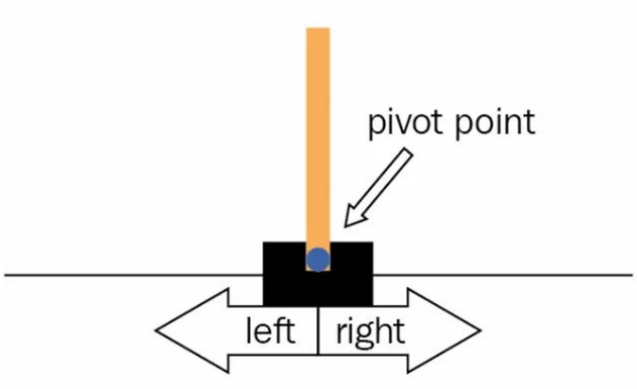
The **Cart-Pole** problem:

- A pole is attached to a cart that moves along a frictionless track.
- The objective is to balance the pole upright by applying forces to the cart, either left or right.



- **State Space**: the cart's position, velocity, the angle of the pole, and the angular velocity.
- **Actions**: push the cart left or push it right.
- **Rewards**: The agent receives a reward of +1 for every time step the pole remains upright. The episode ends when the pole falls beyond a certain angle or the cart moves too far from the center.

$$\text{SmoothL1Loss}(x, y) = \begin{cases} 0.5 \cdot \frac{(x-y)^2}{\beta} & \text{if } |x-y| < \beta \\ |x-y| - 0.5 \cdot \beta & \text{otherwise} \end{cases}$$

Why Use SmoothL1Loss in DQN?

1. **Stability**: It helps stabilize training by reducing the sensitivity to outliers compared to Mean Squared Error (MSE) loss. This is particularly useful in reinforcement learning, where the temporal difference errors can vary significantly.
2. **Gradient Behavior**: The smooth transition between L1 and L2 loss allows for more stable gradients, which can prevent issues like exploding gradients during training.
3. **Performance**: Empirical results have shown that using SmoothL1Loss can lead to better convergence properties in DQN, improving the overall performance of the agent in learning optimal policies.