

The Unified Equation of Intelligence: From dx/dt to Graph Theory — On the Ultimate Form of Agents and World Models

Concepts in the AI field have been exploding recently. Based on this, I wrote this article to share my thoughts.

I. Macro Perspective: The Three Major Schools of the AI World

From a macro perspective, the current development of the AI world can be divided into three major schools: Simulation, Generation, and Cognition.

- 1. The Simulation School:** Represented by NVIDIA, xAI, the autonomous driving camp, and traditional embodied intelligence. They believe that world models must adhere to physical laws, trusting that the world is a massive differential equation: $dx/dt = f(x, t)$. This is **explicit modeling** of the world.
- 2. The Generation School:** Represented by OpenAI and the Google DeepMind team. They believe there is no need to explicitly apply physical laws; instead, models can learn spatiotemporal distribution laws from massive data, allowing the world to be modeled "implicitly." Whether it's VLMs like ChatGPT and Gemini, or video generation models like Sora, Kling, and Google Genie, they essentially believe the world is a massive probability theory: $p(x|text/video)$, practicing the philosophy of "predicting the world with probability." This is **implicit modeling** of the world.
- 3. The Cognition School:** Represented by Fei-Fei Li's team and Yann LeCun's team. They view the world as a massive graph theory problem: $G = (V, E)$, composed of entities (nodes) and relationships (edges). There's no need to know the movement of every atom, only that "the cup (entity) is on the table (entity) (relationship)." It focuses on logical structure and causal reasoning. This is **semi-explicit, semi-implicit modeling** of the world.

II. Unification of the Three Schools: The Essence and State Equation of Agents

The most cutting-edge exploration in the AI field is actually attempting to merge these three. The unification of these three schools is the renowned **Agent**.

If we view an Agent as a dynamic, self-organizing, non-explicit, cyclic, and evolving graph structure $G_t = (V_t, E_t)$, then its dynamic update rule is:

$$G_{t+1} = f(G_t, a_t, o_t)$$

In this equation, **The Cognition School** endows the Agent with "Data Structure" (G_t), **The Generation School** endows the Agent with the "Evolution Function" (f), and **The Simulation School** endows the Agent with the "Interaction Interface" (a_t, o_t).

To understand this more intuitively, let's take an example: Suppose you want a home service robot to put an apple from the dining table into the refrigerator. How does the Agent work?

- 1. Perception & Cognition (Cognitive):** The robot opens its eyes (camera input) and sees a pile of pixels. It cannot just look at pixels; it must understand semantics. It uses visual algorithms (like YOLO + Scene Graph Generation models) to identify entities (nodes={apple, table, refrigerator, robotic arm}) and relationships (edges={applettable, refrigeratorclosed}). At this point, the graph structure G_t is the semantic map of the current scene. The role of the Cognition School is to turn unstructured images into a logical structure capable of reasoning.
- 2. Planning & Generation (Generative):** Next, the robot needs to plan its actions. It quickly simulates in

its "brain": "What happens if I push the apple directly?" It calls its internal World Model, not to solve physical equations (too slow), but to generate the future like dreaming.

- Attempt Plan A: Grab the apple with only one hand → Predicted Result: $P(\text{apple falls}) = 0.8 \rightarrow$ Abandon.
- Attempt Plan B: Open the fridge door first, then grab the apple → Predicted Result: $P(\text{task complete}) = 0.9 \rightarrow$ Adopt.
- In this way, a series of specific High-level actions are generated (e.g., move hand to handle, apply force to pull open). Here, the Generation School's role is low-cost trial and error in latent space, finding the optimal solution through probability models.

3. Execution & Control (Simulation): After deciding to "pull open the door," how much current should be given to the motor joints? How to overcome friction? Here we enter the physical control layer. The Agent calculates the specific inverse dynamics solution or uses a policy network trained in this physical environment to precisely control the trajectory of the robotic arm, ensuring the action conforms to physical laws, using friction to pull the door open. The real physical action a_t changes the real world. Here, the Simulation School's role is to ensure the action is executed and grounded.

III. Core Technique: The Role of Reinforcement Learning (RL)

Everyone should have noticed that Reinforcement Learning (RL) is becoming increasingly popular in 2025. RL is actually the "martial art" common to all three schools; regardless of the school, training through RL is necessary. RL provides the environment, state, action, and reward, which itself is a small implicit world.

- **How does the Simulation School use RL?** Enter a virtual world that fully complies with physical laws, such as NVIDIA Isaac Sim or MuJoCo. Let the robot move randomly inside; if it hasn't grabbed the apple, reward -1; if it has, reward +10. The simulation environment here is both the venue and the coach. Through billions of physical interactions, the Agent learns basic "muscle memory" (Policy) and understands the relationship between force and motion.
- **How does the Generation School use RL?** Use massive video data recorded by robots running wildly in simulators or videos demonstrated by humans as data sources to train a model similar to Sora or Dreamer. Cover the second half of the video and let the AI guess: The previous frame was a hand holding an apple; what happens in the next frame? Make the predicted image distribution $P(x_{t+1})$ as close as possible to the real image. This allows the Agent to learn the Generation School's ability, i.e., predicting consequences without looking at reality.
- **How does the Cognition School use RL?** Add semantic labels during training. When the Agent sees an image of "apple falling on the ground," force its internal state to map to the graph structure {apple, on, ground}. Ensure its "imagination" is not wild thinking but supported by a logical structure.

Finally, if the friction coefficient in the simulation is different from reality, then download this "hybrid brain" into the real robot body. In the real world, the Agent discovers that the force needed to push the door is greater than in the simulation. It quickly updates its $f(G_t, a_t, o_t)$ model through **RL Online Learning** to correct the error. This is the essence of today's hottest buzzwords: Agent, LLM, VLM, VLA, RL, and Embodied AI.

IV. Future Predictions: Three Directions of AI Development

Based on the above framework, we can discern the development directions of AI in the coming years:

1. The Simulation School: Solving the Bottleneck of "Synthetic Data" and "Sim-to-Real" Transfer

In the future, the World Simulator will become infrastructure. Since real-world data acquisition is expensive and scarce, the Simulation School will be dedicated to generating high-quality Synthetic Data to "feed" large models. At the same time, solving the Sim-to-Real transfer problem will be the core, allowing skills learned by AI in the virtual dx/dt world to seamlessly migrate to the complex real physical world, which is key to the landing of embodied intelligence.

2. The Generation School: Bidding Farewell to Brute-Force Scaling, Turning to "Small and Beautiful" and Video Generation

I am most familiar with the Generation School, so let me state the conclusion first: The era of Scaling Law, where significant improvements were achieved purely by stacking parameters, has passed.

Ilya Sutskever, former senior partner of OpenAI and one of the key figures pushing for brute-force scaling early on, also confirmed this in a recent podcast interview. The reason is that the bottleneck of model Scale Up has appeared:

- **Theorem Aspect:** The Scaling Law itself is sub-linear, with diminishing marginal returns.
- **Data Aspect:** Constrained by the classic empirical "Data Scaling Law": if data volume does not keep up with model scale, performance will be limited by data rather than parameters, leading to inefficiency. Currently, global data available for training has nearly all been scraped, and simply providing more unlabeled data to the model can no longer achieve further improvements.
- **Resource Aspect:** As model scale increases, the required computing power, cost, and energy also increase, which very few companies can afford.
- **Technical Aspect:** The larger the model parameters, the greater the difficulty of engineering implementation, requiring handling multi-dimensional parallel engineering problems during training and high-cost service overheads during inference applications.

Therefore, the focus will shift: Everyone is starting to pay more attention to the capability improvement of **small parameter models**, hoping to use algorithmic innovation to counter the stacking of computing power. Small models, combined with fine-tuning + algorithms, can compensate for the gap in model parameters, save resources and costs, achieve edge-side inference, and have more landings and applications. At the same time, the modality will shift. The Scaling Law era in the text field is tired; the GPT and Scaling Law era in the **video field** has just begun. Video generation models will shift from "generating good-looking images" to "videos that understand physical laws," becoming the core component of world models.

3. The Cognition School: From "Fast Thinking" to "Slow Thinking," Logic and Reasoning Return

If the Generation School mimics human "System 1" (intuition, fast thinking), then the Cognition School will lead AI towards "System 2" (logic, slow thinking) in the next few years. Pure probability prediction leads to hallucinations because it lacks logical anchors.

In the future, Neuro-Symbolic AI will revive: the black box of pure neural networks will be opened, incorporating graph theory and symbolic logic. AI must not only "guess" what is highly probable but also be able to "deduce" inevitable causality. Reasoning capability is justice, similar to the idea of the OpenAI o1 series models, using Chain of Thought (CoT) and search trees to let the Agent "think" before answering. Structured memory (RAG + Knowledge Graph) will continue to develop, and the Agent's memory will no longer be a fuzzy vector but a queryable, modifiable knowledge graph. This is the only solution to solving long-term planning problems.

In summary, the second half of AI is not a solo show for any single school, but a deep integration of Simulation (Body), Generation (Brain), and Cognition (Heart). Whoever can first open up the "Ren and Du" meridians (key energy channels) of these three within a machine will define true General Embodied Intelligence.

V. Theoretical Deepening: World Models are the Dawn of Model-based RL

Finally, I want to use the Agent's unified state equation $G_{t+1} = f(G_t, a_t, o_t)$ to further explain: Why can models trained by RL surpass humans, while traditional neural networks, or the upper limit of all traditional supervised learning, cannot?

Traditional Supervised Fine-Tuning (SFT) is essentially a simplified version of the unified equation: $a_t = f^{-1}(o_t)$. It fits the model's policy based on results; it lacks o_t (observation feedback), which equates to lacking the process of **Exploration (Search)**. Fitting human data is merely imitation, which explains why large model SFT can only make the model imitate but cannot truly learn key logic. It only brings the policy space and solution space roughly closer, avoiding situations where RL cannot search for a solution later due to insufficient search computing power and the limited upper limit of the model's own policy space.

The essence of RL is empowering the model. Letting the model control a_t itself, and the reward function r is equivalent to giving the model o_t . The essence of the model getting an a_t is brute-force search. After supervised learning brings the policy space quickly close to the solution space, the probability of searching for a good solution increases. The difference between good and bad solutions gives the model a learnable signal. RL searches more precisely to enable the model to surpass human data—this is why AlphaGo defeated Lee Sedol. Without RL, it is impossible to learn this effect purely from human data.

However, mainstream Model-free RL still hasn't solved the problem of low efficiency in brute-force searching for a_t, o_t . Thus, we need World Models.

Let the model plan logically first, establish cognition through G (cognitive structure), directly avoid bad search paths, and search more smartly rather than brute-force search. Therefore, I believe the development of World Models, or the macro Cognition School, is most important for making RL's acquisition of a_t smarter, thereby accelerating the training process of f .

This is the progression route of **Supervised Learning (Neural Networks) → RL → World Models**. World Models will ultimately be the dawn of Model-based RL!