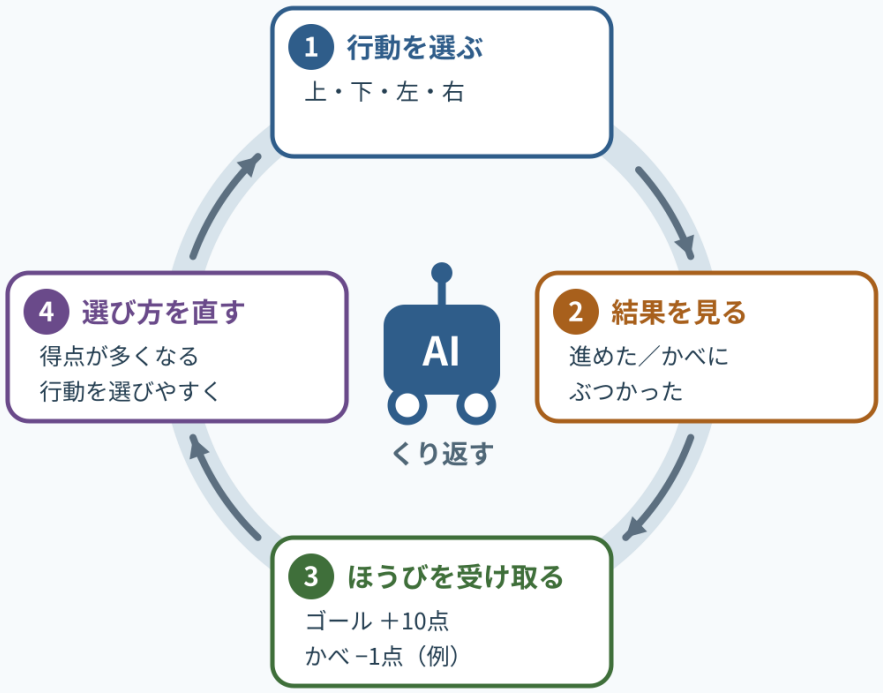
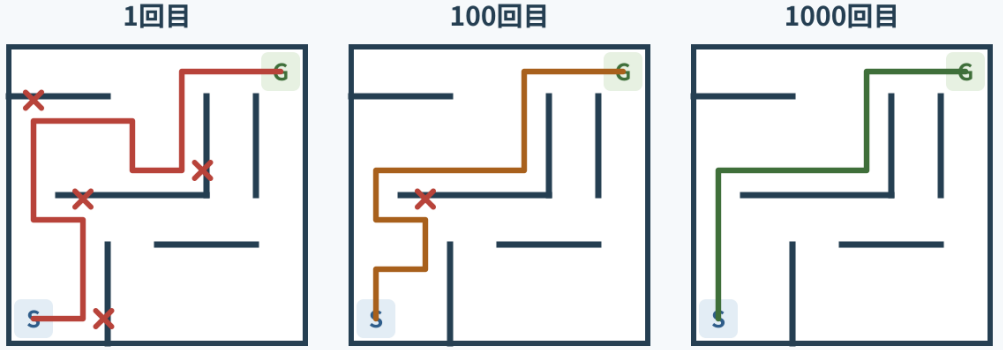
強化学習（試して、ほうびで上手になる）

正解を教わらなくても、AIが上手になるのはどうしてだろう

例：迷路を進むロボット（架空）。ゴールまでの道を、くり返し試して覚える



くり返すほど、道すじがよくなる（模式）



ぶつかる 4回・14歩
いろいろ試す

ぶつかる 1回・12歩
まわり道が減る

ぶつかる 0回・10歩
近い道を選ぶ

学び方のくらべ

教師あり学習	正解つきの例から学ぶ
教師なし学習	正解なしで、似たもの同士をまとめる
強化学習	正解の代わりに、ほうびを手がかりに試して学ぶ

ポイント　その場の得点だけでなく、長い目で見て得点が多くなる行動を学ぶ。どの結果に何点をあげるかは、つくる人が決める

使われている例　囲碁などの対戦／ロボットの歩き方／対話型AIの答え方の調整（人の評価をほうびにする）

※ 例は架空。点数・回数・歩数は説明のための例。