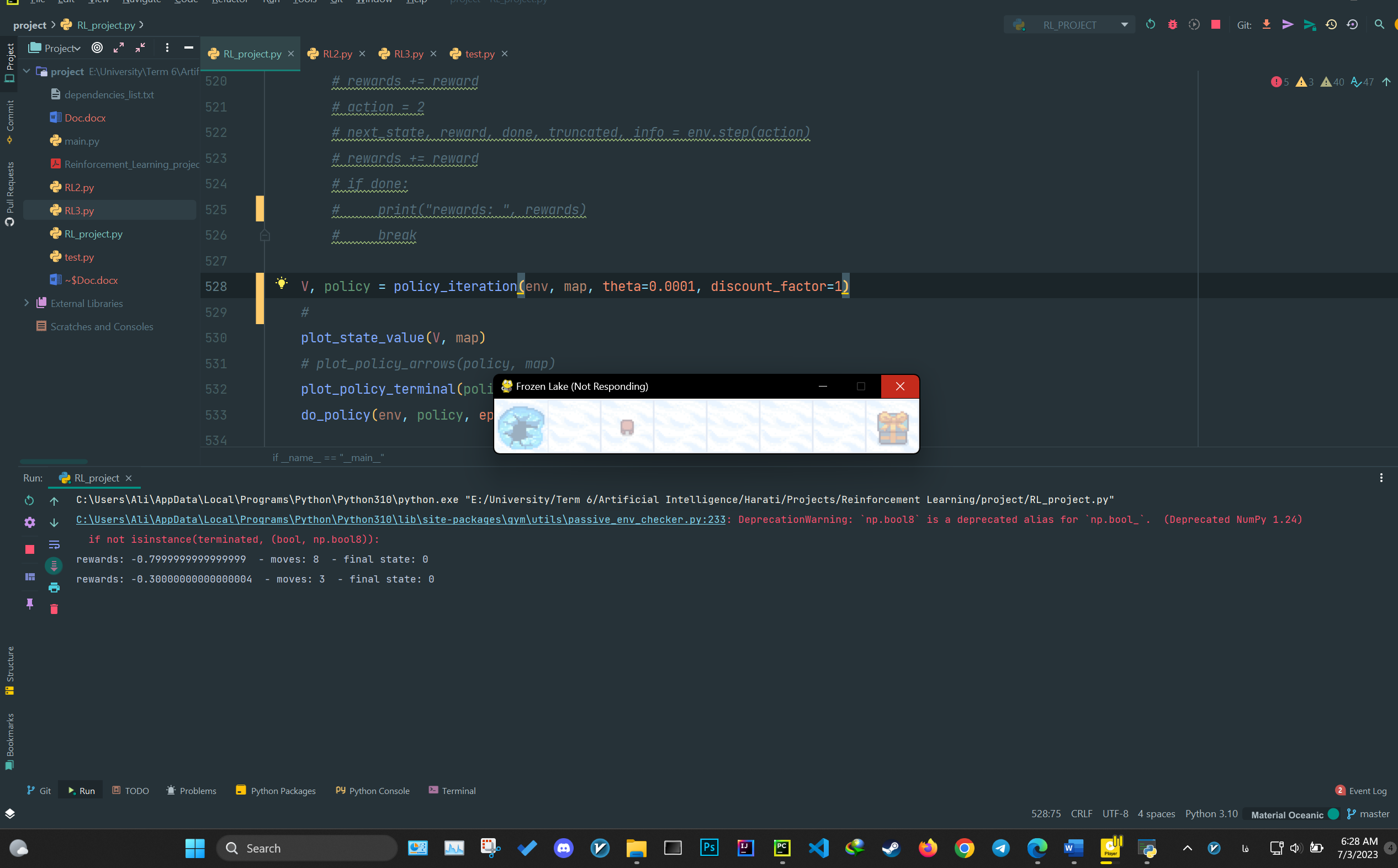
بخش دوم – ارزیابی و بررسی الگوریتم ها

پروژه RL

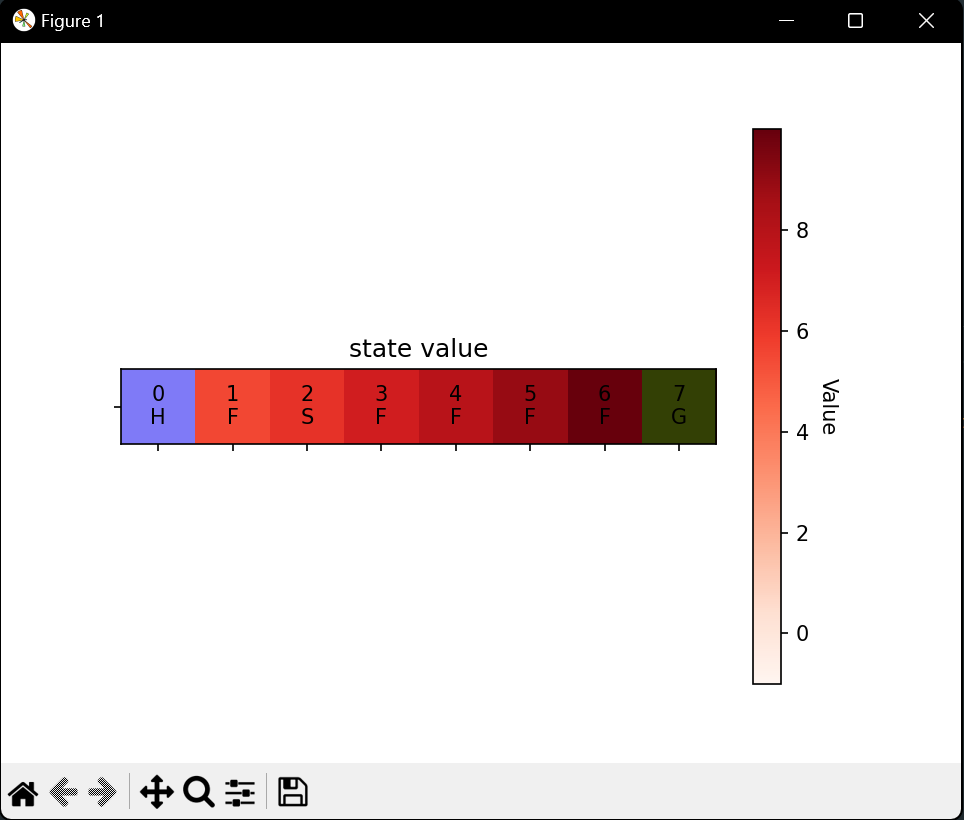
علی حمیدزاده 9912762529 – محمد افشاریان

سناریو 1:

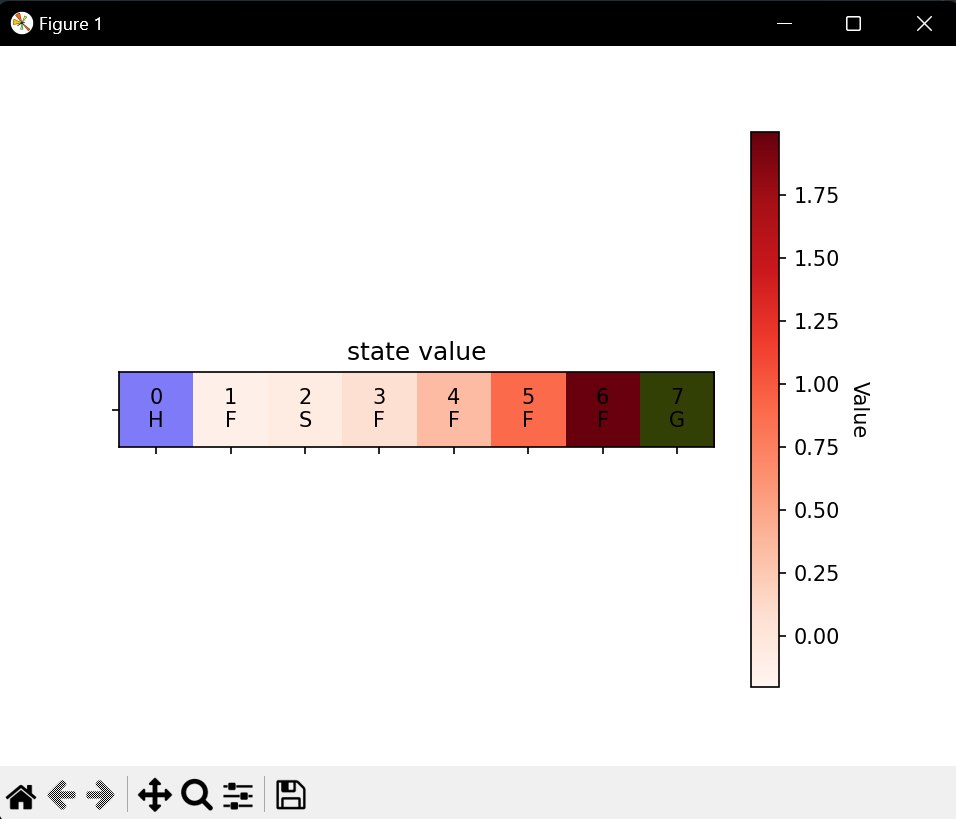
* با مقدار گاما = 1: کرش میکنه و نمودار state-value نمیده!



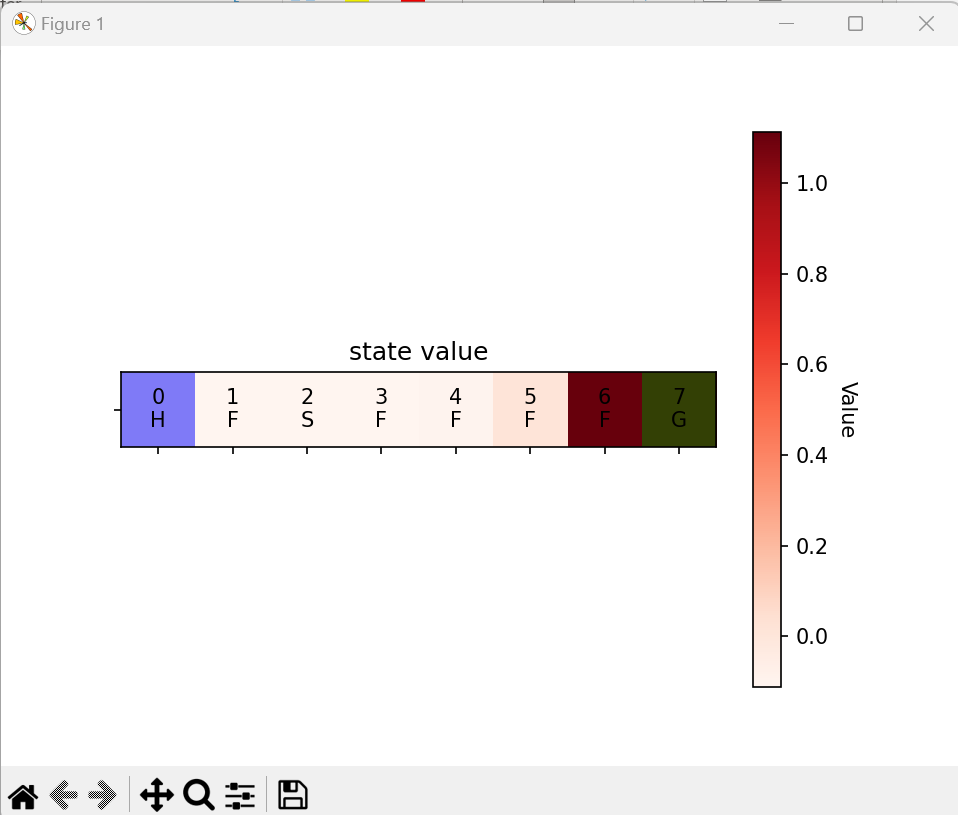
* با مقدار گاما = 0.9: در هنگام محاسبه تابع ارزش در فرآیند ارزیابی سیاست، ضریب تخفیف به پاداش‌های آینده اعمال می‌شود. این ضریب باعث کاهش ارزش پاداش‌های به‌دست آمده در زمان‌های آینده می‌شود. ضریب تخفیف صفر نشان‌دهنده این است که عامل فقط به پاداش‌های فوری توجه می‌کند و از پاداش‌های آینده غافل می‌شود. از سوی دیگر، **ضریب تخفیف برابر با یک به این معنی است که عامل پاداش‌های آینده را با پاداش‌های فوری به‌صورت یکسان ارزیابی می‌کند**.



* با مقدار گاما = 0.5: عامل پاداش های آینده را کمتر از ارزش پاداش های فوری (الان) ارزیابی میکند.



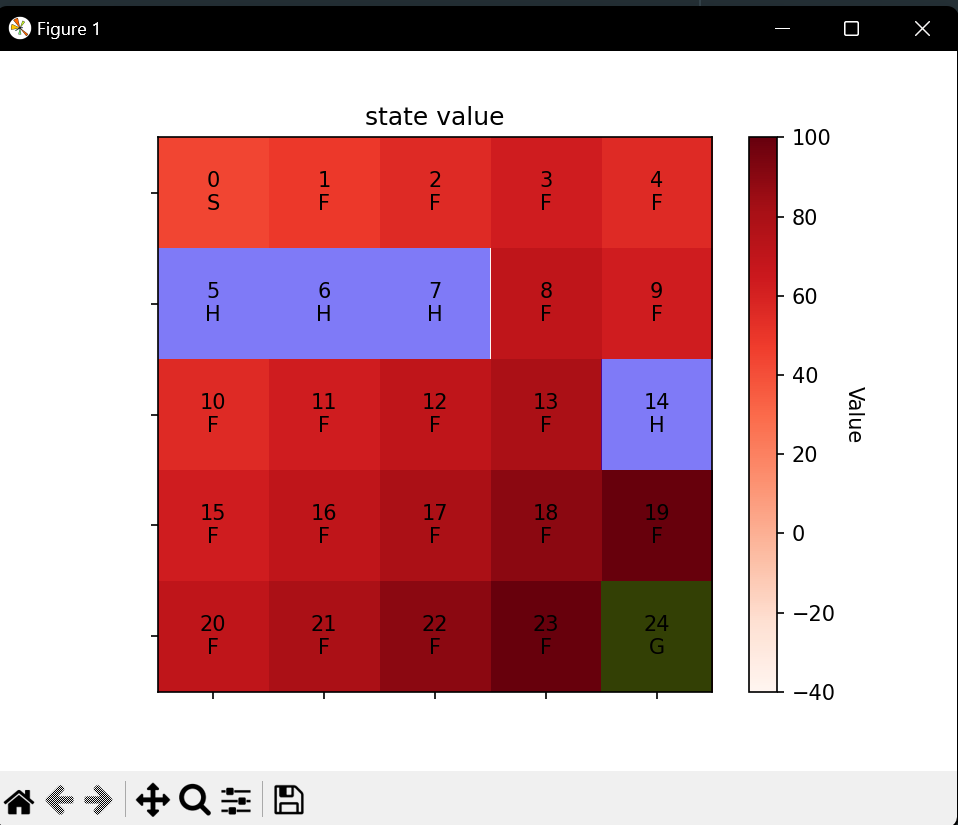
* مقدار گاما = 0.1: کمترین ارزش به گام های اینده تعلق میگیره



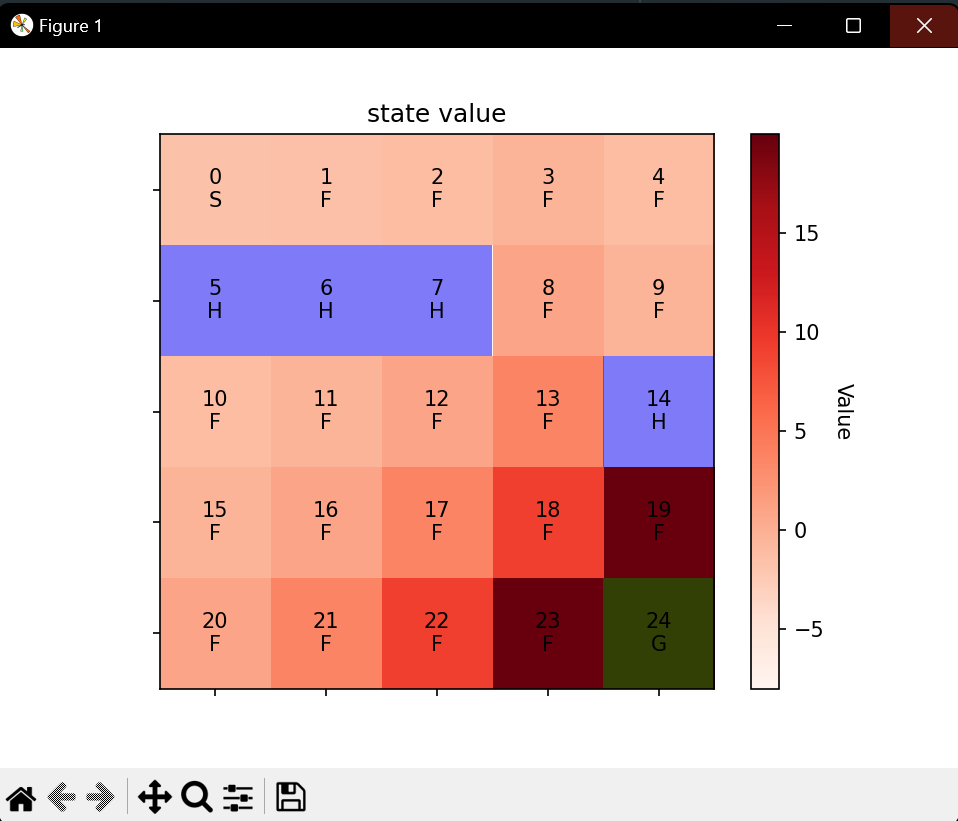
علت تفاوت بین نتایج، ارزشی هست که به گام های آینده در مقابل تجارب حال داده میشود.

سناریو 2:

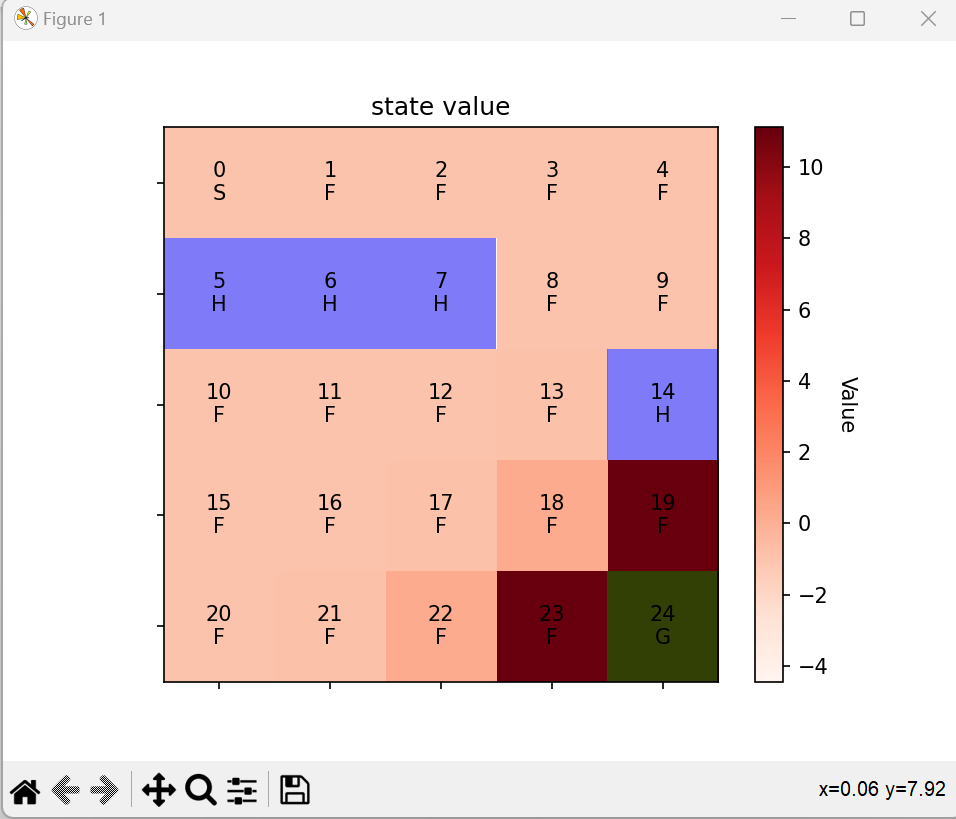
* گاما = 1: کرش میکنه
* گاما = 0.9:



* گاما = 0.5:



* گاما = 0.1: علت تفاوت در مقادیر همانطور که گفته شد به دلیل تفاوت در میزان ارزش گذاری گام های آینده در مقابل گام های حال میباشد. به نظر میاد که گاما = 0.5 میتونه نسبتا بهینه عمل کنه.



سناریو 3: