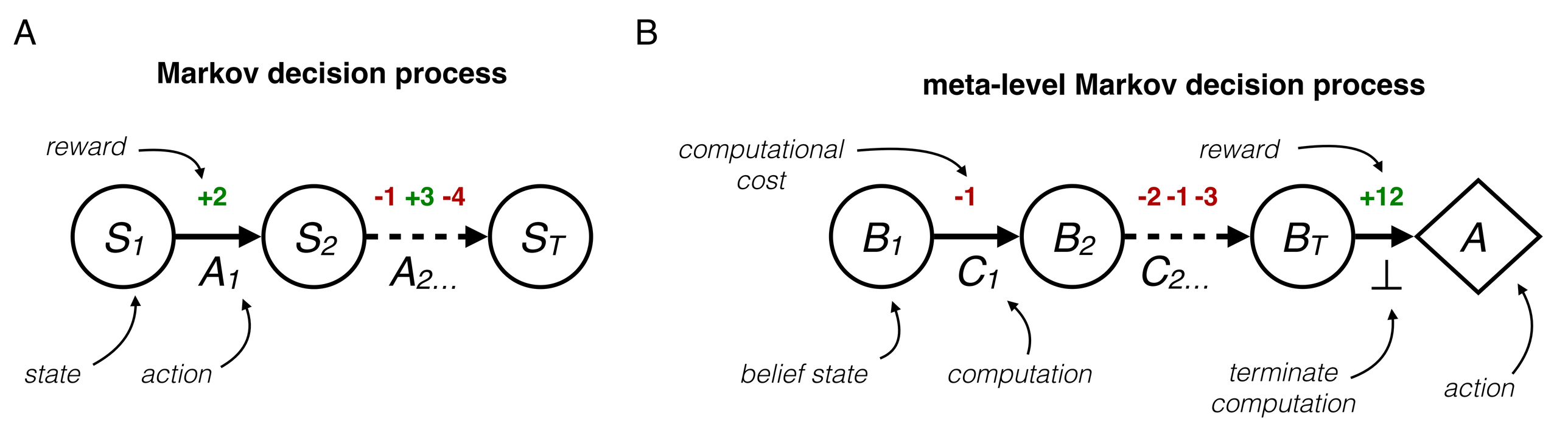
A
Markov decision process
reward
+2
-1 +3 -4
S_1
S_2
S_T
A_1
$A_{2...}$
state
action
B
meta-level Markov decision process
computational cost
reward
-1
-2 -1 -3
+12
B_1
B_2
B_T
A
C_1
$C_{2...}$
$\perp$
belief state
computation
terminate computation
action