1. Reward system 2 sẽ tốt hơn vì ở mỗi bước đi số điểm sẽ bị trừ đi 10, khi đó agent sẽ có cơ sở để tìm ra đường ngắn nhất với số điểm bị trừ ít nhất.

Còn ở system 1 agent sẽ chỉ cứ thế mà brute force đi hết mê cung mà không tìm được đường đi tối ưu.

1. Reward system cho bài toán tìm đường khỏi mê cung hiệu quả hơn ở system trên: agent sẽ có 2 cột điểm, xuất phát với 0 điểm mỗi 1 bước đi sẽ trừ lại 1 điểm cho đến khi đi ra khỏi mê cung và lưu lại số điểm đó. Mỗi khi ra khỏi mê cung với số điểm lớn hơn hoặc bằng số điểm max lần trước agent sẽ được cộng 1 điểm vào cột reward (nếu lớn hơn sẽ được cộng 2 điểm).

* Fully observable: Khi chụp ảnh, ta có thể sử dụng chế độ Auto có thể tự thay đổi độ sáng của bức ảnh ứng với điều kiện môi trường để cho ra bức ảnh mong muốn.
* Partialy observable: những bài toán, trò chơi gỡ mìn, sudoku, … những trò chơi này sẽ cho gợi ý trước về đường nên đi tiếp theo. Với gỡ mìn là những con số sẽ thể hiện số lượng bom xung quanh ô đó, với sudoku là những con số đã có trên cột cho biết những số còn thiếu trên hàng, cột, ô tương ứng.
* Unobservable: những thể loại game Choice matter (game lựa chọn tình huống mà với mỗi lựa chọn sẽ đưa ra cái kết tương ứng) ta sẽ không biết tiếp theo chuyện gì sẽ xảy ra tương ứng với lựa chọn của người chơi (ví dụ game Life is Strange, Detroit become humman, …)
* Episodic: trong bài toán xác suất thả con xúc xắc nhiều lần, mỗi lần con xúc xắc rơi như thế nào không ảnh hướng tới số điểm của xúc xắc phía sau.
* Sequential: ví dụ là những trò chơi cờ, với mỗi bước đi tiếp theo sẽ ảnh hưởng tới toàn bộ quá trình sau đó, một nước đi sai có thể làm mất 1 quân hoặc cả ván cờ.

1. Tìm đường đi từ TP.HCM tới Kiên Giang

* Initial state: TP.HCM
* Actions:
  + S (TP.HCM) = {Tây Ninh, Bến Tre, An Giang}
  + S (Tây Ninh) = {An Giang}
  + S (Bến Tre) = {An Giang, Cần Thơ}
  + S (An Giang) = {Kiên Giang, Bến Tre}
  + S (Cần Thơ) = {Kiên Giang, Bến Tre, Cà Mau, Bạc Liêu}
* Transititon model:
  + Result (TP.HCM, An Giang) = An Giang
  + Result (An Giang, Kiên Giang) = Kiên Giang
* Goal test:
  + Result (TP.HCM, An Giang) = An Giang -> F
  + Result (An Giang, Kiên Giang) = Kiên Giang -> True
* Path cost: số km phải đi qua để đến tỉnh thành Kiên Giang. Cost (TP.HCM, Kiên Giang) = 152 + 54 = 206 (km)