

### 3 Finite Markov Decision Processes

- 3.1 Devise three examples of your own that fit the MPD framework, identifying for each its states, actions and rewards. Make the three examples as *different* from each other as possible. The framework is abstract and flexible and can be applied in many different ways. Stretch its limits in some way in at least one of your examples.

*text*

- A train driving down a train track. This would obviously have to be tested *very extensively* before releasing it into the wild.
  - Environment: The current speed of the train, and a map of the upcoming track (including at least upcoming curves and train stations.
  - Action: gas or break.
  - Reward: 1 if the train is still running, 0 if not. Can be multiplied by the current speed to encourage speedy driving.
- A robot shooting a bow and arrow at a board.
  - Environment: Wind speed and direction. Can be extended with the position if the robot is not in a fixed place.
  - Action: How far to pull back the arrow and where to aim it.
  - Reward: The number of points on the board.
- A machine delivering advertisement content to a website visitor.
  - Environment: The current web page and everything that is known about the customer through e.g. customer data or cookies. Could include age, location, previous website visits, etc.
  - Action: The different pieces of content available to the machine.
  - Reward: 1 if the visitor clicks, 0 if not.