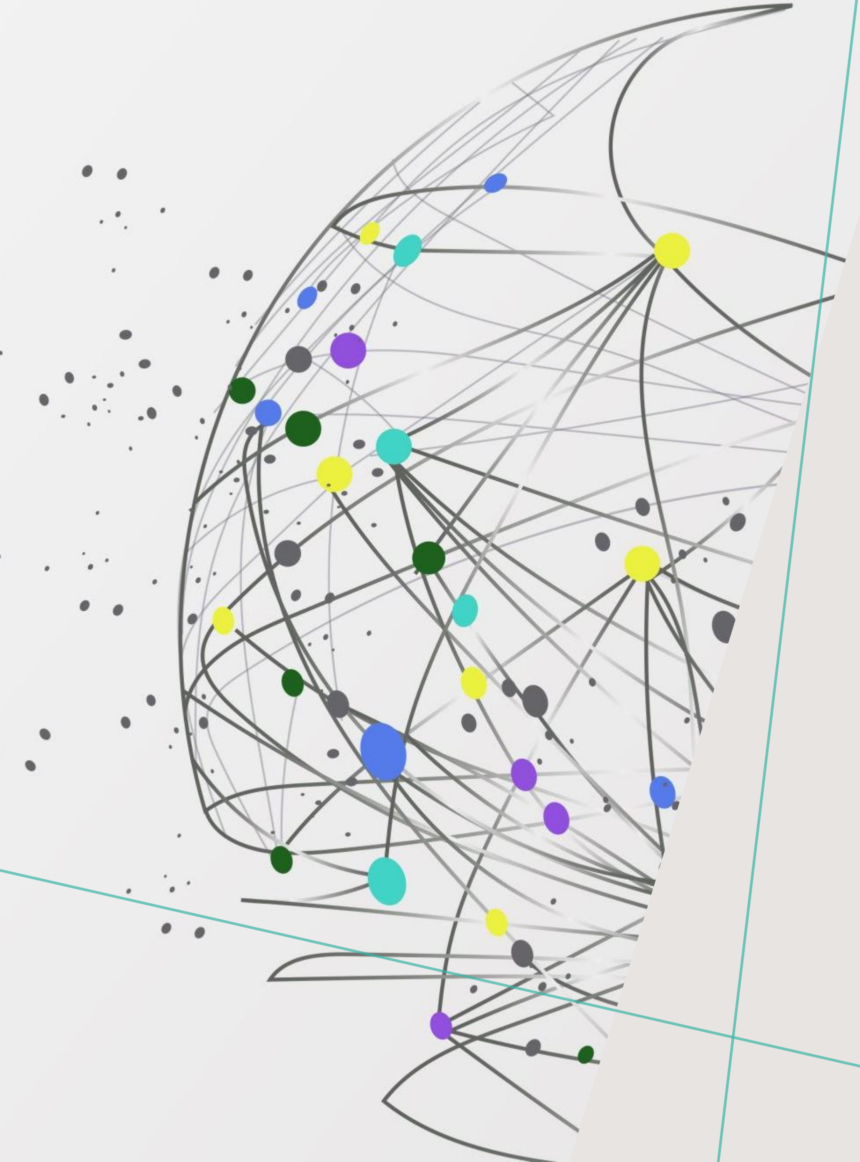


AN INTRODUCTION TO WORLD MODELS

TEACHING MACHINES TO DREAM

TURING AWARD WINNER 2024!

Andrew G. Barto

Richard S. Sutton

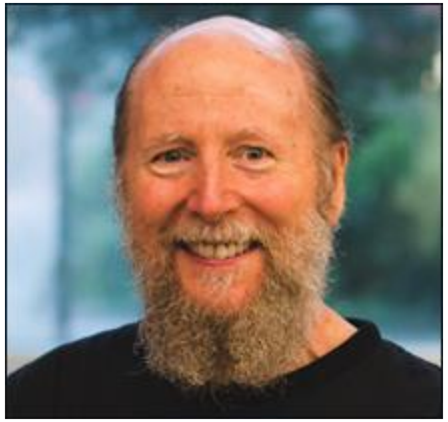
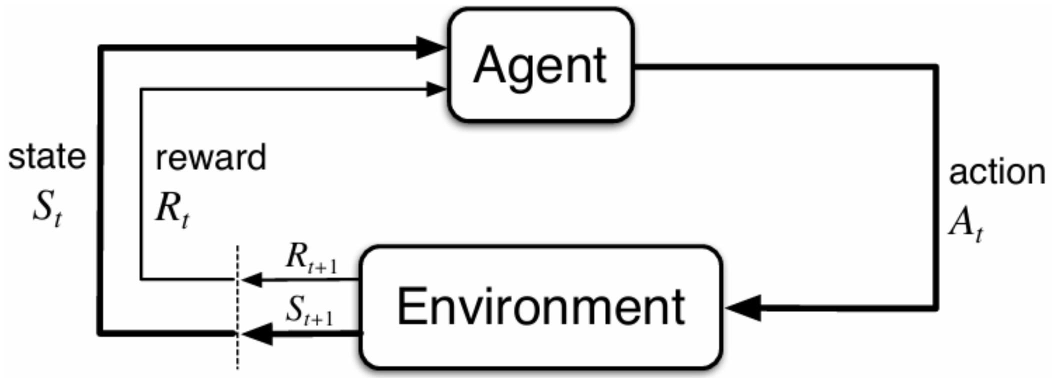


Photo courtesy of ACM – Turing Award 2024

REINFORCEMENT LEARNING

- Another branch of Artificial Intelligence which study the interaction between an agent with the surroudings
- The RL problems are formulated as a **Markov Decision Processes**, where agents has to maximaze the **reward** received by interacting with a stochastic environment



THE ROLE OF WORLD MODEL

- A good World Model allows AI agent to **predict the future** based on the current state of the environment
- Thus, to improve the **decision making**, and how can a decision can **impact the future** state
- Simplify, and make faster the training of new policies by **reducing** the input **dimension**
- Enables **efficient planning**, e.g. in autonomous driving by predicting the traffic flow
- A realization of how living being builds the **mental representation** of the environment.

SOME BACKGROUND...

- The concept of creating a world model is not new in AI
- Back to 1971, a problem solver called **STRIPS**, used world model to represent a set of well formed formulas of the first-order predicate calculus, the aims was to find a suitable representation which satisfies the goal condition
- Another **application** is **on robotics** developed in 1985, where a mobile robot learns to reconstruct the environment with the inaccuracies introduced by sensors, and able to locate itself in the environment
- More recently, World Model is getting attention again with the **Diffusion Models** and its **application in Reinforcement Learning** and **Autonomous Driving**

WHAT IS A WORLD MODEL

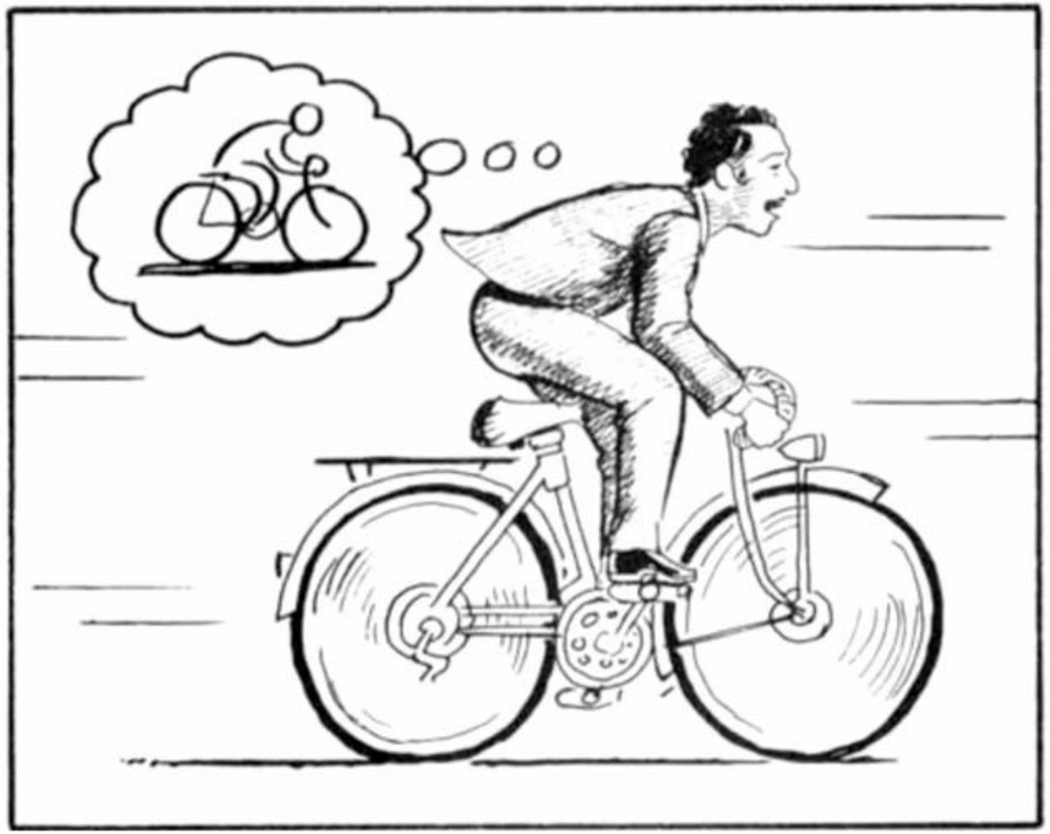
- A simplified **simulation** of the external world
- Yann LeCun has given his definition on a post:

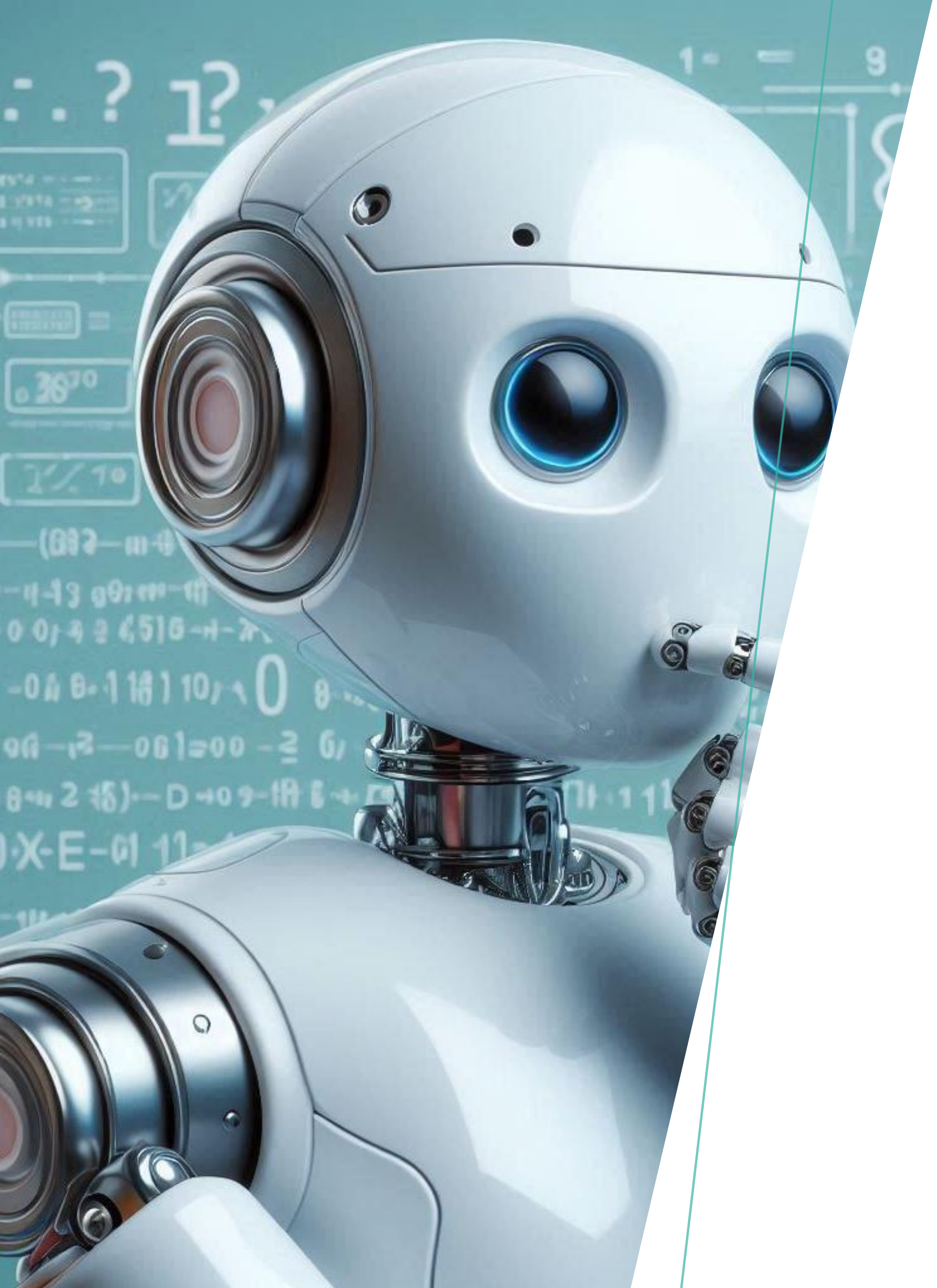
Given:

- An observation $x(t)$
- A previous estimate of the world state $s(t)$
- An action proposal $a(t)$
- A latent variable proposal $z(t)$

A world model computes:

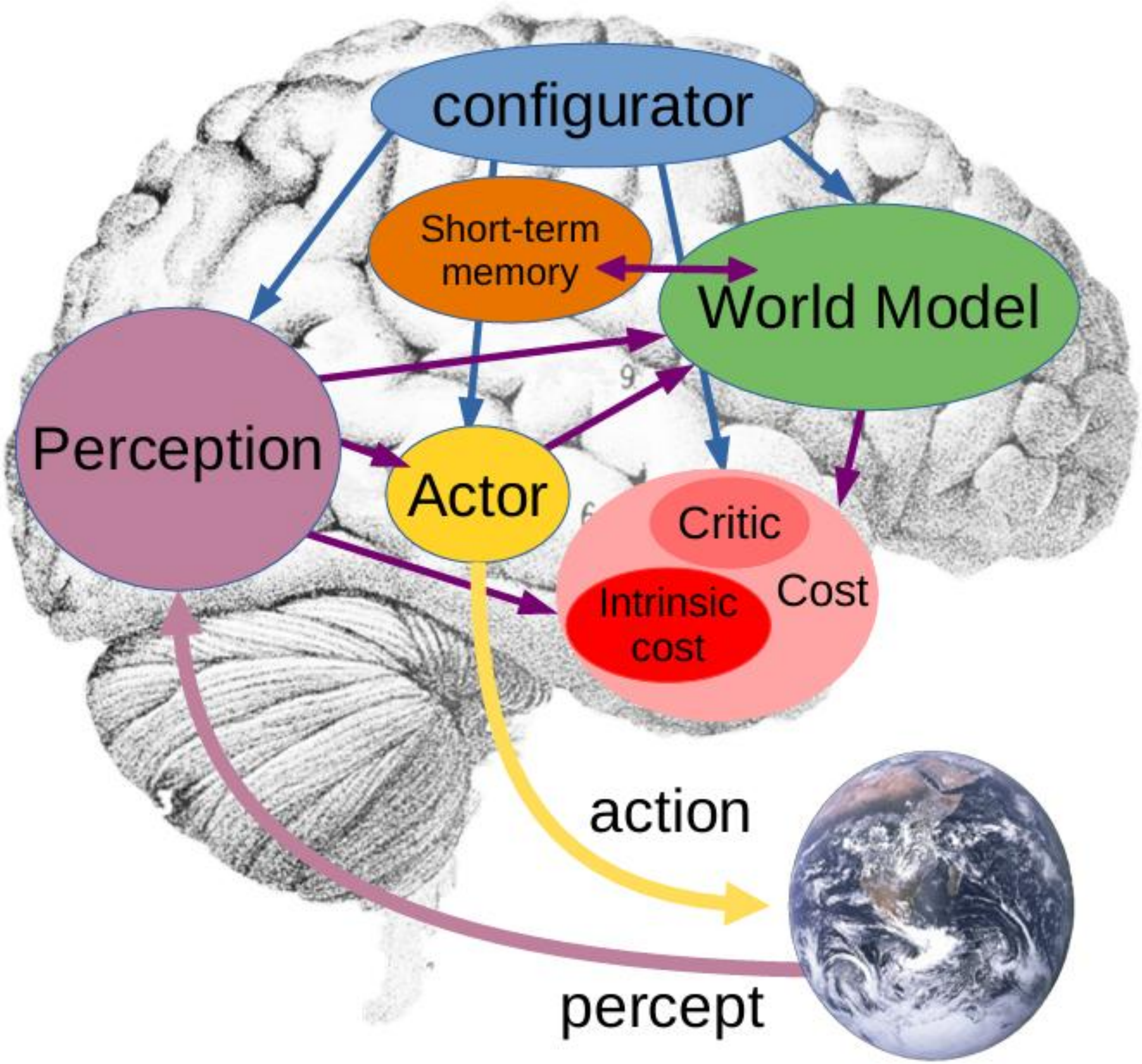
- Representation: $h(t) \; = \; Enc(x(t))$
- Prediction: $s(t+1) \; = \; Pred(h(t), s(t), a(t))$





NOT ONLY…

- World models has no an official definition
- Many ML model could be called to be a world model since it has a concept of the external world, even though it is limited to a specific task:
  - Generative Models
  - Large Language Models
  - Reinforcement Learning algorithms


configurator
Short-term memory
World Model
Perception
Actor
Critic
Cost
Intrinsic cost
action
percept

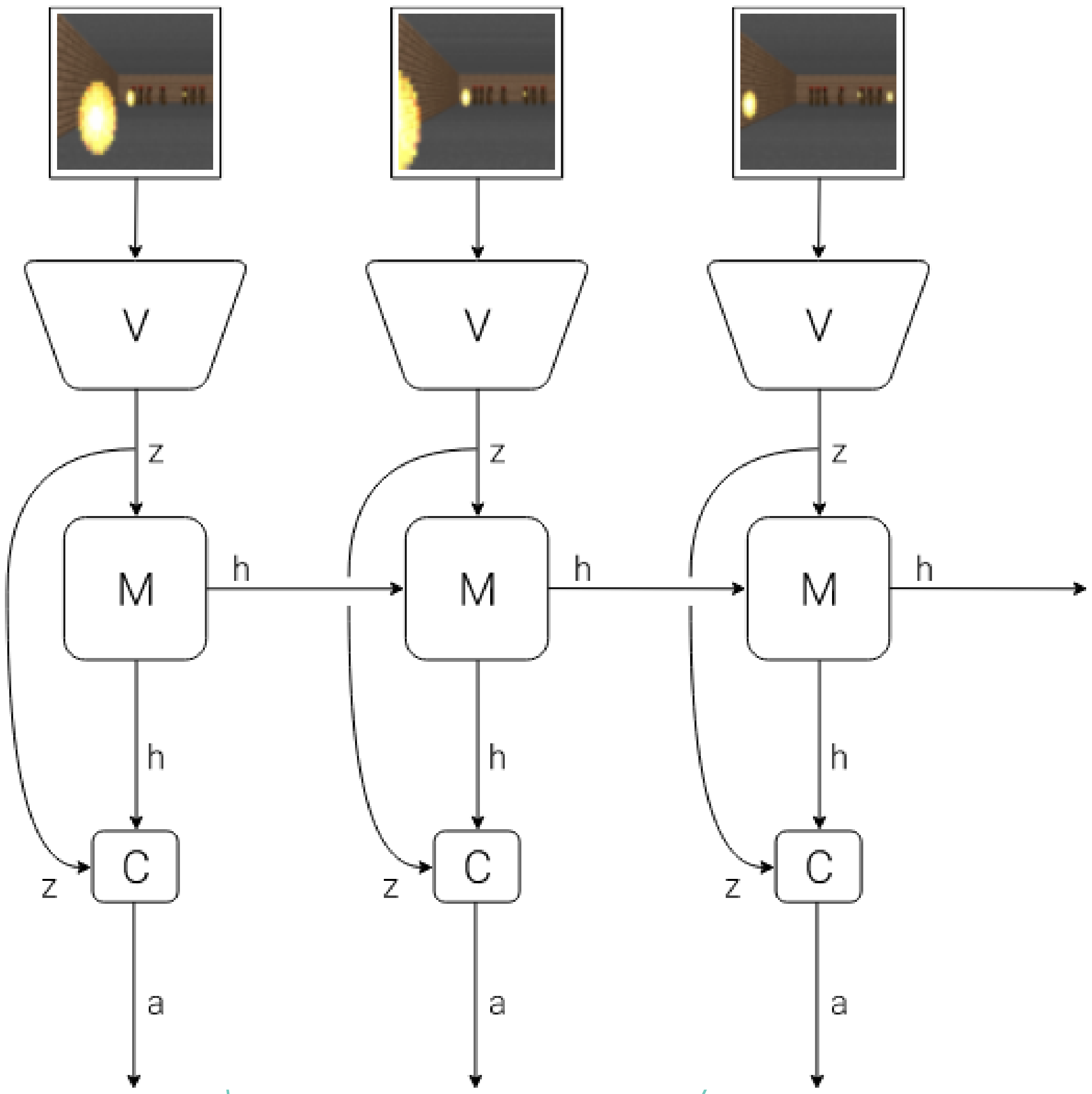
HOW?

- Self supervised learning → Capture inputs' mutual dependencies
- From a given input $\boldsymbol{x}$, there is a infinite number of **compatible** $\boldsymbol{y}$
- **Energy Based Models**, **Diffusion Models**, **Transformers** etc.
  - Maybe learning a hirarchical representation of the world dynamics?
- Handle uncertanty using **latent** variables
- Prevent model to collapse:
  - Regularization
  - Contrastive Learning

SOME ARCHITECTURES AND THEIR APPLICATION

- WORLD MODEL using VAE
- JEPA
- DREAMER
- DIAMOND
- GAIA

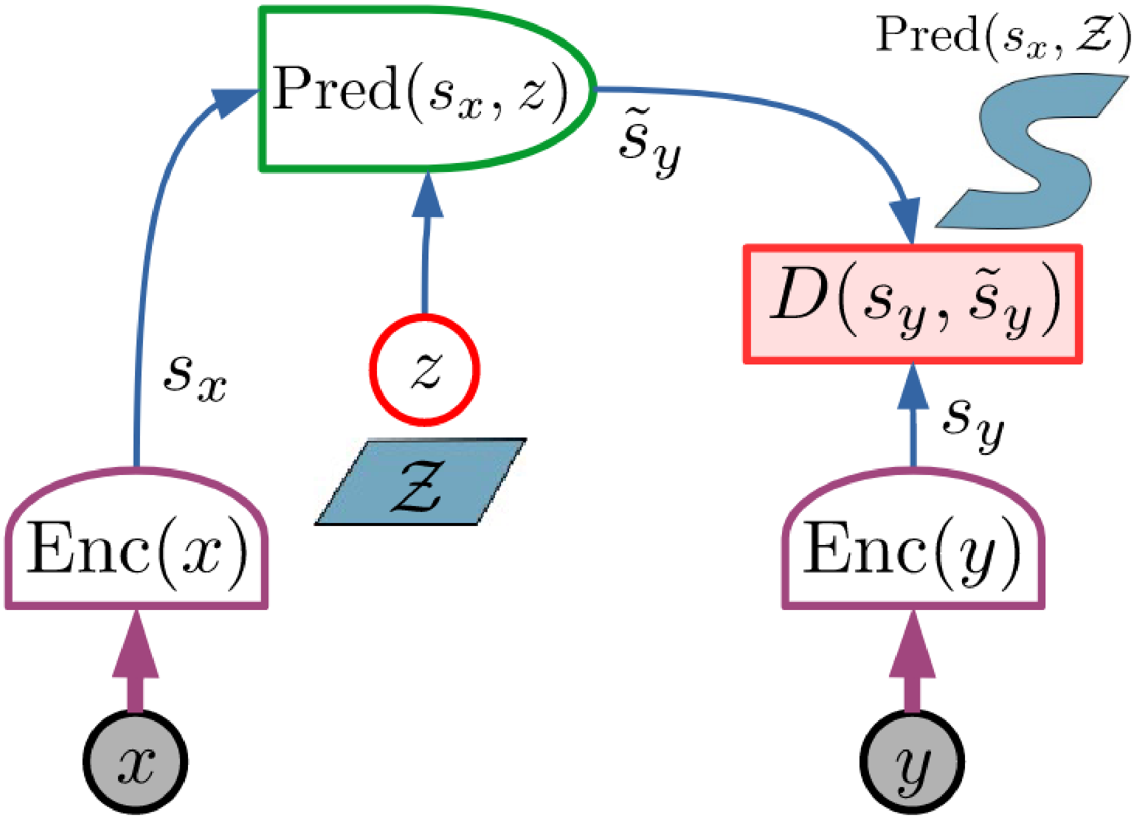
VAE-WORLD MODELLING



VAE-WORLD MODEL

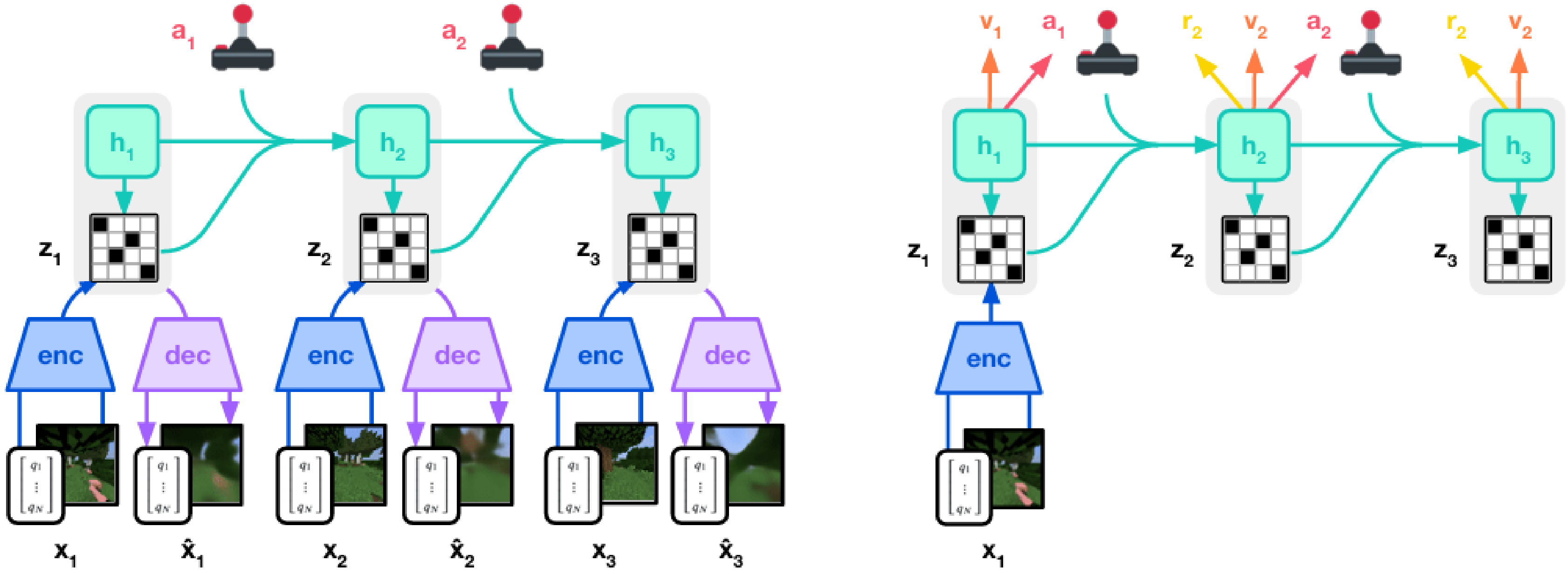
- Inspired by the **human cognitive system**
- The **prediction** of **future** are made **on a latent space** instead of directly operating on the observation space, which is usually very large
- We can **train** agents **inside their dreams**
  - Adding a stochastic component in the predition module to prevent cheating

JOINT EMBEDDING PREDICTIVE ARCHITECTURE (JEPA)



JEPA

- Also here the prediction is on a latent space
- The encoder for y has invariance properties
- Ignoring not relevant information
- Training **criteria**:
  - Maximize information content on observable variables representation
  - Maximize information content on not observable variables representation
  - Make not observable variables easy to predict
  - Minimize the information content in latent variables used for prediction
- Extension: hierarchical JEPA



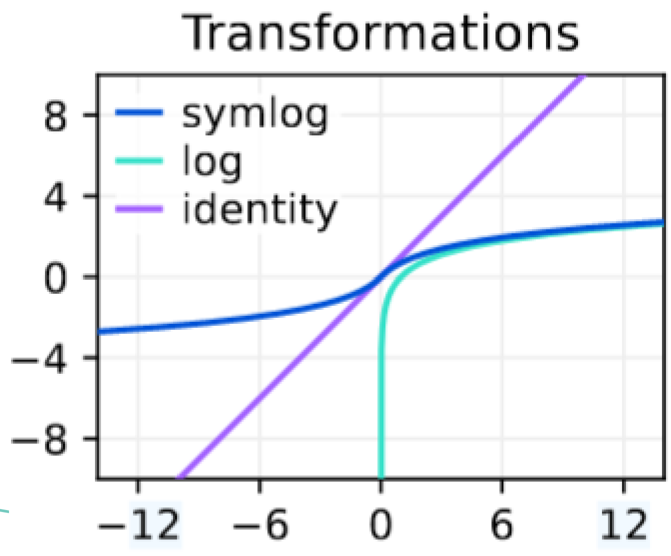
(a) World Model Learning

(b) Actor Critic Learning

DREAMER-V3

DREAMER-V3

- Using symlog as transformation:
  - decoder
  - reward predictor
  - critic



- World model as Recurrent State-Space Model

$$\text{RSSM}\begin{cases}\text{Sequence model:} & h_t = f_\phi(h_{t-1}, z_{t-1}, a_{t-1}) \\ \text{Encoder:} & z_t \sim q_\phi(z_t \mid h_t, x_t) \\ \text{Dynamics predictor:} & \hat{z}_t \sim p_\phi(\hat{z}_t \mid h_t)\end{cases}$$

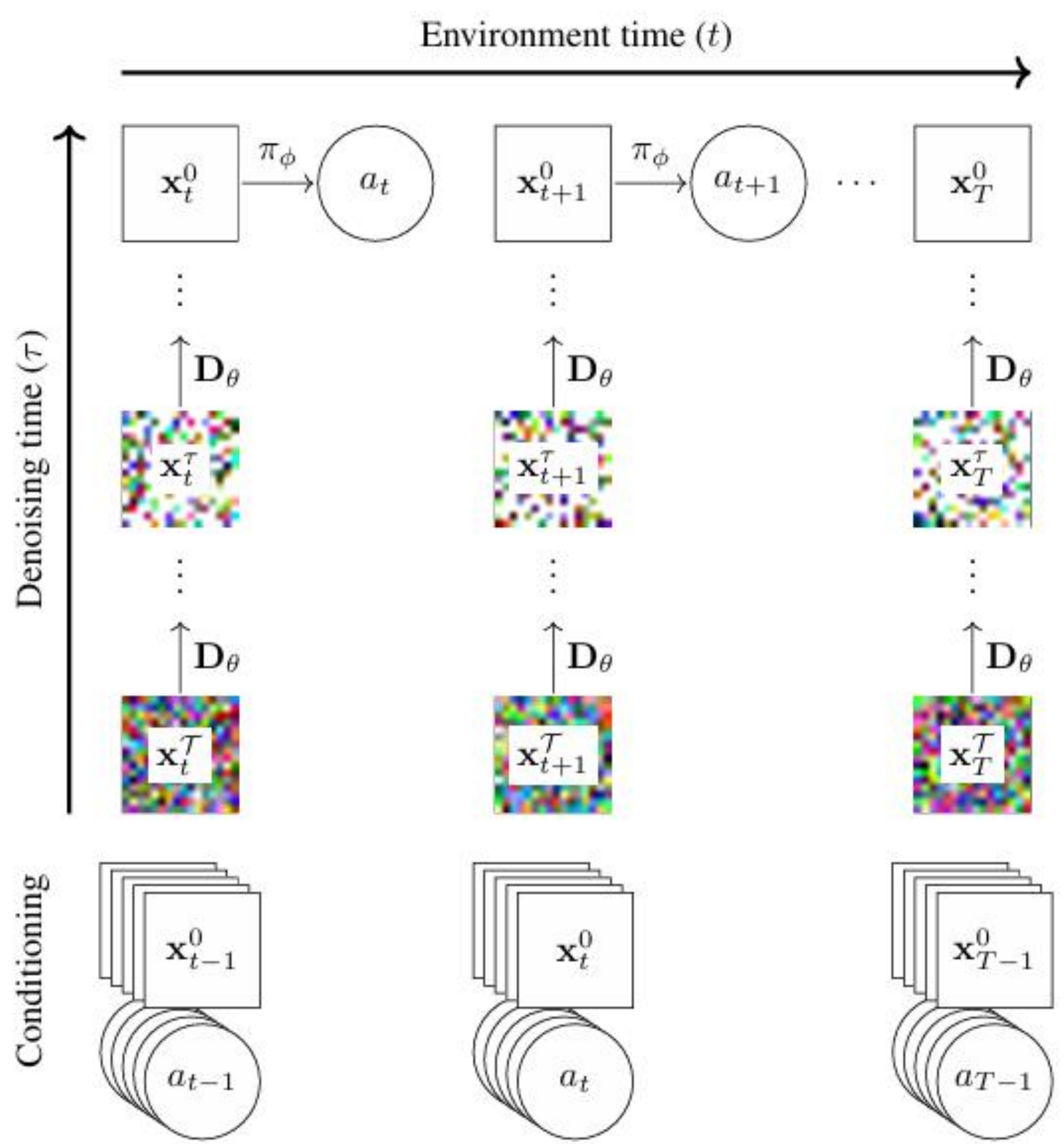
$$\begin{array}{ll}\text{Reward predictor:} & \hat{r}_t \sim p_\phi(\hat{r}_t \mid h_t, z_t) \\ \text{Continue predictor:} & \hat{c}_t \sim p_\phi(\hat{c}_t \mid h_t, z_t) \\ \text{Decoder:} & \hat{x}_t \sim p_\phi(\hat{x}_t \mid h_t, z_t)\end{array}$$

$$\mathcal{L}_{\text{pred}}(\phi) \doteq -\ln p_\phi(x_t \mid z_t, h_t) - \ln p_\phi(r_t \mid z_t, h_t) - \ln p_\phi(c_t \mid z_t, h_t)$$

$$\mathcal{L}_{\text{dyn}}(\phi) \doteq \max\big(1, \text{KL}\big[\operatorname{sg}(q_\phi(z_t \mid h_t, x_t)) \,\big\|\, p_\phi(z_t \mid h_t)\big]\big)$$

$$\mathcal{L}_{\text{rep}}(\phi) \doteq \max\big(1, \text{KL}\big[q_\phi(z_t \mid h_t, x_t) \,\big\|\, \operatorname{sg}(p_\phi(z_t \mid h_t))\big]\big)$$

DIAMOND



Environment time (t)
Denoising time (τ)
Conditioning
$\mathbf{x}_t^0$
π_ϕ
a_t
$\mathbf{x}_{t+1}^0$
π_ϕ
a_{t+1}
$\cdots$
$\mathbf{x}_T^0$
$\mathbf{D}_\theta$
$\mathbf{x}_t^\tau$
$\mathbf{x}_{t+1}^\tau$
$\mathbf{x}_T^\tau$
$\mathbf{x}_t^\mathcal{T}$
$\mathbf{x}_{t+1}^\mathcal{T}$
$\mathbf{x}_T^\mathcal{T}$
$\mathbf{x}_{t-1}^0$
a_{t-1}
$\mathbf{x}_t^0$
a_t
$\mathbf{x}_{T-1}^0$
a_{T-1}

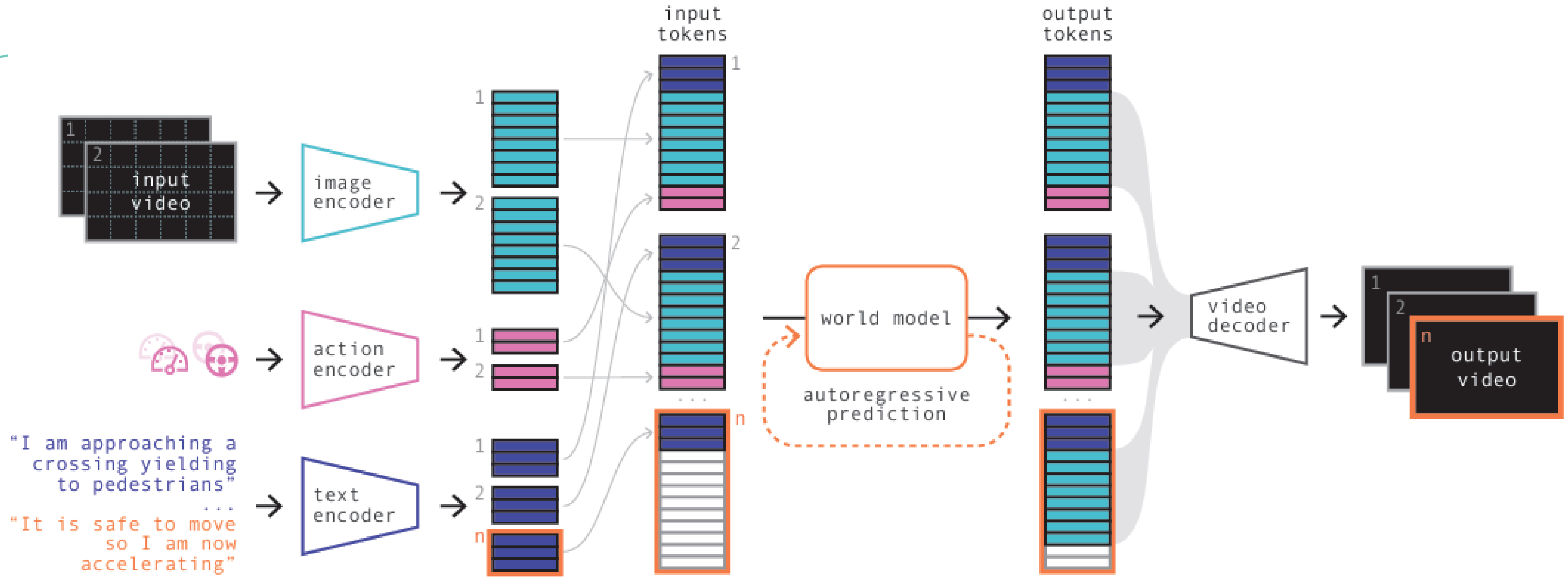
DIAMOND

- Diffusion based model
- The diffusion process is conditioned on past observations and actions

$$\mathcal{L}(\theta) = \mathbb{E}\left[\|\mathbf{D}_\theta(\mathbf{x}_{t+1}^\tau, \tau, \mathbf{x}_{\leq t}^0, a_{\leq t}) - \mathbf{x}_{t+1}^0\|^2\right]$$

$$\mathbf{D}_\theta(\mathbf{x}_{t+1}^\tau, y_t^\tau) = c_{\text{skip}}^\tau \, \mathbf{x}_{t+1}^\tau + c_{\text{out}}^\tau \, \mathbf{F}_\theta\left(c_{\text{in}}^\tau \, \mathbf{x}_{t+1}^\tau, y_t^\tau\right)$$

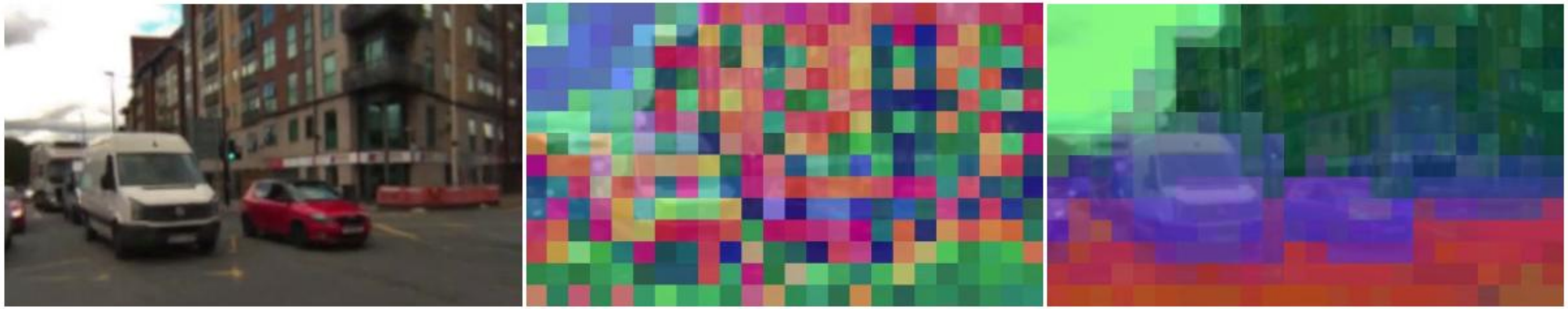
- The world model is completed with a reward model and termination model (similarly in Dreamer-V3)
- Agents trained on imagination with actor-critic networks using REINFORCE


input tokens
output tokens
1
2
input video
image encoder
action encoder
text encoder
"I am approaching a crossing yielding to pedestrians"
...
"It is safe to move so I am now accelerating"
world model
autoregressive prediction
video decoder
n
output video

GAIA-1

GAIA-1

- The world model plays with the token-space
- Tokens give a semantical meaning
- For videos tokens are preferably to be shorter and larger in vocabulary
  - Dividing videos frames in patch
  - Pushed towards DINO-distilled tokens to get a meaningfull semantic



(a) Input image (b) Base VQ-GAN tokens (c) DINO-distilled tokens

GAIA-1

- Decoder is used only to train the encoder
- An image based decoder is not enough to capture the temporal dependencies
- Diffusion based video decoder, trained on different tasks:
  - Image and video generation
  - Autoregressive decoding
  - Video interpolation
- Conditioned with image tokens
- v-parametrization is used to prevent unnatural color shifts e consistency

$$L_{\text{video}} = \mathbb{E}_{\epsilon,t'}\left[\|\epsilon_\theta(\mathbf{x}^{t'}, t', \mathbf{z}, \mathbf{m}) - \epsilon\|_2^2\right]$$

ISSUES

- Is World Models enough/too much?
- There are debates on the different approaches to achieve general intelligence
    - Reward is enough hypothesis
    - Robust agents learn causal world models

CONCLUSIONS

- World Models has found many applications in different area
  - **Autonomous driving**, is essential to predict the behavior of other vehicles or pedestrians, and the self-driving car must make safe and efficient real-time decisions. E.g. the Tesla FSD
  - **3D world simulation**, recently Google DeepMind released the Genie 2, they called it as a Large-scale foundation world model, capable to simulate different environment and be consistent with the action of the agent
  - **Robotics** is the very first area where the application of world models has found success. One of the main focuses was to learn the interaction between the agent and the surroundings (UniSim)
  - **Gaming** is another new field where World Models is becoming very popular. Some recent research focuses on developing game engine by generating game scenarios

THANKS FOR YOUR ATTENTION

[1] Y. LeCun, «A Path Towards Autonomous Machine Intelligence Version 0.9.2, 2022-06-27».
[2] R. S. Sutton e A. G. Barto, «Reinforcement Learning: An Introduction».
[3] «Strips: A new approach to the application of theorem proving to problem solving - ScienceDirect». Consultato: 6 marzo 2025. [Online]. Disponibile su: https://www.sciencedirect.com/science/article/pii/0004370271900105
[4] R. Chatila e J. Laumond, «Position referencing and consistent world modeling for mobile robots», in *1985 IEEE International Conference on Robotics and Automation Proceedings*, mar. 1985, pp. 138–145. doi: 10.1109/ROBOT.1985.1087373.
[5] A. Elfes, «Using occupancy grids for mobile robot perception and navigation», *Computer*, vol. 22, fasc. 6, pp. 46–57, giu. 1989, doi: 10.1109/2.30720.
[6] D. Ha e J. Schmidhuber, «World Models», mar. 2018, doi: 10.5281/zenodo.1207631.
[7] E. Talpes *et al.*, «Compute Solution for Tesla's Full Self-Driving Computer», *IEEE Micro*, vol. 40, fasc. 2, pp. 25–35, mar. 2020, doi: 10.1109/MM.2020.2975764.
[8] D. Silver, S. Singh, D. Precup, e R. S. Sutton, «Reward is enough», *Artificial Intelligence*, vol. 299, p. 103535, ott. 2021, doi: 10.1016/j.artint.2021.103535.
[9] Y. Matsuo *et al.*, «Deep learning, reinforcement learning, and world models», *Neural Networks*, vol. 152, pp. 267–275, ago. 2022, doi: 10.1016/j.neunet.2022.03.037.
[10] D. Hafner, J. Pasukonis, J. Ba, e T. Lillicrap, «Mastering Diverse Domains through World Models», 10 gennaio 2023, *arXiv*: arXiv:2301.04104. doi: 10.48550/arXiv.2301.04104.
[11] A. Hu *et al.*, «GAIA-1: A Generative World Model for Autonomous Driving», 29 settembre 2023, *arXiv*: arXiv:2309.17080. Consultato: 27 settembre 2024. [Online]. Disponibile su: http://arxiv.org/abs/2309.17080
[12] J. Richens e T. Everitt, «ROBUST AGENTS LEARN CAUSAL WORLD MODELS», 2024.
[13] E. Alonso *et al.*, «Diffusion for World Modeling: Visual Details Matter in Atari», 20 maggio 2024, *arXiv*: arXiv:2405.12399. Consultato: 25 ottobre 2024. [Online]. Disponibile su: http://arxiv.org/abs/2405.12399
[14] S. Yang *et al.*, «Learning Interactive Real-World Simulators», 26 settembre 2024, *arXiv*: arXiv:2310.06114. doi: 10.48550/arXiv.2310.06114.
[15] «Genie 2: A large-scale foundation world model», Google DeepMind. Consultato: 6 marzo 2025. [Online]. Disponibile su: https://deepmind.google/discover/blog/genie-2-a-large-scale-foundation-world-model/