

超級馬力歐兄弟是由任天堂於1983年推出的遊戲，截至目前為止，由人類打出來的any%（指不限完成率的情況下速通）成績為4分54.448秒。而Tool-Assisted Speedrun（以下簡稱tas），是指用其他工具，以精度極高的方式完成遊戲，可避免手操上難度要求與運氣因素，其成績為4分54.265秒，是目前人類根據自己的遊戲理解，找出的理論最佳的成績。但我認為20年後的ai，將可能打破這個認知。

事實上，早在2018年，ai就已經能自主學習通關這遊戲，然而，嚴格意義上的速通，在目前為止仍遠不如人類玩家。為了打破tas記錄，我認為ai最關鍵的學習方法為強化學習，因為速通的目標為通過一連串的動作序列，以最短時間完成遊戲，ai能人物位置、速度、當前時間等作為資訊來源，從過關時間得到獎勵，不斷試錯找出最短時間。但單純強化學習，ai仍幾乎不可能達到目標，原因如下：

- 1.行動空間巨大：若單純窮舉，相當於每秒60幀，每幀8個方向與跳躍組合。
- 2.獎勵延遲：ai在過程中無法知道結果好壞，只有通關後看到時間，才能得到「獎勵」。
- 3.操作精度要求極大：某些速通技巧，即使操作時機只差一幀，結果也會有極大區別。

因此，監督式學習也相當重要，先學習現有的tas路徑作為起點，再嘗試優化，找到更短的路線。

設計出簡化模型的方式如下：

將遊戲合理的分成大量的起點和終點（如8-4開始區域至下一個水管區域），並分別對每段做出訓練，嘗試短的時間區間內找到最佳動作序列，以到達終點的時間減少作為獎勵。可測性方面，可將tas成績作為基準，若能找到超越基準的操作方式，即可知道模型成功。

需要的機器學習工具（對此方面不了解，因此下列為chatgpt提供的結果）：

Policy gradient (PPO) 或 Actor-Critic (A2C / IMPALA) 配合 CNN (若用像素)。

強烈建議採用 imitation initialization：用行為克隆 (Behavior Cloning) 學習一條良好但非最優的示範路徑，再用 RL 微調 (加速探索)。

若想要「幀級精細優化」，在訓練後用局部搜索 (CEM / CMA-ES / beam search) 做精調 (fine-tuning) 於最佳策略輸出序列上

若未來真能用ai突破tas記錄，其影響力遠不止在遊戲方面。像是能說明ai可在極小誤差容忍下，找到全域最優解，這在組合爆炸問題、程式最佳化等方面也能取得重大突破。同時說明ai具備改進演算法的能力，不僅能夠學習，還能發明新策略，甚至可能在創造性上超越人類。