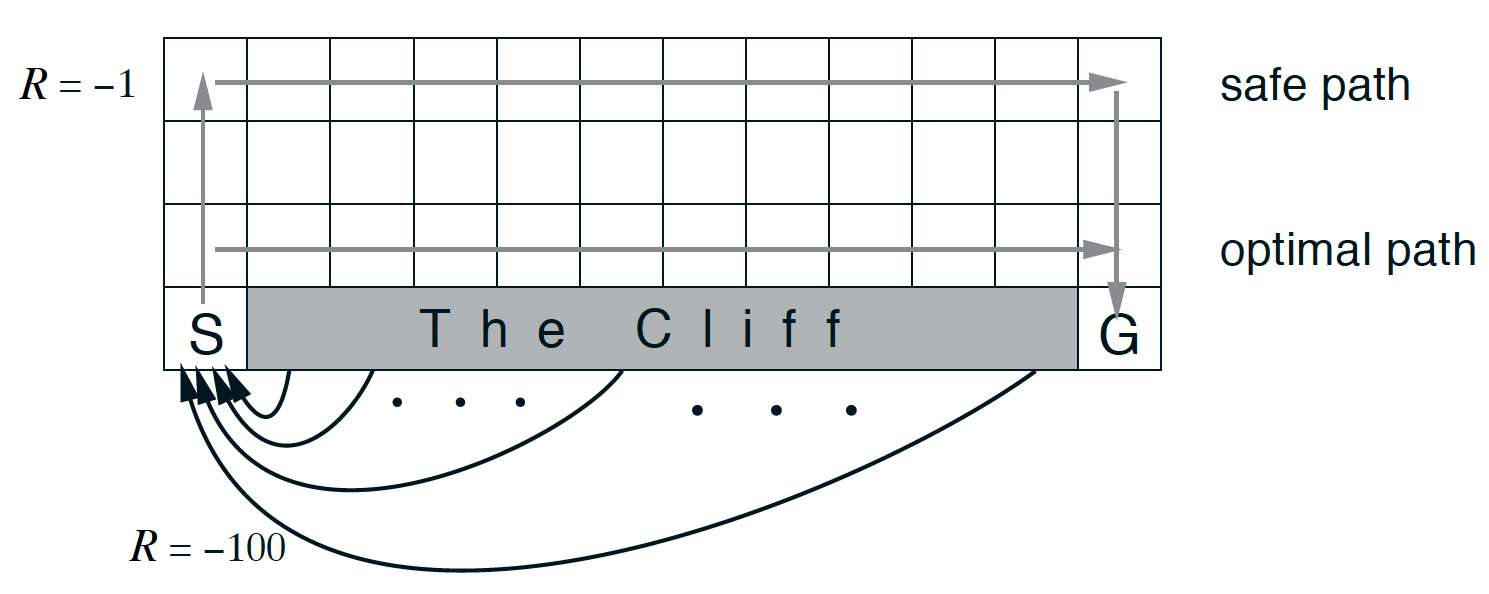
# 1.9 Temporal Difference Methods Summary

**[[](https://classroom.udacity.com/nanodegrees/nd893/parts/8f607726-757e-4ef5-8b64-f2368755b89a/modules/a85374fa-6a60-425b-a480-85b211c5bd5d/lessons/78acdf66-f45c-418b-bbe1-30fda9e42c8c/concepts/b5f852a2-0378-4dbb-afc0-c16da612d66d)](https://classroom.udacity.com/nanodegrees/nd893/parts/8f607726-757e-4ef5-8b64-f2368755b89a/modules/a85374fa-6a60-425b-a480-85b211c5bd5d/lessons/78acdf66-f45c-418b-bbe1-30fda9e42c8c/concepts/b5f852a2-0378-4dbb-afc0-c16da612d66d)**

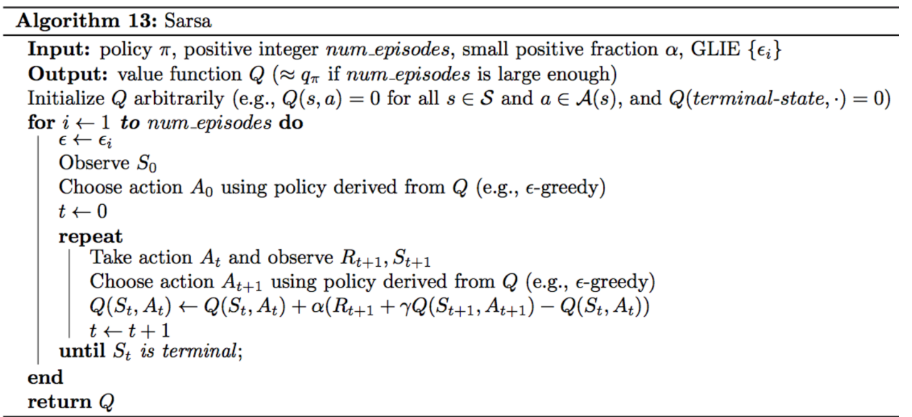
**[The cliff-walking task (Sutton and Barto, 2017)](https://classroom.udacity.com/nanodegrees/nd893/parts/8f607726-757e-4ef5-8b64-f2368755b89a/modules/a85374fa-6a60-425b-a480-85b211c5bd5d/lessons/78acdf66-f45c-418b-bbe1-30fda9e42c8c/concepts/b5f852a2-0378-4dbb-afc0-c16da612d66d)**

### Temporal-Difference Methods

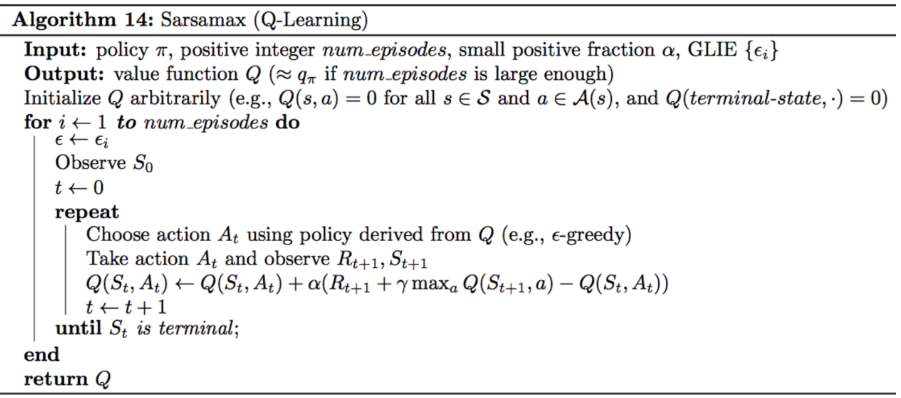
* Whereas Monte Carlo (MC) prediction methods must wait until the end of an episode to update the value function estimate, temporal-difference (TD) methods update the value function after every time step.

### TD Control

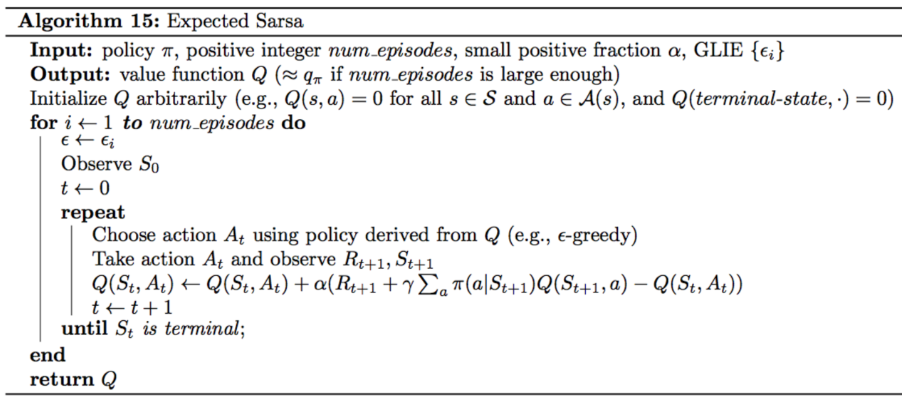
* **Sarsa(0)** (or **Sarsa**) is an on-policy TD control method. It is guaranteed to converge to the optimal action-value function *q*∗​, as long as the step-size parameter *α* is sufficiently small and *ϵ* is chosen to satisfy the **Greedy in the Limit with Infinite Exploration (GLIE)** conditions.



* **Sarsamax** (or **Q-Learning**) is an off-policy TD control method. It is guaranteed to converge to the optimal action value function *q*∗​, under the same conditions that guarantee convergence of the Sarsa control algorithm.



* **Expected Sarsa** is an on-policy TD control method. It is guaranteed to converge to the optimal action value function *q*∗​, under the same conditions that guarantee convergence of Sarsa and Sarsamax.



### Analyzing Performance

* On-policy TD control methods (like Expected Sarsa and Sarsa) have better online performance than off-policy TD control methods (like Q-learning).
* Expected Sarsa generally achieves better performance than Sarsa.