RL has demonstrated its potential in solving goal-oriented sequential tasks. However, with the increasing capabilities of RL agents, ensuring morally responsible agent behavior is becoming a pressing concern. Previous approaches have included moral considerations by statically assigning a moral score to each action at runtime. However, these methods do not account for the potential moral value of future states when evaluating immoral actions. This limits the ability to find trade-offs between different aspects of moral behavior and the utility of the action. In this paper, we aim to factor in moral scores by adding a constraint to the RL objective that is incorporated during training, thereby dynamically adapting the policy function. By combining Lagrangian optimization and meta-gradient learning, we develop an RL method that is able to find a trade-off between immoral behavior and performance in the decision-making process.