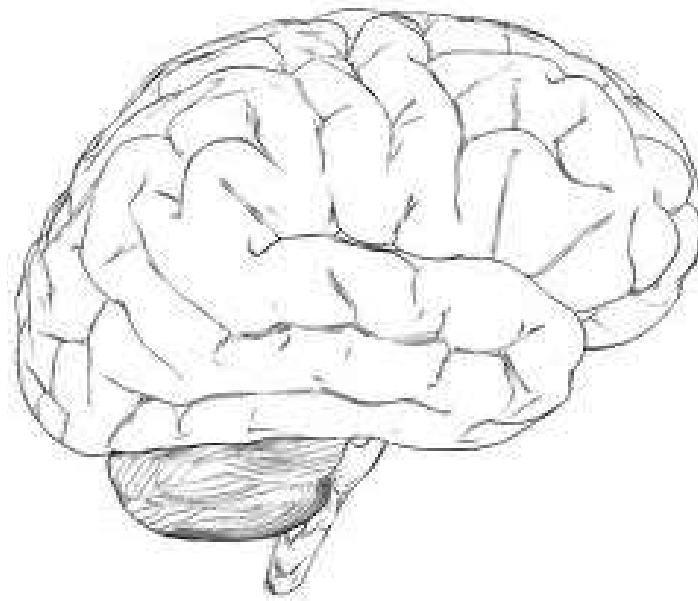


observation

$O_t$



action

$A_t$

reward

$R_t$

