Temporal Difference (TD) learning:

* Pros: Can learn online, does not require the completion of an episode, can learn from incomplete sequences, combines advantages of MC and DP methods.
* Cons: Can be sensitive to initial conditions, may converge to suboptimal solutions, can be unstable, requires tuning of the learning rate and exploration rate.
* Uses: Used in a variety of problems including control and prediction, continuous control problems, and large-scale problems.
* Strengths: Can learn from incomplete sequences, can learn online, and combines advantages of MC and DP methods.
* Weaknesses: Can be unstable and sensitive to initial conditions, may converge to suboptimal solutions.

Monte-Carlo (MC) methods:

* Pros: Can learn from complete episodes, unbiased estimate, easy to understand and implement, can learn offline.
* Cons: Can be computationally expensive, cannot learn online, requires the completion of an episode.
* Uses: Used in a variety of problems including control and prediction, and episodic problems.
* Strengths: Can learn from complete episodes and provides an unbiased estimate.
* Weaknesses: Cannot learn online and can be computationally expensive.

SARSA:

* Pros: Can learn online, can handle stochastic policies, can converge to optimal policy.
* Cons: Slower convergence than Q-learning, may converge to suboptimal solutions, requires the completion of an episode.
* Uses: Used in problems where the policy is stochastic, and control problems.
* Strengths: Can learn online and can handle stochastic policies.
* Weaknesses: Slower convergence than Q-learning and may converge to suboptimal solutions, requires the completion of an episode.

Q-learning:

* Pros: Can learn online, can handle stochastic policies, can converge to optimal policy, has strong convergence properties.
* Cons: Can overestimate the value of actions, can be slow to converge, cannot learn from incomplete sequences.
* Uses: Used in control problems, where the policy is deterministic, and in large-scale problems.
* Strengths: Can learn online, can handle stochastic policies, and has strong convergence properties.
* Weaknesses: Can overestimate the value of actions and can be slow to converge, cannot learn from incomplete sequences.

Overall, each algorithm has its own strengths and weaknesses, and the choice of algorithm depends on the specific problem and its requirements. For example, if the policy is stochastic, SARSA may be a better choice than Q-learning, while if online learning is required, TD learning may be preferred over MC methods.