



Introduction

Interactivity

Intuitive Multi-step

Rich Dynamics Action Order

Planning Action Timing

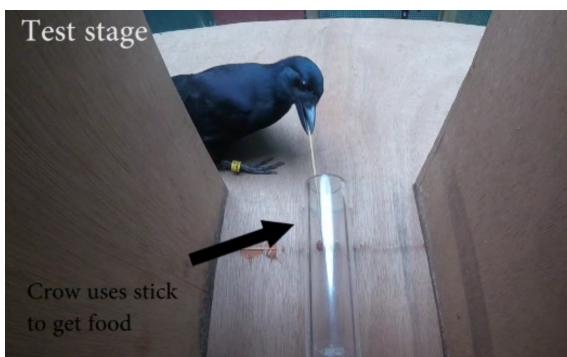
Infants of 4.5-month-old can understand basic physical concepts like support. (Cognition, 1993)



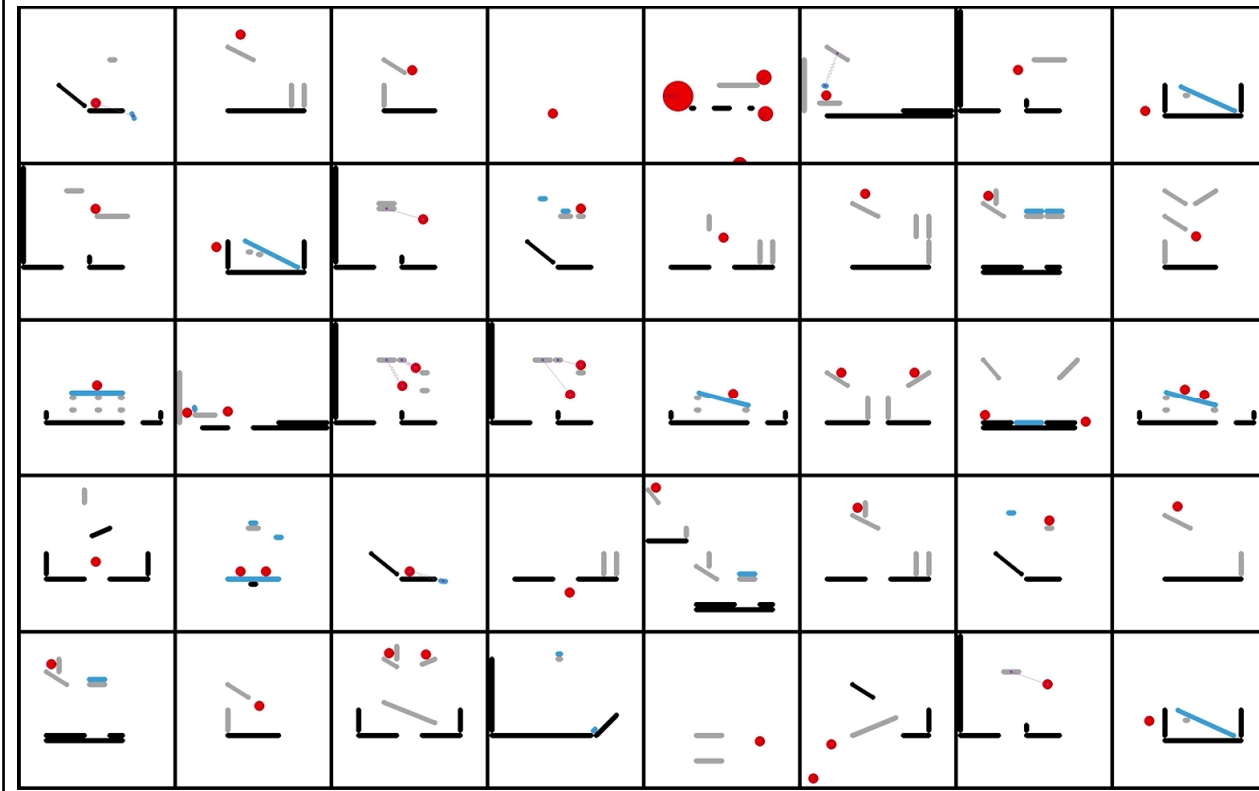
Professional players demonstrate mastery in playing 3D pinball, with records of continuous play 3 hours.



New Caledonian crows use mental representations to solve tool problems. (Current Biology, 2019)



Benchmark



- Intermediate State
- Block Elimination
- Success
- Failure
- Basic Game
- Compositional Game
- Noisy Game
- Multi-ball Game

Task: Drop the red balls into the abyss by eliminating gray blocks

Suboptimal solution: Use 3 steps

Perfect solution: Use 3 steps

Failure: Miss the perfect timing

Failure: Wrong order leads to missing elimination timing for two balls

Perfect solution: Use 2 steps

Suboptimal solution: Disturbed by the noisy block

In-situ adaptation

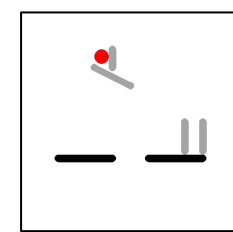
Optimal

Failure: Wrong order

Compositional solution: Long-sequence planning

Planning Strategies

A: Planning in advance



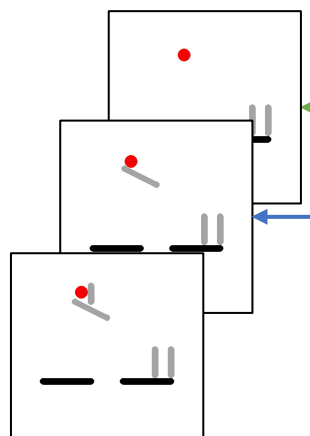
Model

a1	t1
a2	t2
a3	t3
a4	t4

$$A^* = \arg \max_A P(A|s_0)$$

$$= \arg \max_A f(A, s_0; \theta)$$

B: Planning on the fly



Simulator

Model

a1	a2
----	----

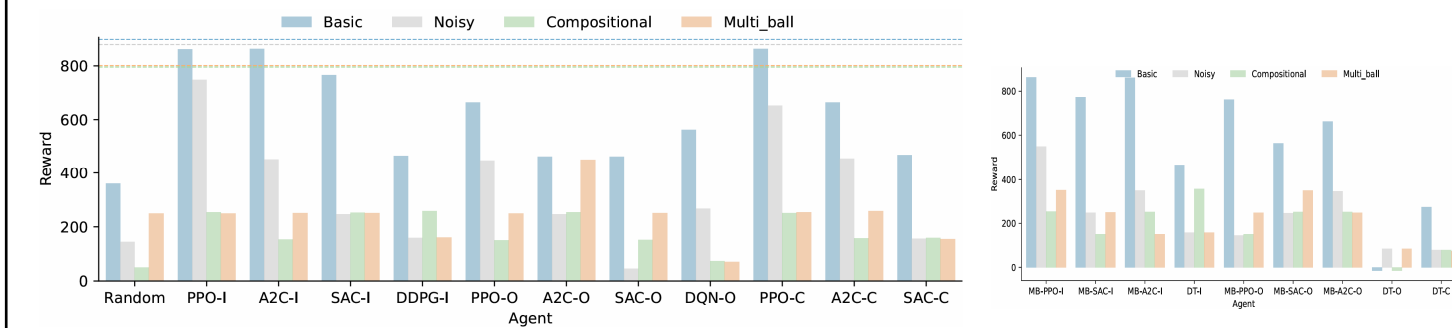
$$a_t^* = \arg \max_{a_t} P(a_t|a_{t-1}, s_{t-1})$$

$$= \arg \max_{a_t} g(a_t, s_t; \phi)$$

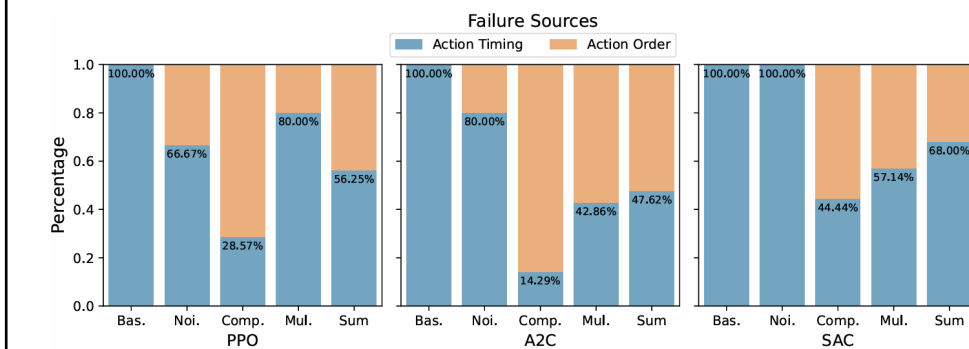
where $s_t = h(s_{t-1}, a_{t-1})$,

Experiments

A: RL agents fall back from human players in generalizing physical reasoning



B: The failure comes from wrong action order and timing



C: Classifiers and GPT4 fail to grasp action timing

Agent	Bas.	Noi.	Comp.	Mul.
Global Fusion	87.5	59.8	57.0	56.7
Object Fusion	71.4	60.1	60.2	54.1
Vision Fusion	86.4	57.5	55.8	59.5

Agent	Bas.	Noi.	Comp.	Mul.
GPT-4	75.17	64.17	-29.10	77.49