Key Approaches to AGI Development

A Rahimi

October 26, 2025 (Version R03)

Artificial General Intelligence (AGI) refers to systems capable of understanding, learning, and reasoning across domains—matching or surpassing human-level adaptability. This paper surveys core approaches toward AGI, organized by conceptual lineage, technical foundation, and research trajectory.

1 Symbolic and Logic-Based Approaches

Goal: Encode intelligence through explicit symbols, logic, and rules.
Core Idea: Reasoning and understanding arise from manipulating structured knowledge representations.

Subtype	**Description**	**Examples / Key Projects**
Classical Symbolic AI	Uses formal logic, ontologies, and expert systems.	Cyc, OpenCog (early versions), Logic Theorist
Logic + Probabilistic Integration	Combines symbolic reasoning with uncertainty modeling.	Probabilistic Graphical Models, Bayesian Logic Networks
Neuro-Symbolic Systems	Neural perception + symbolic reasoning for interpretability and generalization.	DeepProbLog, Logic Tensor Networks, JARVIS, NS-CL

Strengths: Explainability, reasoning, abstraction, compositionality.
Weaknesses: Poor perception, limited scalability, brittle generalization.

2 Connectionist / Neural Approaches

Goal: Build general intelligence through learning representations and patterns from data.
Core Idea: Intelligence emerges from large-scale distributed computation akin to neural networks.

Subtype	Description	Examples / Key Projects
Deep Learning Architectures	Large-scale networks trained end-to-end.	GPT, Gemini, Claude, LLaMA, DeepMind Gato
Transformer-based AGI Candidates	Scaled sequence models with multi-modal reasoning.	GPT-4/5, Gemini 2.0, DeepSeek, xAI's Grok
World Models / Predictive Coding	Agents trained to predict and simulate environments.	DreamerV3, MuZero, EfficientZero
Meta-Learning and Self-Improving Models	Models that learn how to learn.	MAML, Reptile, MetaGPT

Strengths: Flexible, scalable, good perception.
Weaknesses: Weak reasoning, opacity, brittleness.

3 Cognitive Architecture Approaches

Goal: Emulate human cognition's modular structure.
Core Idea: Intelligence results from interaction between perception, memory, attention, and reasoning.

Architecture	Features	Examples
SOAR	Problem-space search, chunking learning	University of Michigan's SOAR
ACT-R	Hybrid symbolic + subsymbolic model	Carnegie Mellon ACT-R
Sigma / LIDA / Global Workspace	Cognitive cycles and conscious attention	Baars, Dehaene, Franklin models

Strengths: Human-like modular reasoning; explainable.
Weaknesses: Limited perception; scalability issues.

4 Evolutionary and Emergent Approaches

Goal: Evolve or self-organize intelligence instead of designing it manually.
Core Idea: Intelligence as emergent optimization, competition, or adaptation.

- **Evolutionary Algorithms / NEAT:** Neural architectures evolved via fitness optimization (NEAT, HyperNEAT, AutoML-Zero).

- **Self-Organizing / Complex Systems:** Emergent intelligence from dynamic interactions (Cellular Automata, Friston's Free Energy Principle).
- **Open-Ended Evolution:** Systems designed for unbounded creative growth (POET, Generative Agents, DeepMind's open-endedness studies).

Strengths: Creativity, novelty, diversity.
Weaknesses: Expensive, unstable, unpredictable.

5 Embodied and Interactive AGI

Goal: Develop intelligence through real or simulated world interaction.
Core Idea: Cognition emerges from sensorimotor coupling and adaptive behavior.

- **Embodied Agents / Robotics:** Learning through physical or simulated bodies (Boston Dynamics, DeepMind Control Suite).
- **Simulation-Based Learning:** Agents trained in virtual environments (XLand, MineDojo, Habitat3D).
- **Social / Multi-Agent Systems:** Intelligence from cooperation and competition (Generative Agents, SocialSim, MAgent).

Strengths: Grounded understanding, causal learning.
Weaknesses: Expensive, slow training, hard abstraction.

6 Hybrid and Layered AGI Architectures

Goal: Integrate complementary paradigms for perception, reasoning, and planning.
Core Idea: Hybridization balances flexibility and explainability.

Hybrid Type	**Description**	**Examples**
Neuro-Symbolic Hybrid	Neural perception + symbolic reasoning	DeepProbLog, Logic Tensor Networks, NSCL, OpenCog Hyperon
Neural + Cognitive Architecture	Deep learning with explicit working memory or planning	AlphaZero + Tree Search, Gato, DeepMind AdA
Multi-Agent + Meta-Learning Systems	Networks of specialized agents that learn collaboratively	AutoGPT, MetaGPT, Voyager

Strengths: Balanced reasoning and perception.
Weaknesses: Complex integration, stability challenges.

7 Theoretical and Mathematical Foundations

Goal: Formalize AGI from first principles of computation and optimality.
Core Idea: Intelligence as algorithmic, logical, or probabilistic optimality.

Framework	Description	Examples
AIXI / Universal Intelligence	Solomonoff induction + reinforcement learning	Hutter's AIXI, Gödel Machines
Algorithmic Information Theory (AIT)	Intelligence as compression or Kolmogorov complexity minimization	Schmidhuber's theory, Universal AI
Mathematical AGI Foundations	Formal proofs and reasoning-based intelligence	Gödel Machines, Z3/SMT solvers

Strengths: Theoretical rigor, provable optimality.
Weaknesses: Intractable in practice, abstraction gap.

8 Emerging "Post-Transformer" Paradigms (2024–2025)

Goal: Move beyond static language models to autonomous, self-reflective systems.
Core Idea: Combine reasoning, planning, memory, and meta-learning.

Approach	Description	Examples
Agentic AI Systems	LLMs with persistent goals, tools, and reflection	AutoGPT, Devin, OpenAI o1/o3, Claude Opus
Neuro-Cognitive Loops	Agents that plan, act, and reflect recursively	OpenDevin, Voyager, MemGPT
World Model + Planner Systems	Internal simulation-driven planning	DeepMind Genie, OpenAI Worldscope
Meta-Reasoning Frameworks	Models that improve their own reasoning loops	Self-Reflective Transformers, Reason+Act+Reflect

9 Integrated Path Toward AGI

Modern consensus suggests no single paradigm suffices. A layered integration appears most promising:

Perception (Neural) → Abstraction (Symbolic) → Reasoning (Logic/Planning) → Self-Improvement (Meta-Learning)

This synthesis—often termed *Neuro-Symbolic Cognitive Architecture*—is a dominant conceptual trajectory toward AGI.

10 References

1. Marcus, G. (2022). *The Next Decade in AI: Hybrid Systems for Robust Intelligence.*
2. Schmidhuber, J. (2015). *Deep Learning in Neural Networks: An Overview.*
3. Hutter, M. (2005). *Universal Artificial Intelligence: Sequential Decisions based on Algorithmic Probability.*
4. Bengio, Y. (2024). *System 2 Deep Learning and Consciousness Prior.*
5. DeepMind & OpenAI Technical Reports (2023–2025).
6. Friston, K. (2023). *Free Energy Principle in Active Inference Frameworks.*

Prepared by A Rahimi — Research Draft R03 (2025-10-26).
Part of the AGI Research Program: Exploring Modern Foundations of General Intelligence.