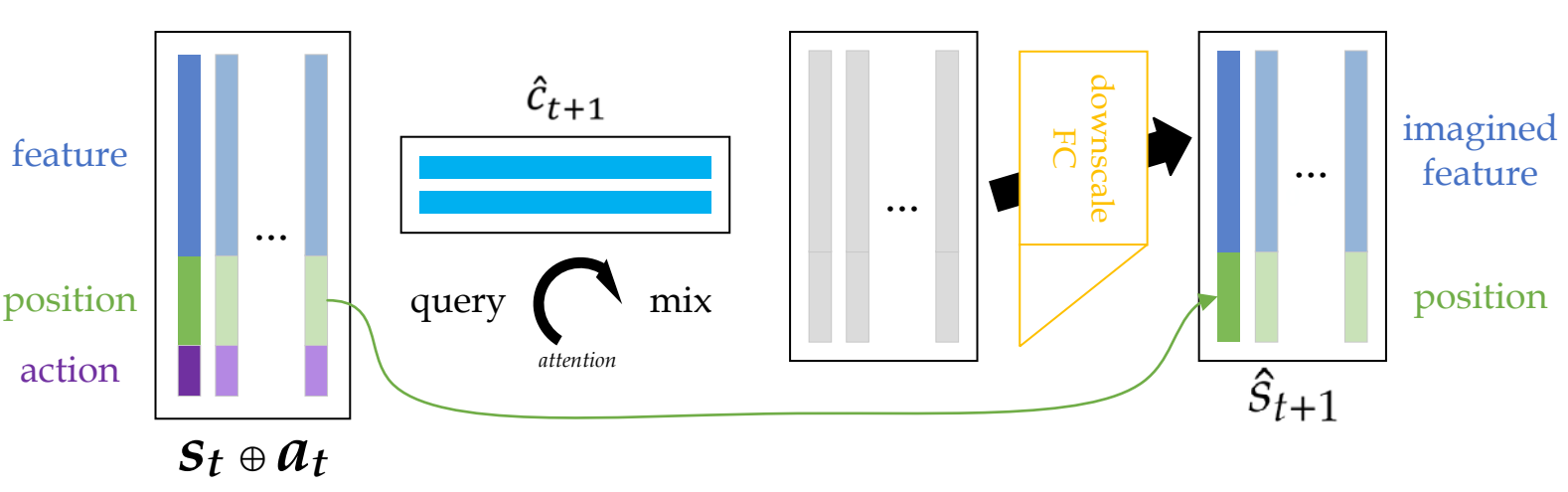
feature
position
action
$s_t \oplus a_t$
$\hat{c}_{t+1}$
query
mix
attention
downscale FC
imagined feature
position
$\hat{s}_{t+1}$