

AI-Based Thinking: Fast and Slow

Adnan Agbaria¹, Yael Dubinsky²

¹Department of Software Engineering, Kinneret Academic College, Zemach, Israel

²Faculty of Technology, Ramat-Gan Academic College, Ramat Gan, Israel

Email: adnan@mx.kinneret.ac.il, yael.d@iac.ac.il

How to cite this paper: Agbaria, A. and Dubinsky, Y. (2026) AI-Based Thinking: Fast and Slow. *Journal of Computer and Communications*, 14, 136-147.
<https://doi.org/10.4236/jcc.2026.146010>

Received: March 2, 2026

Accepted: June 27, 2026

Published: June 30, 2026

Abstract

This paper introduces AI-based Thinking, a dual-process framework inspired by Kahneman's theory of Fast and Slow cognition. We formalize the distinction between intuitive, pattern-driven inference (Fast Thinking) and deliberate, structured reasoning (Slow Thinking) in artificial intelligence systems. We propose a modular architecture that integrates machine learning models with knowledge-based reasoning systems through a meta-cognitive orchestration layer. Unlike traditional AI pipelines, our framework dynamically balances computational efficiency with contextual depth, enabling adaptive reasoning across domains. A case study based on the SmartAI architecture demonstrates the practical implementation of this framework in scene understanding. We further discuss broader implications for medical diagnosis, cybersecurity, and educational AI systems, and identify open challenges in engineering dual-process artificial intelligence.

Keywords

AI-Based Thinking, Dual-Process AI Framework, Neuro-Symbolic Orchestration

1. Introduction

Artificial Intelligence (AI) has achieved remarkable progress in recent years, driven primarily by advances in machine learning (ML), deep neural networks, and large language models (LLMs). These systems demonstrate extraordinary capabilities in pattern recognition, language generation, perception, and decision-making tasks. However, despite their impressive performance, contemporary AI systems remain fundamentally limited in their ability to flexibly combine rapid intuitive inference with structured, context-aware reasoning. This limitation reflects a deeper architectural gap: modern AI excels at fast pattern-based processing, but struggles with deliberate analytical thinking.

Human cognition offers a useful lens through which to understand this imbalance. Daniel Kahneman’s dual-process theory distinguishes between **System 1**, which is fast, automatic, and intuitive, and **System 2**, which is slow, deliberate, and analytical. While System 1 enables rapid judgments and pattern recognition, System 2 supports structured reasoning, abstraction, and contextual evaluation. Effective human intelligence emerges not from either system alone, but from their dynamic interaction.

Current AI systems disproportionately emphasize mechanisms similar to System 1. Deep neural networks and LLMs perform inference through high-dimensional statistical pattern matching, producing rapid responses that resemble intuitive reasoning. Although these models can generate coherent and contextually rich outputs, their reasoning remains largely implicit and opaque. They may hallucinate, overlook logical inconsistencies, or fail in tasks requiring structured multi-step reasoning.

In contrast, the classical Knowledge Representation and Reasoning (KRR) approaches [1] [2] embody characteristics of System 2 thinking. Symbolic reasoning systems, semantic networks, and logic-based frameworks provide explicit representations of knowledge, supporting traceable inference, and structured decision-making. However, these systems typically lack the scalability, adaptability, and perceptual strength of modern neural models.

Generative AI models, such as GPT, have pushed the boundaries of ML by showcasing exceptional capabilities in natural language generation, rapid decision-making, and pattern recognition [3]. **Figure 1(a)** highlights the features achievable by ML models, organized by complexity, effort, and sequence. Despite substantial progress, achieving higher cognitive functions like “wisdom” requires incorporating KRR [4] [5]. **Figure 1(b)** illustrates wisdom at the apex of AI capabilities, underscoring its reliance on structured reasoning, context-awareness, and logical decision-making.

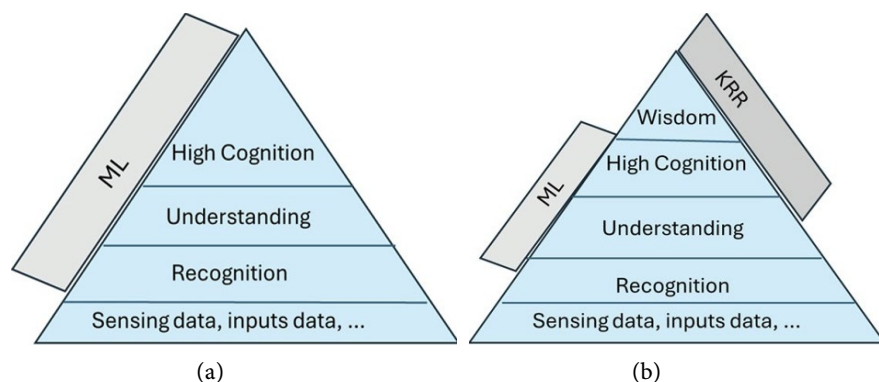


Figure 1. Listed features that can be achieved using AI. (a) Features with ML; (b) For wisdom, KRR mostly required.

This paper proposes AI-based Thinking, a dual-process framework that formalizes the integration of Fast (ML-driven) and Slow (reasoning-driven) mechanisms

within artificial systems. Rather than treating neural and symbolic approaches as competing paradigms, we conceptualize them as complementary cognitive processes. We argue that the next generation of AI systems must incorporate an explicit orchestration layer capable of dynamically balancing intuitive inference with analytical reasoning.

Building on our previous SmartAI architecture [6], which demonstrated the integration of large language models and semantic reasoning for scene understanding, this work generalizes the concept to a broader cognitive model. We extend the architectural perspective into a theoretical framework that 1) maps dual-process theory to AI systems, 2) defines the characteristics of Fast and Slow AI thinking, and 3) proposes mechanisms for meta-cognitive orchestration between them.

We demonstrate how this framework enables improved contextual reasoning in complex tasks and discuss its implications for domains such as healthcare, cybersecurity, education, and autonomous systems. Finally, we outline key research challenges in the engineering of dual-process AI, including system switching criteria, computational efficiency, and evaluation metrics for artificial reasoning.

By formalizing AI-based Thinking, this work aims to bridge the gap between high-performance statistical models and structured symbolic intelligence, advancing the development of AI systems capable of more flexible, context-aware, and cognitively inspired behavior.

2. Related Work

The distinction between fast, intuitive processing and slow, deliberate reasoning originates from Kahneman's dual-process theory [7], which differentiates between automatic, heuristic-based cognition (System 1) and analytical, rule-based reasoning (System 2). Although developed within cognitive psychology, this framework has increasingly influenced artificial intelligence research. Booch *et al.* [1] explicitly discuss the relevance of "Thinking Fast and Slow" in AI, arguing that contemporary systems predominantly emphasize rapid statistical inference while lacking robust mechanisms for structured analytical reasoning. Their discussion highlights the need for architectures capable of integrating reactive and deliberative processes, yet it remains largely conceptual without proposing a concrete modular implementation.

Advances in machine learning, particularly in deep learning and generative models, have significantly strengthened the System 1-like capabilities of AI systems. Neural models demonstrate strong pattern recognition and rapid inference performance in perception and language tasks [3]. However, surveys on knowledge-enhanced neural reasoning emphasize that purely neural approaches often struggle with logical consistency, interpretability, and structured reasoning in complex domains [2]. These limitations reveal the challenges of relying exclusively on data-driven statistical inference when contextual depth and explicit reasoning are required.

KRR provides structured mechanisms for representing and manipulating ex-

plicit knowledge. The foundational work by Brachman and Levesque [5] formalized the theoretical principles underlying symbolic reasoning systems. Semantic networks, as described by Sowa [8], offer expressive frameworks for modeling relationships among entities, while marker-passing techniques and multi-context mechanisms enable efficient inference across interconnected knowledge structures [9] [10]. These approaches align closely with the characteristics of System 2 thinking, focusing on explicit structure, traceable inference, and contextual awareness. Applications of semantic knowledge-based reasoning have demonstrated effectiveness in collaborative and dynamic environments [11], and logic-based neural integration approaches have combined perceptual inputs with structured knowledge to support semantic interpretation tasks [5]. Research on context-aware systems similarly underscores the necessity of structured reasoning to handle dynamic contextual information [12] [13].

In related work, Chain-of-Thought (CoT) prompting has emerged as a promising strategy to improve the reasoning capabilities of large language models (LLMs) [14]. CoT prompting aligns closely with System 2 thinking, emphasizing step-by-step structured reasoning to enable logical progression and traceability. By incorporating intermediate reasoning steps, this method mirrors the deliberate and analytical nature of human cognition. Research frameworks further suggest that the incorporation of systematic search processes, process supervision, and reinforcement learning within LLMs can improve their ability to handle complex reasoning tasks [15]. Although these advancements improve reasoning performance within neural architectures, they remain embedded in statistical sequence modeling and do not explicitly integrate independent symbolic reasoning systems.

Our previous SmartAI framework demonstrated a modular integration of machine learning models and semantic-network-based reasoning for scene understanding. SmartAI is particularly well-suited for scene understanding—a field of computer vision focused on enabling machines to interpret and comprehend visual scenes in a human-like manner [16]. Scenes often involve intricate relationships between objects, spatial contexts, and activities, which require AI systems to extract meaningful insights from images or video sequences. Although traditional ML models excel at recognizing individual objects, they struggle with the deeper reasoning needed to interpret context, relationships, and underlying intentions. SmartAI addresses these limitations by using ML for fast recognition and KRR for in-depth, context-aware reasoning, ensuring both efficiency and accuracy in scene interpretation. Beyond scene understanding, SmartAI advocates for the integration of specialized technologies rather than isolated solutions. Its architecture treats the ML and KRR components as independent modular systems, orchestrated without internal modifications. Building on state-of-the-art methodologies in context-based and knowledge-driven systems [12] [13], SmartAI establishes a foundation for adaptable, context-based AI systems applicable in domains such as healthcare, education, and autonomous systems. The present work extends this integration into a broader dual-process framework, formalizing Fast and Slow

thinking as complementary architectural components, and emphasizing the role of orchestration in coordinating their interaction.

3. AI-Based Thinking: A Dual-Process Framework

Contemporary artificial intelligence systems are predominantly driven by statistical learning mechanisms that approximate rapid and intuitive processing. Although highly effective for perception, language modeling, and pattern recognition, such systems often lack structured deliberation, explicit reasoning, and consistent contextual awareness. This imbalance has been increasingly discussed in the AI community, particularly in relation to Kahneman's dual-process theory of cognition [7], which distinguishes between fast, automatic processing (System 1) and slow, analytical reasoning (System 2). Booch *et al.* [1] argue that modern AI systems are heavily biased toward fast inference while lacking principled mechanisms for deliberate reasoning. Similarly, recent discussions in AI research emphasize the need to integrate reactive neural computation with structured symbolic reasoning to approximate higher-level cognition [3] [17].

Within this context, we propose AI-based Thinking, a dual-process cognitive framework that models artificial cognition as the interaction between Fast and Slow computational processes. Fast Thinking corresponds to data-driven statistical inference implemented through deep neural networks and large language models. These systems operate via high-dimensional function approximation and next-token prediction mechanisms [18], enabling rapid inference and strong pattern generalization. Chain-of-Thought prompting [14] and self-consistency reasoning [19] have demonstrated that LLMs can simulate stepwise reasoning when appropriately guided. More recent work explores structured search and meta chain-of-thought learning to approximate System 2-like behavior within neural architectures [15]. Tool-augmented language models further extend these capabilities by incorporating external computation modules into inference pipelines [20]. Despite these advances, reasoning within purely neural systems remains fundamentally embedded in statistical sequence modeling, often lacking persistent structured world representations and explicit logical consistency mechanisms.

Slow Thinking, on the contrary, relies on explicit knowledge representation and reasoning mechanisms grounded in symbolic AI traditions [4]. Semantic networks [8], logic-based systems, and multi-context reasoning frameworks [9] [10] provide structured representations that support traceable inference and contextual constraint enforcement. Neuro-symbolic approaches attempt to combine perception and structured reasoning [5] [21], while knowledge-enhanced neural reasoning surveys emphasize the importance of integrating symbolic structure into learning systems [2]. These approaches align closely with the characteristics of System 2 cognition, enabling causal reasoning, compositionality, and systematic generalization [17]. However, tightly coupled neuro-symbolic models often embed reasoning within neural architectures, reducing modularity and interpretability boundaries.

The central contribution of AI-based Thinking is the introduction of an explicit meta-cognitive orchestration layer (Section 4) that coordinates interaction between Fast and Slow processes. Rather than embedding symbolic reasoning into neural networks, the framework maintains modular independence between statistical inference and structured reasoning components. This separation enables the reuse of existing machine learning and knowledge-based systems, supports independent scalability, and preserves interpretability. The orchestration layer evaluates the adequacy, uncertainty, and contextual completeness of the Fast Thinking output. When rapid inference is sufficient, the system returns the result directly. When ambiguity, inconsistency, or high complexity of reasoning is detected, Slow Thinking is activated to perform structured analytical processing. This design parallels the adaptive computation and cognitive control mechanisms discussed in both neuroscience and AI literature [22].

Formally, let $F(x)$ be the output of Fast Thinking for input x , and let $S(x)$ be the output of Slow Thinking. A meta-cognitive controller M determines whether deeper reasoning is required. The system output $O(x)$ is defined as:

$$O(x) = \begin{cases} F(x) & \text{if confidence and contextual criteria are satisfied} \\ \text{Integrated}(F(x), G(x)) & \text{Otherwise} \end{cases} \quad (1)$$

This formulation emphasizes the adaptive allocation of computational effort, mirroring the human cognitive control mechanisms in which deliberation is invoked selectively. By conceptualizing Fast and Slow processes as cooperative yet distinct computational modes governed by meta-cognitive control, AI-based Thinking provides a principled foundation for designing systems capable of balancing efficiency with structured reasoning. The framework generalizes beyond tightly integrated neuro-symbolic models by treating intuitive inference and analytical reasoning as modular cognitive agents, coordinated dynamically to achieve flexible and context-aware artificial intelligence.

4. Proposed Architecture

To realize the AI-based Thinking framework in practice, we propose a modular architecture that explicitly separates Fast and Slow computational processes while coordinating them through a dedicated orchestration layer. This architecture is built on the basis of our previous work [6]. Beyond scene understanding, we generalize the architecture for dual-process artificial intelligence systems.

Figure 2 shows that our proposed architecture consists of three primary components: a Fast Thinking module, a Slow Thinking module, and an Orchestrator. Each component is implemented as an independent subsystem with clearly defined interfaces. The guiding design principle is modular separation, ensuring that machine learning models and knowledge-based reasoning systems operate without internal modification while interacting through structured exchanges of information.

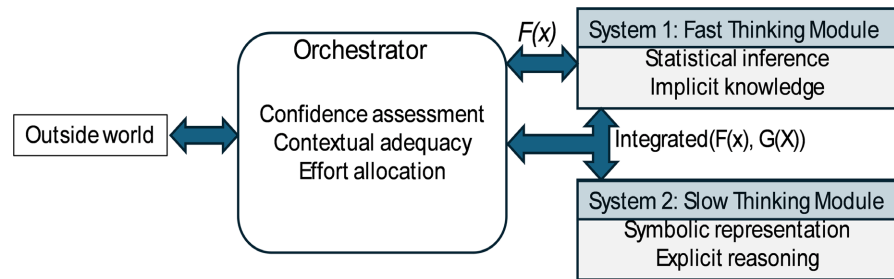


Figure 2. Dual-process AI architecture with meta-cognitive orchestration.

The Fast Thinking module is implemented using machine learning models such as large language models or vision-based neural networks. This component performs rapid inference directly from the input data and produces an immediate interpretation of the task. In visual scenarios, it may generate object detections or textual descriptions of a scene. In language-based tasks, it produces preliminary interpretations or candidate responses. The Fast module leverages statistical generalization learned from large-scale datasets and is optimized for speed and responsiveness. However, its reasoning is implicit within distributed representations and may lack explicit logical structure or contextual depth.

On the other hand, the Slow Thinking module is implemented using a Knowledge Representation and Reasoning (KRR) system. In alignment with SmartAI implementation [6], semantic networks and marker-passing reasoning mechanisms provide structured contextual analysis. The Slow module operates over explicitly represented entities, relations, and constraints, enabling traceable inference and logical consistency. Unlike the Fast module, which implicitly encodes knowledge in learned parameters, the Slow module maintains persistent symbolic structures that support contextual reasoning, relational interpretation, and rule-based validation.

The Orchestrator serves as the coordination mechanism between the two modules. Upon receiving an input query, the Orchestrator forwards it to the Fast module to obtain an initial interpretation. Then, it evaluates the adequacy of this output according to predefined criteria such as completeness, ambiguity detection, domain-specific validation rules, or confidence estimation. If the Fast output sufficiently addresses the query, it is returned directly to the user, ensuring minimal computational overhead. If uncertainty or insufficient contextual depth is detected, the Orchestrator generates a structured query for the Slow module, using the Fast output as contextual input. The final response may consist of the refined output of the Slow module or of an integration of the results from both modules.

This workflow preserves computational efficiency while enabling deeper reasoning when required. The architecture mirrors the allocation of adaptive cognitive effort observed in human decision-making, where intuitive judgments are accepted when reliable, and analytical reasoning is invoked when complexity increases. By selectively locating structured reasoning, the system balances responsiveness with contextual precision.

A critical feature of the proposed architecture is that integration occurs at the orchestration level rather than within the internal mechanisms of the modules. Neural models are not modified to embed symbolic logic, and symbolic systems are not required to replicate perceptual inference. Each component operates within its natural computational paradigm. This separation enhances modularity, scalability, and interpretability, allowing different implementations of Fast and Slow modules to be substituted without redesigning the overall system.

This architecture was applied in [6] to understanding the scene. The Fast module generated a descriptive interpretation of a visual scene, while the Slow module leveraged semantic knowledge to refine contextual interpretation, such as identifying the meaning of a traffic sign through structured reasoning. Initially, human operators fulfilled the orchestration role by transforming Fast outputs into targeted reasoning queries. The present framework formalizes this orchestration as a computational component capable of automated confidence assessment and reasoning activation.

The proposed architecture therefore serves as a domain-agnostic template for implementing dual-process AI systems. By preserving modular independence while enabling structured interaction, it operationalizes the conceptual Fast and Slow framework and provides a practical pathway toward context-aware, adaptable artificial intelligence.

5. Case Study: SmartAI for Scene Understanding

To demonstrate the practical realization of the proposed dual-process architecture, we revisit the SmartAI system [6] as a case study in scene understanding. This example illustrates how Fast and Slow processes interact within a concrete implementation and highlights the benefits of meta-cognitive orchestration.

Scene understanding is a fundamental task in computer vision that requires not only object recognition but also contextual interpretation of relationships, spatial arrangements, and semantic meaning [16]. Although modern machine learning models excel at detecting objects and generating descriptive captions, they often struggle with deeper contextual reasoning, especially when interpretation requires domain-specific knowledge or relational inference.

In the SmartAI implementation, the Fast Thinking module was realized using a large language model capable of generating textual descriptions of input images. Upon receiving a visual scene, the Fast module produced an immediate description summarizing visible objects and environmental context. This output reflects the statistical inference learned from large-scale training data and corresponds to System 1-like processing within the proposed framework.

However, Fast inference alone proved to be insufficient for a precise semantic interpretation in certain scenarios. In an illustrative example, an image depicting a curved forest road with a traffic sign was submitted to the system. The Fast module correctly identified the road, trees, and the presence of a warning sign, yet failed to determine the exact semantic meaning of the sign. Although the descrip-

tion was coherent and plausible, it lacked contextual specificity. The Slow Thinking module, implemented using a semantic-network-based KRR system, addressed this limitation. Using explicitly encoded knowledge about traffic signs and road semantics, the Slow module performed structured reasoning over the initial description generated by the Fast module. Through marker-passing inference across the semantic network, the system identified the correct interpretation of the sign as indicating a slight right curve ahead.

Within this workflow, the Orchestrator played a central role. It evaluated the adequacy of the Fast module's output and detected that the description lacked sufficient semantic precision. Consequently, it activated the Slow module to perform deeper contextual reasoning. The final output integrated the perceptual strengths of the Fast module with the structured inference of the Slow module, producing a more accurate and contextually grounded interpretation.

This case study illustrates several key properties of the proposed architecture. First, it demonstrates modular cooperation without internal modification of either subsystem. The Fast module operated purely as a neural inference engine, while the Slow module maintained independent symbolic reasoning structures. Second, it shows selective activation of structured reasoning, preserving efficiency when Fast inference is sufficient and invoking analytical processing only when needed. Third, it highlights the complementary strengths of statistical and symbolic approaches in complex perceptual tasks.

Importantly, this example extends beyond a single application domain. The same interaction pattern can be generalized to other settings. In medical diagnostics, for example, Fast inference may detect patterns of symptoms from patient data, while Slow reasoning may apply structured differential diagnosis rules. In cybersecurity, anomaly detection can serve as rapid inference, while rule-based reasoning reconstructs attack chains. The SmartAI scene-understanding case thus serves as a concrete instantiation of the broader AI-based Thinking framework.

By grounding the dual-process architecture in a practical implementation, this case study validates the feasibility of separating intuitive and analytical components while coordinating them through orchestration. It demonstrates that dual-process AI is not just a conceptual analogy but a deployable design principle capable of enhancing contextual reasoning in real-world scenarios.

6. Conclusions

This paper introduced AI-based Thinking, a dual-process cognitive framework for AI inspired by Kahneman's distinction between Fast and Slow cognition. Although modern AI systems have achieved remarkable success through large-scale statistical learning, they remain predominantly optimized for rapid pattern-based inference. Such systems often lack structured analytical reasoning, contextual consistency, and adaptive cognitive control. The proposed framework addresses this imbalance by explicitly modeling artificial cognition as the interaction between two complementary processes: Fast Thinking, implemented through ma-

chine learning models, and Slow Thinking, implemented through structured knowledge representation and reasoning mechanisms.

Unlike tightly coupled neuro-symbolic systems, the proposed approach maintains modular independence between statistical inference and symbolic reasoning components, coordinating them through a meta-cognitive orchestration layer. This design preserves scalability and reusability while enabling adaptive allocation of computational effort. By activating deeper reasoning only when necessary, the architecture balances efficiency with contextual depth, reflecting the principles observed in human cognitive control.

The SmartAI system, originally introduced for scene understanding [6], served as a case study demonstrating the feasibility of this dual-process realization. Through the integration of large language models and semantic-network-based reasoning, SmartAI illustrated how intuitive perceptual inference can be refined through structured contextual reasoning. The case study validated the architectural principle that Fast and Slow processes can operate as independent but cooperative modules within a unified system.

Beyond scene understanding, AI-based Thinking provides a general design pattern for context-aware artificial systems. The framework can extend to domains such as healthcare, cybersecurity, education, and autonomous systems, where rapid inference must often be complemented by structured analytical reasoning. By formalizing the interaction between intuitive and deliberative components, this work moves toward AI systems capable of more flexible, explainable, and adaptive behavior.

Several open research challenges remain. First, determining optimal criteria for triggering Slow Thinking requires principled confidence estimation and uncertainty modeling. Second, scalable construction and maintenance of knowledge bases remain a practical challenge for symbolic reasoning components. Third, evaluating dual-process AI systems demands new metrics that capture not only accuracy but also reasoning quality, contextual adequacy, and computational efficiency. Finally, future work may explore automated learning of orchestration policies, enabling systems to dynamically adapt reasoning strategies based on task complexity and environmental feedback.

By bridging statistical learning with structured reasoning under a unified cognitive framework, this framework offers a step toward more balanced and human-like AI. Rather than viewing neural and symbolic paradigms as competing approaches, this work demonstrates that their coordinated integration can yield systems capable of both speed and depth in real-world decision-making.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Booch, G., Fabiano, F., Horesh, L., Kate, K., Lenchner, J., Linck, N., *et al.* (2021)

- Thinking Fast and Slow in AI. *Proceedings of the AAAI Conference on Artificial Intelligence*, **35**, 15042-15046. <https://doi.org/10.1609/aaai.v35i17.17765>
- [2] Chowdhury, T., Ling, C., Zhang, X.C., Zhao, X.J., Bai, G.J., Pei, J., Chen, H.F. and Zhao, L. (2023) Knowledge-Enhanced Neural Machine Reasoning: A Review. <https://doi.org/10.48550/arXiv.2302.02093>
 - [3] Goyal, A. and Bengio, Y. (2020) Inductive Biases for Deep Learning of Higher-Level Cognition. <https://doi.org/10.48550/arXiv.2011.15091>
 - [4] Brachman, R. and Levesque, H. (2004) Knowledge Representation and Reasoning. Elsevier.
 - [5] Donadello, I., Serafini, L. and d'Avila Garcez, A.S. (2017) Logic Tensor Networks for Semantic Image Interpretation. *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence*, Melbourne, 19-25 August 2017, 1596-1602. <https://doi.org/10.24963/ijcai.2017/221>
 - [6] Agbaria, A. and Dubinsky, Y. (2025) SmartAI: Enhancing Scene Understanding by Combining Different AI Technologies. In: *Human Interaction and Emerging Technologies (IHJET-FS 2025): Future Systems and Artificial Intelligence Applications*, AHFE International, 198-205. <https://doi.org/10.54941/ahfe1005968>
 - [7] Kahneman, D. (2011) Thinking, Fast and Slow. Macmillan.
 - [8] Sowa, J.F. (1987) Semantic Networks. In: Shapiro, S.C., Ed., *Encyclopedia of Artificial Intelligence*, Wiley-Interscience, 1493-1511.
 - [9] Fahlman, S.E. (2011) Using Scone's Multiple-Context Mechanism to Emulate Human-Like Reasoning. *Advances in Cognitive Systems. Papers from the 2011 AAAI Fall Symposium (FS-11-01)*, Arlington, 4-6 November 2011, 98-105.
 - [10] Fahlman, S.E. (2012) Beyond Idiot-Savant AI. *Advances in Cognitive Systems*, **1**, 15-22.
 - [11] Akkaladevi, S.C., Plasch, M., Hofmann, M. and Pichler, A. (2021) Semantic Knowledge Based Reasoning Framework for Human Robot Collaboration. *Procedia CIRP*, **97**, 373-378. <https://doi.org/10.1016/j.procir.2020.05.253>
 - [12] Adomavicius, G. and Tuzhilin, A. (2011) Context-Aware Recommender Systems. In: Ricci, F., Rokach, L., Shapira, B. and Kantor, P., Eds., *Recommender Systems Handbook*, Springer US, 217-253. https://doi.org/10.1007/978-0-387-85820-3_7
 - [13] Liu, S., Zhang, X.Y., Sun, X.D., Yin, X.H., Tang, R.C. and Jiang, Q.Y. (2019) Survey on Context-Aware Systems and Their Applications. *2019 IEEE 9th International Conference on Electronics Information and Emergency Communication (ICEIEC)*, Beijing, 12-14 July 2018, 234-237. <https://doi.org/10.1109/iceiec.2019.8784545>
 - [14] Yu, Z.H., He, L., Wu, Z., Dai, X.Y. and Chen, J.J. (2023) Towards Better Chain-of-Thought Prompting Strategies: A Survey. <https://arxiv.org/abs/2310.04959>
 - [15] Xiang, V., Snell, C., Gandhi, K., Albalak, A., Singh, A., Blagden, C., Phung, D., et al. (2025) Towards System 2 Reasoning in LLMs: Learning How to Think with Meta Chain-of-Thought. <https://arxiv.org/abs/2501.04682>
 - [16] Zhou, B., Zhao, H., Puig, X., Xiao, T., Fidler, S., Barriuso, A., et al. (2019) Semantic Understanding of Scenes through the ADE20K Dataset. *International Journal of Computer Vision*, **127**, 302-321. <https://doi.org/10.1007/s11263-018-1140-0>
 - [17] Lake, B.M., Ullman, T.D., Tenenbaum, J.B. and Gershman, S.J. (2016) Building Machines That Learn and Think like People. *Behavioral and Brain Sciences*, **40**, e253. <https://doi.org/10.1017/s0140525x16001837>
 - [18] Brown, T.B., et al. (2020) Language Models Are Few-Shot Learners. *Advances in Neu-*

ral Information Processing Systems 33: Annual Conference on Neural Information Processing Systems 2020, NeurIPS 2020, 6-12 December 2020, 1877-1901.

- [19] Wang, X., Wei, J., Schuurmans, D., Le, Q.V., Chi, E.H., Narang, S., Chowdhery, A. and Zhou, D. (2023) Self-Consistency Improves Chain of Thought Reasoning in Language Models. arXiv preprint arXiv:2203.11171. <https://arxiv.org/pdf/2203.11171>
- [20] Schick, T., Dwivedi-Yu, J., Dessi, R., Raileanu, R., Lomeli, M., Hambro, E., *et al.* (2023) Toolformer: Language Models Can Teach Themselves to Use Tools. *Advances in Neural Information Processing Systems* 36, New Orleans, 10-16 December 2023, 68539-68551. <https://doi.org/10.52202/075280-2997>
- [21] Garcez, A., *et al.* (2015) Neural-Symbolic Learning and Reasoning: Contributions and Challenges. *AAAI Spring Symposia*, Palo Alto, 23-25 March 2015, 18-21.
- [22] Graves, A. (2016) Adaptive Computation Time for Recurrent Neural Networks. <https://arxiv.org/abs/1603.08983>