A ﬁnite-horizon Markov Decision Process (MDP) can be speciﬁed as M=<S,A,P(·,·),R(·,·),T>, where S is the state space, A Is the action space, P:S×A7→Pr(S) is the transition function that maps states and actions to a probability density over subsequent states, R : S×A7→R is a reward function over the state and action space, and T is the time-horizon.

HRL considers a process, possibly stochastic, which transitions between a set of MDPs {M1,...,Mk}.