 times during different training iterations, each on independent systems initialized with a random agent after warm-up. We evaluated the performance of our agents using metrics such as test reward, throughput loss, total energy savings, and energy savings per part percentage compared to the *ALL ON* policy over an 8-hour window. Fig. 3 presents the comparison for a single set of hyperparameters except s and γ, while Table 1 shows the performance and average learned policy distribution for specific γ values. These results demonstrate the efficacy of our introduced modifications. Hyperparameters significantly influence agent performance, highlighting the importance of proper tuning for notable improvements and effective control. This underscores the potential of our proposed methodology for EEC applications. Importantly, the framework and formalism of active inference agents exhibit notable promise for non-stationary scenarios, where model-free agents may struggle to adapt swiftly. Future work will concentrate on extending experiments and tailoring the methodology for such non-stationary scenarios.

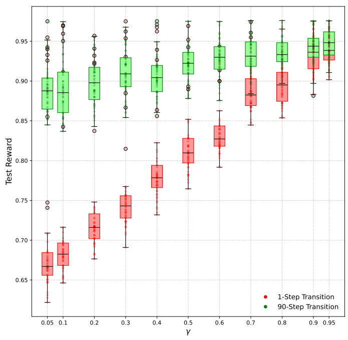

Figure 3. Results

Performance Metrics (%)				
γ	**Reward**	**TH Loss**	**EN Saving**	**EN/Part Saving**
0.05	89.05±0.86	8.88±0.45	10.78±0.50	2.13±0.10
0.3	91.11±0.80	6.71±0.40	8.83±0.47	2.30±0.12
0.6	92.88±0.75	4.84±0.35	7.94±0.42	3.29±0.15

Policy Probabilities (A_n: Keeping n Machines *ON*)							
γ	$P(A_0)$	$P(A_1)$	$P(A_2)$	$P(A_3)$	$P(A_4)$	$P(A_5)$	$P(A_6)$
0.05	0.00±0.00	0.00±0.00	0.07±0.00	0.18±0.01	0.22±0.01	0.27±0.02	0.26±0.01
0.3	0.00±0.00	0.00±0.00	0.05±0.00	0.13±0.01	0.17±0.01	0.38±0.02	0.27±0.01
0.6	0.00±0.00	0.00±0.00	0.03±0.00	0.07±0.01	0.10±0.01	0.57±0.02	0.23±0.01

Table 1. Performance and the Learned Policy for 90-Step Transition.

References

Fountas, Z., Noor Sajid, Pedro A. M. Mediano, Karl J. Friston. 2020. Deep active inference agents using monte-carlo methods. *ArXiv* **abs/2006.04176**.

Friston, Karl. 2010. The free-energy principle: a unified brain theory? *Nature reviews neuroscience* **11**(2) 127–138.

Loffredo, Alberto, Marvin Carl May, Louis Schäfer, Andrea Matta, Gisela Lanza. 2023. Reinforcement learning for energy-efficient control of parallel and identical machines. *CIRP Journal of Manufacturing Science and Technology* **44** 91–103.

Pezzato, Corrado, Carlos Hernández Corbato, Stefan Bonhof, Martijn Wisse. 2023. Active inference and behavior trees for reactive action planning and execution in robotics. *IEEE Transactions on Robotics* **39**(2) 1050–1069.