گزارش تمرین برنامه نویسی (۵)

بازی Flappy Bird با استفاده از Reinforcement Learning

هلیا شمس زاده

۴۰۰۵۲۱۴۸۶

- **policy**: تابعی که به عنوان استراتژی انتخاب عمل در هر حالت (state) استفاده می‌شود.

یک استیت را به عنوان ورودی گرفته و اکشن بیشترین Q را به عنوان policy برمی‌گرداند.

- **get_all_actions**: تمام اکشن‌های ممکن در هر حالت (بالا یا پایین حرکت کردن پرنده) را برمی‌گرداند.
- **convert_continuous_to_discrete**: تبدیل حالت پیوسته به حالت گسسته. نمی‌توانیم به ازای تمام حالت‌های پیوسته، استیت داشته باشیم، پس باید به گسسته تبدیل کنیم.
- **compute_reward**: محاسبه پاداش بر اساس امتیاز به دست آمده در حالت جدید. در صورتی که رد کرده بود ۱۰۰ و در صورت حرکت اشتباه ۱۰۰- برمی‌گرداند.
- **get_action**: انتخاب عمل با توجه به استراتژی انتخاب تصادفی یا بر اساس استراتژی Q-learning. در این روش، با احتمال حرکت رندوم انجام می‌دهیم، یا برحسب تجربیات قبلی حرکت می‌کنیم. به احتمال اپسیلون حرکت رندوم، و به احتمال اپسیلون – ۱ حرکت بر اساس policy کنونی انجام می‌دهیم.
- **maxQ**: برگرداندن عملی که مقدار Q بیشینه دارد.
- **max_arg**: برگرداندن عمل متناظر با مقدار Q بیشینه.
- **update**: بروزرسانی مقدار Q با توجه به پاداش و حالت بعدی. فرمول این بروزرسانی به شکل زیر است:

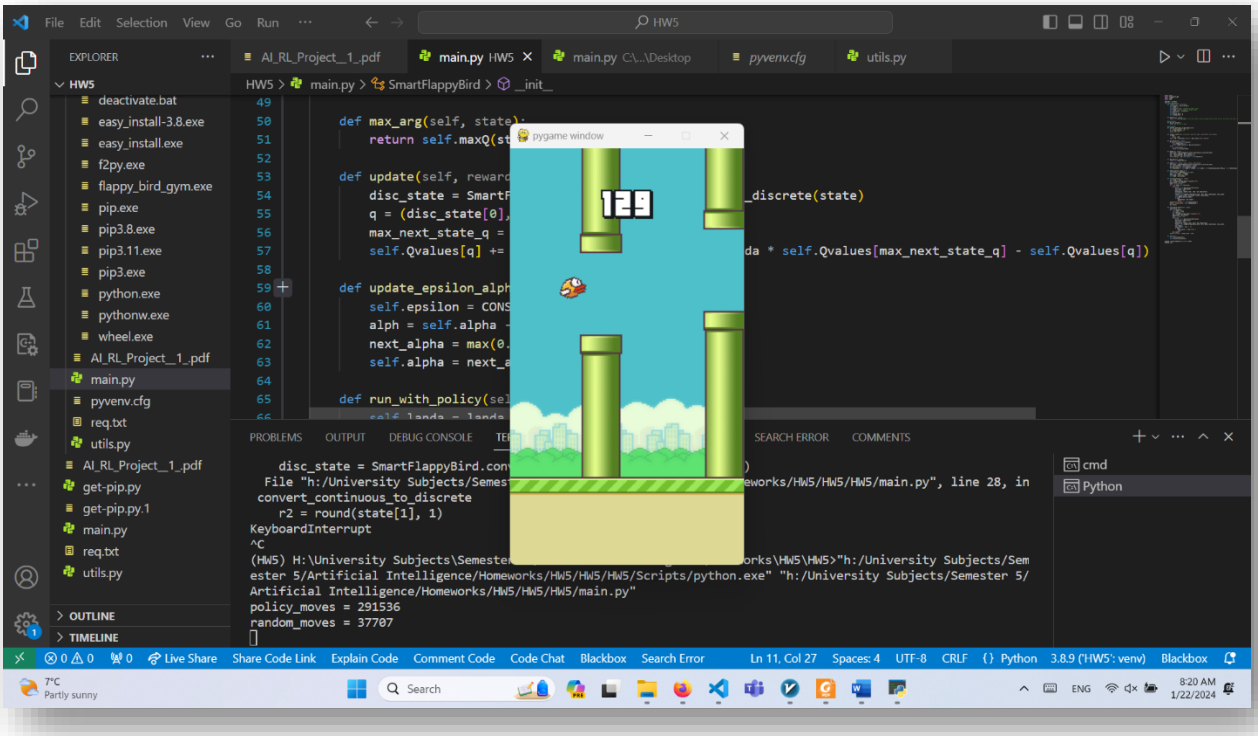
$$Q(s,a) = (1-\alpha)\cdot Q(s,a) + \alpha \cdot (r + \gamma \cdot \max_{a'} Q(s',a'))$$

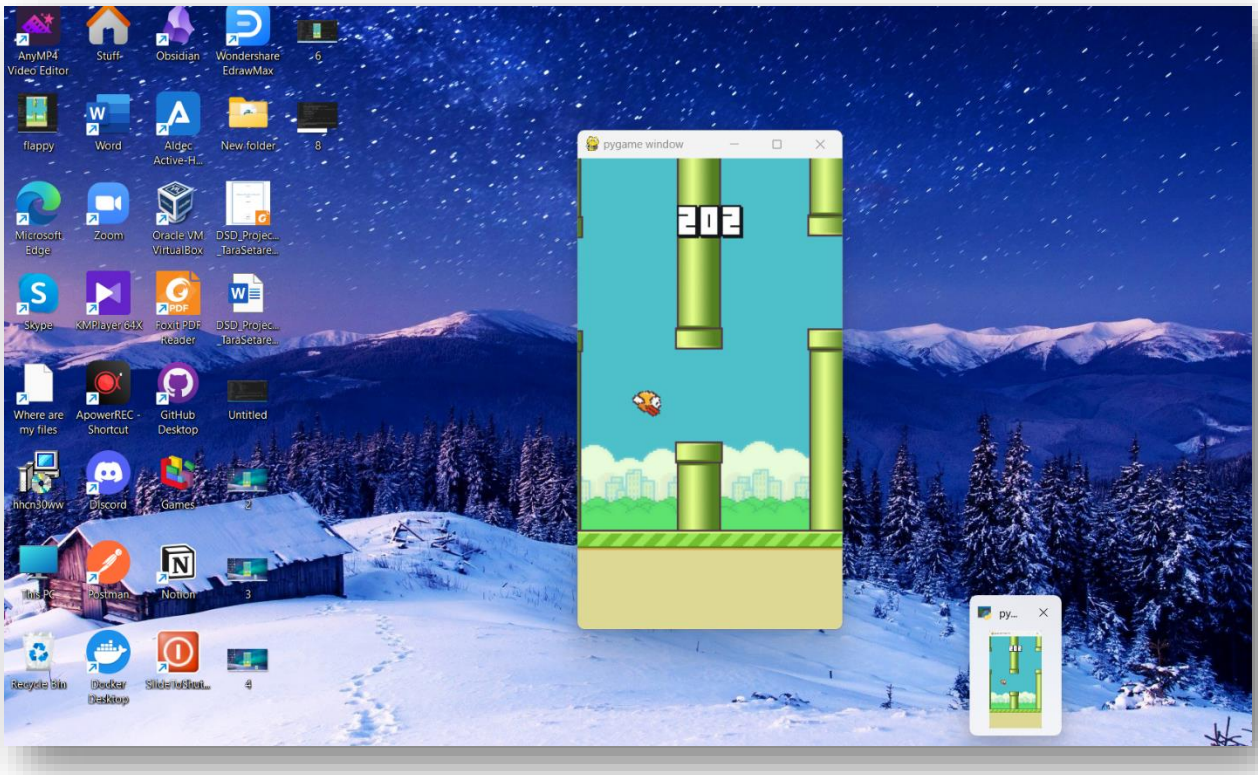
که α در اینجا نرخ یادگیری، و λ نرخ تخفیف می‌باشد.

- **update_epsilon_alpha**: بروزرسانی مقدار epsilon و alpha در هر مرحله.

مقدار اپسیلون به ازای هر حرکت با سیاست ۰.۹۹۹۹۷۳ برابر شده و مقدار آلفا هر بار ۰.۰۰۰۰۰۶ کاهش می‌یابد.

❍ اجرای کد:


pygame window
129
def max_arg(self, state):
return self.maxQ(state)[2]
def update(self, reward, state, action, next_state):
disc_state = SmartFlappyBird.convert_continuous_to_discrete(state)
q = (disc_state[0], disc_state[1], action)
max_next_state_q = self.maxQ(next_state)
self.Qvalues[q] += self.alpha * (reward + self.landa * self.Qvalues[max_next_state_q] - self.Qvalues[q])
def update_epsilon_alpha(self):
self.epsilon = CONSTANT ** self.move
alph = self.alpha - 0.000005
next_alpha = max(0.01, alph)
self.alpha = next_alpha
def run_with_policy(self, landa):
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS POLYGLOT NOTEBOOK SEARCH ERROR COMMENTS
convert_continuous_to_discrete
r2 = round(state[1], 1)
KeyboardInterrupt
^C
(HW5) H:\University Subjects\Semester 5\Artificial Intelligence\Homeworks\HW5\HW5>"h:/University Subjects/Semester 5/Artificial Intelligence/Homeworks/HW5/HW5/HW5/Scripts/python.exe" "h:/University Subjects/Semester 5/Artificial Intelligence/Homeworks/HW5/HW5/HW5/main.py"
policy_moves = 291536
random_moves = 37707
average = 74.0


AnyMP4 Video Editor
Stuff
Obsidian
Wondershare EdrawMax
6
flappy
Word
Aldec Active-H...
New folder
8
Microsoft Edge
Zoom
Oracle VM VirtualBox
DSD_Projec... _TaraSetare...
Skype
KMPlayer 64X
Foxit PDF Reader
DSD_Projec... _TaraSetare...
Where are my files
ApowerREC - Shortcut
GitHub Desktop
Untitled
hhcn30ww
Discord
Games
2
This PC
Postman
Notion
3
Recycle Bin
Docker Desktop
Slide To Shut...
4
pygame window
202
py...