**به نام خدا**

**تمرین سری 5 یادگیری ماشین**

**محمد میرزانژاد – 830596027**

* **سیاست بهینه :**

**با طول 56**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |  | **W** | |  |  |  |  |  |  |  |  |  |
|  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |  | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 |
| 0 | **S** |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | **G** |
| 1 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 2 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 3 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 4 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 5 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 6 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 7 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 8 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 9 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 10 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 11 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 12 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 13 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 14 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 15 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 16 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 17 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 18 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| 19 |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

**W= Wall**

**S=Start point**

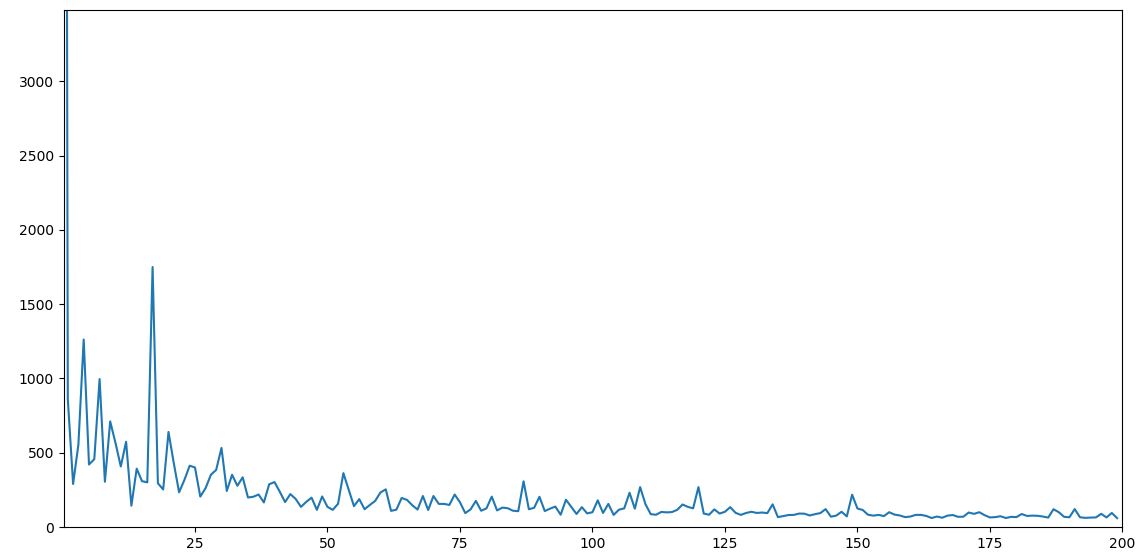
**G= Goal**

**Selected path:**

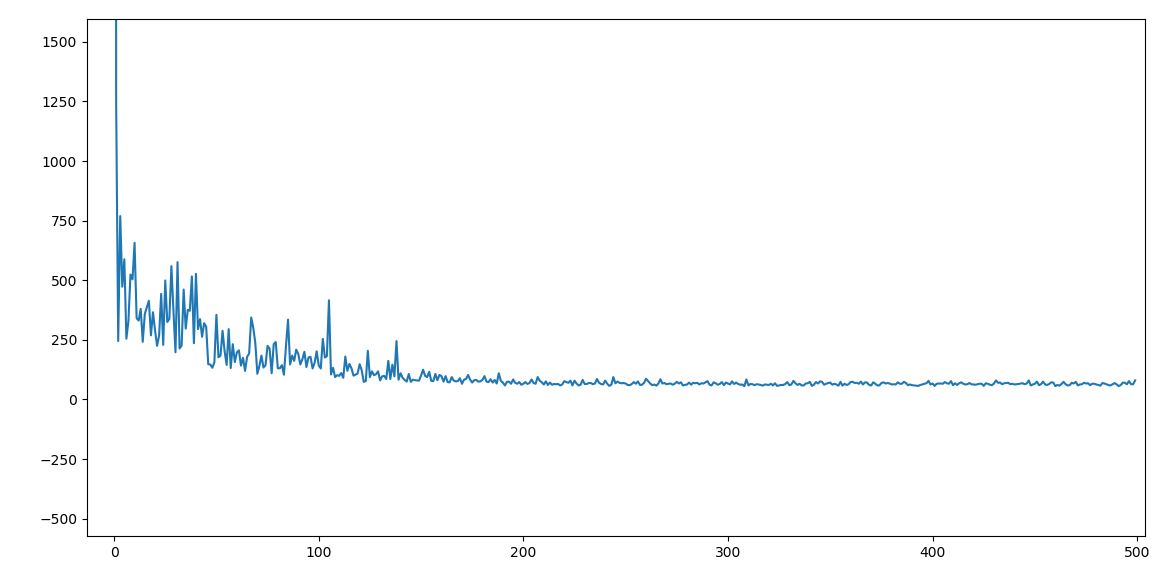
**[(0, 0), (1, 0), (2, 0), (3, 0), (3, 1), (3, 2), (3, 3), (3, 4), (3, 5), (3, 6), (3, 7), (3, 8), (4, 8), (5, 8), (6, 8), (6, 9), (7, 9), (8, 9), (9, 9), (10, 9), (11, 9), (12, 9), (13, 9), (14, 9), (15, 9), (16, 9), (17, 9), (18, 9), (18, 10), (17, 10), (16, 10), (15, 10), (14, 10), (13, 10), (13, 11), (12, 11), (11, 11), (10, 11), (10, 12), (10, 13), (9, 13), (8, 13), (7, 13), (6, 13), (5, 13), (4, 13), (4, 14), (3, 14), (3, 15), (2, 15), (1, 15), (0, 15), (0, 16), (0, 17), (0, 18), (0, 19)]**

**len : 56**

* **نمودار تغییر طول اپیزودها با افزایش تجربه عامل :**

****

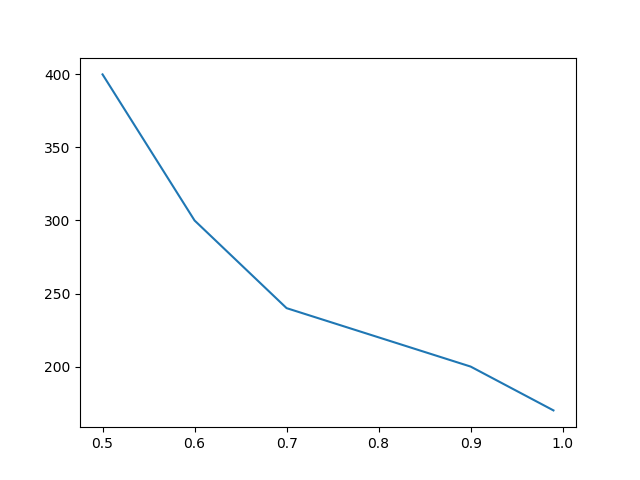
هر چه تعداد اپیزودها پیش میرود به تبع آن مقادیر جدول Q آپدیت میشود و عامل میتواند انتخاب های بهینه تری انجام داده و گامهای کمتری رابه سمت هدف طی کند، اما در این بین هر چند روند کلی تعداد گامهای هر اپیزود به صورت نزولی پیش میرود اما همانطور که در نمودار قابل مشاهده است نوسانهایی وجود دارد یعنی تعداد گامهای اپیزود بالا و پایین میرود که این امر به دلیل احتمال 15 درصدی در نظر گرفته شده برای انتخاب حرکت هایی غیر از حرکت بهینه ( یعنی مقدار Q بیشتر ) جهت رعایت اصل Exploaration و آپدیت تمام مسیرهای جدول است که در فاز آموزش مد نظر قرار میگیرد. این نمودار برای نرخ یادگیری 0.8 و 200 تکرار رسم شده است که منجر به تولید مسیر بهینه با طول 56 شده است . البته در صورت اجرای چند باره با توجه به احتمال انتخاب تصادفی مسیر ممکن است به ازای تعداد تکرار بیشتر ( حدودا بیشتر از 10 یا 20 ) به همگرایی برسیم و یا طول مسیر بهینه بیشتر از 56 باشد که در اجراهای انجام شده مشاهده گردید.

****

**نمودار به ازای 500 اپیزود**

* **تاثیر مقادیر مختلف گاما بر همگرایی الگوریتم :**

هر چه مقدار گاما به صفر نزدیک تر باشد میزان اپیزود های لازم برای همگرایی افزایش پیدا میکند . در واقع عامل بیشتر به پاداش های آنی متمرکز میشود و هر چه مقدار گاما به یک نزدیک شود عامل به پاداش های با delay وزن بیشتری اختصاص میدهد. در زیر نمونه هایی از مقادیر مختلف گاما و تعداد اپیزود ها آمده است .



|  |  |  |  |
| --- | --- | --- | --- |
| **گاما** | **تعداد اپیزود** | **تعداد کمترین حرکت در یک اپیزود** | **همگرایی** |
| **0.99** | **100** | **103** | **خیر** |
| **0.99** | **150** | **75** | **خیر** |
| **0.99** | **170** | **60** | **بله** |
| **0.9** | **170** | **59** | **خیر** |
| **0.9** | **200** | **58** | **بله** |
| **0.8** | **220** | **59** | **بله** |
| **0.7** | **240** | **58** | **بله** |
| **0.6** | **260** | **57** | **خیر** |
| **0.6** | **300** | **55** | **بله** |
| **0.5** | **380** | **57** | **خیر** |
| **0.5** | **400** | **55** | **بله** |
| **0.3** | **500** | **103** | **خیر** |