Lunar Landing v2

Environment:

Rocket trajectory optimization is a classic topic in Optimal Control. According to Pontryagin's maximum principle it's optimal to fire engine full throttle or turn it off. That's the reason this environment is OK to have discreet actions (engine on or off).

The landing pad is always at coordinates (0,0). The coordinates are the first two numbers in the state vector. Reward for moving from the top of the screen to the landing pad and zero speed is about 100 to 140 points.

If the lander moves away from the landing pad it loses reward. The episode finishes if the lander crashes or comes to rest, receiving an additional -100 or +100 points. Each leg with ground contact is +10 points. Firing the main engine is -0.3 points each frame. Firing the side engine is -0.03 points each frame.

Solved is 200 points.

Landing outside the landing pad is possible. Fuel is infinite, so an agent can learn to fly and then land on its first attempt.

Source code for environment:

gym/lunar_lander.py at master · openai/gym (github.com)

States: