XGEN R&D

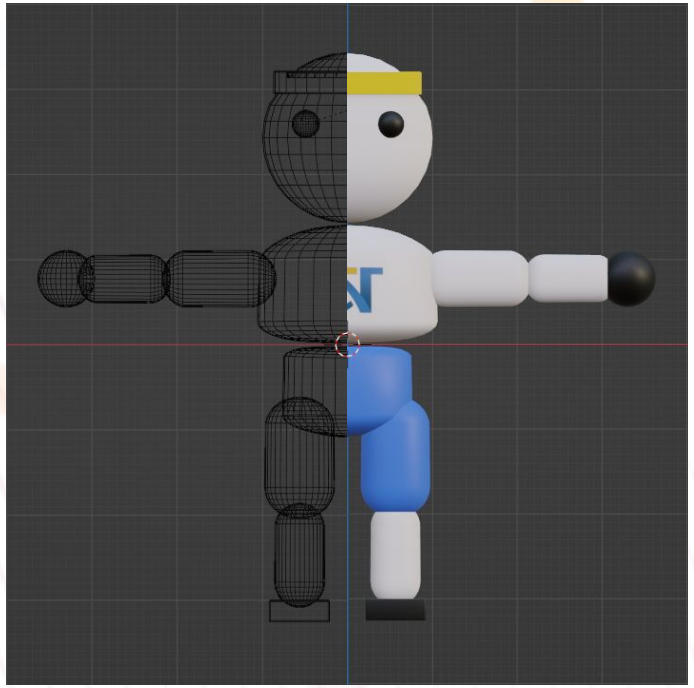
A Game Engine is all you need

Author: Radu Ciobanu
Supervisor: Prof. Dr. Daniela Zaharie
Affiliation:West University of Timisoara

XGEN R&D

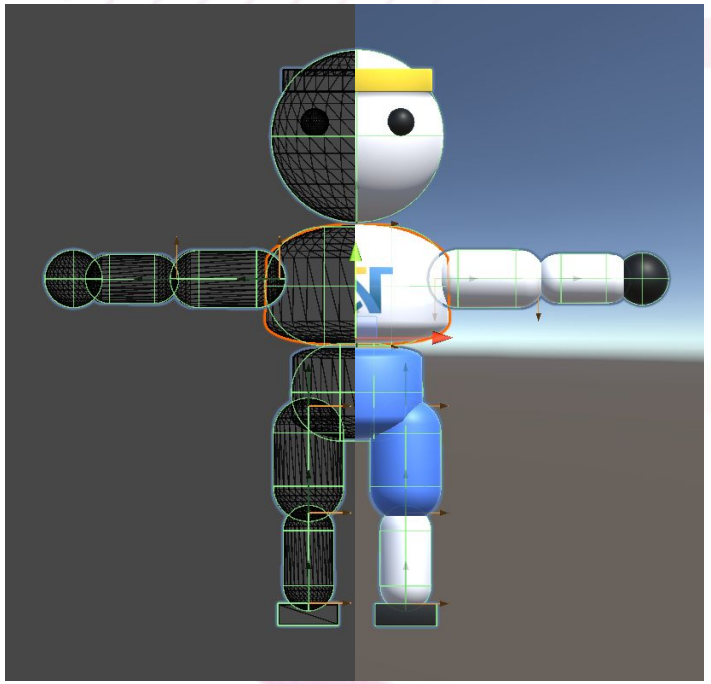
Creating the agent

Mesh, Materials, Textures

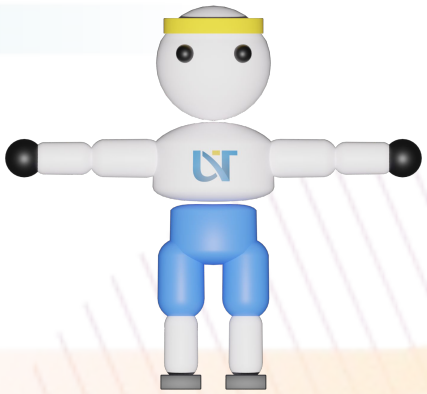


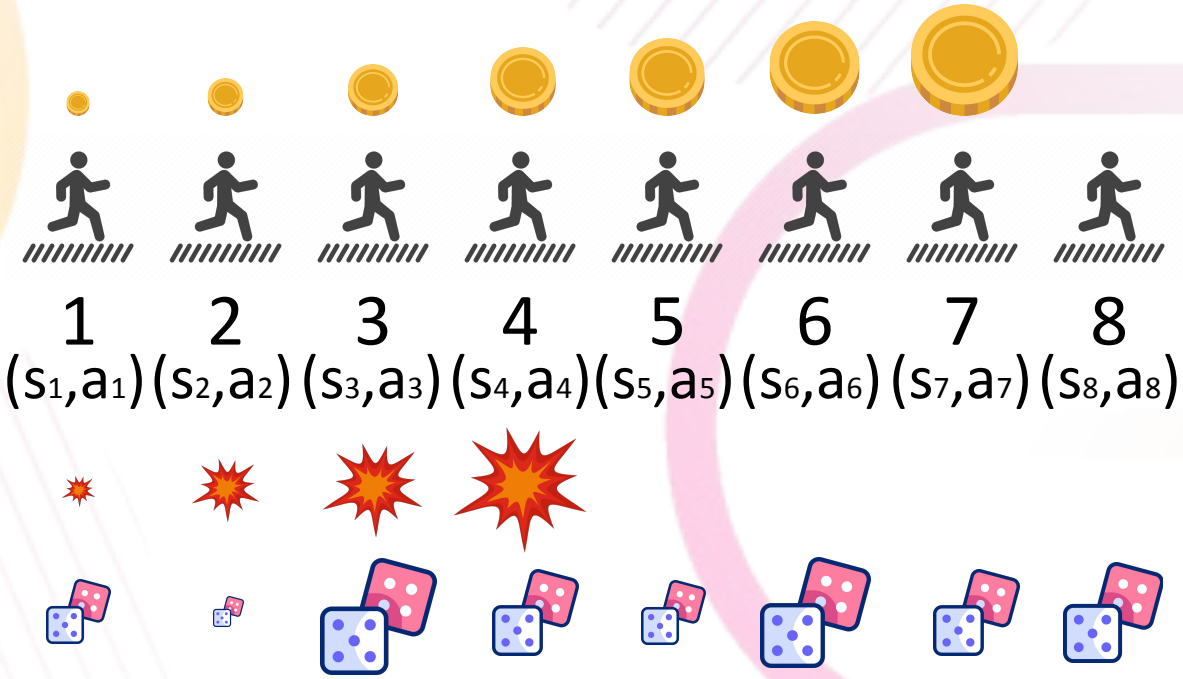
FBX/GLB

Colliders, Rigidbodies, Joints



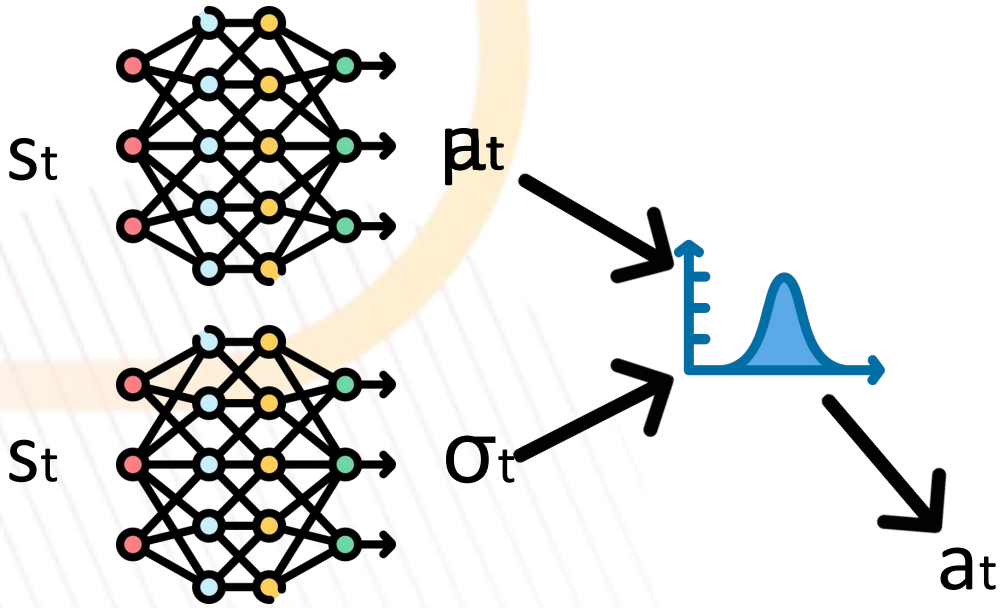
Collecting experience



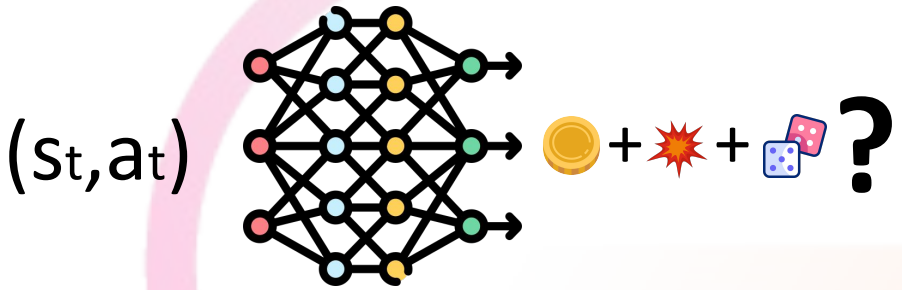


Agent’s behaviour

The Policy (π)



The Critic (Q)



The objective function

Vanilla Policy Gradient (VPG) maximizes

$$\hat{\mathbb{E}}_t\left[\log \pi_\theta(a_t \mid s_t)\hat{A}_t\right]$$

Proximal Policy Optimization (PPO) maximizes

$$\hat{\mathbb{E}}_t\left[\min(r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t)\right] \qquad r_t(\theta) = \frac{\pi_\theta(a_t|s_t)}{\pi_{\theta_{old}}(a_t|s_t)}$$

Soft Actor-Critic (SAC) maximizes

$$\mathbb{E}\left[\min_{i=1,2} Q_{\phi_i}(s, \tilde{a}_\theta(s)) - \beta \log \pi_\theta(\tilde{a}_\theta(s)|s)\right]$$

Line search
(like gradient ascent)

Trust region

XGEN
R&D

Results