DSC291: Machine Learning with Few Labels

World Model

Zhiting Hu

Lecture 17, May 27, 2025

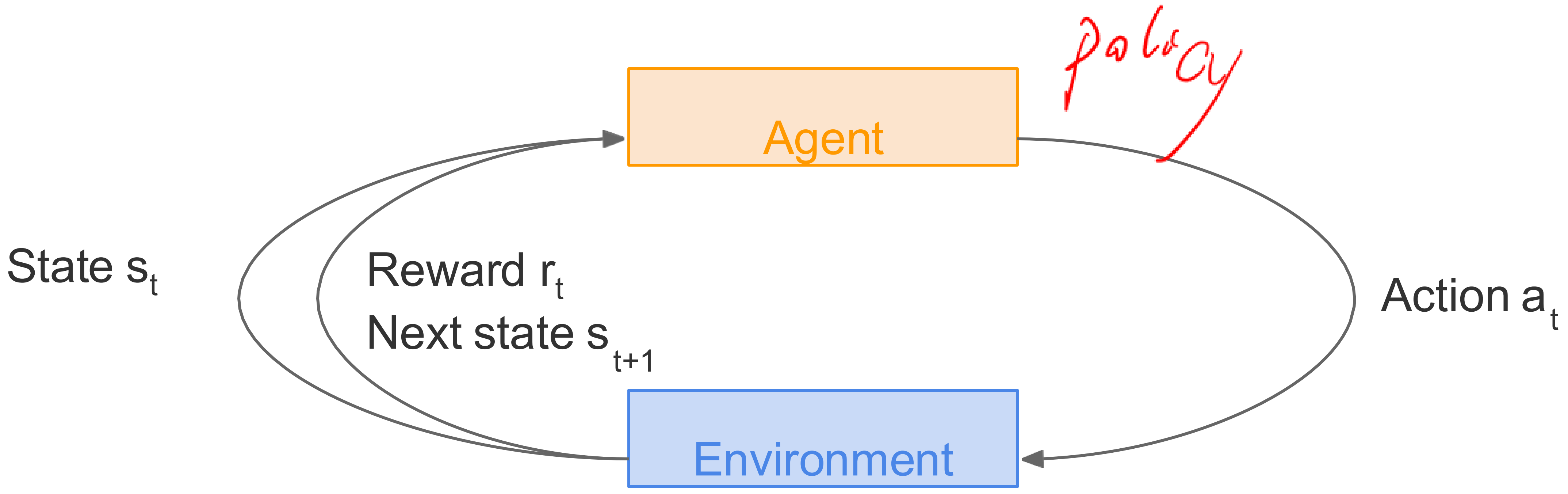
UC San Diego
HALICIOĞLU DATA SCIENCE INSTITUTE

Outline

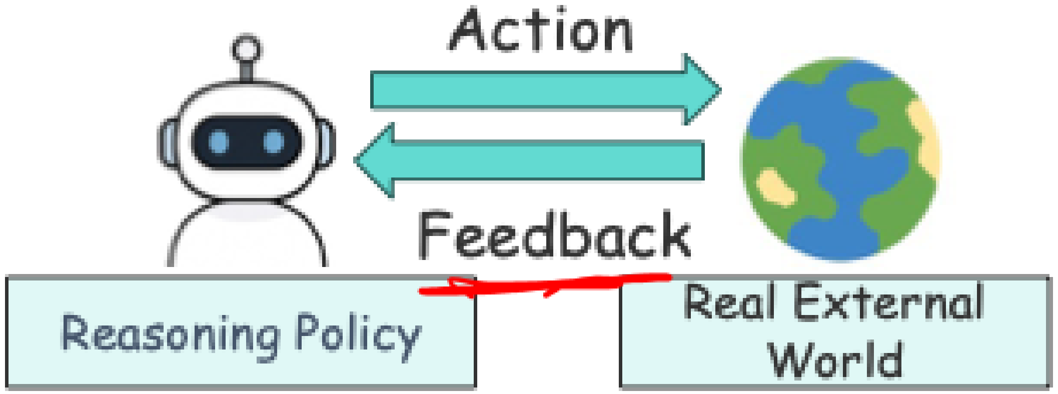
- World Model

- Paper presentation:
    - Sijin Lyu, Tianhao Zhou: "Improving noisy student training for low-resource languages in End-to-End ASR using CycleGAN and inter-domain losses”

Reinforcement Learning



Reinforcement Learning

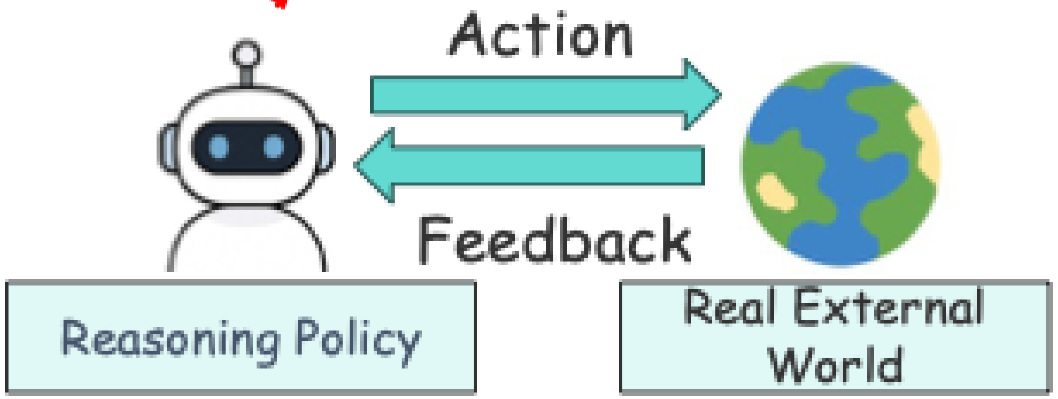


- Deployed in the real world
- Expensive, slow to get feedback

① human actions supervised ← teleoperate

Reinforcement Learning

② RL



$\pi(a|s)$

- Deployed in the real world
- Expensive, slow to get feedback

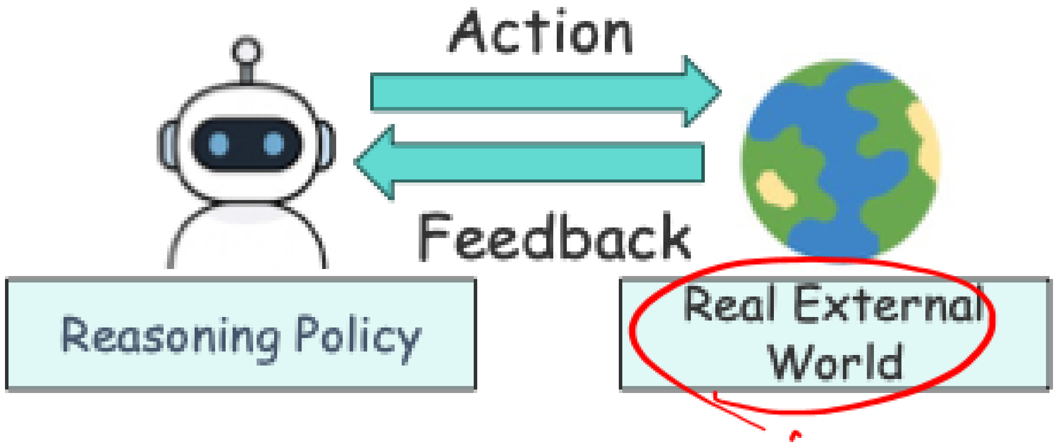
Human data collection farm

Tesla

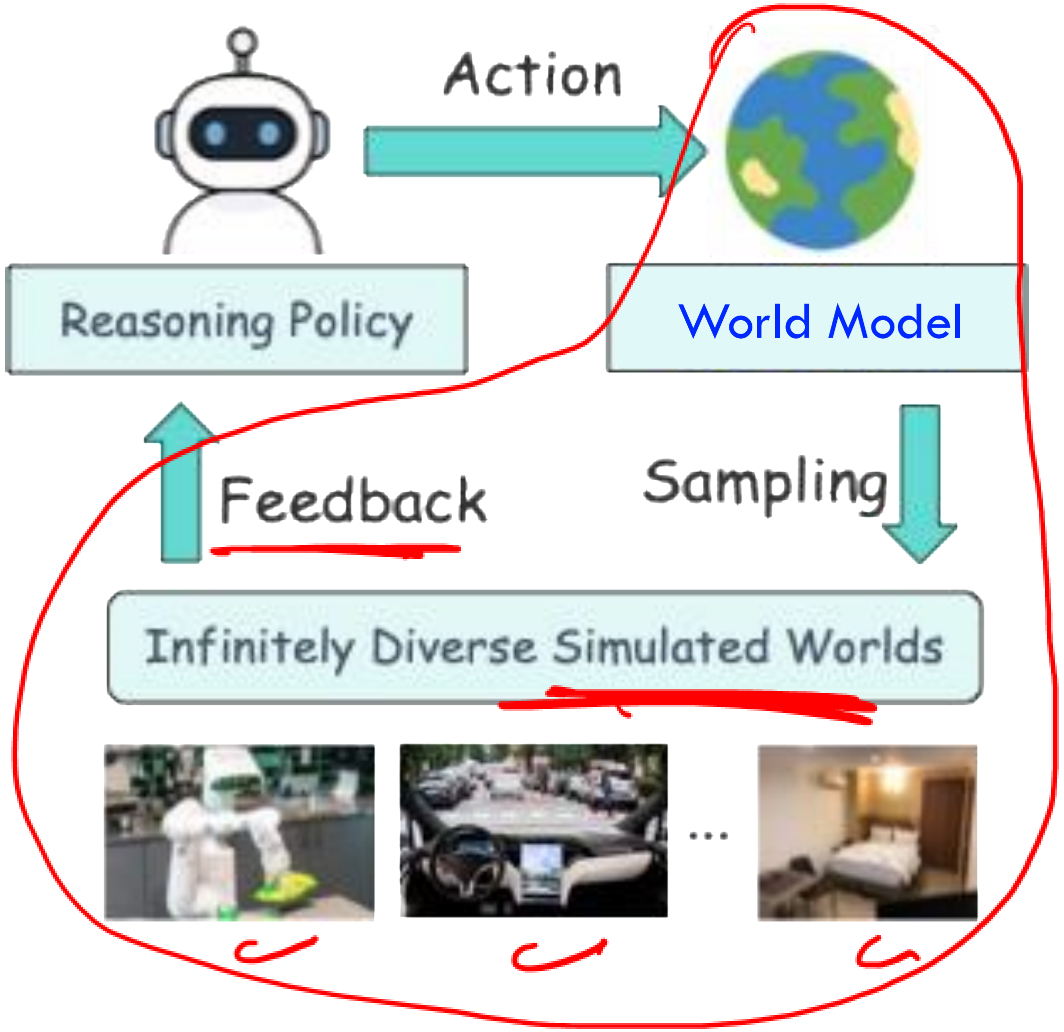
Baidu

Reinforcement Learning

- Deployed in infinitely diverse simulated worlds
- Cheap, fast to get feedback



- Deployed in the real world
- Expensive, slow to get feedback



World model

- State transition probabilities
- Next "world" prediction

~ next-word pred: LLM.

$$P(s'|s,a)$$

next state

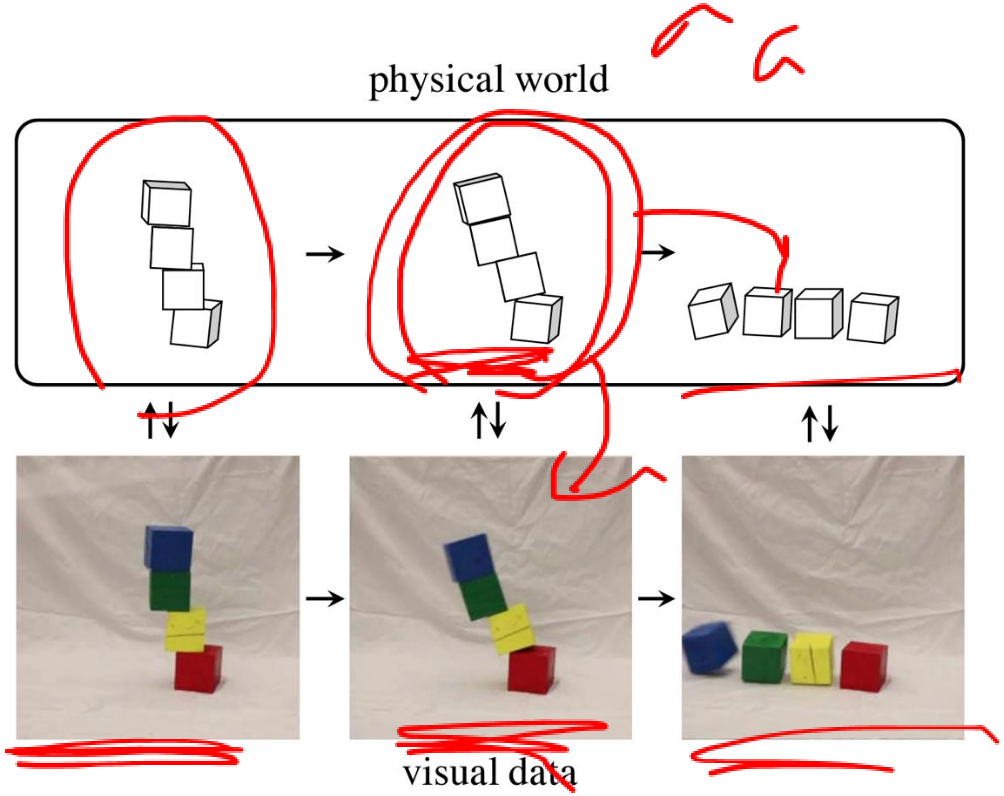
current state

action

$r(s,a)$

World model

- Next "world" prediction $P(s'|s,a)$
- Prior research built domain-specific world models
  - Primarily in robotics and embodied AI



(i) Computer vision: model-based physical scene understanding

Wu et al. (2017)

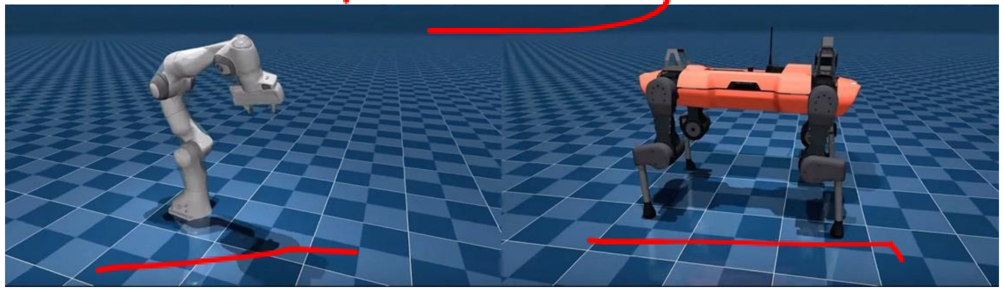
World model

Sim-to-real

- Next "world" prediction $P(s'|s,a)$
- Prior research built domain-specific world models
  - Primarily in robotics and embodied AI

(ii) Physics engines / embodied simulators

MuJoCo



Todorov et al. (2012)

AI2-THOR

Kolve et al. (2017)

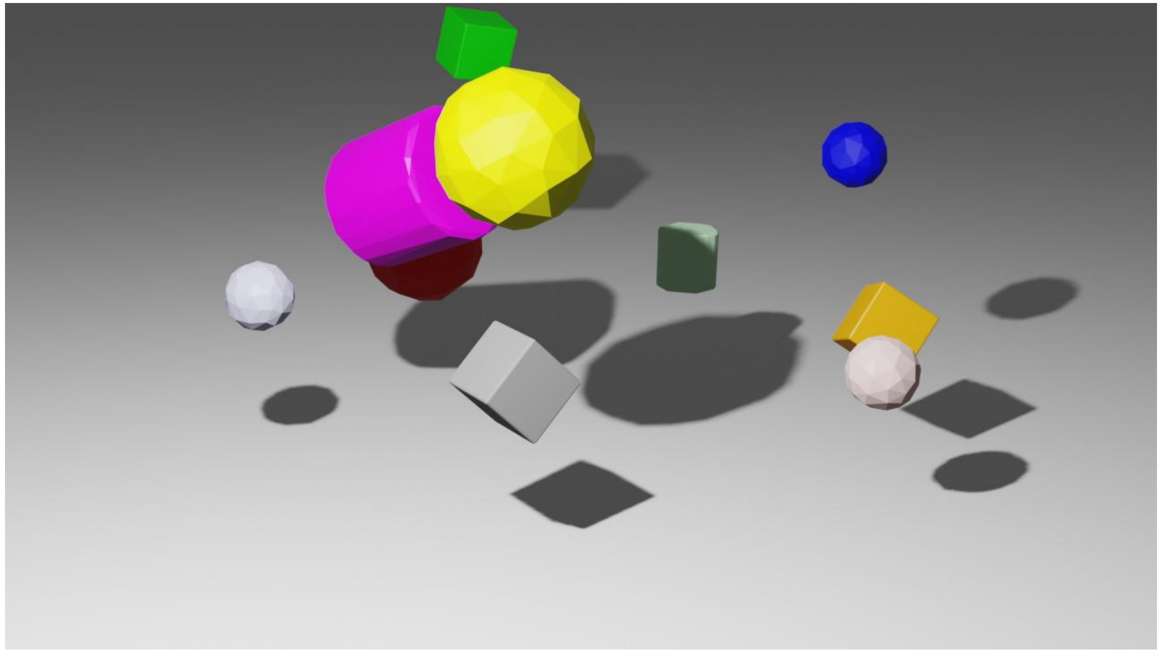
Habitat 2.0

Szot et al. (2021)

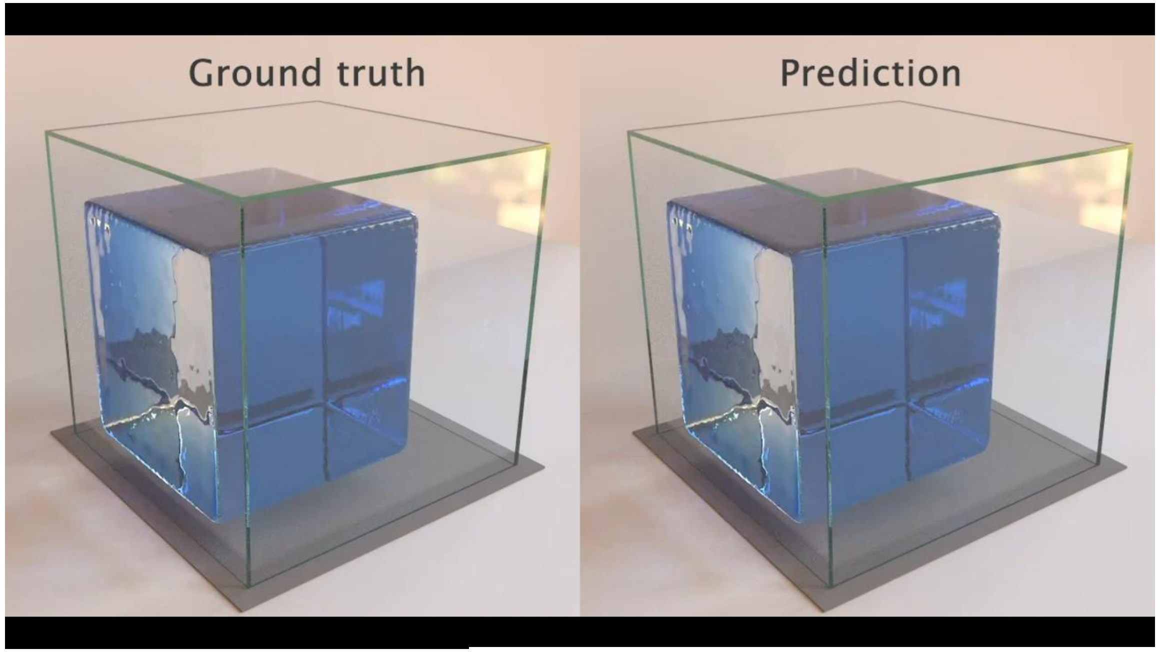
World model

- Next "world" prediction $P(s'|s,a)$
- Prior research built domain-specific world models
  - Primarily in robotics and embodied AI

(iii) Learned neural physics engines



Allen et al. (2023)



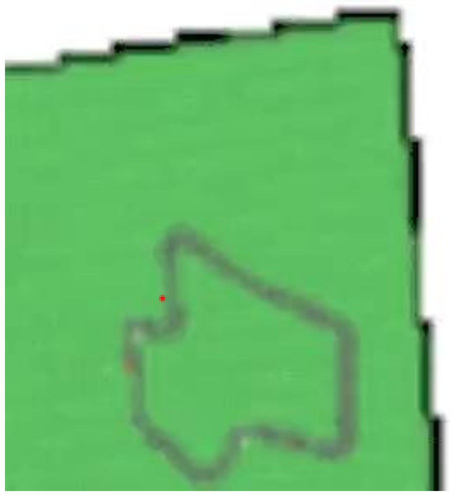
Sanchez-Gonzalez et al. (2020)

World model

- Next "world" prediction $P(s'|s,a)$
- Prior research built domain-specific world models
  - Primarily in robotics and embodied AI

(iv) Video prediction models

Ground-truth



Synthesis

Ha & Schmidhuber (2018)

World model

- Next "world" prediction $P(s'|s, a)$
- Prior research built domain-specific world models
  - Primarily in robotics and embodied AI

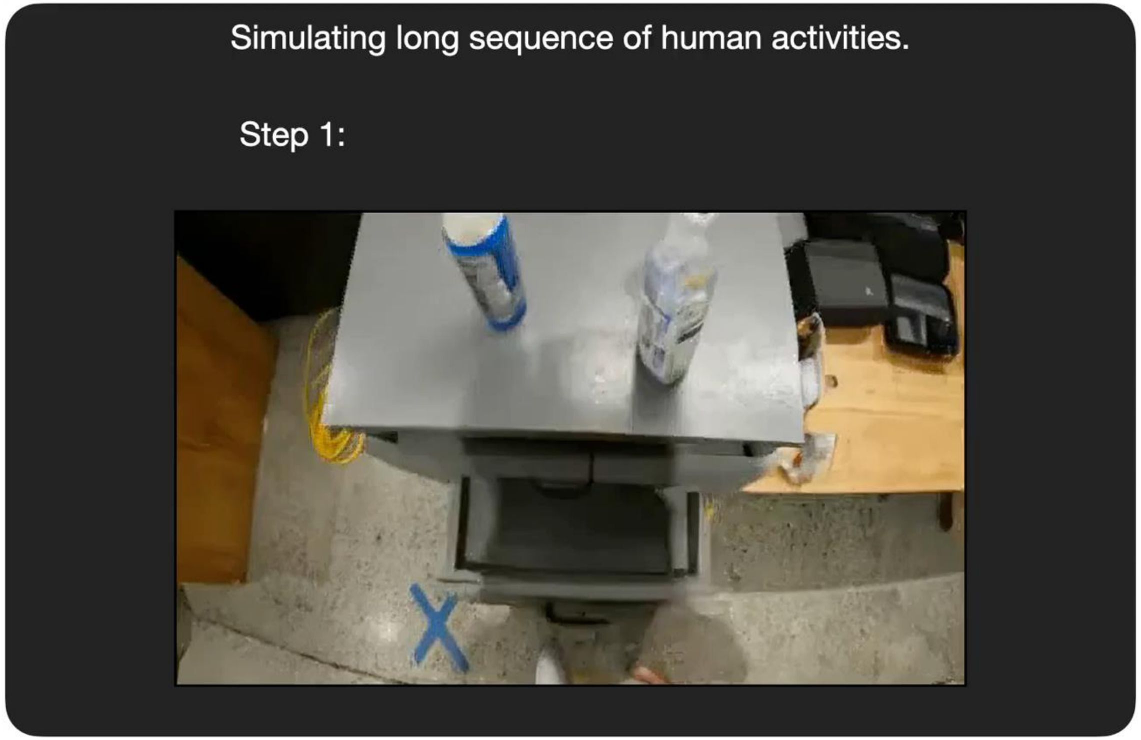
(iv) Video prediction models

GAIA-1

World model

- Next "world" prediction $P(s'|s, a)$
- Prior research built domain-specific world models
  - Primarily in robotics and embodied AI

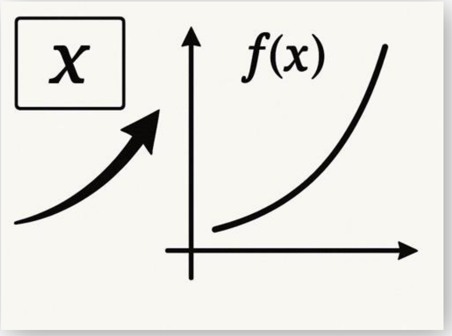
(iv) Video prediction models



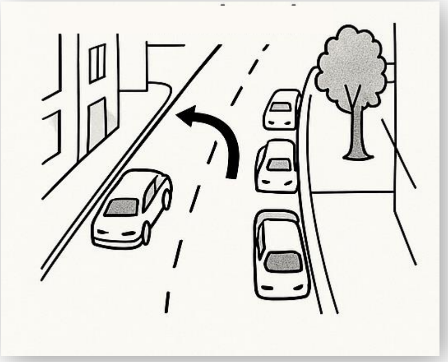
[Yang et al., 2023]

World model

- Next "world" prediction $P(s'|s,a)$
- Prior research built domain-specific world models
  - Primarily in robotics and embodied AI
- The scope of simulation defines the capability of reasoning
  - "More simulation, more intelligence"

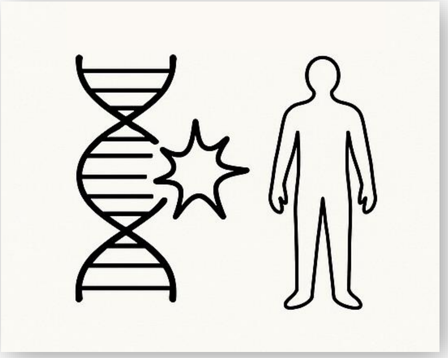


Would the value of a function increase if I changed this variable?



What would happen on the street if I turned the vehicle left

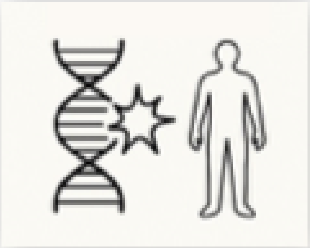
How might a business grow if a specific policy were applied



What biological effects would arise from a specific genetic mutation

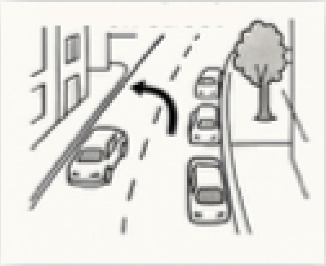
The scope of WM simulation
⇒ the capability of reasoning

Scientific reasoning
(drug discovery)

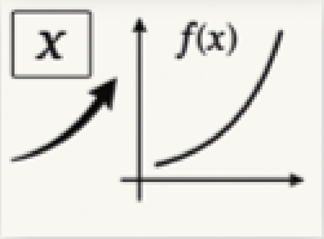


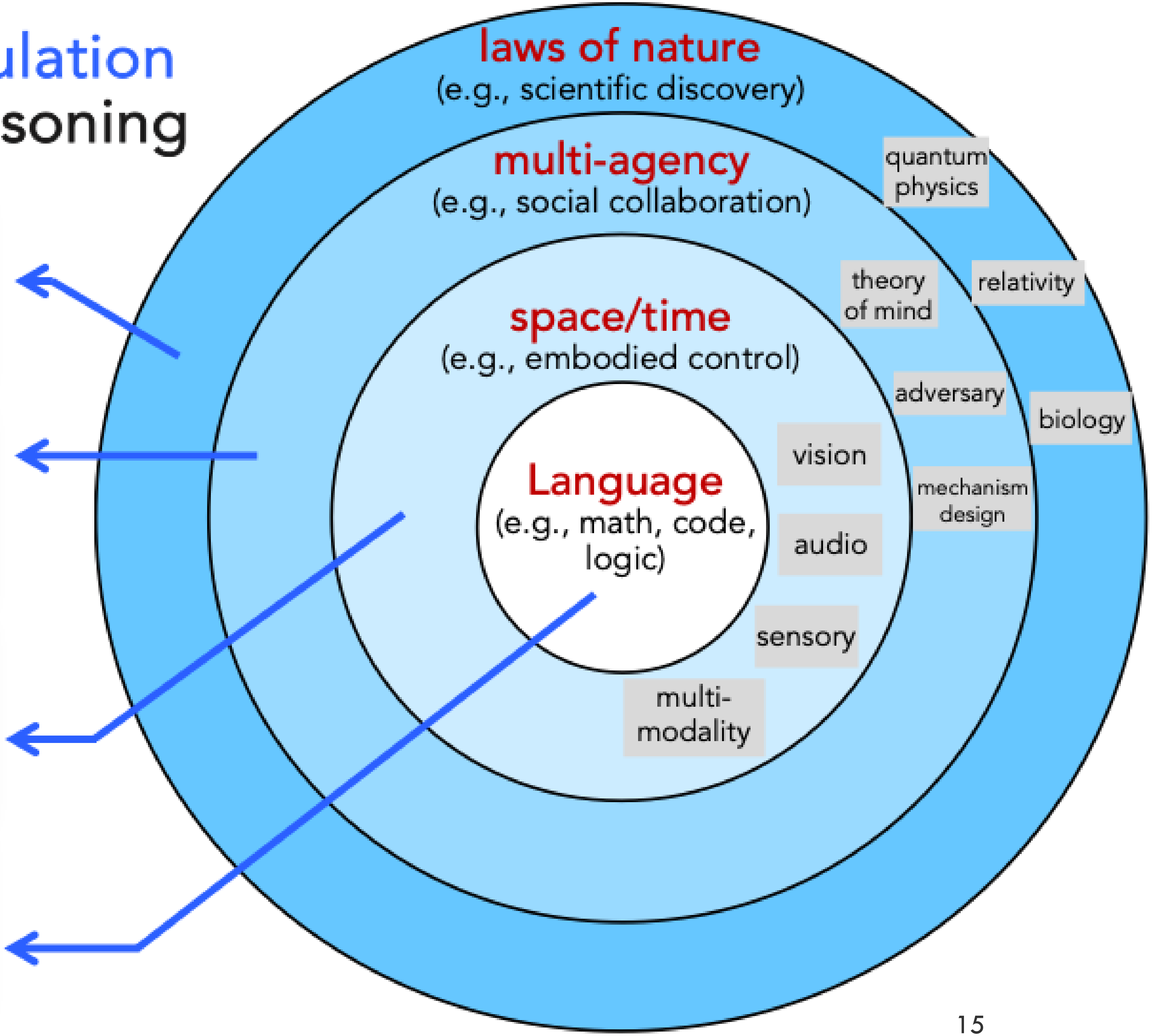
Social/business reasoning

Physical reasoning
(auto-driving, robot)



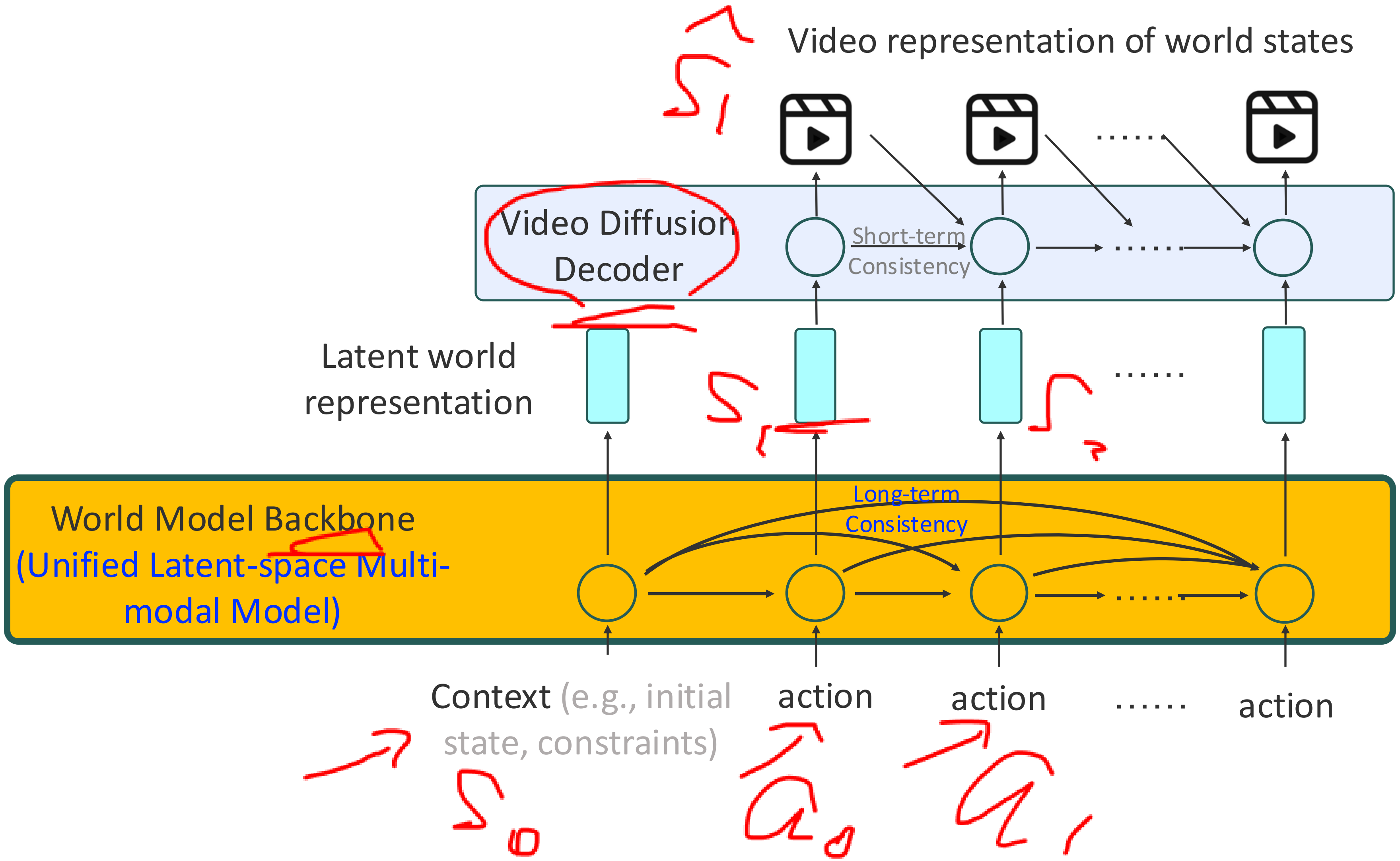
Symbolic reasoning
Computer use

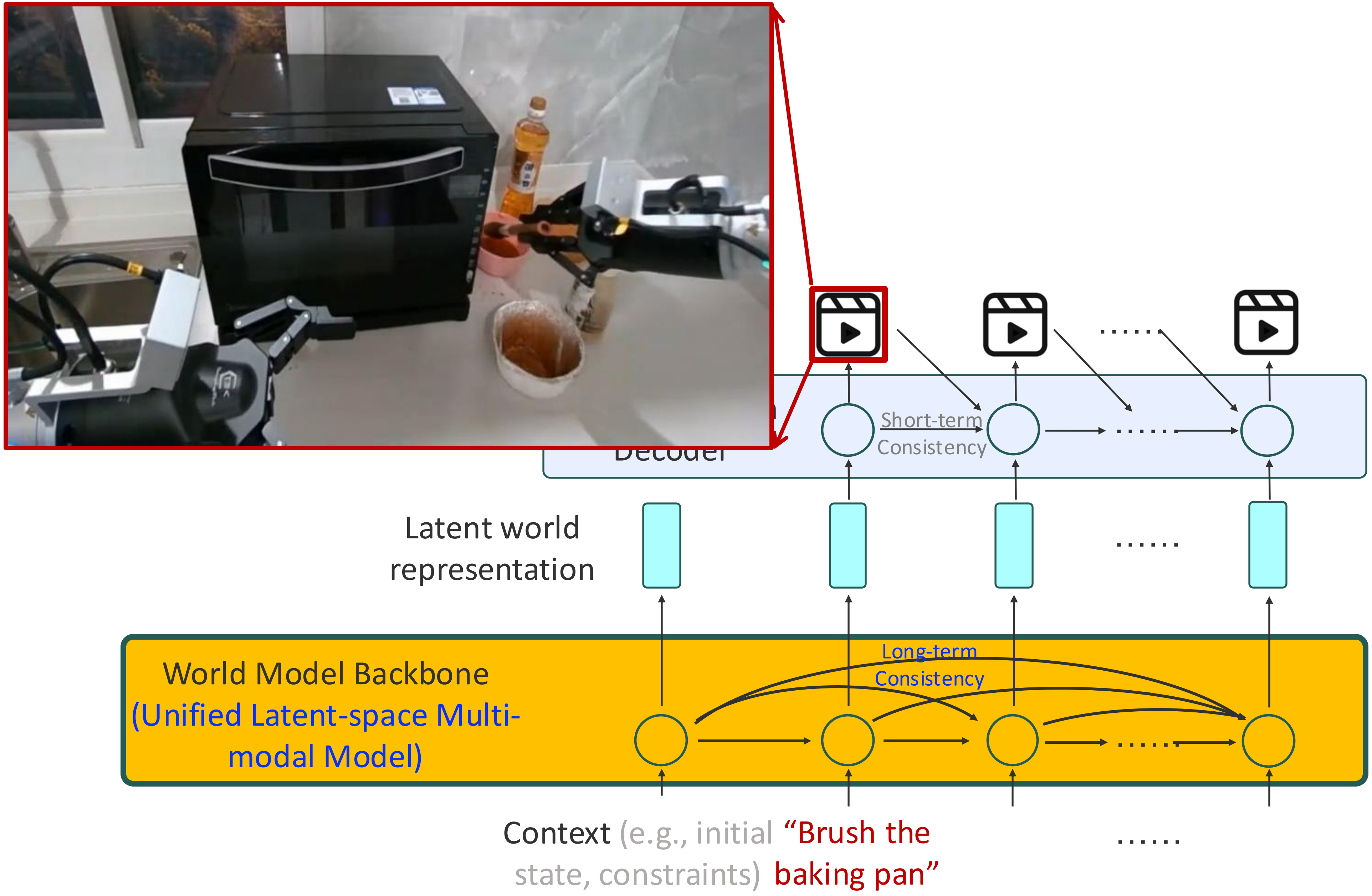


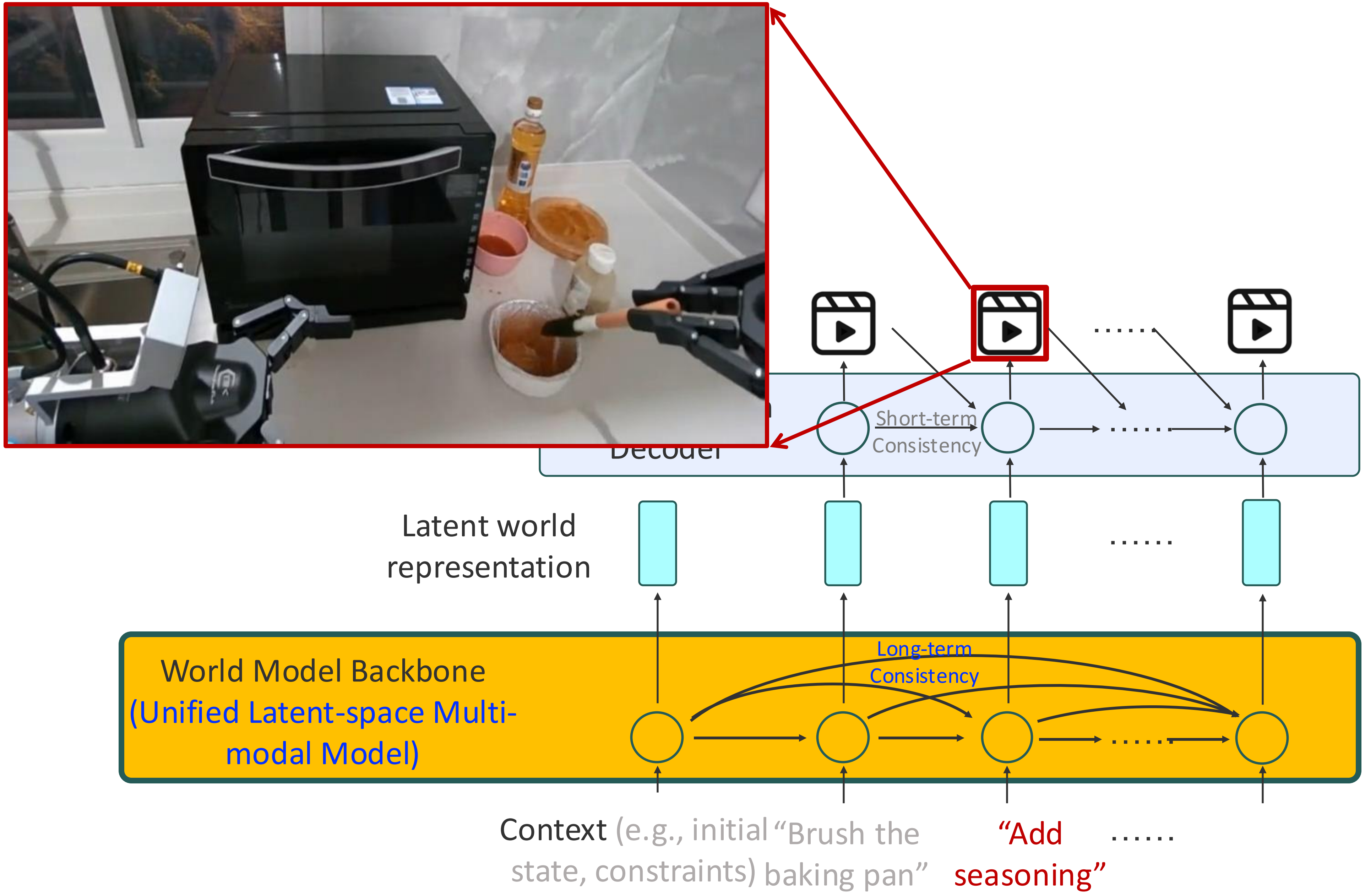


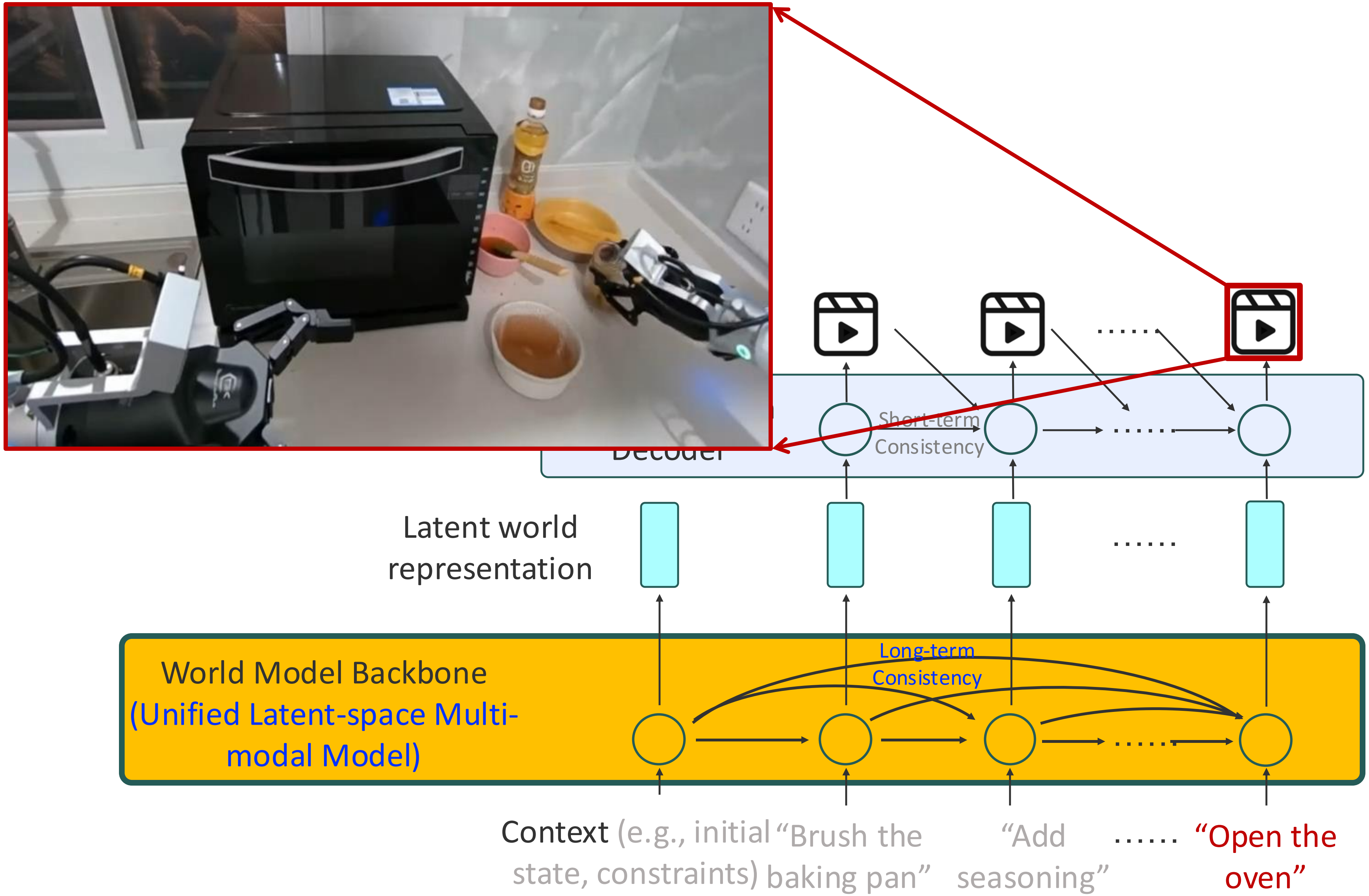
PAN World Model

(Physical, Agentic, Nested)

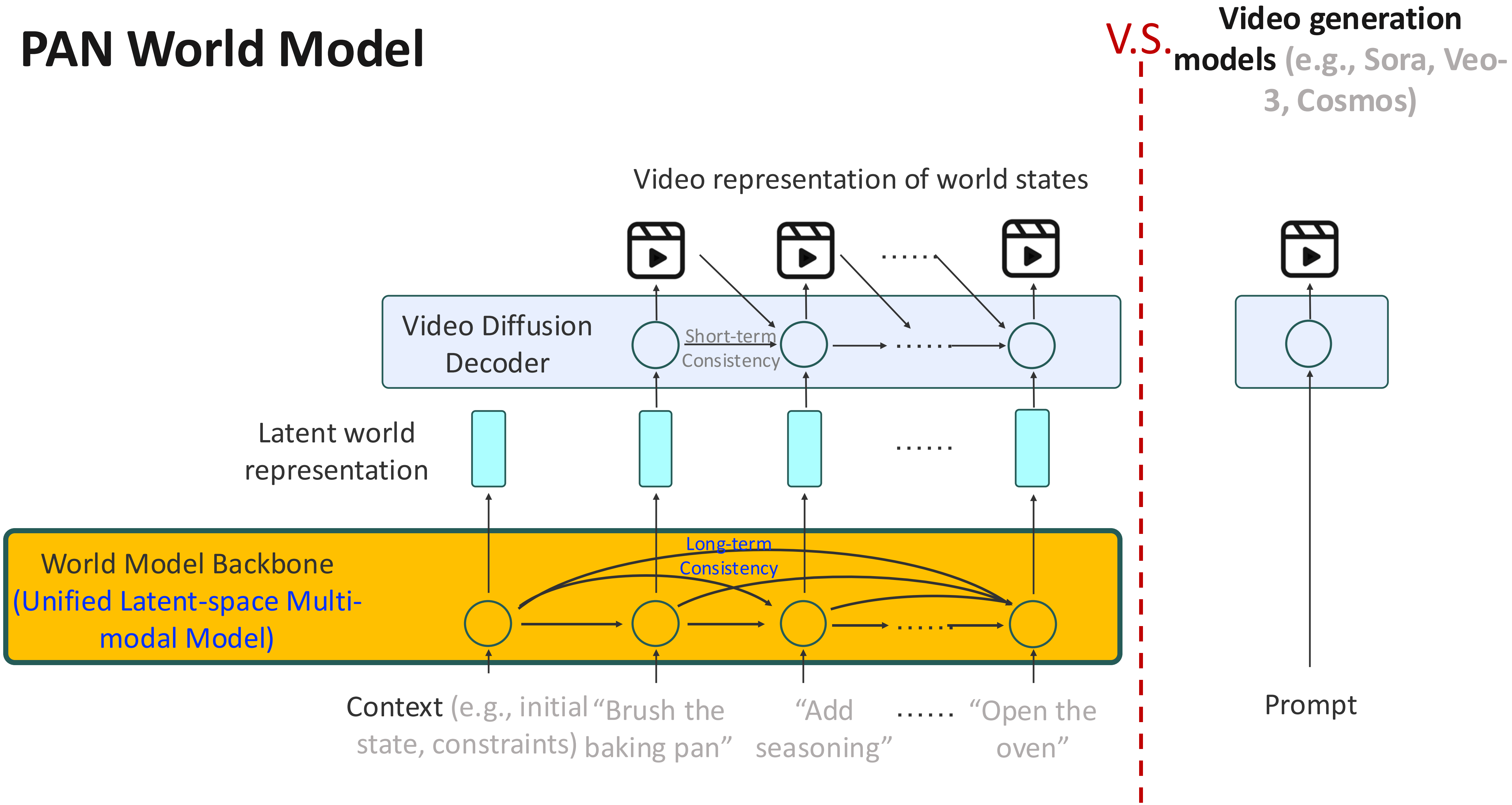



Decoder
Short-term
Consistency
Latent world representation
World Model Backbone
(Unified Latent-space Multi-modal Model)
Long-term
Consistency
Context (e.g., initial state, constraints)
"Brush the baking pan"
......


Short-term
Consistency
Latent world
representation
......
World Model Backbone
(Unified Latent-space Multi-
modal Model)
Long-term
Consistency
Context (e.g., initial
state, constraints)
“Brush the
baking pan”
“Add
seasoning”
......


Short-term
Consistency
Latent world
representation
......
World Model Backbone
(Unified Latent-space Multi-
modal Model)
Long-term
Consistency
Context (e.g., initial
state, constraints)
"Brush the
baking pan"
"Add
seasoning"
......
"Open the
oven"

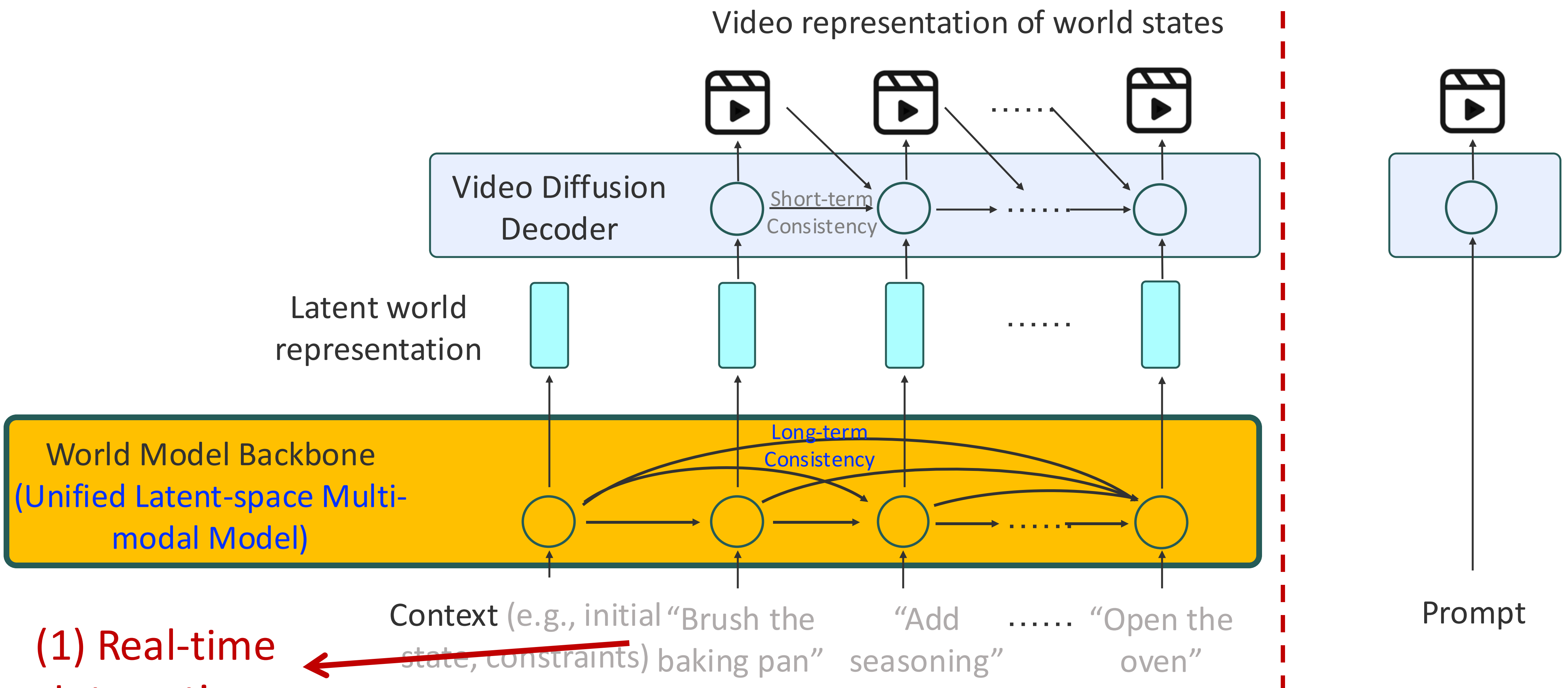
PAN World Model



PAN World Model

$P(s'|s,a)$ V.S.

Video generation models (e.g., Sora, Veo-3, Cosmos)

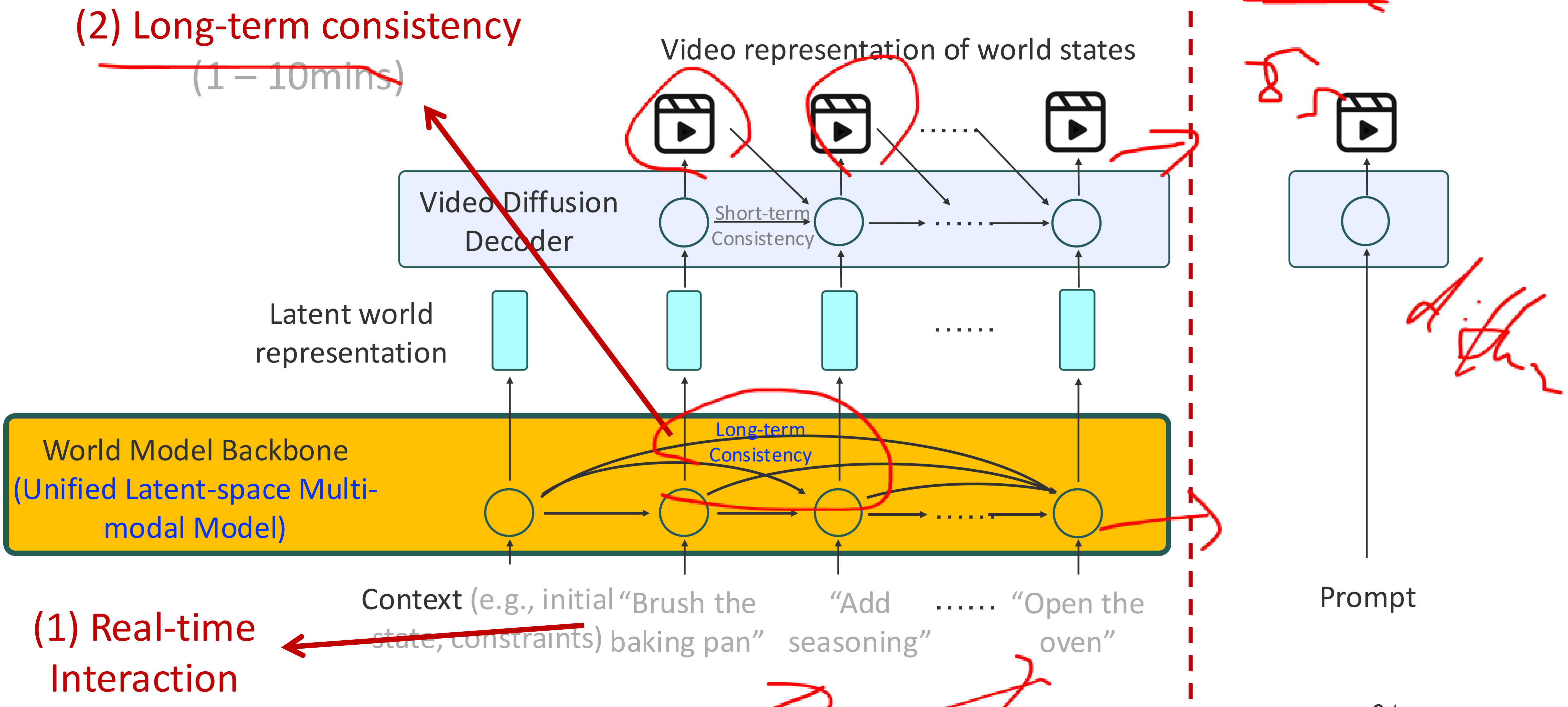


(1) Real-time Interaction

PAN World Model

Auto-regressive V.S.

Video generation models (e.g., Sora, Veo-3, Cosmos)



PAN World Model

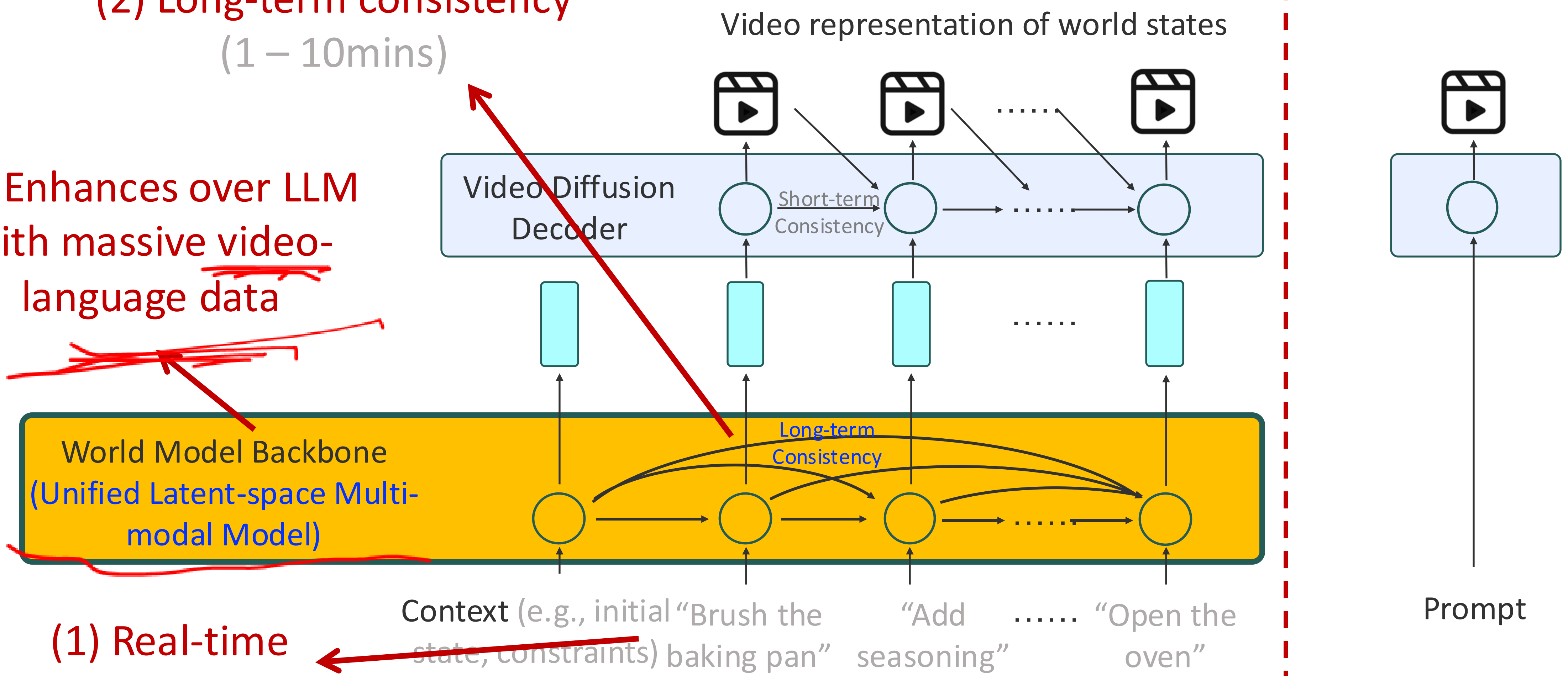
(2) Long-term consistency
(1 – 10mins)

(3) Enhances over LLM with massive video-language data

(1) Real-time Interaction

V.S.

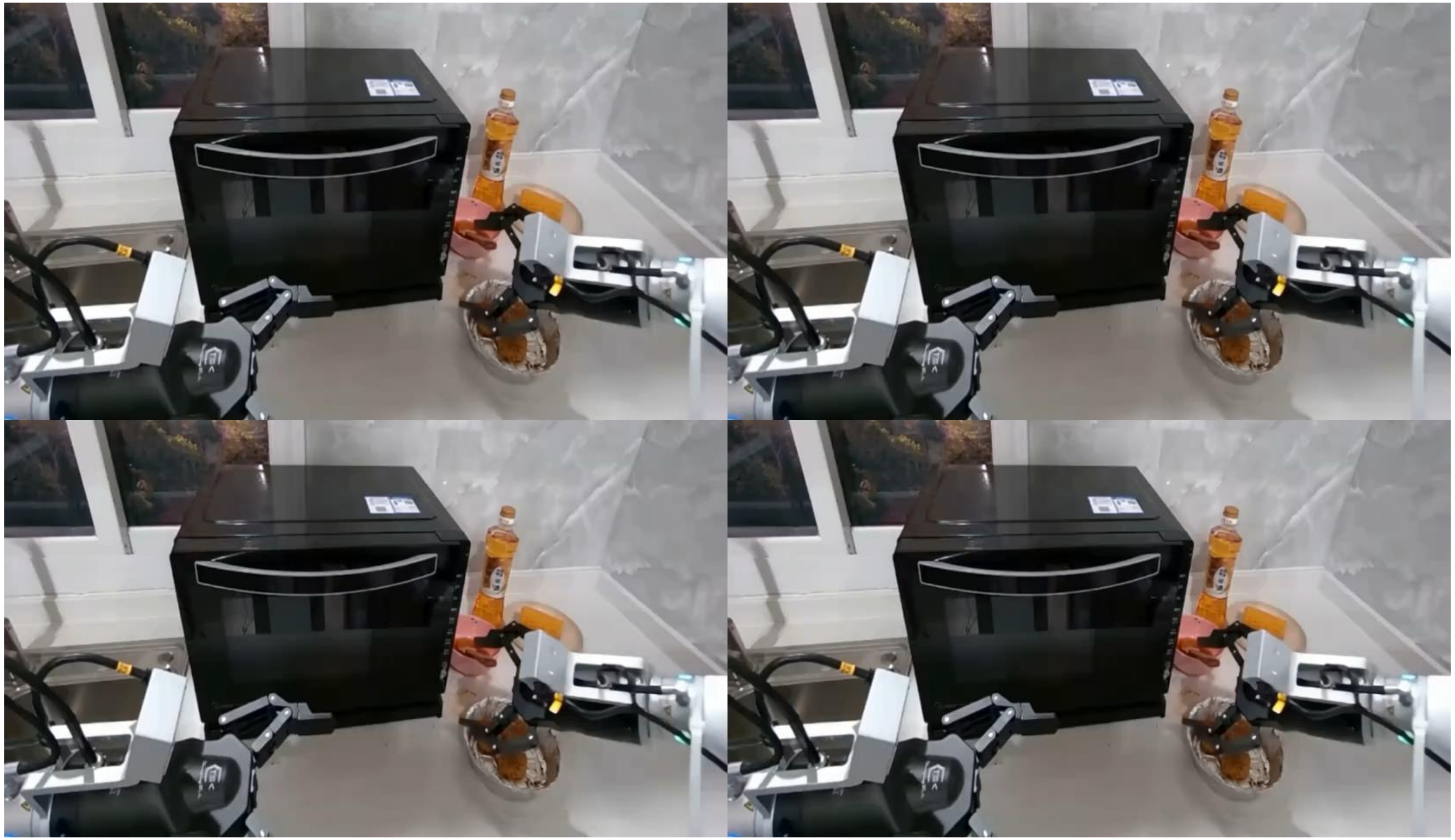
Video generation models (e.g., Sora, Veo-3, Cosmos)



PAN World Model: Simulation Results

Robot: Complex manipulation

Locomotion



PAN World Model: Simulation Results

Robot: Complex manipulation

Locomotion

Input actions:

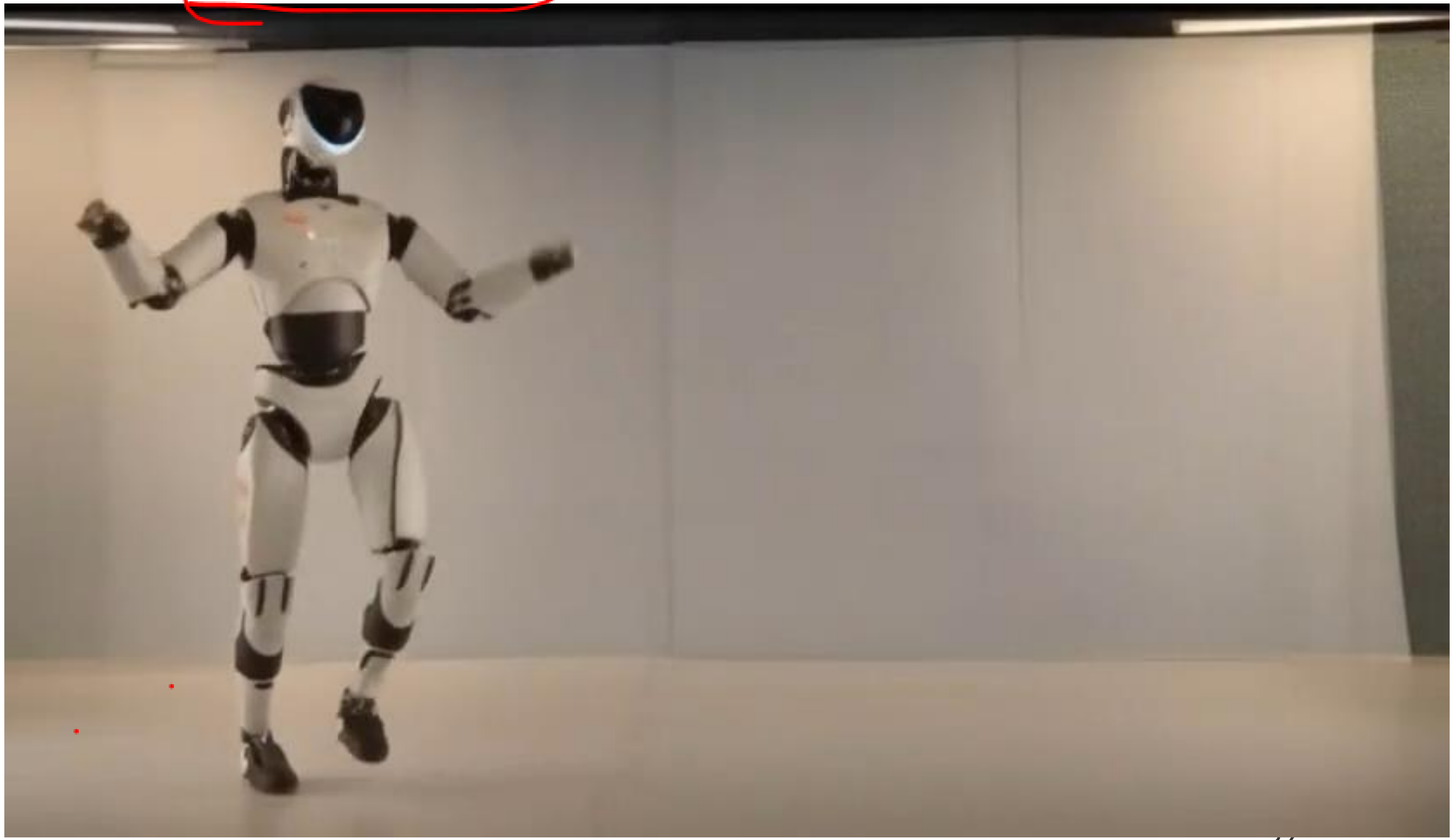
> Wave arms and jump

> Dance dance dance!

> Grasp a rose behind and show to the audience

> …

> Make a heart shape with hands



PAN World Model: Simulation Results

Driving: Dangerous situations

PAN World Model: Simulation Results

Complex environments in various styles

PAN World Model: Simulation Results

Complex environments in various styles

PAN World Model: Simulation Results

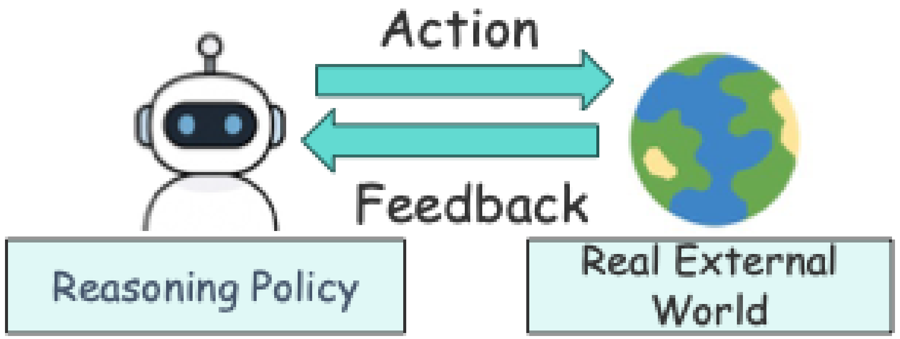
Complex environments in various styles

Summary so far

wake

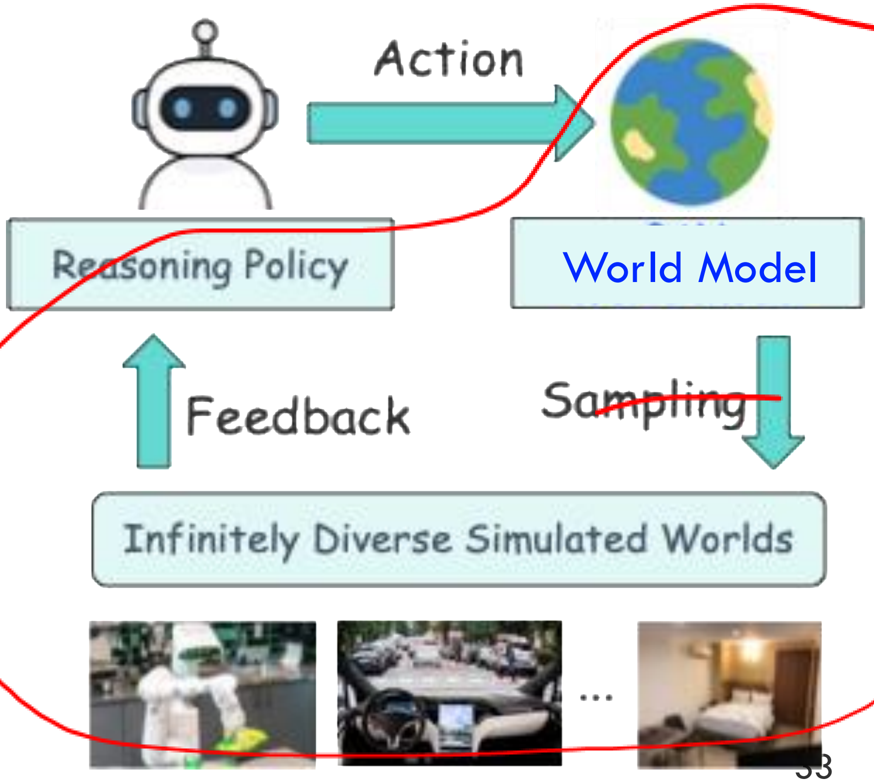
Traditional Reinforcement Learning

- Deployed in the real world
- Expensive, slow to get feedback

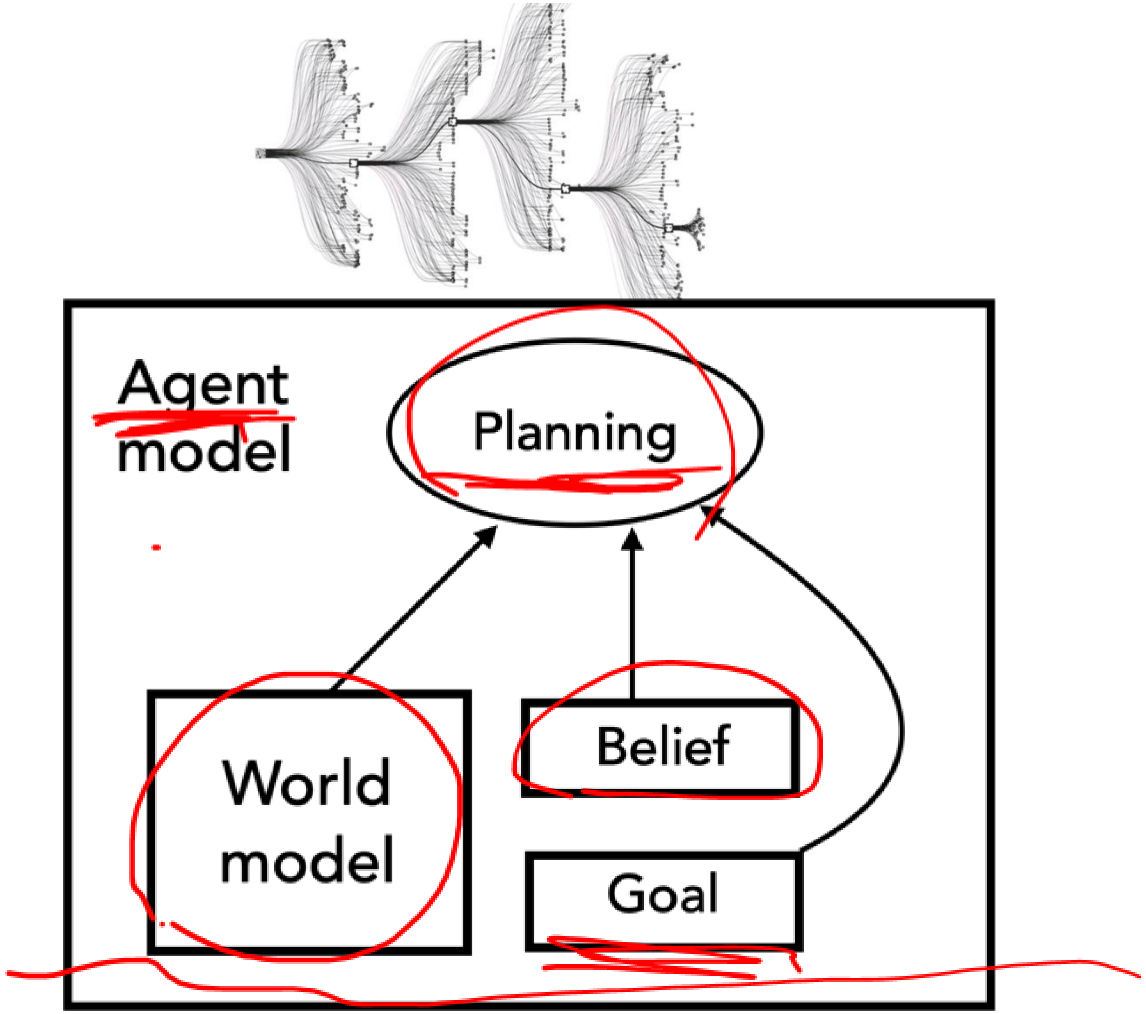


"Dream"-time learning

- Deployed in infinitely diverse simulated worlds
- Cheap, fast to get feedback

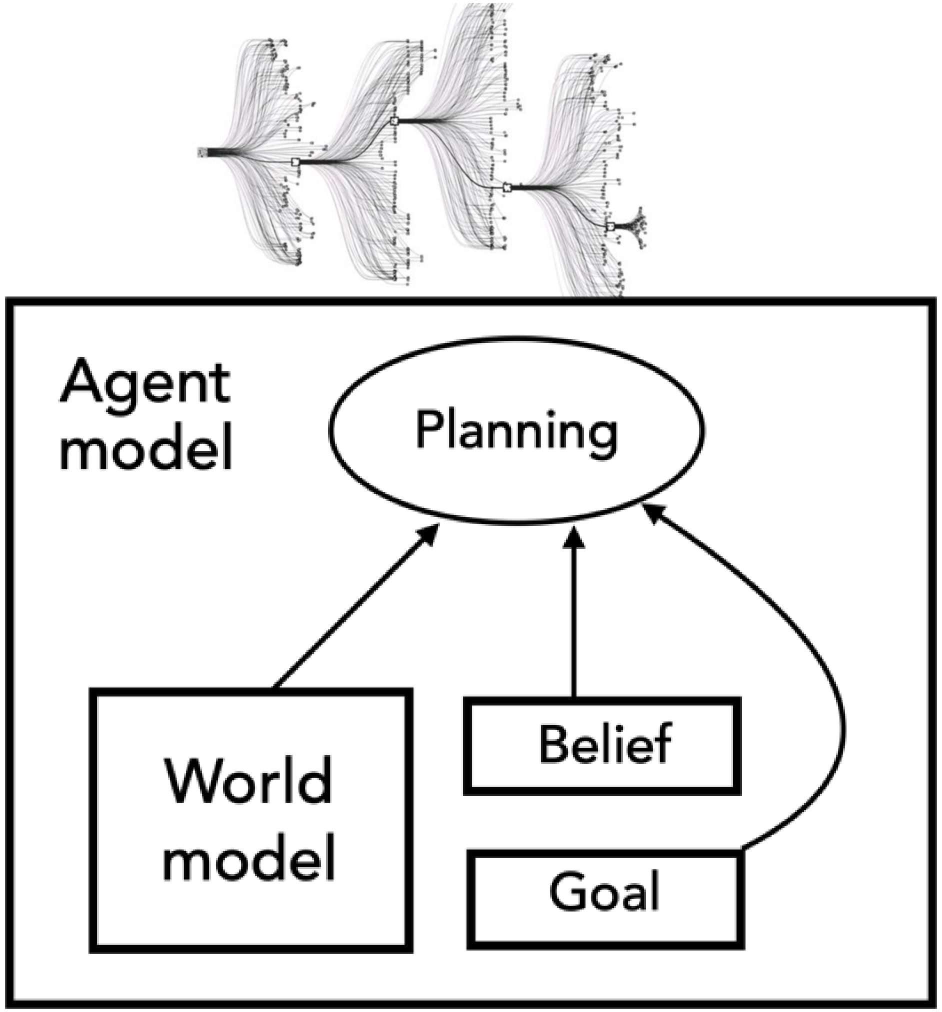


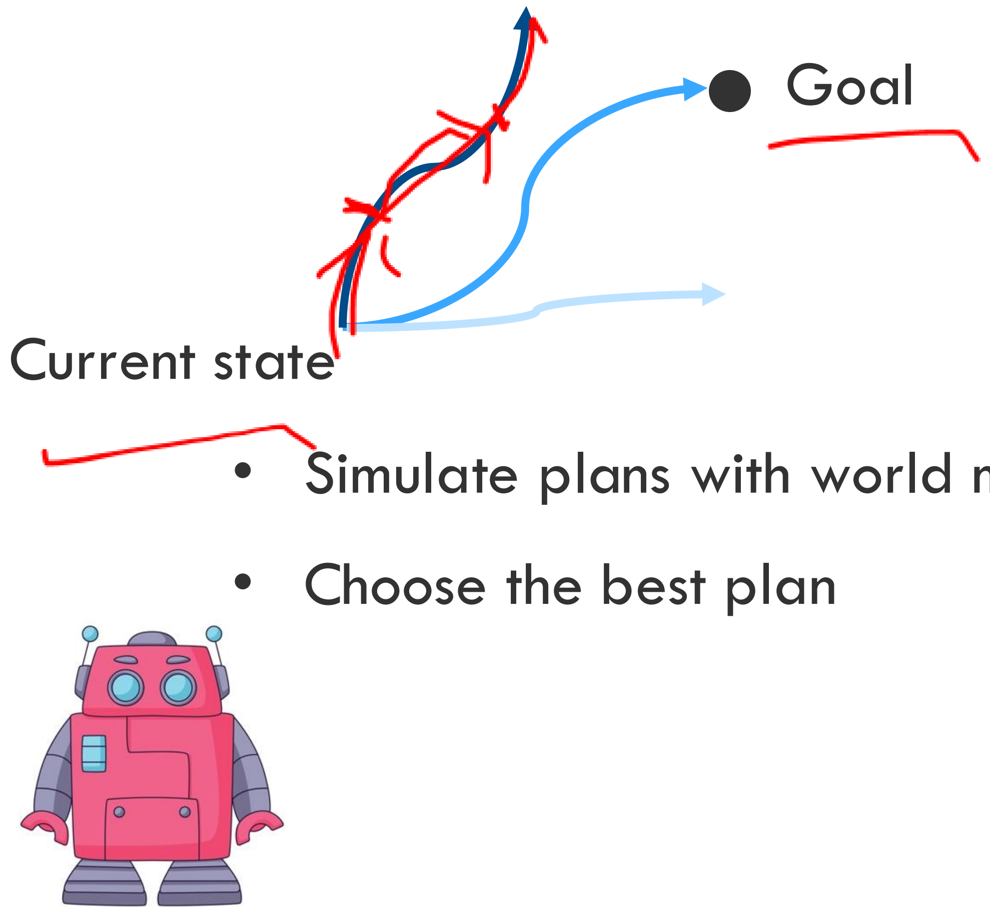
World Model for Inference-Time Planning



$$P(s'|s,a)$$

World Model for Inference-Time Planning



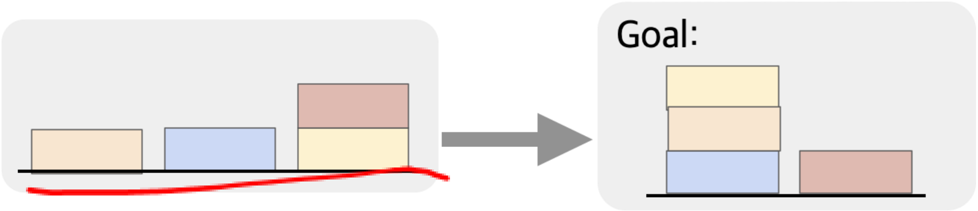


- Simulate plans with world model
- Choose the best plan

$$P(s'|s,a)$$

World Model for Inference-Time Planning

How to move the blocks to the goal state?



LLMs: Autoregressive plan generation

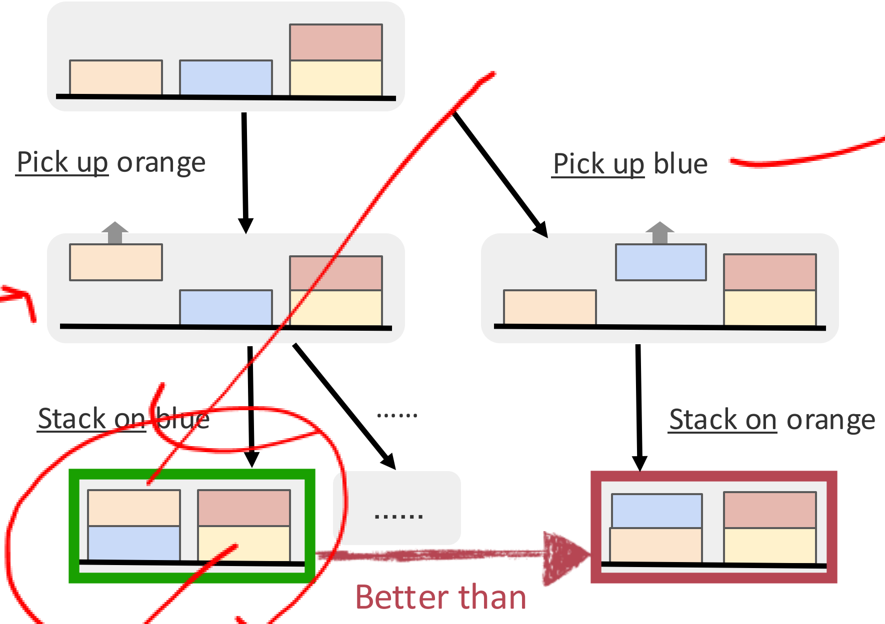
GPT-4

Invalid Action!
The yellow block is still under the red one.

1. Pick up the orange block.
2. Stack it on the blue block.
3. Pick up the yellow block. ✕
4. Stack it on the orange block.
5. Pick up the red block.
6. Put it on the table.

Human: strategic planning

- Internal world model to predict states
- Simulation of alternative plans
- Assess outcomes to refine/pick the best



How Nature Works

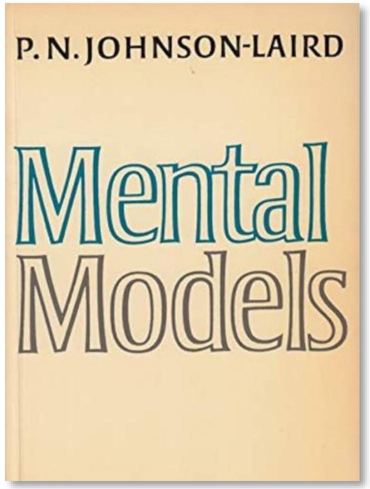
Simulates possibilities recursively; complexity emerges

How Nature Works

Simulates possibilities recursively; complexity emerges

Example 1: Human reasoning

- Humans “reason by thinking about what’s possible”

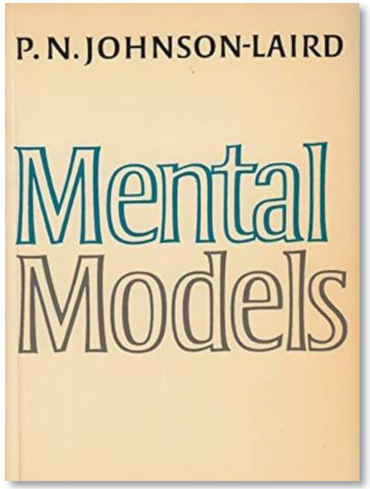


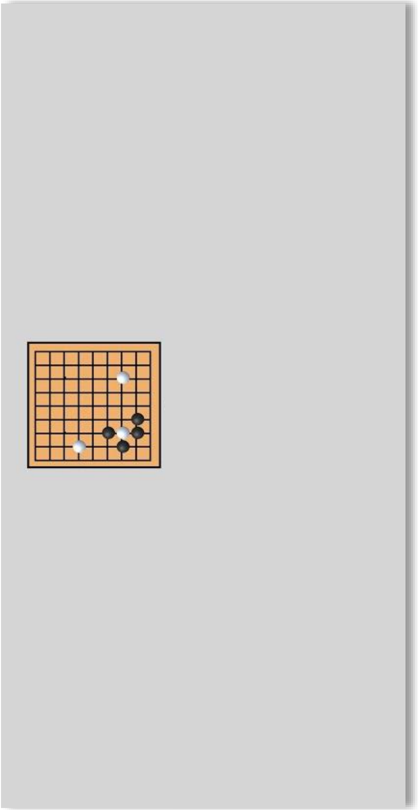
How Nature Works

Simulates possibilities recursively; complexity emerges

Example 1: Human reasoning

- Humans “reason by thinking about what’s possible”



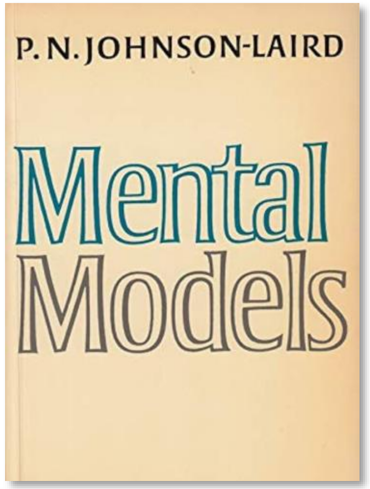


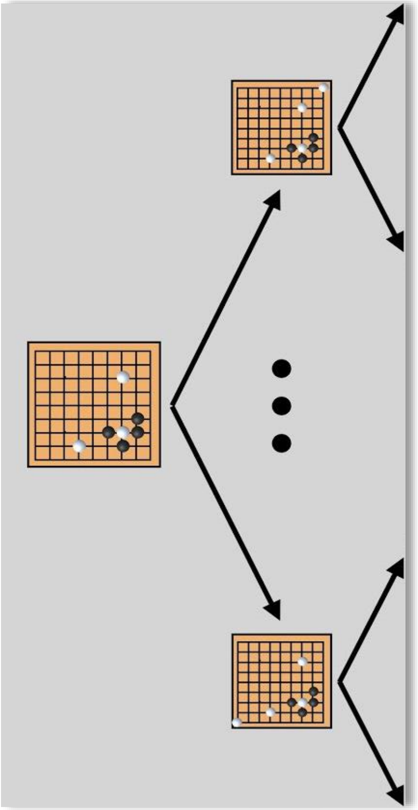
How Nature Works

Simulates possibilities recursively; complexity emerges

Example 1: Human reasoning

- Humans “reason by thinking about what’s possible”





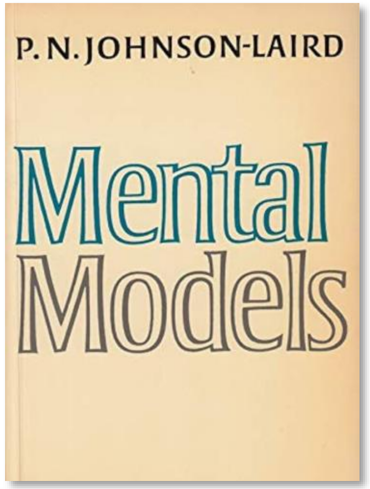
- Simulate alternative hypothetical worlds with **mental models**

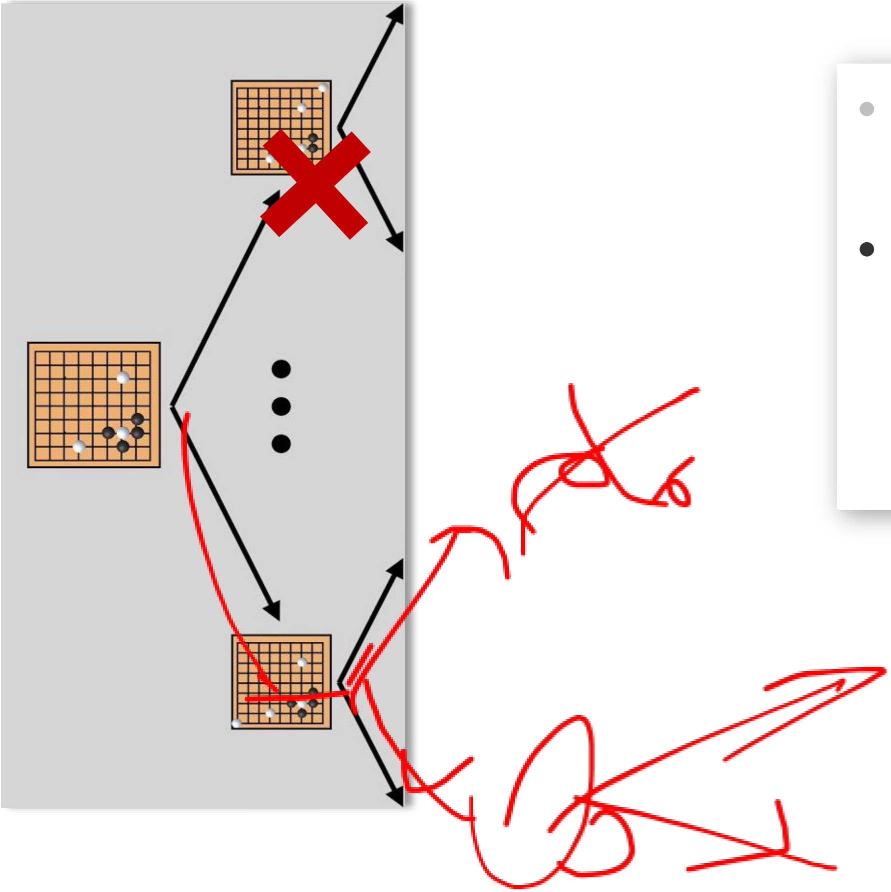
How Nature Works

Simulates possibilities recursively; complexity emerges

Example 1: Human reasoning

- Humans “reason by thinking about what’s possible”





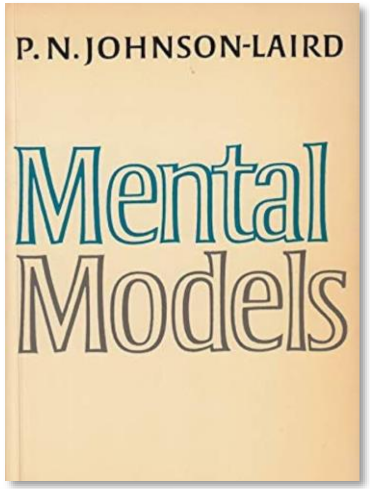
- Simulate alternative hypothetical worlds with **mental models**
- Rule out possibilities that do not fit context, knowledge, or goals

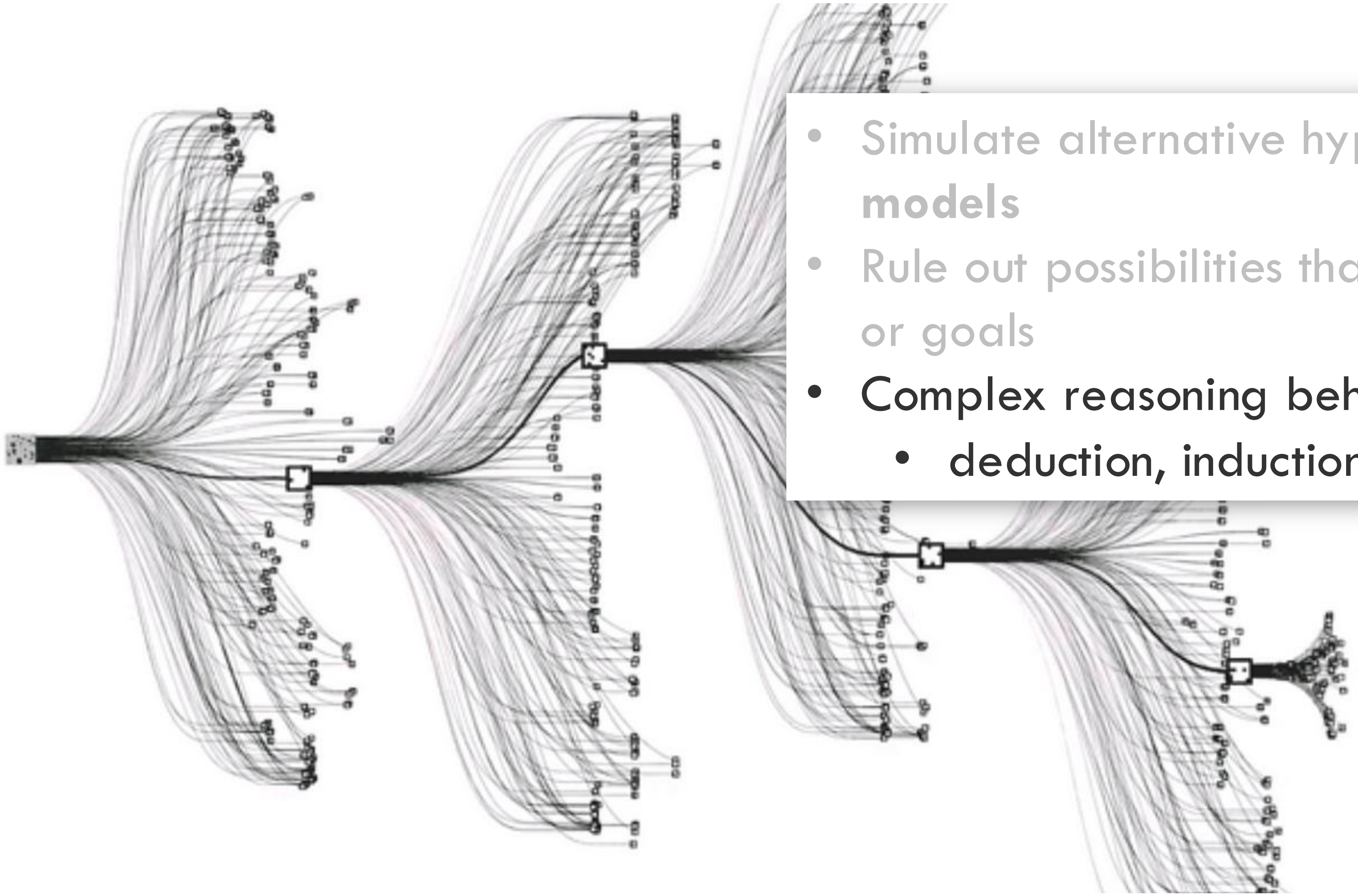
How Nature Works

Simulates possibilities recursively; complexity emerges

Example 1: Human reasoning

- Humans "reason by thinking about what's possible"

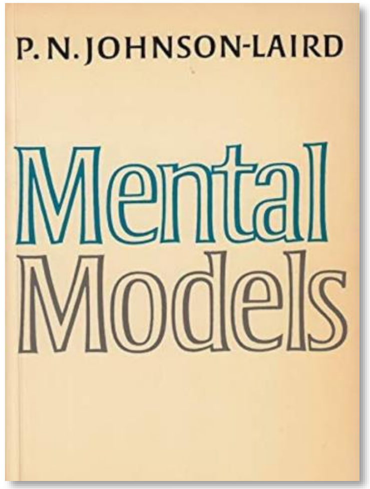




- Simulate alternative hypothetical worlds with **mental models**
- Rule out possibilities that do not fit context, knowledge, or goals
- Complex reasoning behaviors emerge
  - deduction, induction, abduction, …

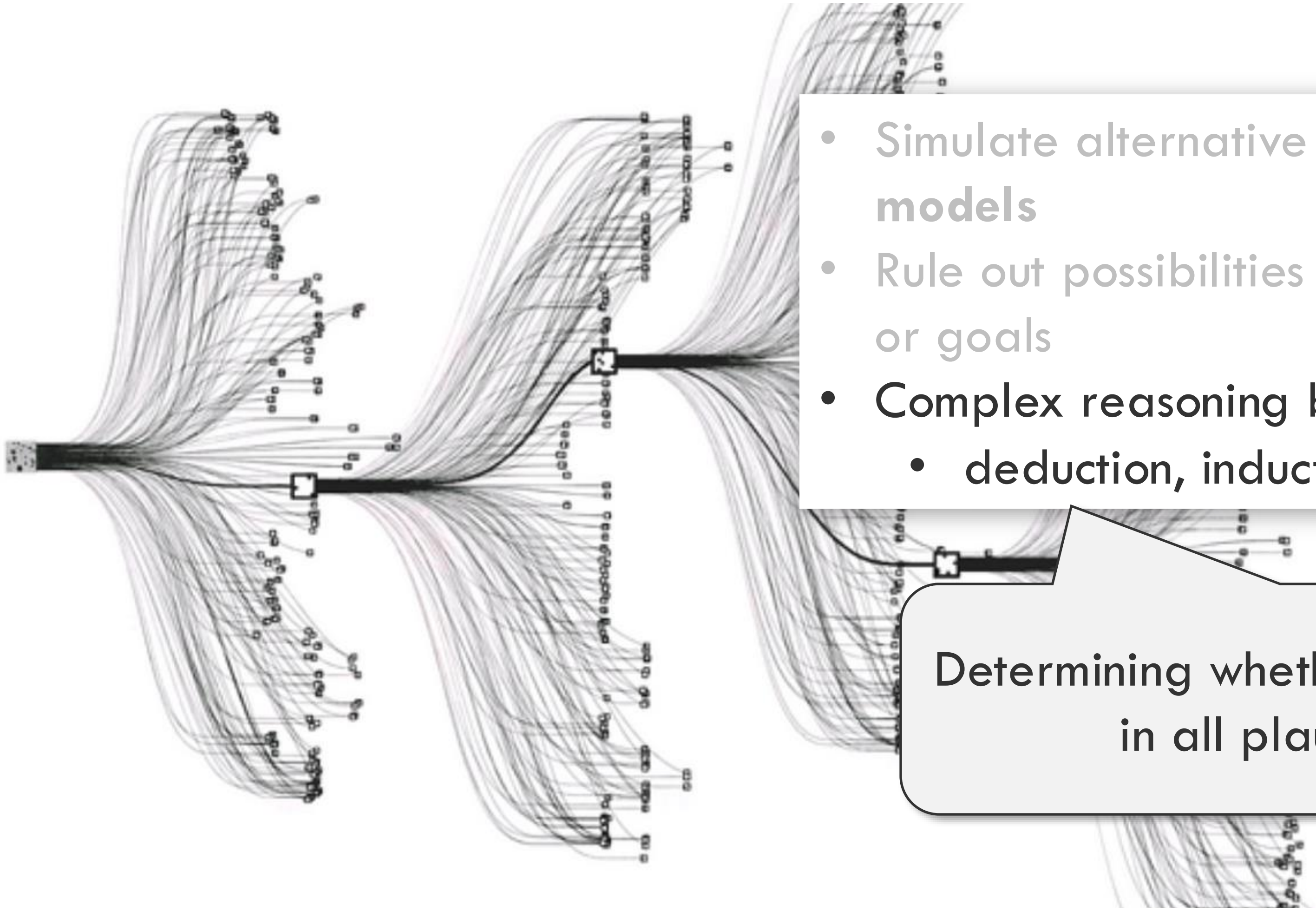
How Nature Works

Simulates possibilities recursively; complexity emerges



Example 1: Human reasoning

- Humans "reason by thinking about what's possible"



- Simulate alternative hypothetical worlds with **mental models**
- Rule out possibilities that do not fit context, knowledge, or goals
- Complex reasoning behaviors emerge
  - deduction, induction, abduction, …

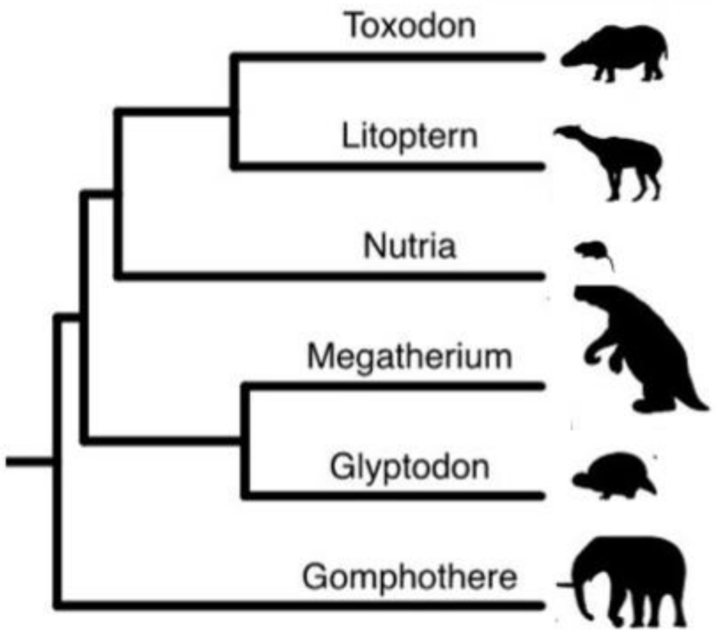
Determining whether a conclusion holds in all plausible worlds

How Nature Works

Simulates possibilities recursively; complexity emerges

Example 2: Natural evolution

- Generate mutations with molecular genetic mechanisms (genotype -> phenotype)

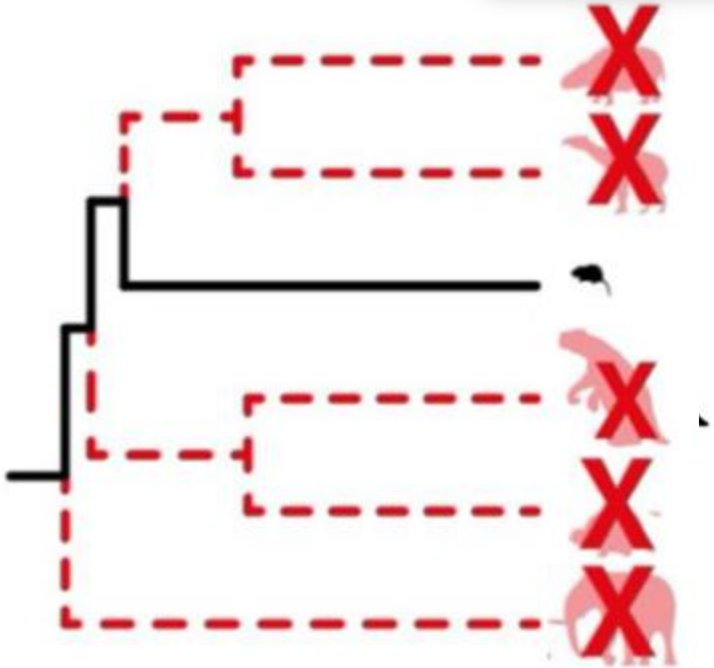


How Nature Works

Simulates possibilities recursively; complexity emerges

Example 2: Natural evolution

- Generate mutations with molecular genetic mechanisms (genotype -> phenotype)
- Rule out possibilities that do not fit the environment

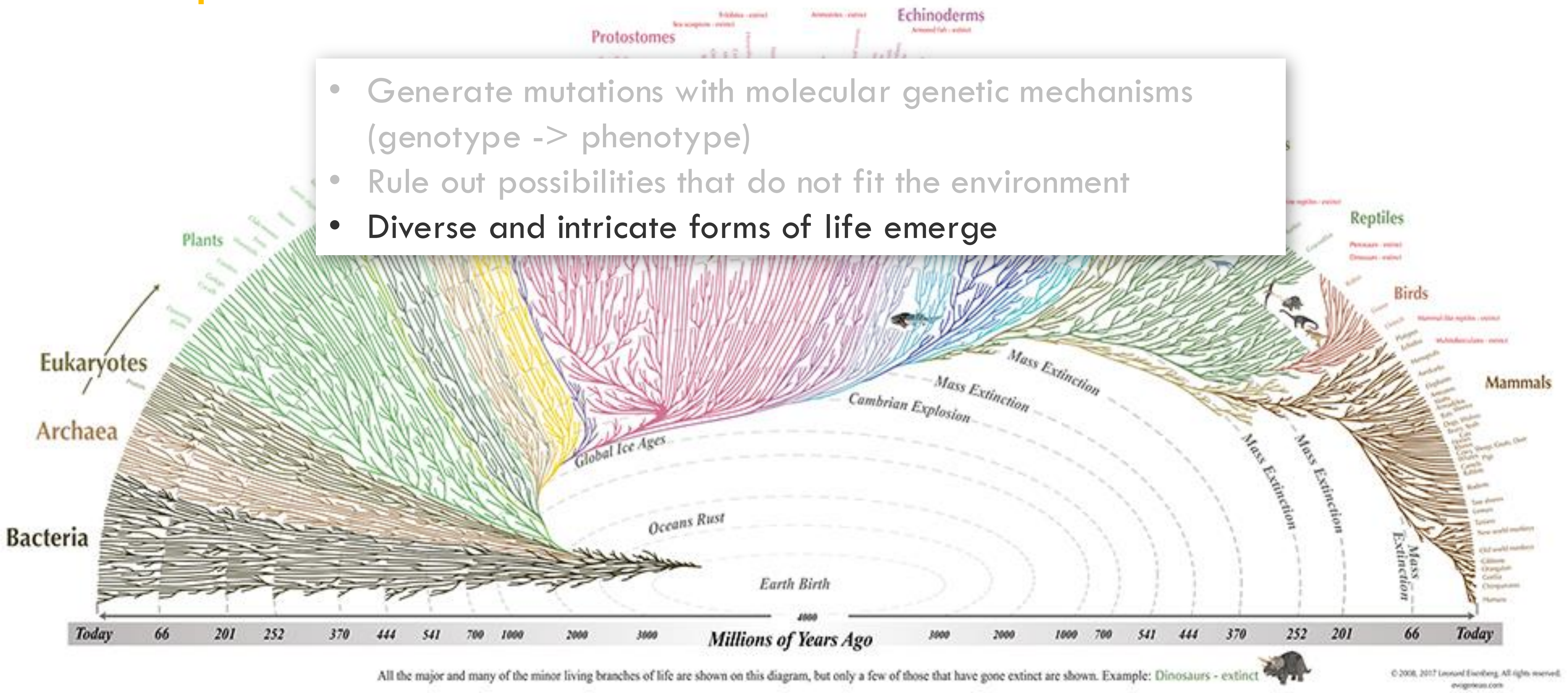


How Nature Works

Simulates possibilities recursively; complexity emerges

Example 2: Natural evolution

- Generate mutations with molecular genetic mechanisms (genotype -> phenotype)
- Rule out possibilities that do not fit the environment
- Diverse and intricate forms of life emerge



How Nature Works

Simulates possibilities recursively; complexity emerges

Human reasoning

Natural evolution

Mental models

Molecular genetic mechanisms

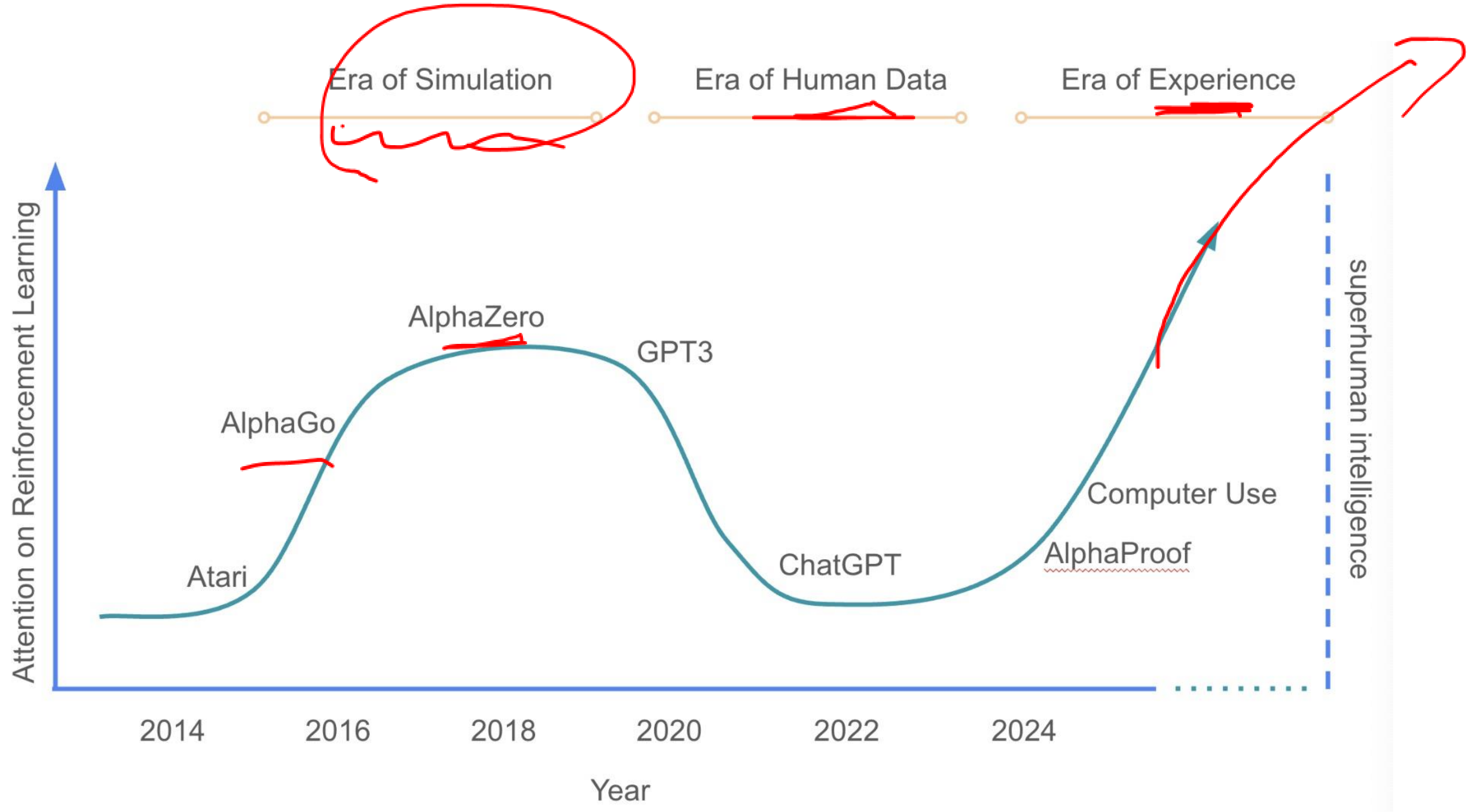
Simulating possible consequences of an action

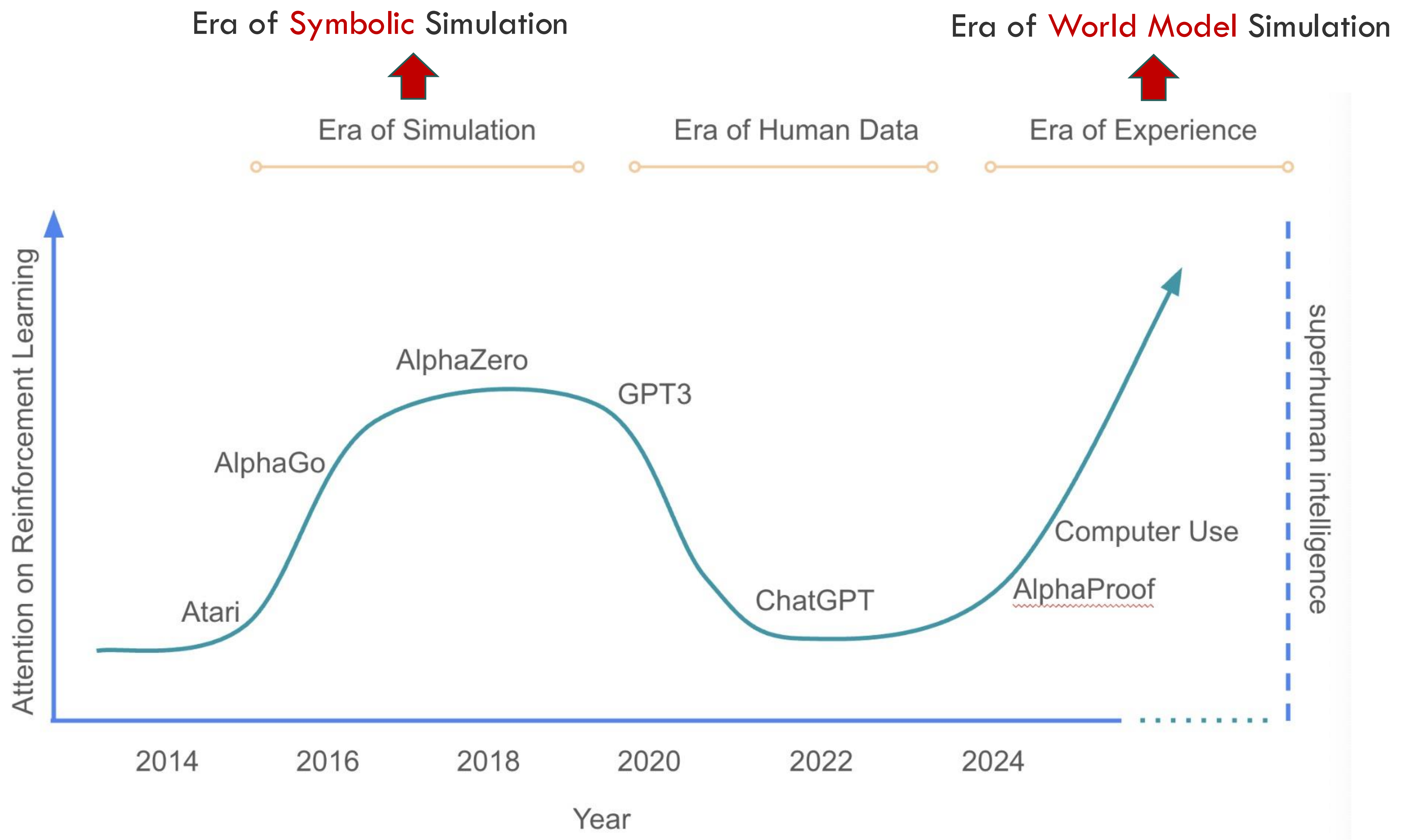
$$P(s'|s, a)$$

→ "World Model"

Welcome to the Era of Experience

David Silver, Richard S. Sutton*





Era of Symbolic Simulation
Era of World Model Simulation
Era of Simulation
Era of Human Data
Era of Experience
Attention on Reinforcement Learning
Atari
AlphaGo
AlphaZero
GPT3
ChatGPT
AlphaProof
Computer Use
superhuman intelligence
2014
2016
2018
2020
2022
2024
Year

Questions?