

# 1. What is the world?

## Sec II. A.

### Object

- Physical
- Agent
- Environmental
- Virtual, ...

### Attribute

- Shape
- Size
- Weight
- Material, ...

### Spatial relation

- Topological
- Directional
- Distance, ...

### Temporal relation

- Sequential
- Simultaneous
- Causal, ...

# 2. Why world modelling is important?

## Sec II. B.

Partial and noisy observation

World modelling

Environmental dynamics and common-sense regularities

# 3. How to model the world?

## Sec II. C. Competing Perspectives

- Video generation?
- Action-conditioned generation?
- Action-conditioned prediction?

## Sec II. D. Revisiting Modern Models

- LLMs, VLMs & VLA
- Video Generation Models

## Sec III. Overview of World Models

### Paradigms

- Implicit World Modeling
- Latent Dynamics Modeling
- Video Generation

### Task Scope

- Action Predictions
- Policy Learning
- Video Generation, ...

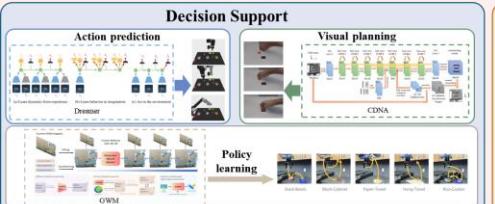
### Architectures

- Flat architecture
- Hierarchical architecture

### World Observation and Representation

- World Dimensionality
- Observation Viewpoint
- World Representation

## Sec IV. Functions of World Models



# 4. How far are we from fully realized world models?

## Sec V. Techniques & Challenges

### Data Limitation



### Long-horizon Reasoning



### Spatiotemporal Consistency



Instruction Understanding and Following



## Sec VI. Core Components and Capabilities

## Sec VIII. B. Future research directions

### Multimodal Perception



### Collaboration



### Imagination



### Physics Awareness



### Counterfactual Reasoning



### Security

