

Abstract

The present application provides a dynamical symmetry modeling based offline reinforcement learning method and apparatus, where the method includes: obtaining an offline data set; processing an original state-action pair in the offline data set by an encoder and a bidirectional dynamical model to obtain a low-dimensional state-action pair and a current time inversion state variation; and inputting the low-dimensional state-action pair to a value network, and obtaining a value of the state-action value function corresponding to the original state-action pair.

FIG. 1