Learning Based on Individual Intermediate Reward