The OaK Architecture: A Paradigm Shift in Artificial General Intelligence

Rich Sutton's Vision for Experience-Based Superintelligence

Author Analysis Based on:

- The OaK Architecture Lecture (August 2025)
- The Alberta Plan for AI Research (2022)
- Settling the Reward Hypothesis (2023)
- Reward is Enough (2021)

Date: September 30, 2025

Executive Summary

Rich Sutton, co-winner of the 2024 Turing Award and a pioneer of reinforcement learning, has proposed a radical reimagining of how we should approach artificial general intelligence. His OaK (Options and Knowledge) architecture represents a fundamental departure from the current paradigm dominated by large language models trained on massive text corpora.

The core thesis: True intelligence emerges not from absorbing vast amounts of human-generated data, but from agents learning continually through direct interaction with their environment, guided solely by reward signals. This isn't just a technical proposal—it's a philosophical stance on the nature of intelligence itself and a critique of the current trajectory of AI research.

Why this matters now:

- The AI industry has reached a potential inflection point where scaling laws may be hitting diminishing returns
- Current AI systems cannot learn continuously without catastrophic forgetting
- LLMs lack true world models and genuine understanding
- Safety and alignment challenges require agents that learn from experience, not just pattern matching

---

Part I: The Problem - Why AI Has "Lost Its Way"

1.1 The Current Paradigm's Limitations

The Scaling Hypothesis Dominance

The past decade of AI has been dominated by a simple but effective strategy: make models bigger, train them on more data, and performance improves. This approach has produced remarkable results with GPT-4, Claude, and similar systems. However, Sutton argues this approach has fundamental limitations:

Lack of Continual Learning: Current AI systems have distinct phases: a massive training period followed by deployment where they cannot learn new things without losing what they already know (catastrophic forgetting). This is fundamentally unlike biological intelligence, where learning never stops and new experiences continuously refine existing knowledge.

Absence of World Models: Large language models are fundamentally predictive text systems. They lack explicit representations of how the world works—the physics, causality, and dynamics that would allow them to truly plan and reason about novel situations. They pattern-match rather than understand.

Inability to Learn from Experience: LLMs learn from human-generated text descriptions of the world, not from direct interaction. This is like trying to learn to ride a bicycle by reading about it—fundamentally incomplete. The richest form of knowledge comes from doing, failing, and adapting.

Computational Waste: The current approach requires retraining from scratch to incorporate new information, wasting enormous computational resources. A truly intelligent system should be able to update its knowledge incrementally and efficiently.

1.2 The Deeper Philosophical Issue

Sutton's critique goes beyond technical limitations to question what intelligence actually is. The current approach assumes intelligence is fundamentally about language and knowledge representation—that if we can predict text well enough, intelligence will emerge.

Sutton argues the opposite: intelligence is fundamentally about **achieving goals in complex environments**. Language, knowledge, perception, and social intelligence are all **tools that evolved to serve reward maximization**, not ends in themselves.

This connects to the "Bitter Lesson" Sutton articulated in 2019: throughout AI's history, approaches that leverage computation and learning consistently outperform those that try to embed human knowledge and insights. We should prioritize methods that scale with compute, not human effort.

Part II: The Solution - The OaK Architecture

2.1 Architectural Overview

The OaK architecture is a **model-based reinforcement learning system** with three distinguishing characteristics:

Feature 1: Universal Continual Learning

Every component learns on every time step. There is no distinction between "training time" and "deployment time." This temporal uniformity means:

- The perception system continually refines how it processes observations
- Value functions continually update their predictions
- The world model continually improves its understanding of dynamics
- Meta-parameters (like learning rates) continually adapt

This is analogous to how biological brains work—you never stop learning, and the mechanisms that enable learning are always active.

Feature 2: Per-Weight Meta-Learning

Each learned parameter has its own dedicated step-size (learning rate) that is **meta-learned using online cross-validation**. This means:

- The system learns how fast to learn for each piece of knowledge
- Some features might need rapid adaptation (for changing aspects of the environment)
- Others might need slow, stable learning (for permanent structures)
- The system automatically discovers the right learning dynamics for each component

This solves one of the most challenging problems in continual learning: the stability-plasticity dilemma. How do you remain plastic enough to learn new things while stable enough not to forget important old things?

Feature 3: The FC-STOMP Progression

The most innovative aspect of OaK is how it creates hierarchical abstractions through a five-step cycle:

Feature Construction (FC):

- The system identifies patterns in its sensory experience
- Creates new features that capture regularities
- These features become the building blocks for higher-level reasoning

SubTask (S):

- High-value features become goals in themselves
- "If I could reliably achieve this feature state, what would that enable?"
- Subtasks respect the main reward signal while pursuing intermediate objectives

Option (O):

- Learn policies (behaviors) to achieve each subtask
- Options are temporally extended actions—sequences of primitive actions
- They have initiation conditions (when to start) and termination conditions (when to stop)

Model (M):

- Learn predictive models of what happens when you execute each option
- "If I execute this option in this state, where will I end up and what reward will I get?"
- These models enable planning without trial-and-error

Planning (P):

- Use the learned option models to simulate possible futures
- Evaluate different sequences of options before committing to actions
- Update policies and value functions based on simulated experience

The Critical Feedback Loop: The genius of FC-STOMP is that it's a **self-reinforcing cycle**. Features that prove useful for planning get reinforced and elaborated. New features lead to new subtasks, which lead to new options, which lead to better models, which enable better planning, which reveals the need for even better features.

This creates an **open-ended process of abstraction creation** that is limited only by computational resources. Each cycle builds on the previous one, creating increasingly sophisticated representations and behaviors.

2.2 Why This Architecture Could Lead to Superintelligence

Compounding Returns on Abstraction: Just as compound interest creates exponential wealth growth, compounding abstraction creates exponential capability growth. Each new layer of abstraction makes the next layer easier to discover and more powerful.

Emergent Specialization: The system doesn't need to be told what abilities it needs—knowledge, perception, social intelligence, language—these emerge naturally as tools for reward maximization in complex environments. In a rich enough environment with appropriate rewards, an OaK agent would discover it needs something like language to coordinate with other agents, something like episodic memory to learn from rare events, and something like causal reasoning to plan effectively.

Unbounded Scalability: Unlike current approaches that require architectural innovations for each capability, OaK provides a single framework that scales naturally with computational resources. More compute means:

- More features to explore
- More options to consider
- Deeper planning horizons
- Finer-grained abstractions

True Transfer Learning: Because knowledge is grounded in experience and organized around achieving goals, it transfers naturally to new situations. The agent learns general principles about how the world works, not just pattern associations specific to training data.

Part III: The Theoretical Foundation

3.1 The Reward Hypothesis

The OaK architecture rests on a profound claim known as the **Reward Hypothesis**:

> "All of what we mean by goals and purposes can be well thought of as maximization of the expected value of the cumulative sum of a received scalar signal (reward)."

This seems impossibly reductive. Can all the complexity of human values, purposes, and goals really be reduced to maximizing a single number? The "Settling the Reward Hypothesis" paper provides a rigorous answer.

3.2 What "Settling the Reward Hypothesis" Actually Proved

The paper doesn't claim the reward hypothesis is always true or always false. Instead, it **completely specifies the conditions under which it holds**.

Key Contribution: The Temporal γ-Indifference Axiom

The paper introduces a new mathematical axiom that captures what it means for preferences to be consistent with reward maximization. Combined with von Neumann-Morgenstern rationality axioms (completeness, transitivity, independence, continuity), this axiom is both **necessary and sufficient** for preferences to be expressible as Markov reward functions.

What This Means in Plain Language:

If your preferences about outcomes satisfy certain rationality conditions—basically, you're consistent and don't have cyclical preferences—then there exists a reward function that captures your preferences. Moreover, this reward function is "Markov," meaning the reward depends only on the immediate observation and action, not on the entire history.

The Deep Insight:

The form of the objective (discounted reward, average reward, episodic reward) is **not chosen by the designer**—it emerges from the structure of the preferences themselves. The temporal indifference axiom essentially encodes how much you care about delayed consequences versus immediate ones, and this determines the discount factor.

3.3 Why This Is Mathematically Beautiful

The result provides an elegant answer to a fundamental question: "When can a single scalar signal guide an agent to achieve complex, multi-faceted goals?"

Answer: "When the preferences those goals express satisfy rational consistency conditions."

This is similar to how expected utility theory answers the question "When can you represent preferences with a utility function?" The answer isn't "always" or "never," but "when preferences satisfy these axioms."

Practical Implications:

1. **For AI Safety:** We now know precisely when reward functions are sufficient and when they're not. This helps us understand when we need more sophisticated objective specifications.
2. **For Architecture Design:** If we ensure our agents' learned preferences satisfy these axioms, we can use simple scalar rewards even for complex goals.
3. **For Understanding Intelligence:** The result suggests biological reward systems (dopamine, etc.) could indeed be sufficient to produce all observed intelligent behavior, given the right learning mechanisms.

3.4 The "Reward is Enough" Hypothesis

The 2021 paper by Silver, Singh, Precup, and Sutton makes an even bolder claim:

> "Intelligence, and its associated abilities, can be understood as subserving the maximisation of reward."

This hypothesis argues that abilities like knowledge, learning, perception, social intelligence, language, generalization, and imitation all emerge as **instrumental capabilities** for reward maximization. They're not separate objectives requiring separate mechanisms—they're tools that an agent discovers it needs in order to maximize reward in a complex environment.

Why This Is Controversial:

Critics argue that:

- Some goals are inherently multi-objective (safety vs. performance)
- Human values are too complex for single reward signals
- Emergent goals might not align with specified rewards
- The hypothesis doesn't address how to specify rewards in the first place

Sutton's Response (Implicit in OaK):

The architecture itself addresses many of these concerns:

- Subtasks create implicit multi-objective optimization
- Meta-learning discovers appropriate optimization dynamics
- Continual learning enables ongoing alignment
- Experience-based learning grounds rewards in actual consequences

---

Part IV: The Alberta Plan - A Roadmap to AGI

4.1 Why We Need a Plan

Most AI research proceeds opportunistically: researchers work on whatever seems promising given current techniques and applications. This has produced impressive results, but it lacks coherence and direction toward general intelligence.

The Alberta Plan is explicitly designed as a **12-step roadmap** that frontloads the hardest problems and builds systematically toward complete AI systems.

4.2 The 12 Steps - Strategic Progression

Steps 1-2: Master Continual Supervised Learning Before tackling reinforcement learning, first solve continual learning in the simpler supervised setting. Learn to:

- Meta-learn step sizes per parameter
- Normalize features effectively
- Track non-stationary distributions
- Discover useful feature combinations

This is the foundation—if you can't do continual learning with labeled data, you certainly can't do it with reward signals.

Steps 3-4: Extend to Prediction and Control Progress to:

- Generalized Value Functions (predicting cumulative future signals)
- Actor-critic architectures (learning policies and value functions together)
- Off-policy learning (learning from any behavior)

These steps master the core RL algorithms in their basic form.

Steps 5-6: Handle Continuing Problems Extend to:

- Average reward settings (no terminal states)
- Continuing control (infinite horizon problems)

Most real intelligence problems don't have clear endpoints. You don't "complete" life and stop. This requires different mathematical tools and algorithms.

Steps 7-9: Add Planning and Search Introduce:

- Planning with average reward
- Model learning
- Search control (deciding what to plan about)

This is where the "model-based" part of model-based RL gets serious. The agent builds an internal simulator of the world.

Steps 10-11: The STOMP Progression and Full OaK Culminate in:

- Temporal abstraction (options)
- The complete FC-STOMP cycle
- The option keyboard (composing options)
- Continual evaluation and replacement of abstractions

This is where everything comes together into the full OaK architecture.

Step 12: Intelligence Amplification Finally, demonstrate how AI can enhance human intelligence, creating:

- Computational exo-cerebellum (predictive assistance)
- Computational exo-cortex (planning and decision support)

4.3 Why This Ordering Matters

Frontloading Challenges: The plan deliberately encounters hard problems (continual learning, meta-learning) in their simplest settings first. This means solutions developed early are robust and general, not hacks that work only in specific contexts.

Avoiding the Parts-Whole Problem: You can't build a complete AI until you have working components, but you don't know what components you need until you try to build complete systems. The Alberta Plan acknowledges this chicken-and-egg problem and works on both fronts simultaneously, with each step informing the next.

Systematic Knowledge Building: Each step builds on previous ones. Step 5 assumes Steps 1-4 are solved. This creates a cumulative research program where progress is measured and certain.

Testable Milestones: Unlike vague goals like "make AI more intelligent," each step has clear success criteria. Did you solve continual supervised learning with meta-learned step sizes? Move to Step 2. This makes the research program concrete and accountable.

Part V: Why This Matters - Implications and Importance

5.1 For the Future of AI Research

A Course Correction:

The current AI industry is dominated by massive corporations training ever-larger models on ever-more data. This approach has produced impressive results but shows signs of diminishing returns. The OaK architecture and Alberta Plan offer an alternative path that:

- **Doesn't require massive datasets** - learns from experience
- **Scales with compute, not data** - more compute means better learning, not larger models
- **Enables continual improvement** - systems that never stop getting better
- **Produces genuine understanding** - world models, not just pattern matching

Bringing Back First Principles:

Sutton's work represents a return to fundamental questions:

- What is intelligence?
- How should agents learn?

- What is the simplest architecture that could produce general intelligence?

This kind of first-principles thinking has historically led to the biggest breakthroughs in AI (backpropagation, deep learning, transformers). The current moment may require another such paradigm shift.

5.2 For Understanding Intelligence

Biological Plausibility:

The OaK architecture is much more aligned with what we know about biological intelligence than current AI systems:

- **Continual learning** - brains never stop learning
- **Reward-based** - dopamine and related systems provide learning signals
- **Model-based planning** - hippocampus and prefrontal cortex enable mental simulation
- **Hierarchical abstraction** - cortical hierarchy creates increasingly abstract representations
- **Meta-learning** - sleep and other processes tune learning parameters

Unified Theory of Intelligence:

If the reward hypothesis is correct and the OaK architecture successfully implements it, we would have a **unified explanation** for diverse intelligent abilities. We wouldn't need separate theories for:

- Language acquisition
- Social intelligence
- Perceptual learning
- Motor control
- Reasoning

They would all be understood as emergent capabilities for reward maximization in complex environments.

5.3 For AI Safety and Alignment

Interpretability:

OaK systems would be more interpretable than current black-box models because:

- The option hierarchy makes behavior modular and hierarchical
- Subtasks are explicit and can be inspected
- Planning processes can be visualized
- The reward signal makes objectives explicit

Alignment by Design:

Experience-based learning provides natural alignment opportunities:

- Agents learn from actual consequences, not human descriptions
- Reward signals can be shaped through experience, not just specified
- Continual learning enables ongoing correction and alignment
- Hierarchical goals create natural safeguards (subtasks must respect main objectives)

Catastrophic Risk Reduction:

Unlike systems trained once on massive datasets and deployed, continual learning systems:

- Can be corrected through experience
- Don't have a single catastrophic failure mode
- Learn incrementally, allowing monitoring and intervention
- Develop capabilities gradually, not suddenly

5.4 For Artificial General Intelligence

A Concrete Path Forward:

The Alberta Plan provides the most concrete and detailed roadmap to AGI published by any mainstream AI researcher. It's not hand-waving about "emergent capabilities" or "scaling laws"—it's a systematic, step-by-step engineering plan.

Testable and Falsifiable:

Each step in the plan produces testable systems with measurable capabilities. This makes the approach scientifically rigorous. If Step 3 can't be solved, that tells us something important. If it can, we make measurable progress toward AGI.

Resource Efficiency:

If successful, OaK-based AGI would be vastly more efficient than current approaches:

- No need for massive pre-training on internet-scale data
- Learning happens online, using actual experience
- Computational resources used for continual improvement, not one-time training
- Natural curriculum learning through subtask hierarchy

Part VI: Challenges and Open Questions

6.1 Technical Challenges

Computational Requirements:

While OaK promises better sample efficiency than current methods, the continual planning and meta-learning components are computationally intensive. Every decision potentially involves:

- Simulating multiple possible futures

- Updating thousands or millions of meta-parameters
- Creating and evaluating new feature combinations
- Maintaining and using complex world models

This requires significant computational resources, though arguably less than training GPT-5.

Credit Assignment:

Even with sophisticated algorithms, assigning credit correctly in complex environments with delayed consequences remains extremely challenging. The FC-STOMP progression helps by creating intermediate objectives, but fundamental credit assignment problems persist.

Exploration:

How does an agent explore effectively in open-ended environments? Random exploration is hopelessly inefficient in high-dimensional spaces. The architecture needs sophisticated exploration strategies, possibly curiosity-driven or uncertainty-based.

Reward Specification:

The reward hypothesis says reward is sufficient, but it doesn't tell us how to specify rewards for complex goals. This remains a critical challenge for applying OaK to real-world problems. The subtask mechanism helps but doesn't fully solve the problem.

6.2 Theoretical Questions

Convergence and Stability:

With everything learning simultaneously and continuously, does the system converge to good solutions? Or does it oscillate or diverge? Theoretical analysis of the full OaK system's dynamics is needed.

Abstraction Hierarchy Depth:

How deep should the hierarchy of abstractions go? Is there a natural limit, or does it grow indefinitely? How do we prevent the system from getting lost in too many layers of abstraction?

Option Discovery:

The FC-STOMP progression describes how options are created from features, but how are the right features discovered in the first place? This bootstrapping problem is fundamental.

6.3 Philosophical Concerns

Is Reward Really Enough?

Critics rightly point out that:

- Multi-objective problems may not reduce to scalar rewards
- Value alignment involves more than reward maximization

- Human preferences are often incomplete or inconsistent
- Some goals may be fundamentally non-reducible

The "Settling the Reward Hypothesis" paper addresses this mathematically, but practical concerns remain.

The Meaning of Intelligence:

Does reward maximization really capture what we mean by intelligence? Or is it just one type of intelligence? Could there be "non-reward-based intelligence"?

Emergence vs. Engineering:

How much should emerge from the learning process vs. being engineered into the architecture? The Alberta Plan specifies many architectural details—is this consistent with the goal of learning everything from experience?

Part VII: Comparison with Current Approaches

7.1 OaK vs. Large Language Models

Aspect	Large Language Models	OaK Architecture
Learning Paradigm	One-time pre-training + fine-tuning	Continuous online learning
Knowledge Source	Internet text (human descriptions)	Direct environmental interaction
World Model	Implicit in parameters	Explicit, learnable models
Planning	None (next token prediction)	Central component
Scalability	Data and model size	Computational resources
Adaptability	Fine-tuning required	Automatic continual adaptation
Interpretability	Black box	Hierarchical, modular
Resource Efficiency	Enormous pre-training cost	Pay-as-you-go learning

7.2 OaK vs. Model-Free Deep RL

Aspect	Model-Free Deep RL	OaK Architecture
Sample Efficiency	Poor (millions of episodes)	Better (learning models enables planning)
Transfer Learning	Limited	Natural via abstraction hierarchy
Continual Learning	Catastrophic forgetting	Core design principle
Meta-Learning	Separate training phase	Integrated, continuous
Temporal Abstraction	Flat action space	Options hierarchy
Exploration	Often random or heuristic	Guided by uncertainty and curiosity

7.3 OaK vs. Traditional Model-Based RL (e.g., Dyna)

Aspect	Traditional Model-Based RL	OaK Architecture
Model Learning	Often hand-designed	Fully learned
Planning	Background, occasional	Continuous, integrated
Abstractions	Fixed or manual	Discovered through FC-STOMP
Meta-Learning	Rare	Per-parameter, continuous
Scalability	Limited by state space	Unlimited via abstraction

Part VIII: Path Forward and Timeline

8.1 Near-Term (1-3 years)

Expected Progress:

- Completion of Alberta Plan Steps 1-4
- Demonstrations of continual supervised learning with meta-learning
- Actor-critic systems with advanced continual learning
- Publications on novel algorithms for each step

Key Milestones:

- Robust meta-learning of step-size parameters
- Effective feature discovery in continual settings
- Scalable GVF prediction with function approximation

8.2 Medium-Term (3-7 years)

Expected Progress:

- Completion of Alberta Plan Steps 5-9
- First Prototype-AI systems (Steps 8-9)
- Demonstrations of planning with learned models
- Early versions of the STOMP progression

Key Milestones:

- Working systems in continuing environments
- Effective search control mechanisms
- Integration of learning, planning, and acting
- Proof-of-concept temporal abstraction

8.3 Long-Term (7-15 years)

Expected Progress:

- Full OaK architecture implementation (Steps 10-11)
- Intelligence amplification applications (Step 12)
- Systems exhibiting emergent abilities
- Potential AGI capabilities

Key Milestones:

- Self-sustaining abstraction hierarchies
- Emergent language and social intelligence
- Superhuman performance in complex, open-ended domains
- Demonstrated intelligence amplification for humans

Part IX: Why This Is Important - The Stakes

9.1 The AGI Race

We are likely in the final decades before artificial general intelligence. The approach we take matters enormously:

Current Trajectory Risks:

- Massive computational waste and environmental impact
- Brittle systems that fail unpredictably
- Lack of interpretability and control
- Misalignment between capabilities and values
- Concentration of power in a few large corporations

OaK Trajectory Benefits:

- More efficient path to AGI
- Interpretable, controllable systems
- Natural alignment opportunities
- Democratized approach (doesn't require massive resources)
- Scientifically rigorous progress

9.2 Understanding Ourselves

If intelligence is fundamentally about reward maximization in complex environments, this has profound implications:

For Neuroscience:

- Validates reward-based theories of brain function
- Provides computational models of learning and development
- Explains the emergence of cognitive abilities

For Psychology:

- Unified theory of motivation and goal-directed behavior
- Explains skill acquisition and expertise development
- Provides framework for understanding mental health

For Philosophy:

- Addresses questions about consciousness and qualia
- Informs debates about free will and determinism
- Clarifies the nature of intelligence and rationality

9.3 The Future of Human Potential

Intelligence amplification (Step 12) could be transformative:

Individual Enhancement:

- AI assistants that truly understand your goals
- Tools that amplify your reasoning and decision-making
- Personalized learning systems that adapt to you

Collective Intelligence:

- Enhanced collaboration through AI intermediaries
- Better group decision-making
- Accelerated scientific discovery

Societal Benefits:

- More effective governance and policy-making
- Better resource allocation
- Solutions to complex coordination problems

Conclusion: A Vision Worth Pursuing

Rich Sutton's OaK architecture and the supporting research represent more than just another AI technique—they represent a **fundamental reconceptualization** of the path to artificial general intelligence.

Key Takeaways:

1. **Intelligence emerges from reward maximization** in complex environments, not from absorbing human knowledge
2. **Continual learning is essential** - systems must learn continuously without forgetting
3. **Hierarchical abstraction** creates open-ended capability growth through self-reinforcing cycles
4. **Experience grounds understanding** - learning from interaction produces genuine world models
5. **Systematic progress is possible** - the Alberta Plan provides a concrete roadmap with testable milestones

Why This Matters:

- **Scientific**: A rigorous, first-principles approach to AGI
- **Practical**: More efficient and effective AI systems
- **Philosophical**: Deep insights into the nature of intelligence
- **Societal**: Safer, more controllable path to advanced AI
- **Personal**: Potential for dramatic intelligence amplification

The Bottom Line:

While current AI approaches have produced impressive results, they may have hit fundamental limits. The OaK architecture offers an alternative path that is:

- More theoretically grounded
- More biologically plausible
- More computationally efficient
- More interpretable and controllable
- More aligned with the goal of genuine general intelligence

Whether or not OaK succeeds in its ambitious goals, Sutton has made an invaluable contribution by articulating a clear vision, identifying the key challenges, and proposing a systematic path forward. This work will shape AI research for the next decade and potentially determines how we achieve artificial general intelligence.

The question is not whether reward-based, continually learning, abstraction-building agents will lead to AGI—Sutton has made a compelling case that they will. The question is whether we, as a research community and society, will pursue this path with the rigor, resources, and urgency it deserves.

References

1. Sutton, R. S., Bowling, M., & Pilarski, P. M. (2022). The Alberta Plan for AI Research. arXiv:2208.11173
2. Bowling, M., Martin, J. D., Abel, D., & Dabney, W. (2023). Settling the Reward Hypothesis. Proceedings of ICML 2023.
3. Silver, D., Singh, S., Precup, D., & Sutton, R. S. (2021). Reward is enough. Artificial Intelligence, 299, 103535.
4. Sutton, R. S. (2025). The OaK Architecture: A Vision of SuperIntelligence from Experience. Presented at RLC 2025 and AGI 2025.
5. Vamplew, P., et al. (2022). Scalar reward is not enough: A response to Silver, Singh, Precup and Sutton (2021). Autonomous Agents and Multi-Agent Systems, 36, 41.

Report prepared: September 30, 2025
Word Count: ~8,000 words
Reading Time: ~30 minutes

This report synthesizes complex research into accessible insights about one of the most important ideas in current AI research. The OaK architecture may well represent the future of artificial intelligence—a future built not on scaling models, but on scaling intelligence itself.