Meta-Adaptive Decision Transformers Inspired by Human Fast and Slow Cognitive Mechanisms

Matteo D. Rhodes
Department of Computer Science, George Mason University, Fairfax, VA, USA.
mdrhodes@gmu.edu

Matteo Bennett
Department of Computer Science and Engineering, University of Nevada, Reno, Reno, NV, USA.
hellomatteo@unr.edu

Abstract

The deployment of artificial intelligence in complex socio-technical systems requires decision-making architectures that balance rapid responsiveness with deliberative reasoning. This paper introduces a meta-adaptive framework for decision transformers that emulates the dual-process theory of human cognition, distinguishing between fast, intuitive, and slow, analytical modes of processing. The proposed architecture integrates an adaptive meta-controller that dynamically modulates between a lightweight transformer operating in a fast inference mode and a deeper, computationally intensive transformer dedicated to reflective reasoning. This design is motivated by the cognitive science literature on bounded rationality and the limitations of monolithic neural architectures when faced with shifting environmental demands, adversarial perturbations, or resource constraints. We examine the structural trade-offs inherent in such hybrid systems, including latency-accuracy profiles, memory overhead, and energy consumption in edge and cloud deployments. From a governance perspective, the framework offers interpretability advantages by separating intuitive outputs from reasoned justifications, thereby facilitating audit and oversight mechanisms. Fairness implications are addressed through the meta-controller's capacity to allocate cognitive resources equitably across diverse user populations and contexts. We further analyze robustness against distributional shift and strategic manipulation, drawing on insights from adversarial machine learning and causal inference. The paper situates the meta-adaptive approach within the broader landscape of infrastructure-scale AI systems, discussing deployment strategies for distributed sensor networks, autonomous fleets, and critical infrastructure monitoring. Sustainability considerations related to computational carbon footprint and hardware lifecycles are evaluated. Finally, we outline policy implications for regulatory frameworks that require adaptive compliance across jurisdictions and operational domains. This work contributes a systems-level blueprint for next-generation decision transformers that are not only performant but also trustworthy, equitable, and sustainable.

Keywords

dual-process cognition, decision transformer, meta-adaptive architecture, fast and slow thinking, socio-technical systems, AI governance, fairness, robustness, sustainability.

1. Introduction

Artificial intelligence systems are increasingly entrusted with autonomous decision-making in domains ranging from autonomous driving and healthcare diagnosis to energy grid management and financial trading. These environments demand both rapid reflexes to avoid imminent hazards and careful deliberation for long-term planning. Yet most contemporary deep reinforcement learning and transformer-based decision models operate under a monolithic inference paradigm, processing every input with uniform computational effort regardless of urgency or complexity. This one-size-fits-all approach leads to suboptimal resource utilization, brittleness under distributional shift, and opacity in reasoning chains. The human mind, by contrast, has evolved a dual-process architecture that delegates cognitive tasks to two distinct systems: System 1, which is fast, automatic, and associative, and System 2, which is slow, deliberate, and analytical [1]. Inspired by this biological blueprint, we propose a meta-adaptive decision transformer that dynamically switches between a fast inference path and a slow reasoning path based on contextual signals.

The core contribution of this paper is a system-level framework that integrates a meta-controller capable of arbitrating between two transformer-based modules: a lightweight, low-latency transformer optimized for pattern matching and habitual responses, and a deep, high-capacity transformer that performs structured causal reasoning and counterfactual simulation. The meta-controller itself is trained via a reinforcement learning loop that learns to trade off speed, accuracy, and computational cost across diverse task distributions. This design enables the overall system to achieve better Pareto-optimality along multiple dimensions of performance and efficiency than either module alone.

From an architectural perspective, the meta-adaptive decision transformer presents novel challenges in system integration, memory sharing, and asynchronous execution. We discuss how the fast and slow modules can share a common latent state representation to minimize redundancy while maintaining independent processing pipelines. The meta-controller may operate at different granularities, from per-token decisions in sequence generation to per-episode mode selection in reinforcement learning settings. We also consider the implications for hardware acceleration, noting that fast modules can be deployed on edge devices while slow modules reside in cloud clusters, creating a hierarchical computing infrastructure.

The paper proceeds as follows. Section 2 reviews related work in dual-process cognitive architectures and decision transformers. Section 3 details the proposed meta-adaptive framework, including the meta-controller design, module specifications, and training regimen. Section 4 analyzes structural trade-offs and performance characteristics. Section 5 addresses robustness, fairness, and governance implications. Section 6 discusses deployment and infrastructure considerations, including sustainability and policy. Section 7 concludes with a synthesis and future research directions.

2. Related Work

The dual-process theory of cognition has been a cornerstone of cognitive science for decades, with Kahneman and Tversky [2] providing foundational empirical evidence for the distinction between intuitive and deliberative reasoning. In artificial intelligence, this dichotomy has inspired numerous hybrid architectures, such as the work on metacognitive loops [3] and bounded optimality frameworks [4]. More recently, decision transformers [5] have emerged as a powerful paradigm that frames sequential decision-making as a sequence modeling problem, leveraging the transformer architecture's ability to capture long-range dependencies through self-attention. However, standard decision transformers apply uniform computational resources to each timestep, ignoring the varying cognitive load required across different contexts.

Several attempts have been made to introduce adaptive computation in transformers. Adaptive computation time [6] allows recurrent networks to adjust the number of processing steps per input. Mixture-of-experts models [7] route different inputs to different subnetworks, achieving sparsity and efficiency. Yet these approaches lack a clear separation into fast and slow reasoning pathways with distinct cognitive roles. Cognitive architectures such as ACT-R [8] and SOAR [9] incorporate explicit production rules and working memory buffers, but their rigid modularity limits scalability to high-dimensional continuous control tasks.

The concept of meta-adaptation has been explored in meta-reinforcement learning [10], where agents learn to adjust their learning strategies based on task history. Our work extends this idea to the architectural level, where the meta-controller not only chooses policies but also selects which cognitive subsystem to activate. Recent work on dual-system decision transformers includes the framework proposed by Dou et al. [13], which integrates fast and slow thinking into a decision-making pipeline using adaptive inference. Their approach demonstrates improved performance in tasks requiring both immediate reaction and strategic planning. However, the present paper focuses on system-level implications such as infrastructure, governance, and sustainability, which are not addressed in that work.

Other relevant research includes studies on interpretability in transformer models [11], fairness-aware reinforcement learning [12], and energy-efficient AI hardware [14]. The meta-adaptive framework provides a natural interface for incorporating these considerations, as the fast module can be designed for low-power execution while the slow module includes explicit fairness constraints and explanation generators.

3. Meta-Adaptive Decision Transformer Framework

The proposed architecture consists of three primary components: a fast decision transformer module, a slow decision transformer module, and a meta-controller that orchestrates their operation. The fast module is a distilled version of a full transformer, using fewer layers, reduced hidden dimensions, and approximate attention mechanisms such as linear or low-rank attention. Its training objective mirrors the original decision transformer loss, but with an emphasis on minimizing inference latency and maximizing throughput. The slow module is a full-scale transformer with deep encoder-decoder stacks, capable of performing iterative self-critique, counterfactual reasoning, and value decomposition. Both modules share a common input embedding layer and output projection head to maintain consistency in state and action spaces.

The meta-controller is a lightweight neural network that observes a window of recent context—including input features, previous actions, internal state statistics of both modules, and external environmental signals such as time pressure or resource availability—and outputs a gating signal. This gating signal can be binary (fast vs. slow mode) or continuous, allowing soft blending of the two modules' outputs. Training the meta-controller involves a reward function that penalizes latency, rewards accuracy, and includes a regularization term for energy consumption. The meta-controller is trained jointly with the two transformer modules using a multi-objective reinforcement learning algorithm, where the overall objective is to maximize expected cumulative reward subject to constraints on computational budget.

One critical design decision is whether the meta-controller should be local (per-timestep) or global (per-episode). Local modulation allows the system to switch rapidly within a single decision sequence—for example, responding automatically to sudden braking in traffic but then engaging slow reasoning for route planning. Global modulation is simpler and suits domains without rapid temporal variation. We advocate a hierarchical scheme: a global meta-controller sets a high-level policy for each episode, and a local meta-controller refines decisions at each timestep based on real-time signals. This hierarchical approach mirrors the brain's dual-process interaction, where System 1 can override System 2 during emergencies.

Memory and state management pose additional challenges. The fast and slow modules may maintain separate hidden states, but integrating them requires careful synchronization to avoid inconsistent world models. We propose a shared episodic memory buffer that records salient experiences, accessible by both modules. The slow module can asynchronously update the buffer with refined representations, which the fast module can then leverage for quick pattern matching. This asymmetric update mechanism resembles consolidation processes in human memory.

4. Structural Trade-Offs and Performance Analysis

The meta-adaptive architecture introduces several trade-offs that must be carefully managed. The most obvious is between latency and accuracy. Fast module outputs are available within milliseconds but may be suboptimal in novel or adversarial situations. Slow module reasoning can take orders of magnitude longer but achieves higher decision quality in complex scenarios. The meta-controller's accuracy in detecting when to switch is therefore paramount. Over-reliance on the fast module can lead to catastrophic failures under distributional shift; excessive invocation of the slow module degrades real-time performance.

Empirical evaluations from simulated autonomous driving and grid control environments show that a well-trained meta-controller achieves near-optimal accuracy while reducing average inference time by 40-60% compared to using the slow module exclusively [13]. However, these gains depend on the diversity of training tasks. If the task distribution is narrow, the meta-controller may overfit to specific switching patterns, reducing robustness.

Memory overhead is another critical factor. Maintaining two transformer models in memory can double storage requirements. However, weight sharing and low-rank approximations can compress the fast module to as little as 10% of the slow module's parameters. Deployment-wise, the fast module can reside on edge devices with limited RAM, while the slow module is hosted on cloud servers, leading to network latency for offloaded reasoning. The meta-controller must therefore incorporate communication delays into its switching decisions. In high-stakes applications like autonomous driving, offloading to the cloud may be unacceptable, necessitating a fully on-board slow module that requires more powerful hardware.

Energy consumption is a growing concern for large-scale AI deployments. The fast module's computational efficiency translates directly to reduced carbon footprint. A meta-adaptive system can dynamically allocate more energy-intensive reasoning only when necessary, aligning with principles of green AI [15]. However, the meta-controller itself consumes energy; if it is too complex, the overhead may offset the savings. A lightweight meta-controller with binary output and a single hidden layer typically adds less than 1% to total energy use.

From a performance standpoint, the system excels in environments with temporal burstiness—periods of low complexity interspersed with critical decision points. In uniformly difficult tasks, the fast module offers little benefit, and the overhead of switching may negate any advantage. Conversely, in uniformly simple tasks, the slow module is rarely needed, and the system reduces to the fast module alone. The meta-adaptive approach is most valuable in heterogeneous, unpredictable environments.

5. Robustness, Fairness, and Governance

Robustness is a defining challenge for AI systems deployed in the wild. A monolithic decision transformer can be fooled by adversarial perturbations that exploit its uniform processing [16]. The meta-adaptive architecture provides inherent robustness through diversification: an attack that succeeds against the fast module may fail against the slow module, and the meta-controller can learn to rely on the slower reasoning when it detects anomalies. For example, the fast module might be susceptible to small input perturbations that cause aberrant attention patterns, while the slow module's iterative self-attention can filter out such noise. The meta-controller can be trained with adversarial examples to improve its discriminative capability.

Fairness considerations arise because the fast module, being a distilled model, may systematically underperform for underrepresented groups if the training data is imbalanced. The slow module, with higher capacity, can correct such biases by performing more deliberative reasoning on sensitive features. However, this creates a fairness-efficiency trade-off: the system may be less likely to allocate slow reasoning to minority users due to resource constraints, exacerbating disparities. To address this, the meta-controller's reward function can include a fairness term that penalizes differential allocation based on demographic attributes. Alternatively, the meta-controller can enforce a minimum budget of slow reasoning for all inputs, ensuring equitable access to deliberative processing.

Governance of such systems requires transparency in decision processes. The dual-module architecture naturally supports explainability: fast decisions can be accompanied by a default justification, while slow decisions can include a full trace of reasoning steps, causal graphs, and counterfactual alternatives. This aligns with emerging regulations such as the EU AI Act, which mandates explanation for high-risk systems [17]. Auditors can examine the ratio of fast to slow decisions and the contexts in which switches occur, providing insight into the system's behavior under stress.

Another governance concern is alignment with human values. The meta-controller's reward function must be carefully specified to avoid unintended optimization of proxy metrics. For instance, if latency is heavily penalized, the system might choose fast decisions even when catastrophic failures are likely. Incorporating safety constraints as hard limits rather than soft penalties helps mitigate this risk. The framework also enables dynamic policy adaptation: when new regulations are issued (e.g., stricter fairness requirements), the meta-controller's objective can be updated without retraining entire modules, by adjusting the reward weights. This makes the system more responsive to evolving governance landscapes.

6. Deployment, Infrastructure, and Sustainability

Deploying a meta-adaptive decision transformer at scale involves multiple layers of infrastructure. At the edge, low-power microcontrollers or mobile GPUs can run the fast module with minimal latency. When the meta-controller determines that slow reasoning is needed, it can either trigger an on-device slow module (if hardware permits) or send a compressed representation to a cloud server for deeper analysis. This edge-cloud synergy requires reliable network connectivity and low-latency data transmission. For mission-critical applications such as autonomous drones or medical robots, edge-only deployment with a trimmed slow module may be required, at the cost of reduced reasoning depth.

Federated learning can be used to train the meta-controller and both modules across distributed devices without centralizing sensitive data [18]. This is especially relevant in healthcare and finance, where data privacy regulations restrict data movement. The meta-controller's switching policy can be personalized per device, adapting to local conditions (e.g., network bandwidth, battery level) while still benefiting from collective knowledge.

Sustainability is a pressing concern for AI infrastructure. The computational cost of training large transformers is enormous, but the meta-adaptive approach can mitigate operational energy consumption. By reducing the average number of floating-point operations per decision, the carbon footprint over the system's lifetime decreases significantly. Moreover, the fast module can be trained with knowledge distillation from the slow module, which itself may be trained once and reused across deployments. This amortizes the training cost over many inference instances.

Hardware design can further exploit the dual-module architecture. Specialized chips such as neuromorphic processors are well-suited for fast, energy-efficient inference, while traditional GPUs or TPUs handle the slow reasoning. The meta-controller can act as a hardware scheduler, waking up the high-power accelerator only when needed. This approach extends the lifetime of hardware components by reducing thermal stress from continuous high-load operation.

Policy implications for sustainability include incentives for using adaptive computation. Carbon tax schemes for AI operations could be designed to reward systems that dynamically adjust complexity. Standardization bodies such as the IEEE may develop benchmarks for energy-aware decision transformers, and our framework provides a template for meeting such standards.

7. Conclusion

This paper has presented a meta-adaptive decision transformer inspired by human fast and slow cognitive mechanisms, offering a comprehensive system-level analysis of its architecture, trade-offs, robustness, fairness, governance, and sustainability. By introducing a meta-controller that dynamically switches between a fast, intuitive transformer and a slow, deliberative transformer, the framework achieves superior performance efficiency in heterogeneous environments while maintaining interpretability and accountability. The analysis highlights that the success of such a system hinges on careful design of the meta-controller's reward function, memory architecture, and deployment infrastructure. Future research directions include extending the framework to multi-agent contexts where each agent has its own dual modules and coordination mechanisms, as well as incorporating online learning for the meta-controller to adapt to never-before-seen distributions. Empirical validation on large-scale benchmarks is a necessary next step, along with real-world pilot studies in autonomous driving, smart grid management, and clinical decision support. The meta-adaptive approach represents a significant step toward AI systems that are not only intelligent but also responsible, efficient, and aligned with human cognitive architecture.

References

1. Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux.

2. Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185(4157), 1124–1131.

3. Cox, M. T. (2005). Metacognition in computation: A selected research review. Artificial Intelligence, 169(2), 104–141.

4. Russell, S. J., & Wefald, E. H. (1991). Do the right thing: Studies in limited rationality. MIT Press.

5. Chen, L., Lu, K., Rajeswaran, A., Lee, K., Grover, A., Laskin, M., ... & Mordatch, I. (2021). Decision transformer: Reinforcement learning via sequence modeling. Advances in Neural Information Processing Systems, 34, 15084–15097.

6. Graves, A. (2016). Adaptive computation time for recurrent neural networks. arXiv preprint arXiv:1603.08983.

7. Shazeer, N., Mirhoseini, A., Maziarz, K., Davis, A., Le, Q., Hinton, G., & Dean, J. (2017). Outrageously large neural networks: The sparsely-gated mixture-of-experts layer. arXiv preprint arXiv:1701.06538.

8. Anderson, J. R. (1996). ACT: A simple theory of complex cognition. American Psychologist, 51(4), 355–365.

9. Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987). SOAR: An architecture for general intelligence. Artificial Intelligence, 33(1), 1–64.

10. Wang, J. X., Kurth-Nelson, Z., Tirumala, D., Soyer, H., Leibo, J. Z., Munos, R., ... & Botvinick, M. (2016). Learning to reinforcement learn. arXiv preprint arXiv:1611.05763.

11. Jain, S., & Wallace, B. C. (2019). Attention is not explanation. Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics, 3543–3556.

12. D'Amour, A., Srinivasan, H., Atwood, J., Baljekar, P., Sculley, D., & Halpern, Y. (2020). Fairness is not static: Deeper understanding of long-term fairness via simulation studies. Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, 525–534.

13. Dou, Z., Cui, D., Yan, J., Wang, W., Chen, B., Wang, H., ... & Zhang, S. (2025). Dsadf: Thinking fast and slow for decision making. arXiv preprint arXiv:2505.08189.

14. Horowitz, M. (2014). Computing's energy problem (and what we can do about it). 2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers, 10–14.

15. Schwartz, R., Dodge, J., Smith, N. A., & Etzioni, O. (2020). Green AI. Communications of the ACM, 63(12), 54–63.

16. Goodfellow, I. J., Shlens, J., & Szegedy, C. (2015). Explaining and harnessing adversarial examples. International Conference on Learning Representations.

17. European Commission. (2021). Proposal for a regulation laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). COM/2021/206 final.

18. McMahan, B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). Communication-efficient learning of deep networks from decentralized data. Artificial Intelligence and Statistics, 1273–1282.

19. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. Neural Networks, 61, 85–117.

20. Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798–1828.

21. Sutton, R. S., & Barto, A. G. (2018). Reinforcement learning: An introduction (2nd ed.). MIT Press.

22. Neys, W. D. (2006). Dual processing in reasoning: Two systems but one reasoner. Psychological Science, 17(5), 428–433.

23. Mitchell, T. M. (1997). Machine learning. McGraw-Hill.

24. Stone, P., & Veloso, M. (2000). Multiagent systems: A survey from a machine learning perspective. Autonomous Robots, 8(3), 345–383.

25. Russell, S. J. (2019). Human compatible: Artificial intelligence and the problem of control. Viking.