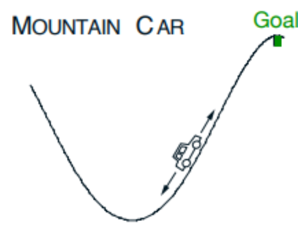
Mountain Car Task
Taken from the book: Reinforcement Learning: An Introduction (second edition)
By: Richard S. Sutton and Andrew G. Barto
Page: 244-245

Consider the task of driving an underpowered car up a steep mountain road. The diculty is that gravity is stronger than the car's engine, and even at full throttle the car cannot accelerate up the steep slope. The only solution is to first move away from the goal and up the opposite slope on the left. Then, by applying full throttle the car can build up enough inertia to carry it up the steep slope even though it is slowing down the whole way. This is a simple example of a continuous control task where things have to get worse in a sense (farther from the goal) before they can get better. Many control methodologies have great diculties with tasks of this kind unless explicitly aided by a human designer.



The reward in this problem is 1 on all time steps until the car moves past its goal position at the top of the mountain, which ends the episode. There are three possible actions: full throttle forward (+1), full throttle reverse (1), and zero throttle (0). The car moves according to a simplified physics. Its position, x_t, and velocity, $\dot{x}_t$, are updated by:

$$
\begin{aligned}
x_{t+1} &\doteq \text{bound}\,[x_t + \dot{x}_{t+1}] \\
\dot{x}_{t+1} &\doteq \text{bound}\,[\dot{x}_t + 0.001 A_t - 0.0025\cos(3x_t)]
\end{aligned}
$$

where the bound operation enforces $1.2 \leqslant x_{t+1} \leqslant 0.5$ and $-0.07 \leqslant \dot{x}_{t+1} \leqslant 0.07$. In addition, when x_{t+1} reached the left bound, $\dot{x}_{t+1}$ was reset to zero. When it reached the right bound, the goal was reached and the episode was terminated. Each episode started from a random position $x_t \in [-0.6,\ -0.4]$ and zero velocity.