Reinforcement Learning for Spacecraft in Stochastic Environments

Jordan Murphy
Computing/DSSI
Jason Bernstein
Michael Schneider

LLNL-PRES-812956
This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under contract DE-AC52-07NA27344. Lawrence Livermore National Security, LLC

Lawrence Livermore National Laboratory

Stochastic Environments

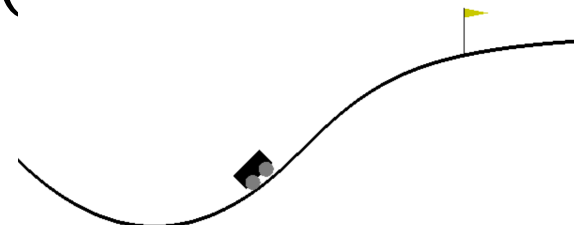
- Stochastic Differential Equations (SDE)

$$\boldsymbol{dX = f\big(X(t)\big)dt + g(}$$

- Mountain Car

$$\boldsymbol{d\vec{r} = \vec{v}}$$

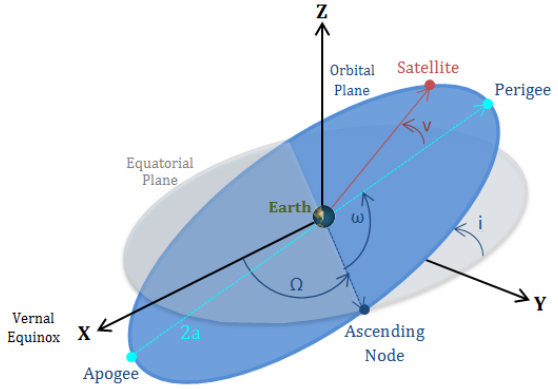
$$\boldsymbol{d\vec{v} = \overrightarrow{a_{grav}} + \frac{\vec{u}}{m} + \sigma dW(t)}$$



- Orbits

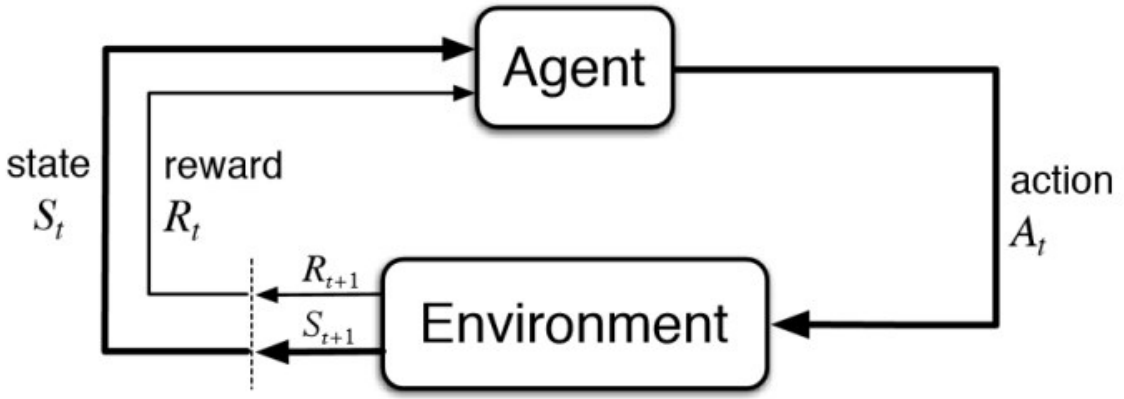
$$\boldsymbol{d\vec{r} = \vec{v}}$$

$$\boldsymbol{d\vec{v} = \frac{\mu}{r^3}\hat{r} + \overrightarrow{a_{pert}} + \frac{\vec{u}}{m} + \sigma dW(t)}$$



Reinforcement Learning

- Agent Seeks to Accumulate Reward



- Bellman Function

$$V^{\pi} = R\big(s, \pi(s)\big) + \gamma \sum_{s'} P(s'|s,a) V^{\pi *}(s')$$

Conclusions & Future Work

- Noisy Environments
  - Stochastic Value Gradient (Heess et al., 2015)
- Sparse Rewards
  - Potential-Based Reward Shaping
  - $F(s, a, s') = \gamma\phi(s') - \phi(s)$
- Reward Design in Orbital Environment
  - Keplerian vs. Cartesian
- Future Work
  - More Realistic & Complex Orbit Environment
  - Augmented Experience Replay
  - Incorporating Approximations of Fokker-Planck

LLNL-PRES-812956