

Resource-Rational Cooperation

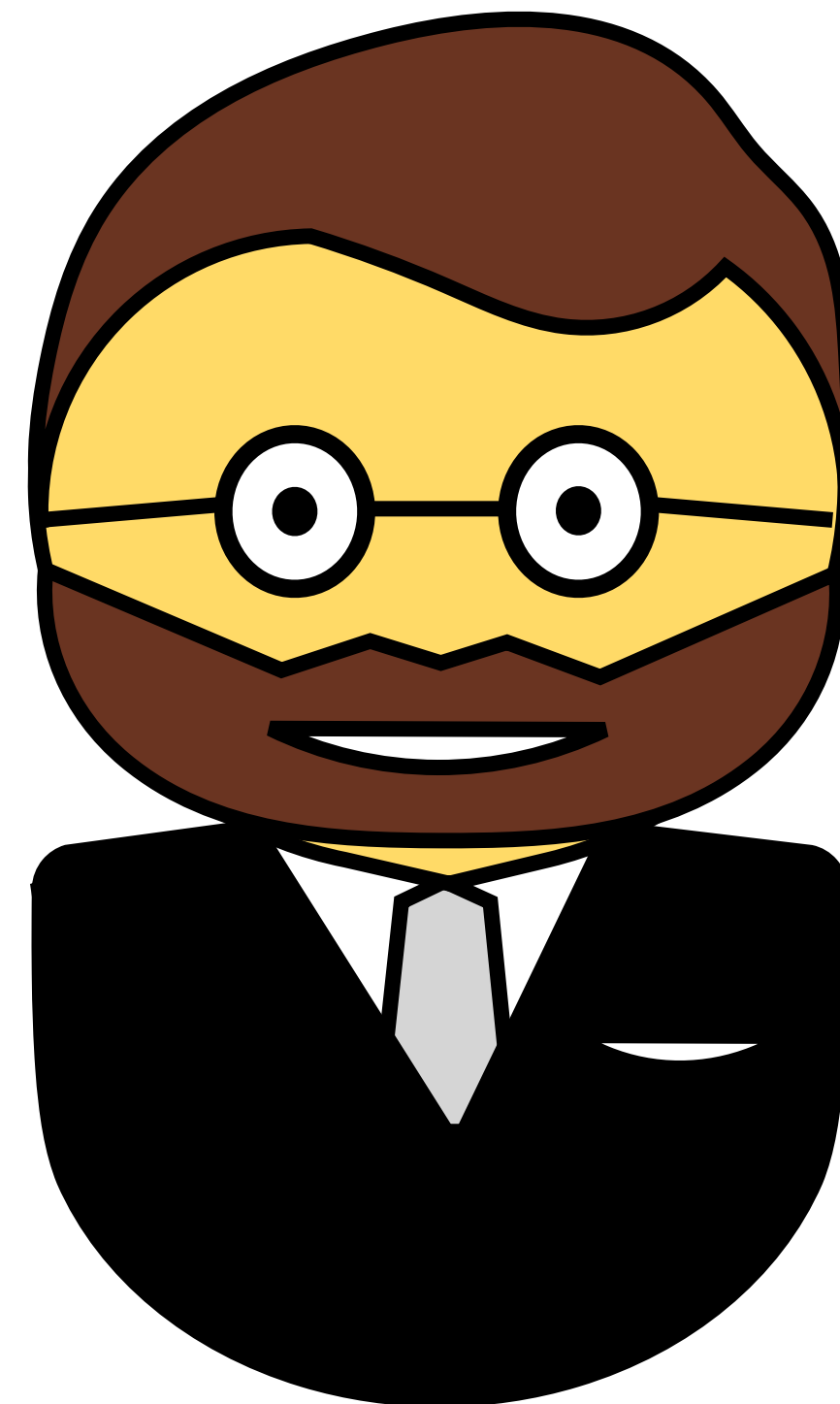
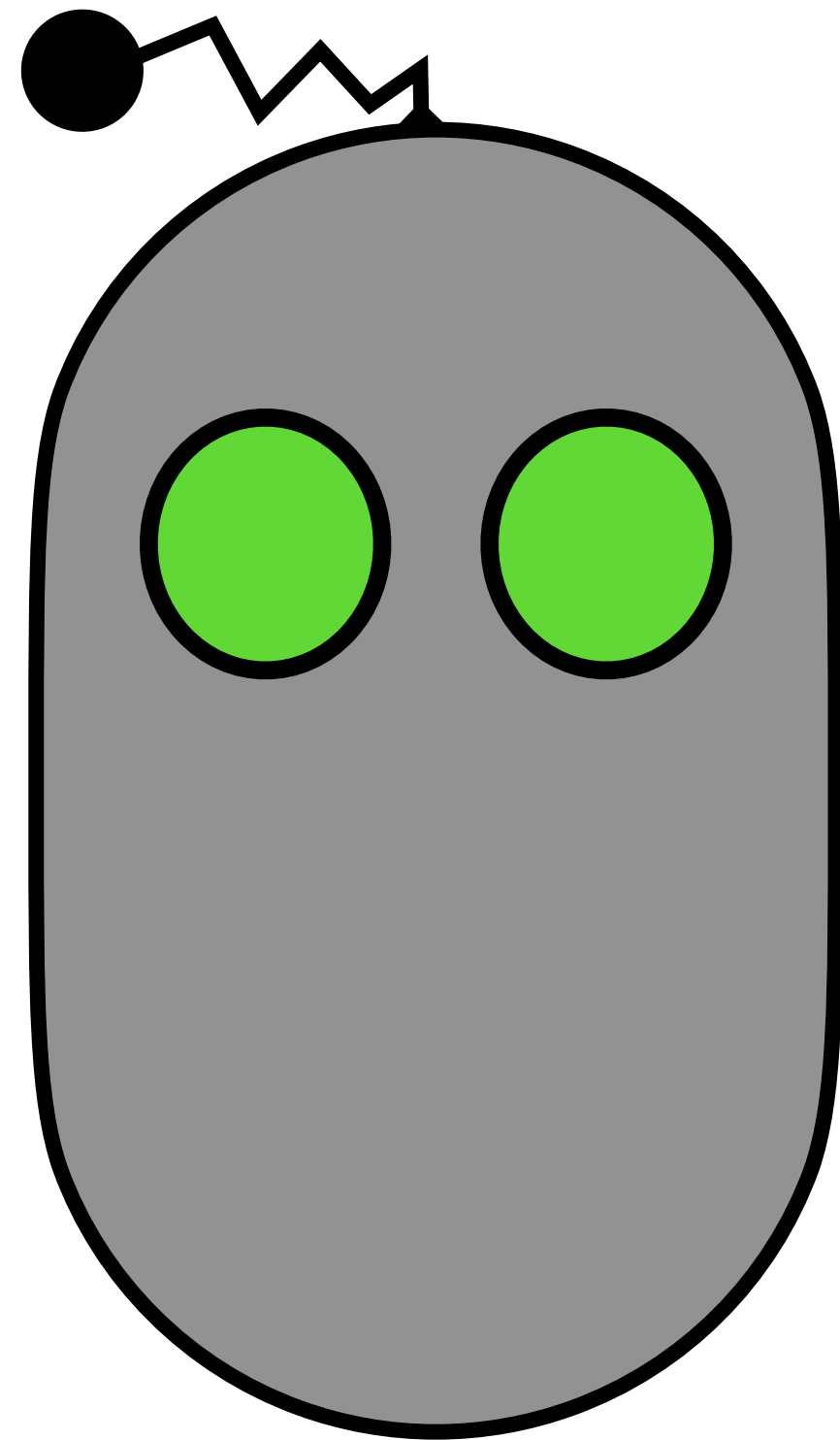
Cam Allen



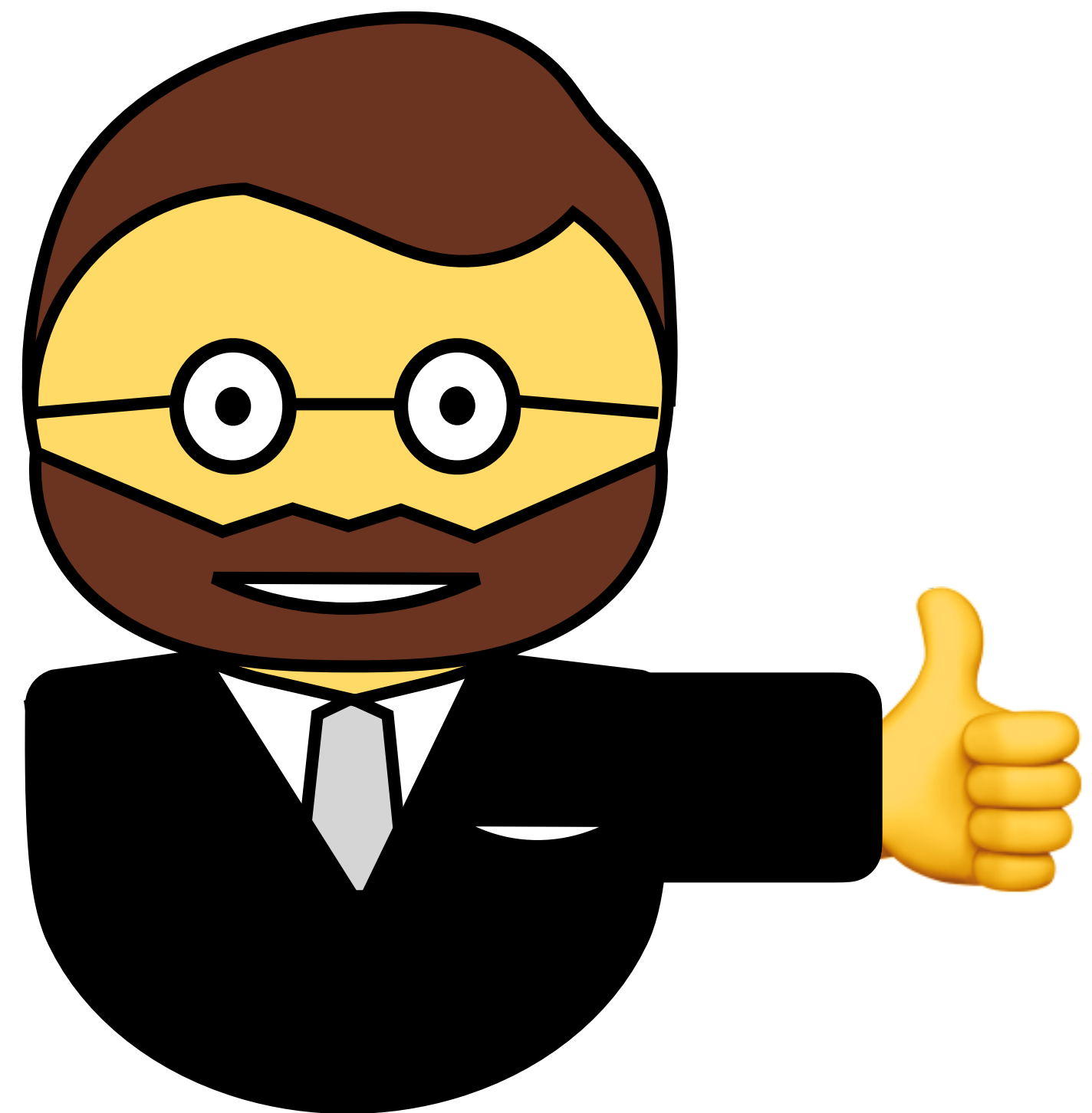
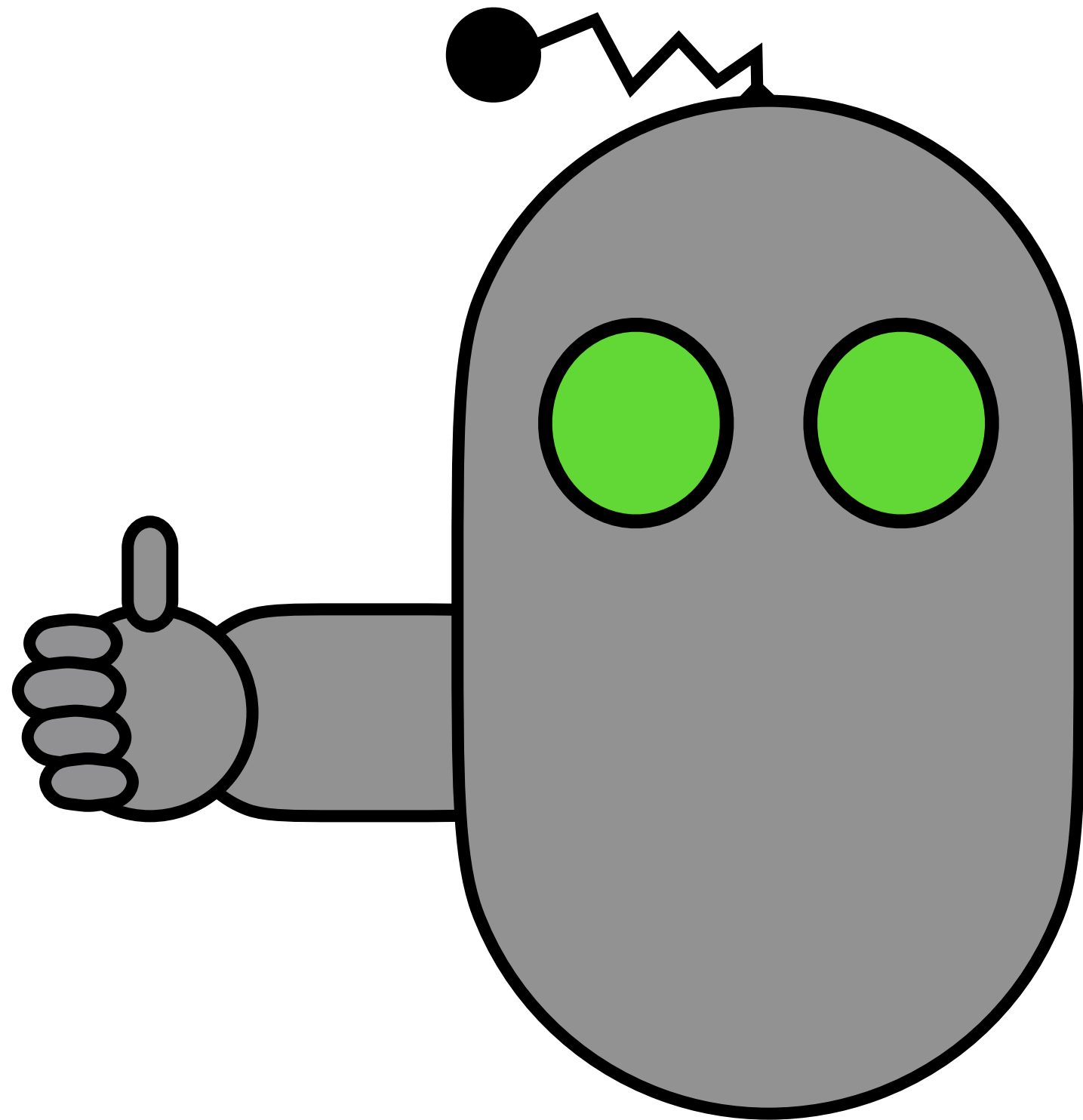
Center for
Human-Compatible
Artificial
Intelligence

Berkeley
UNIVERSITY OF CALIFORNIA

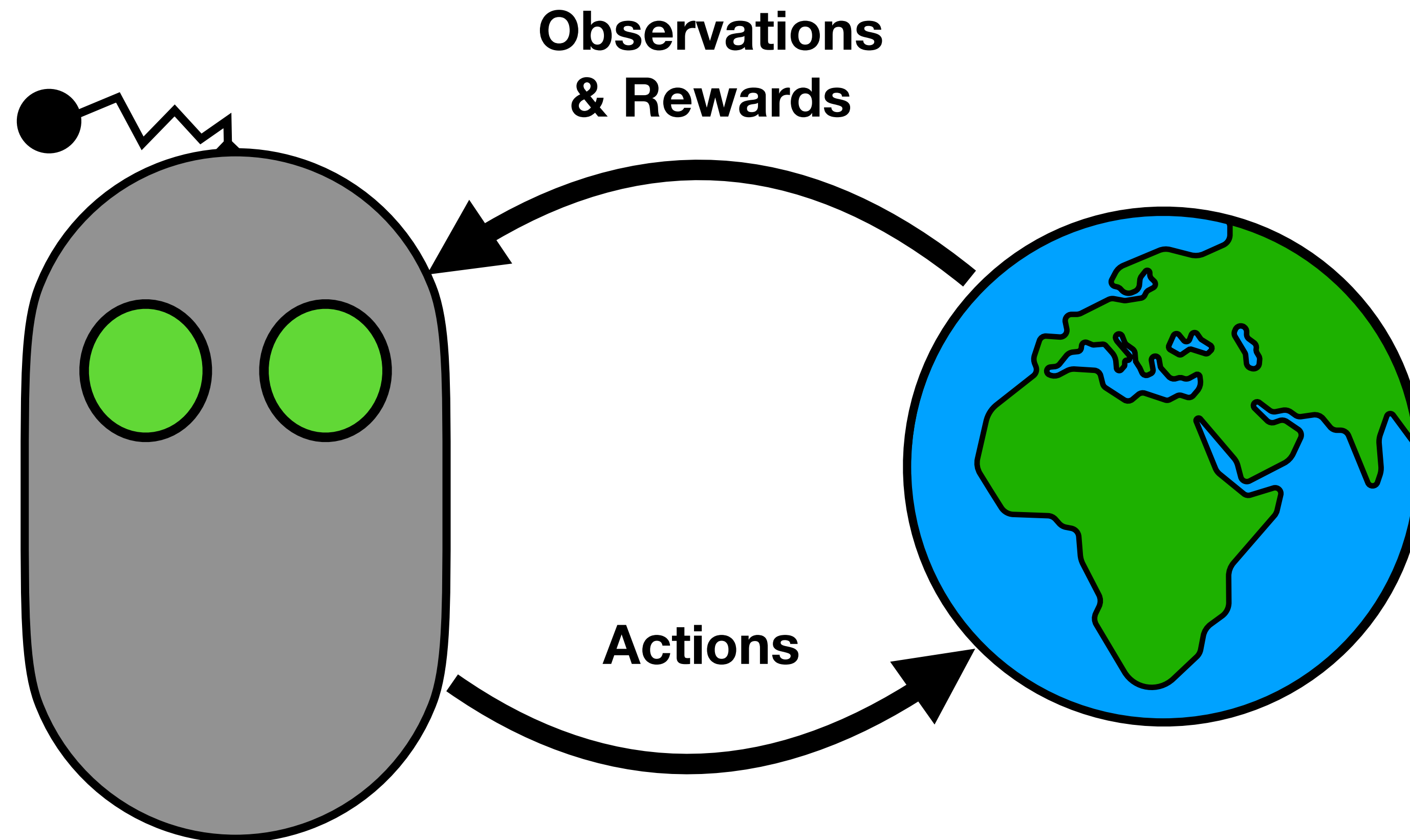
Designing Human-Compatible AI



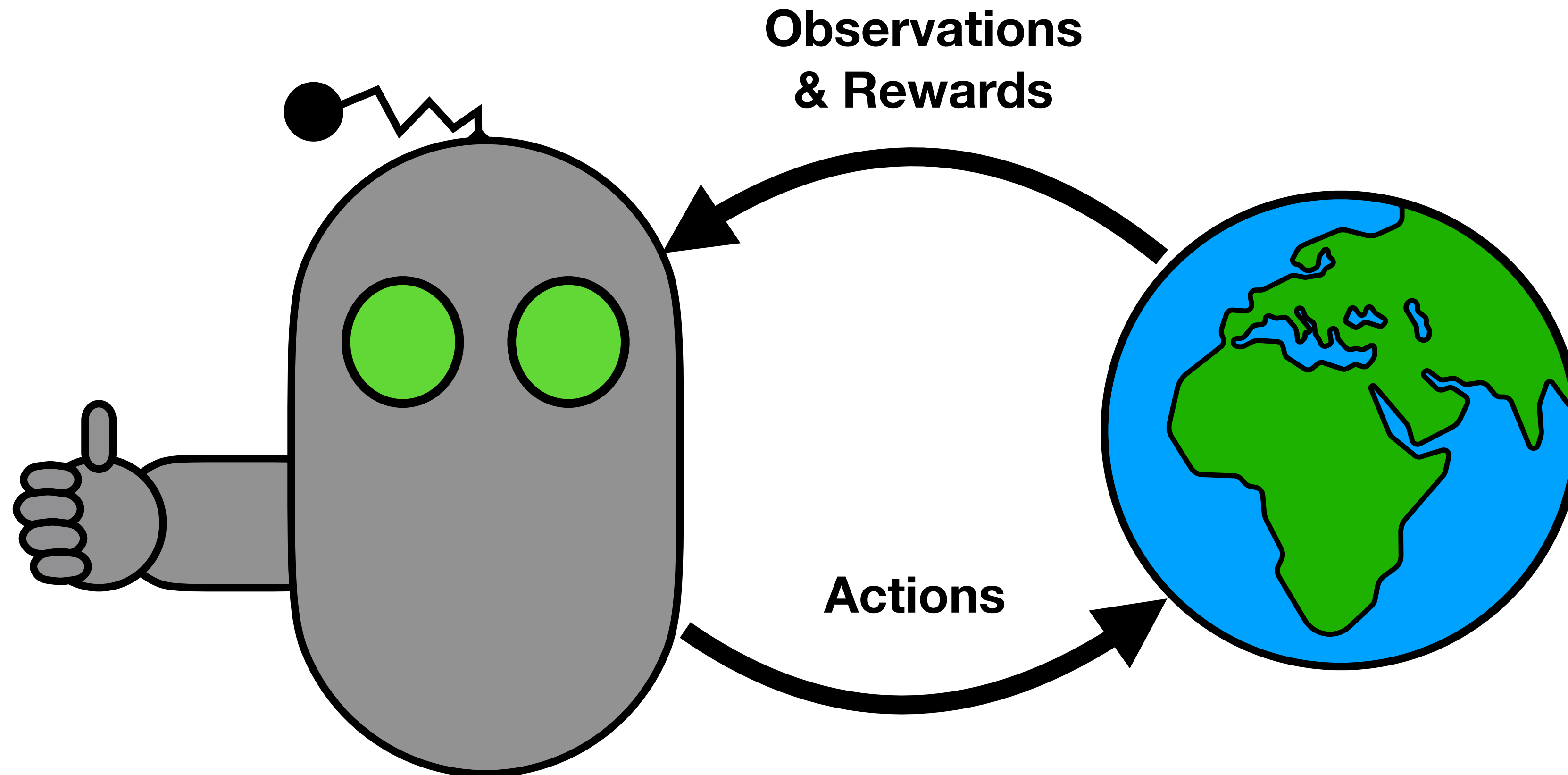
Designing Human-Compatible AI



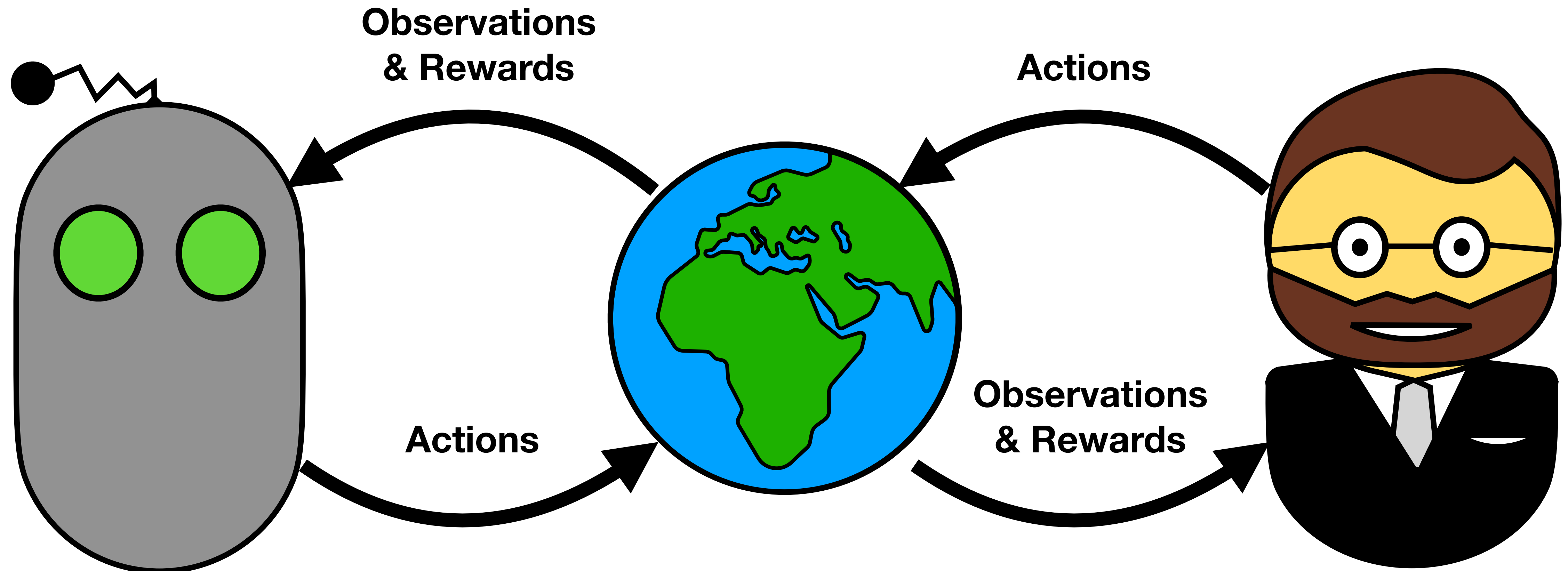
Reinforcement Learning



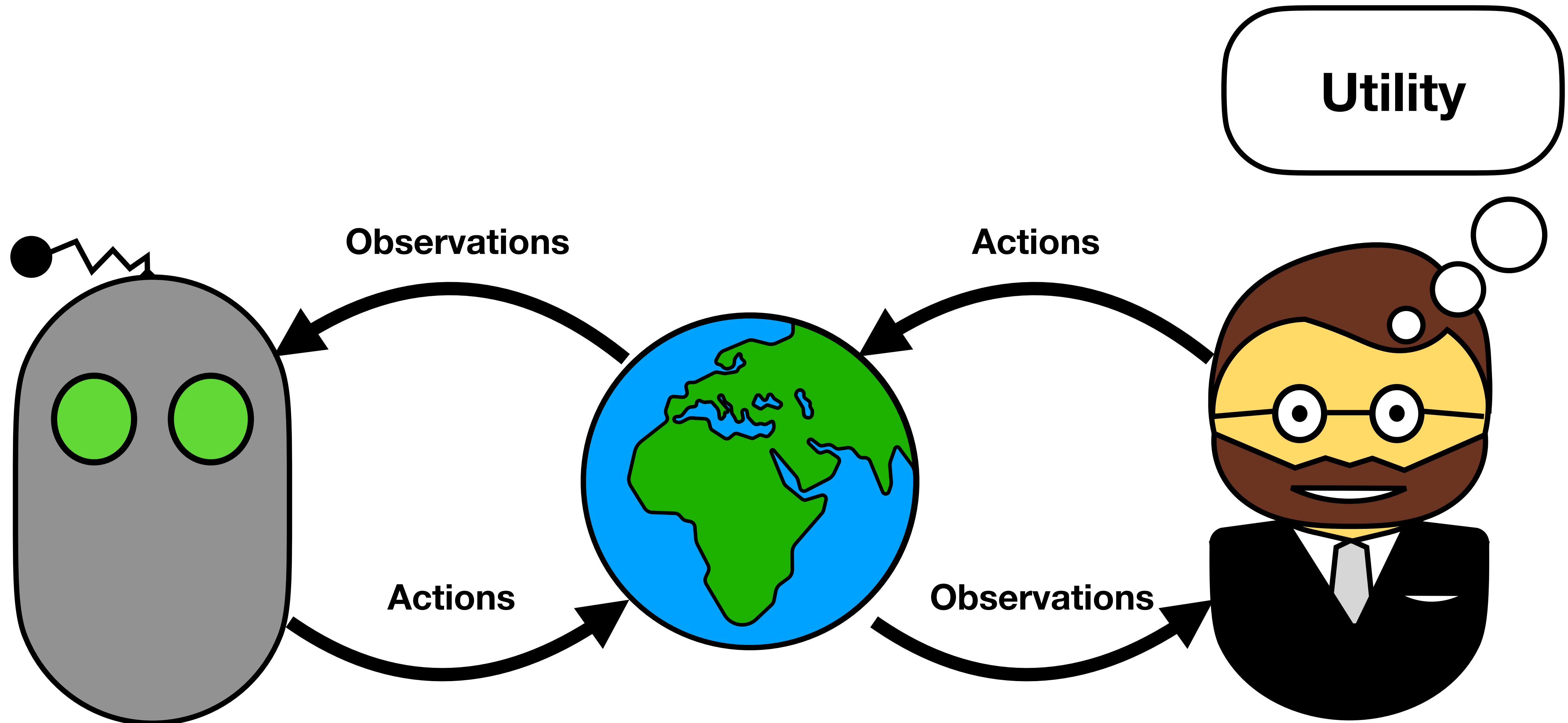
Reinforcement Learning



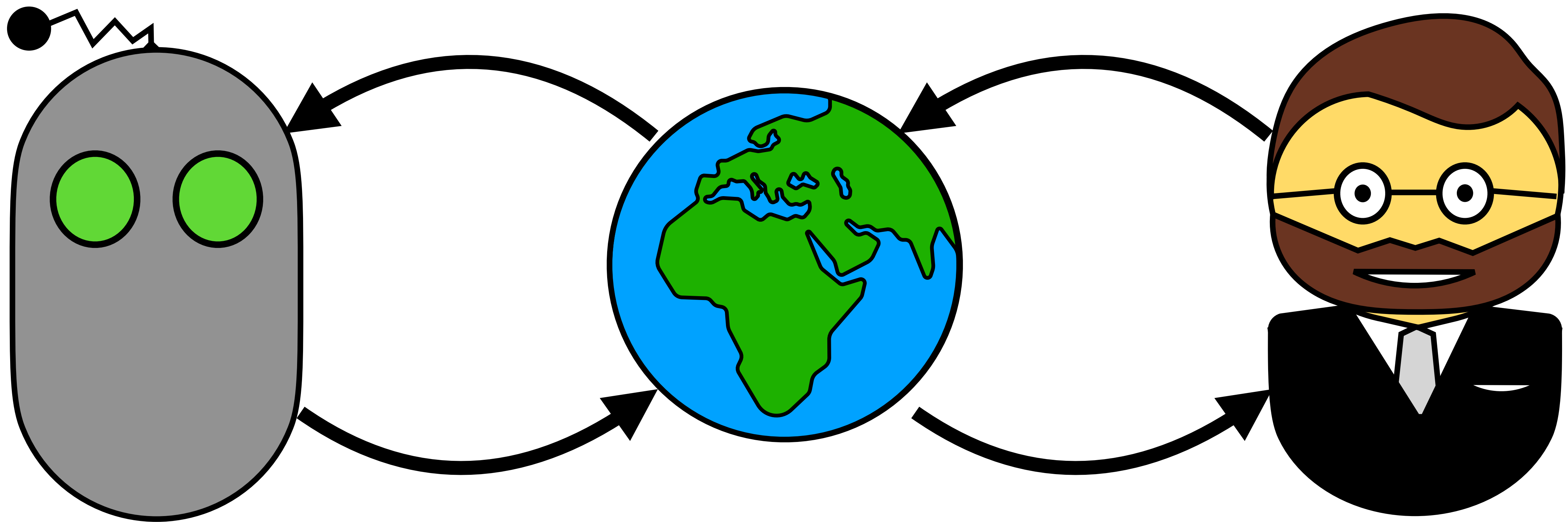
Game Theory



Assistance Games



Rationality Assumption

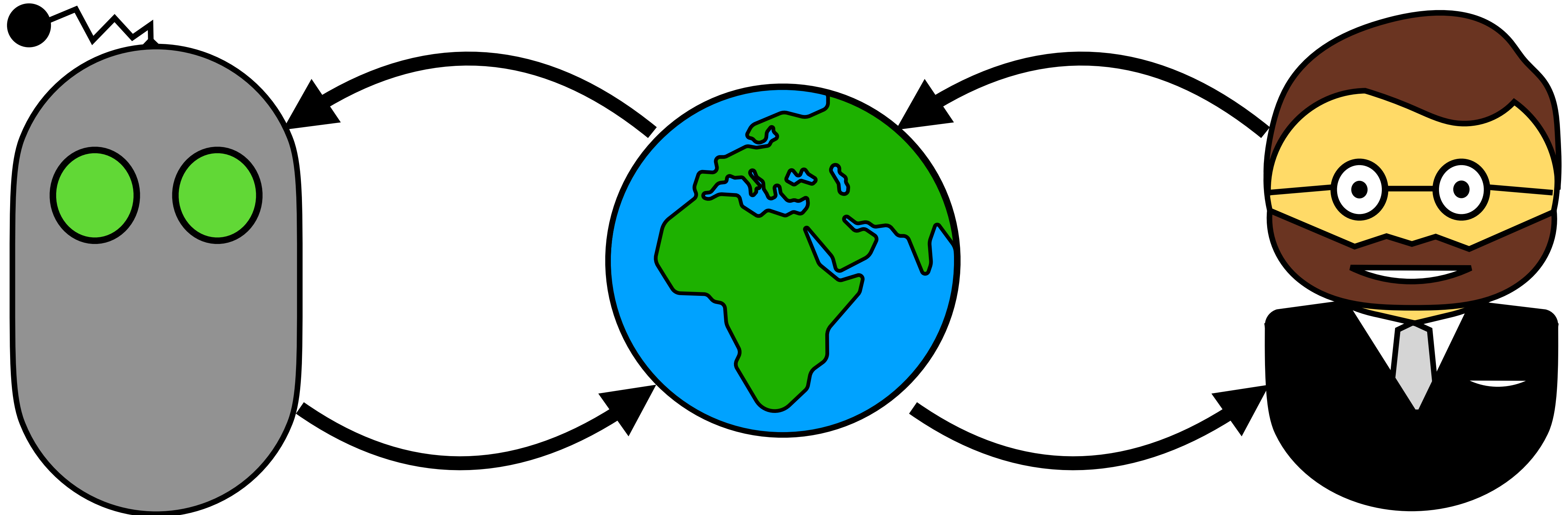


Rationality Assumption

$\times \infty$



$\times \infty$

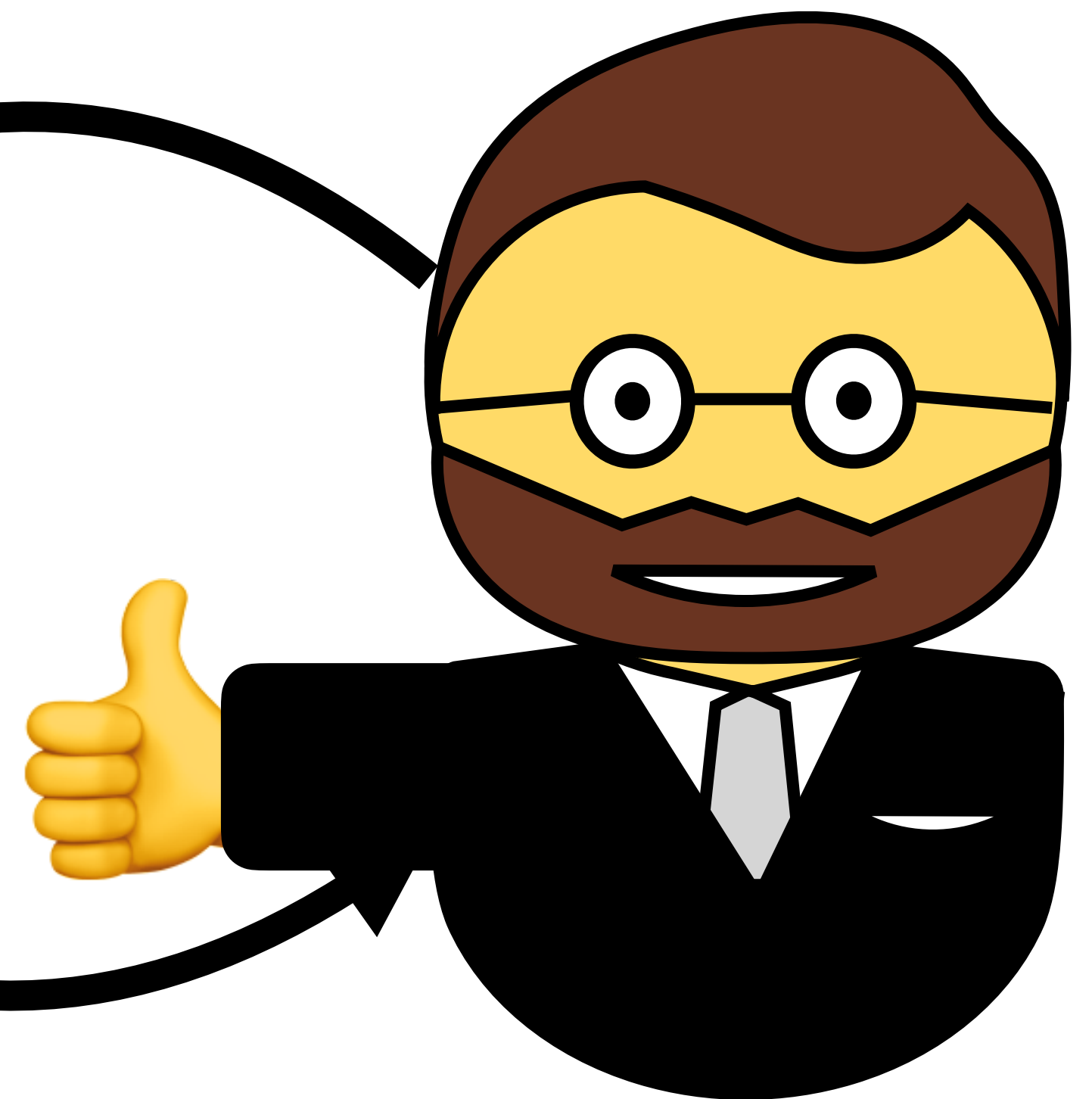
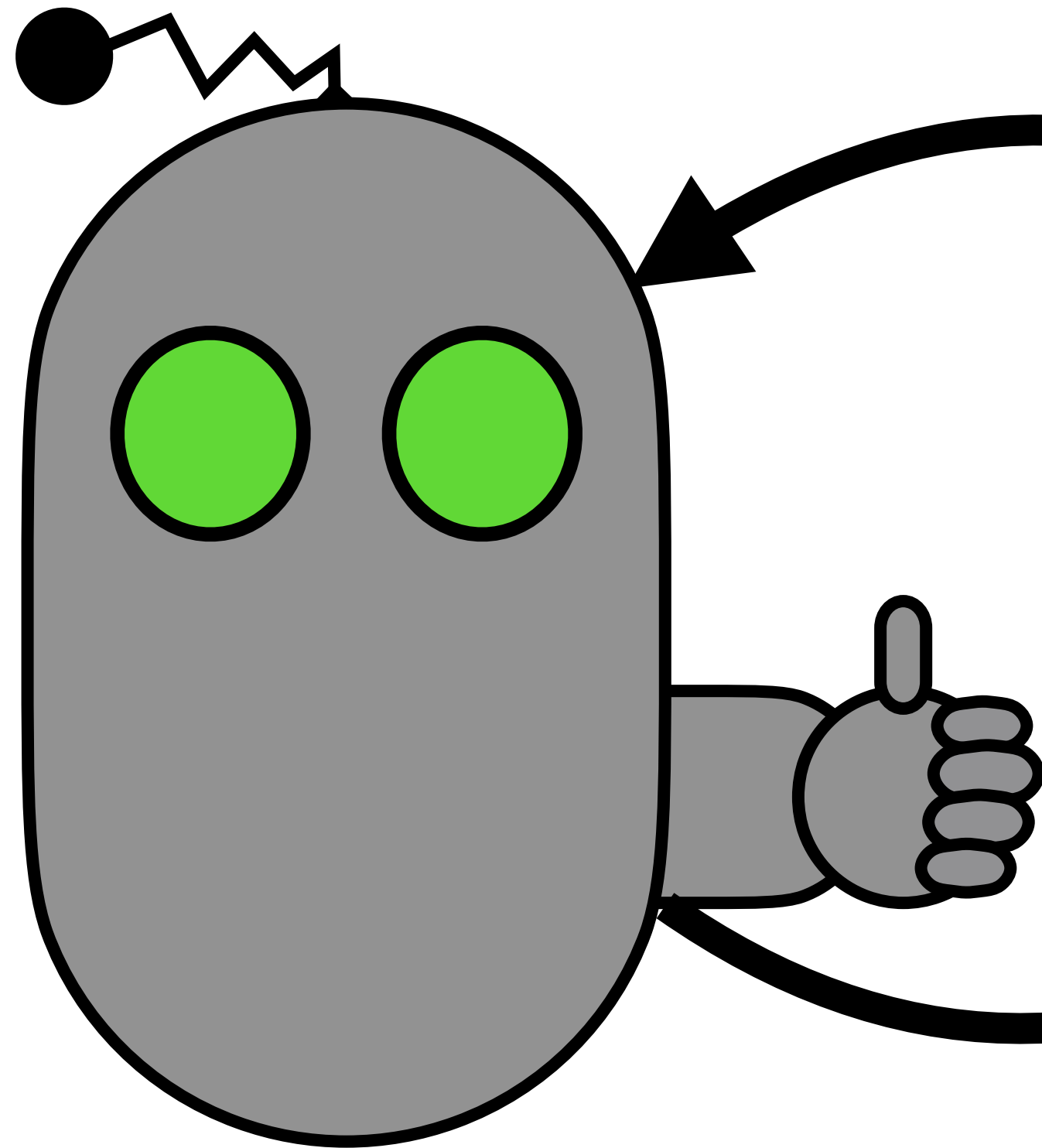


Rationality Assumption

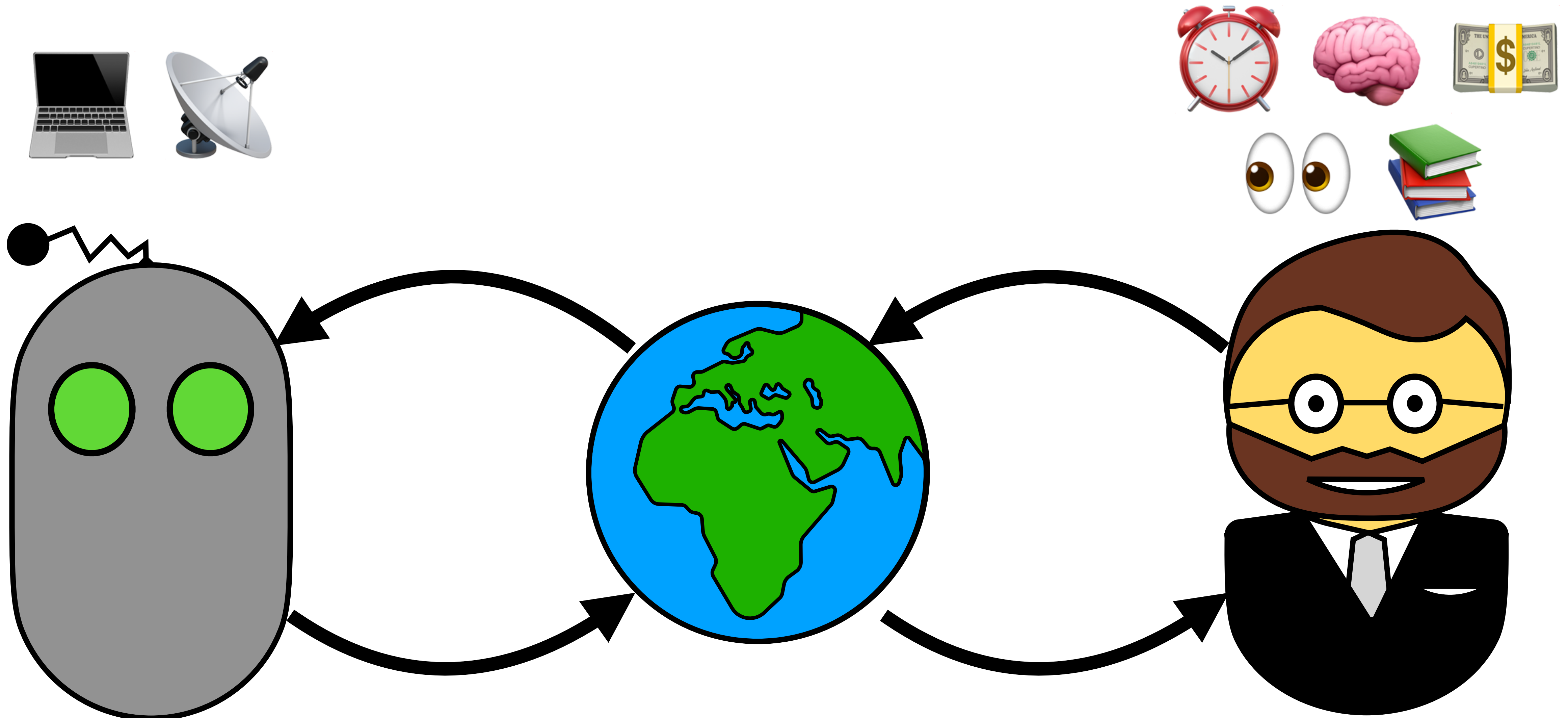
$\times \infty$



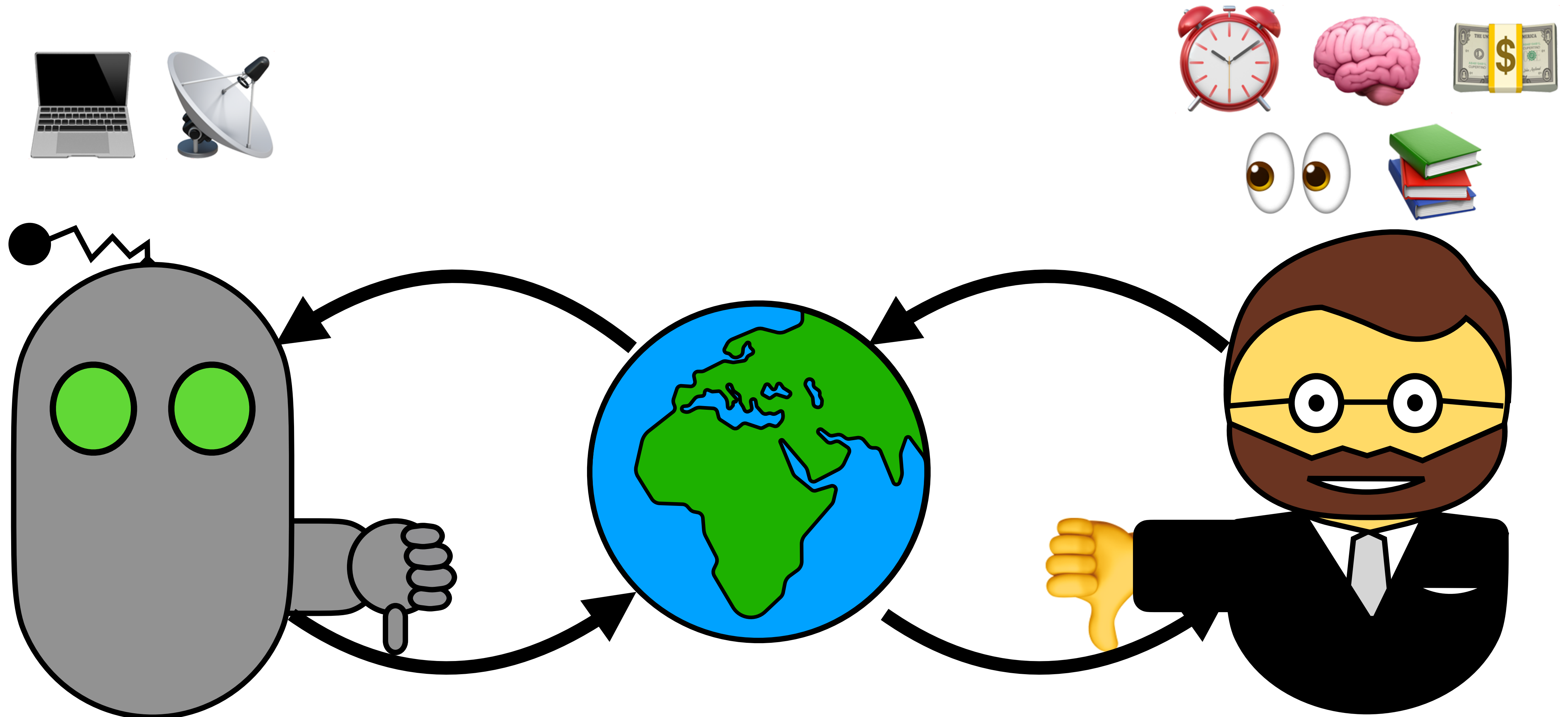
$\times \infty$



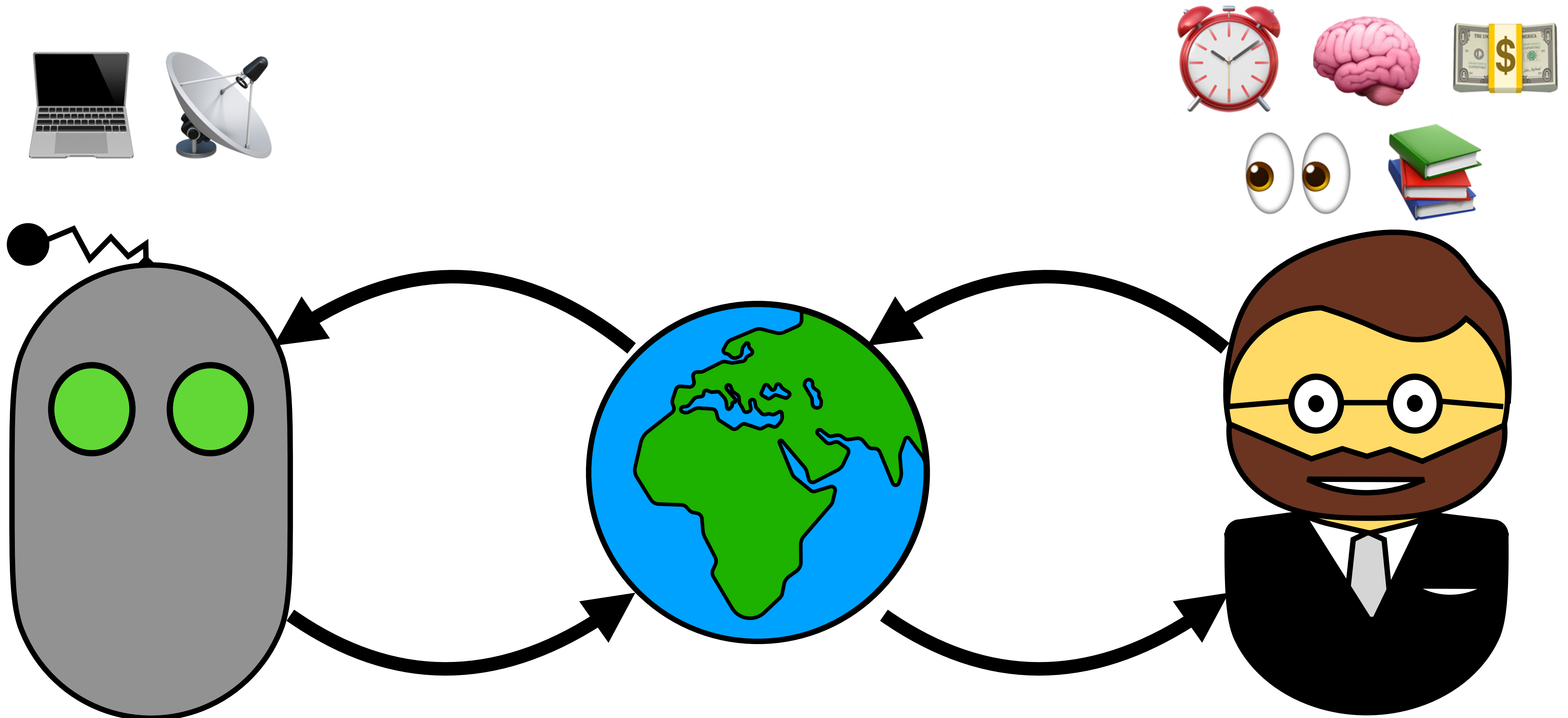
Resource-Rationality



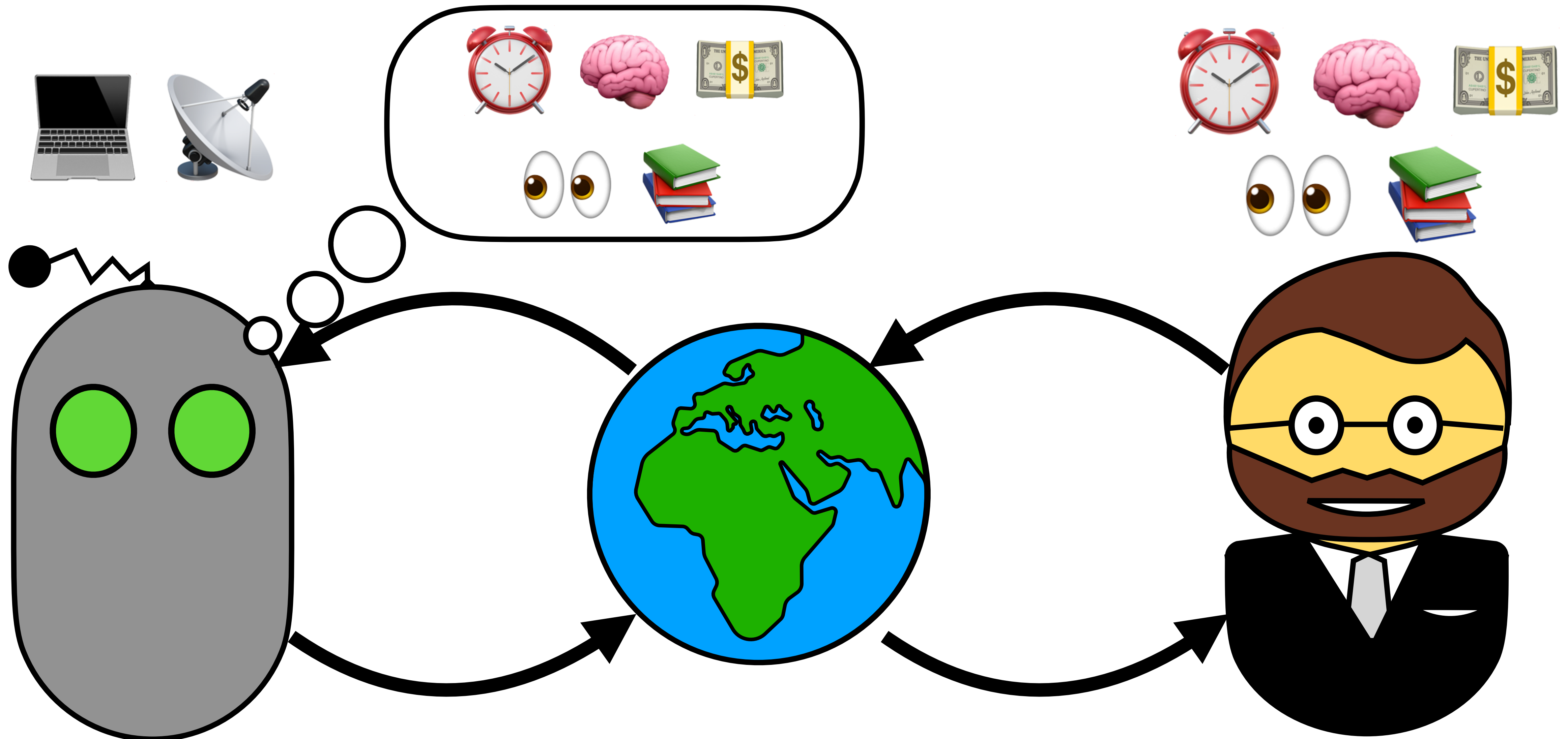
Resource-Rationality



Resource-Rationality



Resource-Rationality



Resource-Rationality

