Reinforcement Learning

Jack Lanchantin

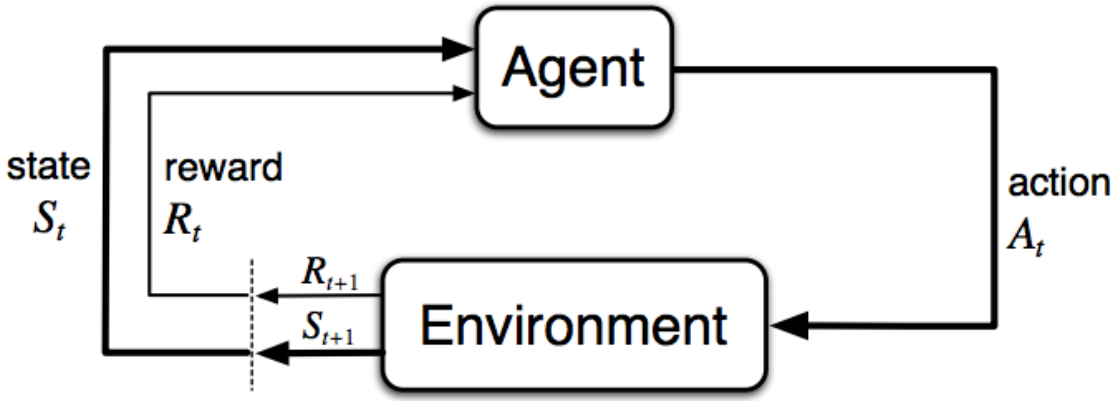
November 30, 2015

Overview

Reinforcement Learning

- **Framing of the problem of learning from interaction to achieve a goal.**
- **Agent**: learner and decision maker
- **Environment**: what the learner interacts with (everything outside the agent)
- Agent selects actions and the environment responds to those actions and presents new situations

Reinforcement Learning



- At each time step t, the agent receives the environment **state** $S_t \in \mathcal{S}$, and the agent then selects an **action** $A_t \in \mathcal{A}(S_t)$
  - $\mathcal{S}$ is the set of possible states
  - $\mathcal{A}(S_t)$ is set of actions available in state S_t
- One time step later, the agent receives a **reward**, $R_{t+1} \in \mathcal{R} \subset \mathbb{R}$, and ends up in a new state S_{t+1}

Reinforcement Learning

- At each time step, the agent implements a mapping π_t from states to probabilities of selecting each possible action, where π_t is called a **policy**
  - $\pi_t(a|s)$ = probability that $A_t = a$ if $S_t = s$

Reinforcement Learning Objective

The agent's goal is to maximize the total amount of reward it receives over the long run by changing its policy as a result of its experience

Reinforcement Learning

- Let the sequence of rewards after time step t is $R_{t+1}, R_{t+2}, R_{t+3}, \ldots$, then we want to maximize the return G_t
- The agent chooses A_t to maximize the discounted return:

$$G_t = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1} \tag{1}$$

  where γ is the discount rate and $0 \geq \gamma \leq 1$
- The closer γ is to 1, the more the agent accounts for future rewards

Markov Decision Processes

Multiple Columns

Heading

1. Statement
2. Explanation
3. Example

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Integer lectus nisl, ultricies in feugiat rutrum, porttitor sit amet augue. Aliquam ut tortor mauris. Sed volutpat ante purus, quis accumsan dolor.

Table

Treatments	Response 1	Response 2
Treatment 1	0.0003262	0.562
Treatment 2	0.0015681	0.910
Treatment 3	0.0009271	0.296

Table: Table caption

Theorem

Theorem (Mass–energy equivalence)

$E = mc^2$

Verbatim

Example (Theorem Slide Code)

```
\begin{frame}
\frametitle{Theorem}
\begin{theorem}[Mass--energy equivalence]
$E = mc^2$
\end{theorem}
\end{frame}
```

Figure

Uncomment the code on this slide to include your own image from the same directory as the template .TeX file.

Citation

An example of the `\cite` command to cite within the presentation:

This statement requires citation [Smith, 2012].

References

John Smith (2012)

Title of the publication

Journal Name 12(3), 45 – 678.