Artificial Intelligence

Homework 5

Reinforcement learning

姓名: 劉育辰

學號:110303585

系級:機械4C

1. the 3D mesh plot of the finished Q-table

一張含有 圖表, 文字, 行, Rectangle 的圖片

AI 產生的內容可能不正確。

上圖為Q-TABLE的3D圖，可以發現真的被用到的狀態空間只有一小塊，因為在對e和de的linespace切割的不夠精細，又e和de的數值取到小數點後好幾位，真的有變化的是小數點後6-7位數的部分，導致離散化後只集中在特定幾個狀態空間內，所以只有這幾個狀態空間對應的Q-TABLE真實被用到。

1. the control and simulation code

一張含有 文字, 螢幕擷取畫面, 字型 的圖片

AI 產生的內容可能不正確。

這段先對state = (error, delta error) 狀態空間作定義，和Q-TABLE作初始化。並定義動作空間action(u(k))。Discretize這個function是用來離散化輸入的value，將value插入對應的linspace內。

一張含有 文字, 螢幕擷取畫面, 字型, 軟體 的圖片

AI 產生的內容可能不正確。

該函示對輸入的參數X = [ I, omega ], u 做計算到下一步Q(S, a’)的數值，在實驗中有發現有數值浮動很大的情況，所以用np.clip()來對數值作約束，防止數值爆炸。

一張含有 文字, 螢幕擷取畫面, 軟體, 作業系統 的圖片

AI 產生的內容可能不正確。

這段是根據PPT上Q-learning得步驟來做RL的模擬，先根據每一次的的episodes做1000步的計算，一開始先根據現在的state(e, delta\_e) 得到Q(s,a)，並從現在的Q-TABLE中選一個最好的ACTION。這裡在選擇策略上是根據ε-greedy ，會有10%機會隨機給ACTION，90%的機會選擇Q(S,a)值最大的ACTION來做。

一張含有 文字, 螢幕擷取畫面 的圖片

AI 產生的內容可能不正確。

這一段落，會根據剛剛ε-greedy選到的ACTION來給motor\_modle(x,u)，這函式會根據現在狀態的電流、omega和電壓(u)，來得到下一步的狀態Q(S,a’)，最後根據公式更新Q-TABLE上的值和現在的狀態S(e,delta\_e)。同時我有紀錄每次步數omega的變化，以利於之後畫圖分析omega得變化。

1. 一張含有 文字, 行, 繪圖, 圖表 的圖片

   AI 產生的內容可能不正確。simulation of controlling results

從結果圖可以知道馬搭大約花到0.05秒後才會穩定達到2.0rad/s的轉速，且成長數度很穩定，趨近線性成長，我在模擬的結束機制是「如果e的值小於0.01時代表收斂」。所以當輸出差異都是在2.0正負0.01值左右代表馬達的轉速已趨於穩定。

1. Discussions

用上圖的結果來討論，我總共跑50次Episode，每10次紀錄omega的變化，可以看出第一個10 Episode(藍色線)，花了約0.07秒才達到穩定狀態，但隨著模擬次數增加，Q-TALBE不斷跟新，策略選擇的ACTION越來越精準，到第50次只需要0.05秒，比第10次快0.02秒。但還有一個值得注意的是第30次，這次其實就能在0.05秒達到目標轉速，但是卻花了0.01秒左右才趨於穩定，可以知道機器找到最快速達到目標的ACTION策略，但在後續的穩定上還不夠好，要再繼續更新Q-TABLE內的值。

再來是系統變數上我有自己做改動，J = 0.01，如果用原本的數值J = 2，會導致慣量很大，要很長時間才能轉起來，每次的omega都是如下圖，可能要花更多的episodes和計算時間，我感覺我的電腦可能會起飛，為了減少計算時間所以修改成0.01。

一張含有 文字, 螢幕擷取畫面, 字型 的圖片

AI 產生的內容可能不正確。

1. Conclusions

就這次實驗作業比上次的還更要花時間來對模型環境去做設定，在對狀態空間去做大小的定義就花了很多時間，為了要讓狀態空間能被有效利用，但是從Q-TABLE的圖可以看出來還是不夠好，爾且也不是說花更多episodes，就能被有效利用，因為最終機器會找到最好的ACTION策略來做，花的時間只會更少，收斂的時間更快。