

*Policy* –  $a_t \sim \pi(x_t)$

