

3D Path Finder where gravity works

Decision Transformer



Introduction members



Idea from



Perseverance, observing Mars robot, by NASA

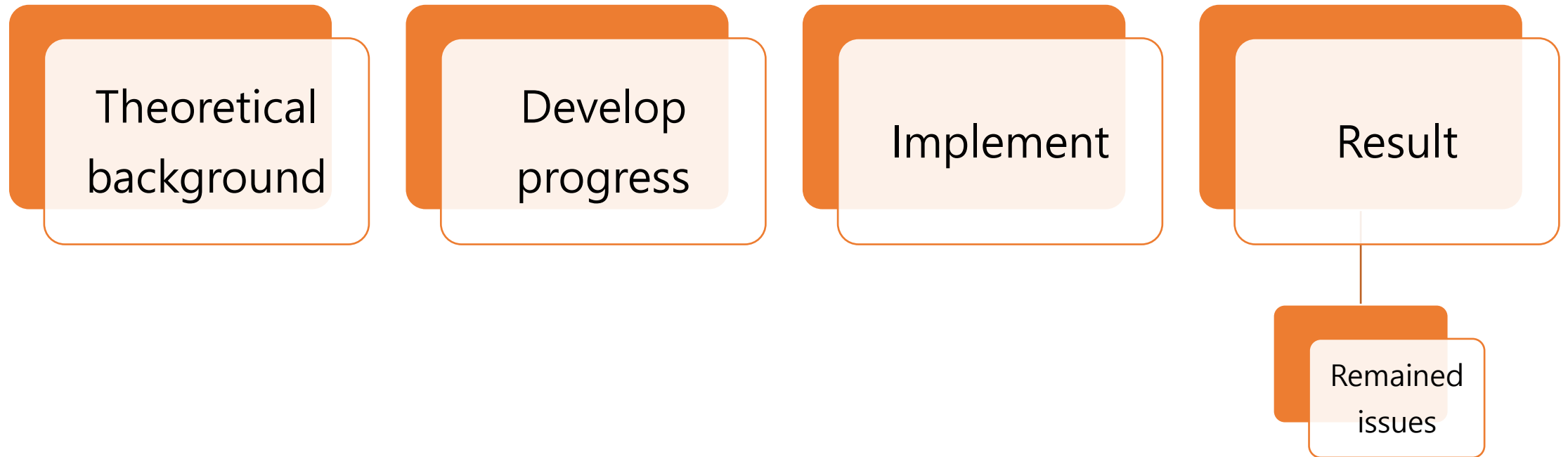


Object

- Making path finder on 3-Dimensional space
- Constraints:
 - **Gravity** works with $9.8m/s^2$
 - Agent has own **stamina** and **HP** so that something can happen if the attributes were starved.
 - Agent behaviors run, splint, jump, climbing, and gliding(use parachute). Agent can also just wait.
- Finding paths where the shortest time would be spent.



Contents

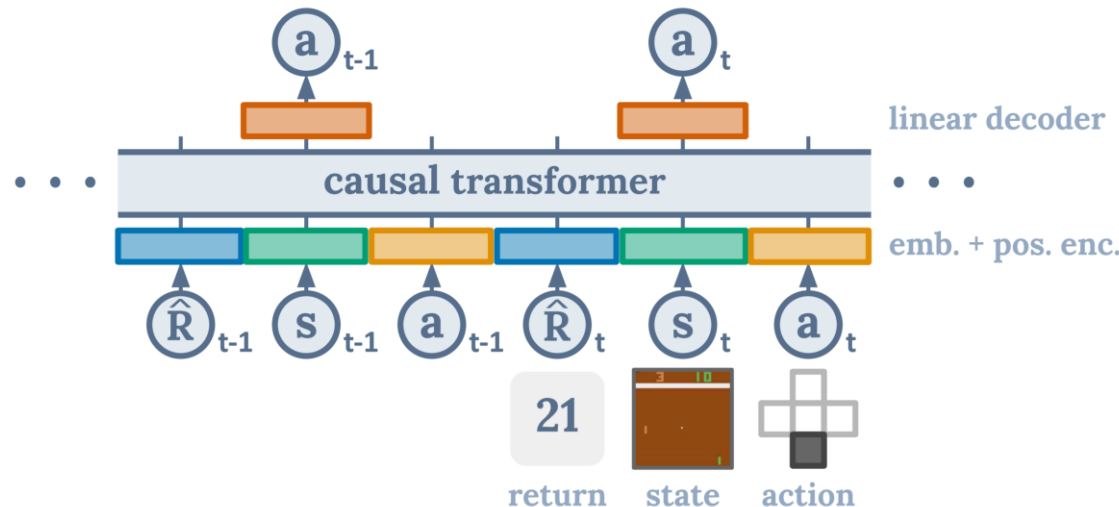


Theoretical background

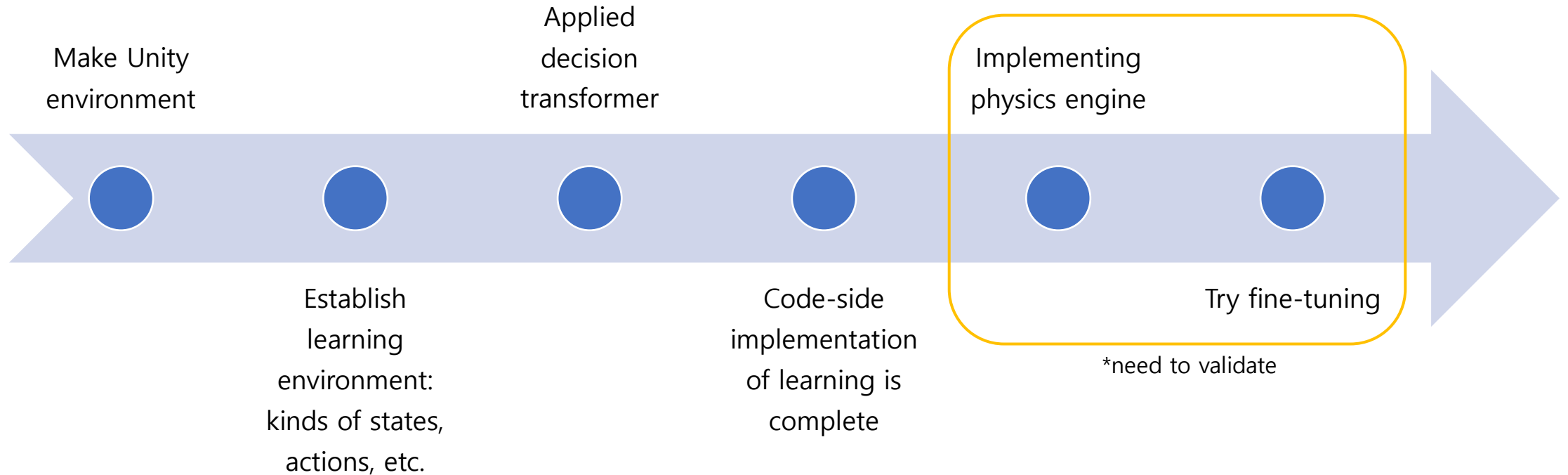
- ***Decision Transformer:** Reinforcement Learning via Sequence Modeling

- **Model-Free** : Just use record tuple sequence!
- Records for reinforcement learning are given as sequence of tuples, (r_t, s_t, a_t, t)
- 'return' is obtained by sum of expected results in the future from timestep t

$$\longrightarrow \widehat{R}_t = \sum_{k=t}^T r_k$$



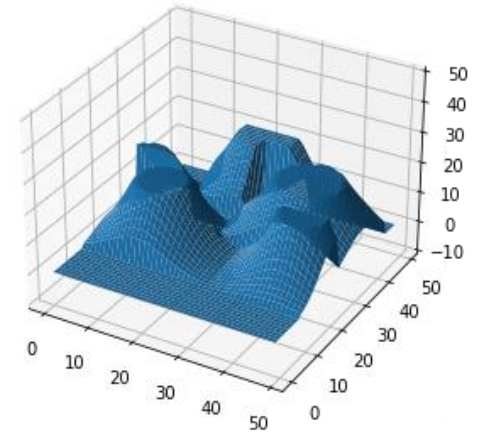
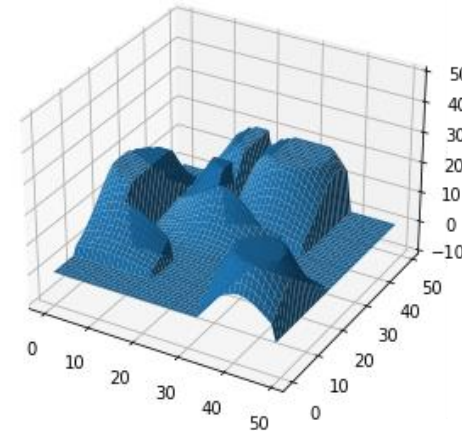
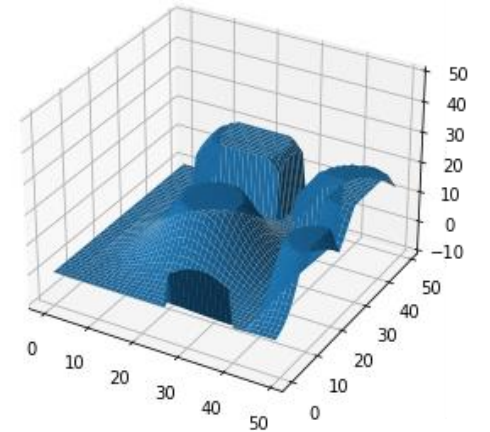
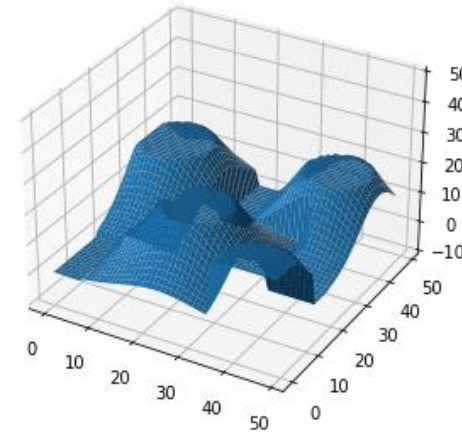
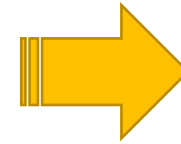
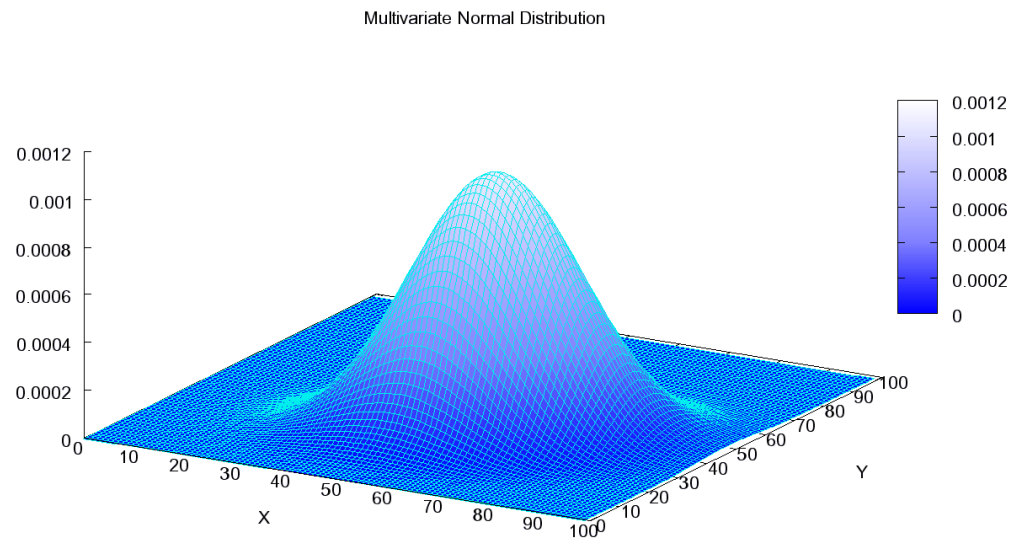
Develop progress



Implement



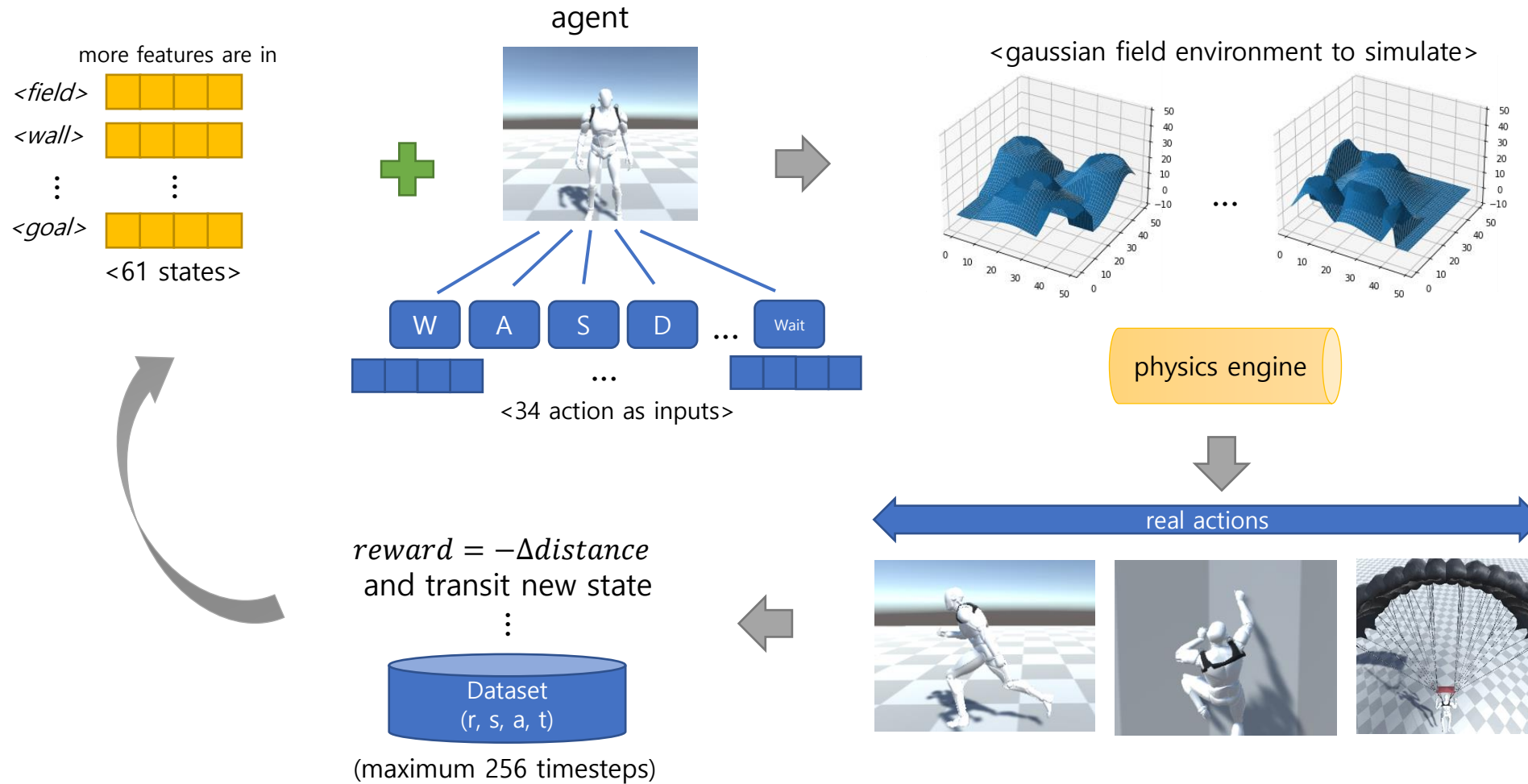
Map designer



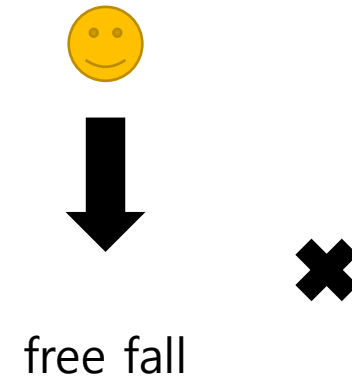
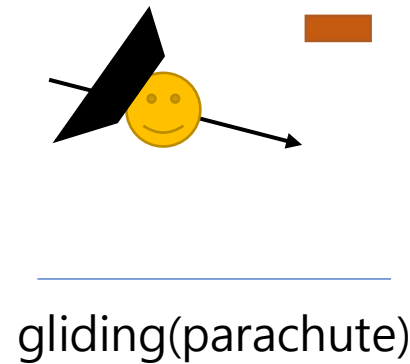
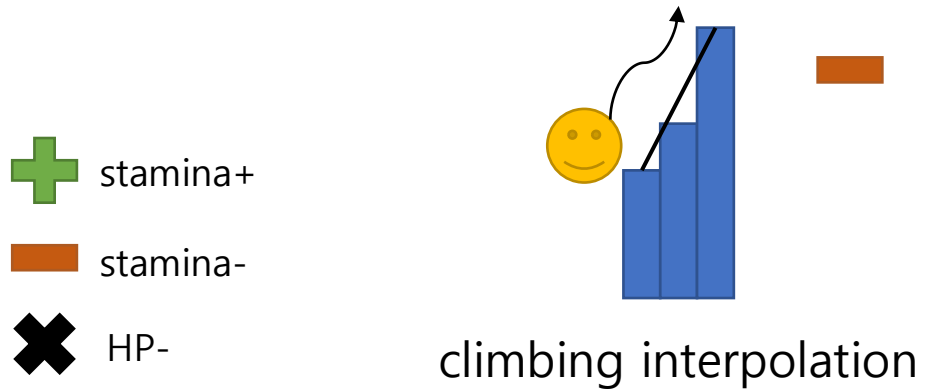
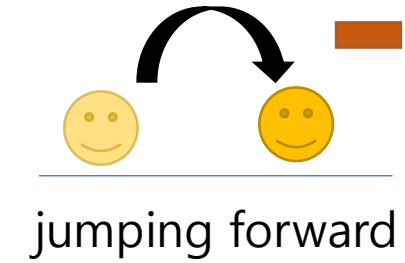
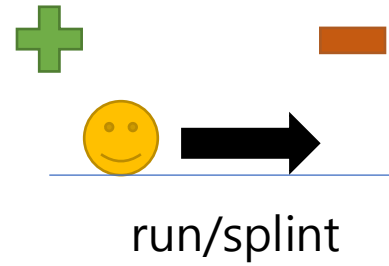
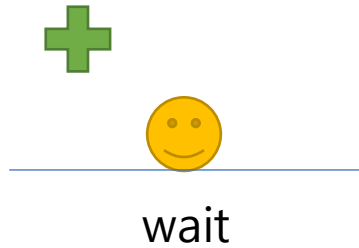
Width	Height	Depth	n	# of maps
50	50	20	5	12



Design of learning environment

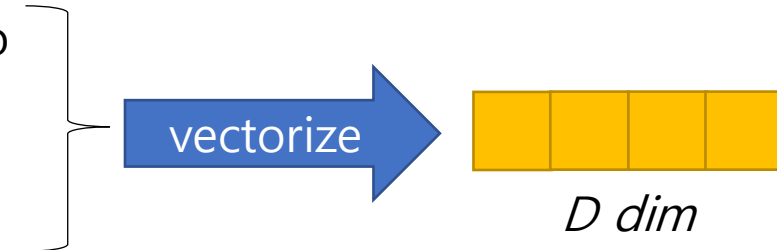


Physics engine



Others

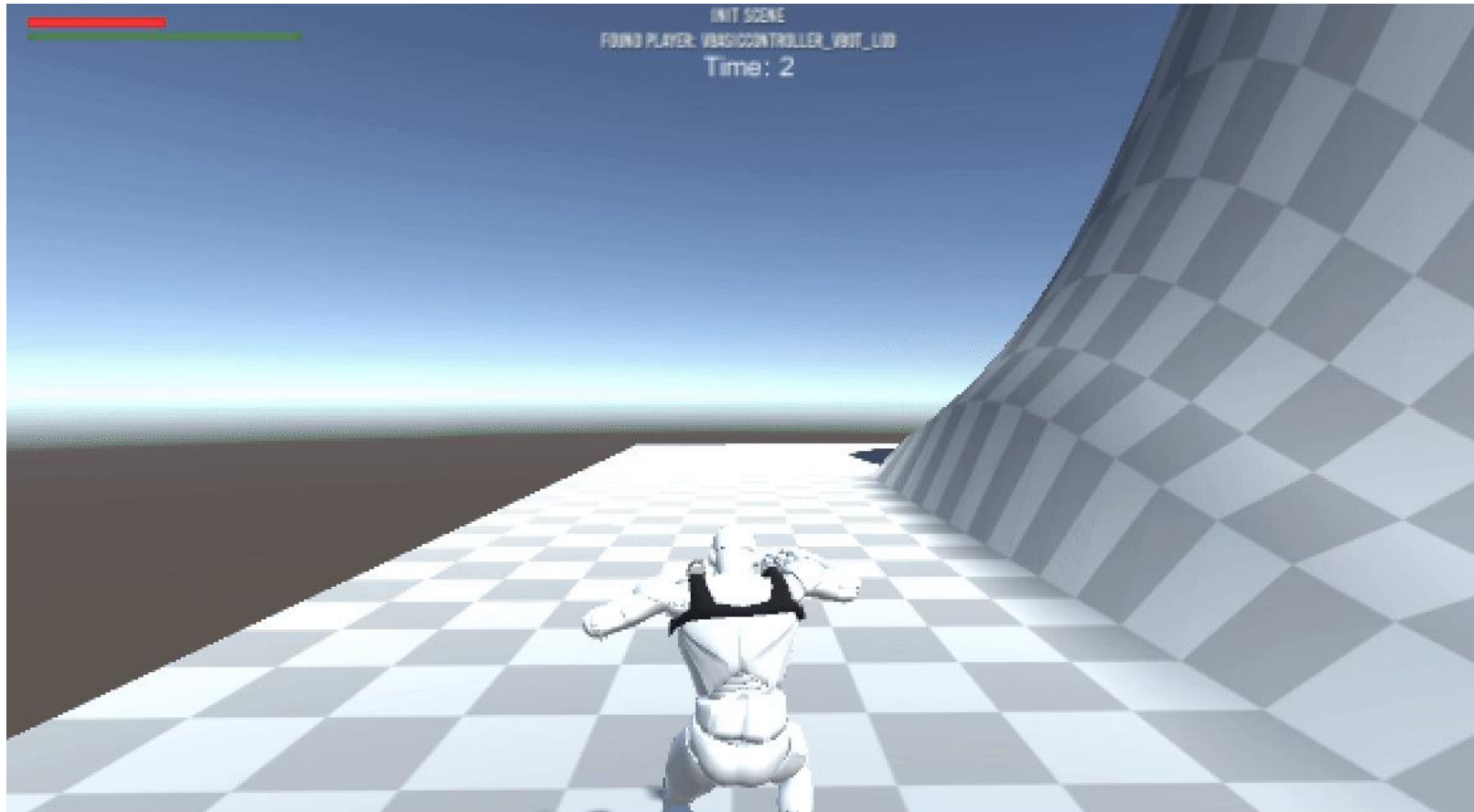
- agent.py – representing several attributes of agent.
 - HP, stamina, viewing direction vector, present position vector, doing action
- action.py – representing motion of each action.
 - id, input key, velocity vector, stamina consume info
- state.py – representing the state of the agent.
 - id, no, remained distance, spend time
- basic_math.py – representing vector operations
 - rotation, projection, size, normalize, distance



Result



Traversal example



Remained Issues

- Improve the strategy of reward
- Multiprocessing upgrade
 - each child process traverses each environment
 - the speed of generation will be faster if they do job in same environment
- Fine tuning portion of the learning
 - Hyper parameters we referred were not fit our model



Thank you

