exploit your knowledge, and choose the arm with the highest value of $\hat{R}(a_j, r)$ on all future trials; or

explore further, by trying some or all actions more times, hoping to get better estimates of the $R(a_j, r)$ values.