**Vivekanand Education Society’s Institute of Technology**

**Department of AIDS Engineering**



**Subject: Reinforcement Learning**

**Class: D16AD**

| **ROLL NO: 30** | **NAME: SUHANEE KANDALKAR** | | |
| --- | --- | --- | --- |
| **EXP NO: 09** | T**ITLE:**   **Temporal Difference Learning** | | |
| **DOP:** |  | **DOS:** |  |
| **GRADES:** | **LOs MAPPED:** | | **SIGNATURE:** |

## 

## Aim**:**

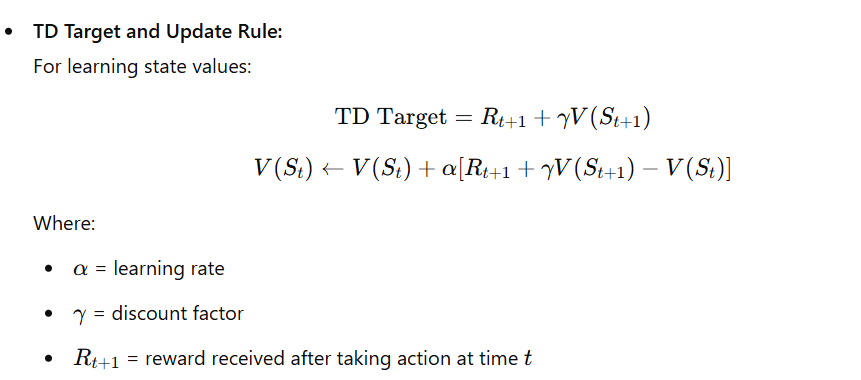
Temporal Difference Learning

## Theory**:**

Temporal Difference (TD) Learning is a fundamental approach in Reinforcement Learning that combines elements of **Monte Carlo methods** and **Dynamic Programming**. It enables agents to learn value functions incrementally from raw experience, even before episodes are completed.

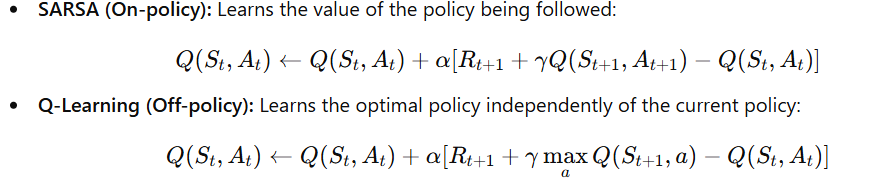
#### Key Concepts:

* **Bootstrapping:** TD methods update current value estimates using other estimates (not final outcomes), allowing faster learning.



#### TD Variants:

* **TD(0):** Updates based on the next state’s value only.

****

* **TD(λ):** Uses **eligibility traces** to balance between TD(0) and Monte Carlo methods.

#### ✅ Advantages:

* Learns from **incomplete episodes**
* More **data-efficient** than Monte Carlo
* **Online learning**: suitable for dynamic or streaming environments

[RL\_EXP9\_30.ipynb](https://colab.research.google.com/drive/1LHlKR-pOuDI7urmlKZNAZ91DL_wVbRik?usp=sharing#scrollTo=n1We-0Jk6zcx)

# Conclusion :

Temporal Difference (TD) Learning blends the advantages of Monte Carlo methods and Dynamic Programming. It updates value estimates based on partially observed returns using bootstrapping, allowing learning to happen at every time step without waiting for the end of an episode. This makes TD methods like TD(0), SARSA, and Q-learning highly effective for real-time and continuous tasks, enabling faster and more efficient policy evaluation and improvement in Reinforcement Learning environments.