**POLICY GRADIENT ALGORITHMS**

Preparing a parsing of reinforcement learning methods of the Policy Gradient family, such as vanilla PG, PPO, TRPO and others (the number of algorithms is chosen by the speaker, but the more extensive the study, the better) and putting them into practice and then comparing their effectiveness (Atari games or OpenAI engine from Gym are suggested to be used as testing ground)

Reinforcement Learning is a type of machine learning algorithm that allows an agent to learn from its environment by trial and error and use feedback from its actions to learn the best strategy for achieving its goals. It works by assigning rewards for successful actions and punishments for unsuccessful ones, allowing the agent to adjust its strategy accordingly.

PG Policy Gradient algorithm   
and its underlying principles:

The Policy Gradient (PG) algorithm is a type of reinforcement learning algorithm used to improve the performance of an agent in a particular task. It works by using a policy to decide how the agent should act in a given situation. The agent then receives feedback on its actions and uses the feedback to adjust its policy so that it can perform better in future situations.

At the core of the PG algorithm is the idea of the policy gradient, which is the rate at which a policy is adjusted in response to an observed reward. The policy gradient is calculated by taking the derivative of the expected return with respect to the policy parameters. This gradient is then used to update the policy parameters so that the agent can maximize its expected return.

The agent then uses this updated policy to take actions and observe rewards, and the process repeats until the agent has learned an optimal policy. This process of learning is called policy iteration. The key idea behind PG is that the agent is encouraged to explore more and take risks in order to maximize its expected return, since it can learn from the rewards it receives.

PG algorithms are widely used in robotics, video games, and other AI applications. They are also used in research fields such as natural language processing and reinforcement learning.