- **Continuing tasks** are tasks that continue forever, without end.
- **Episodic tasks** are tasks with a well-defined starting and ending point.
    - In this case, we refer to a complete sequence of interaction, from start to finish, as an **episode**.
    - Episodic tasks come to an end whenever the agent reaches a **terminal state**.

The Reward Hypothesis

- **Reward Hypothesis**: All goals can be framed as the maximization of (expected) cumulative reward.

Goals and Rewards

- (Please see **Part 1** and **Part 2** to review an example of how to specify the reward signal in a real-world problem.)

Cumulative Reward

- The **return at time step** t is $G_t := R_{t+1} + R_{t+2} + R_{t+3} + \dots$
- The agent selects actions with the goal of maximizing expected (discounted) return. (*Note: discounting is covered in the next concept.*)

Discounted Return

- The **discounted return at time step** t is $G_t := R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \dots$.
- The **discount rate** γ is something that you set, to refine the goal that you have the agent.
    - It must satisfy $0 \leq \gamma \leq 1$.
    - If $\gamma = 0$, the agent only cares about the most immediate reward.
    - If $\gamma = 1$, the return is not discounted.
    - For larger values of γ, the agent cares more about the distant future. Smaller values of γ result in more extreme discounting, where - in the most extreme case - agent only cares about the most immediate reward.

MDPs and One-Step Dynamics

- The **state space** $\mathcal{S}$ is the set of all (*nonterminal*) states.