

ما یک کلاس wampus تعریف می کنیم. در آن حرکت ها و اشیا را به صورت ثابت تعریف میکنیم تا اگر نیازی به تغییر شد راحت در کل کد عوض شود و نیازی به تغییر هر بخش نباشد.

ما یک طول و عرض داریم برای جهانمان که چند در چند بودن جهان را تعریف میکند همچنین مکان های چاه را به عنوان مانع و خوناشام ها و خودمان و گنج را میگیریم.

جهان را تشکیل میدهیم پس از آن Q-learning را روی آن فرا میخوانیم به این صورت که هر اپیزود و هر حرکت مستقل بررسی می شوند و در جدول طبق فرمول آورده شده آپدیت می شوند.

بایستی دقت شود که Explore در مقابل Exploit مان مقدار کافی داشته باشد.

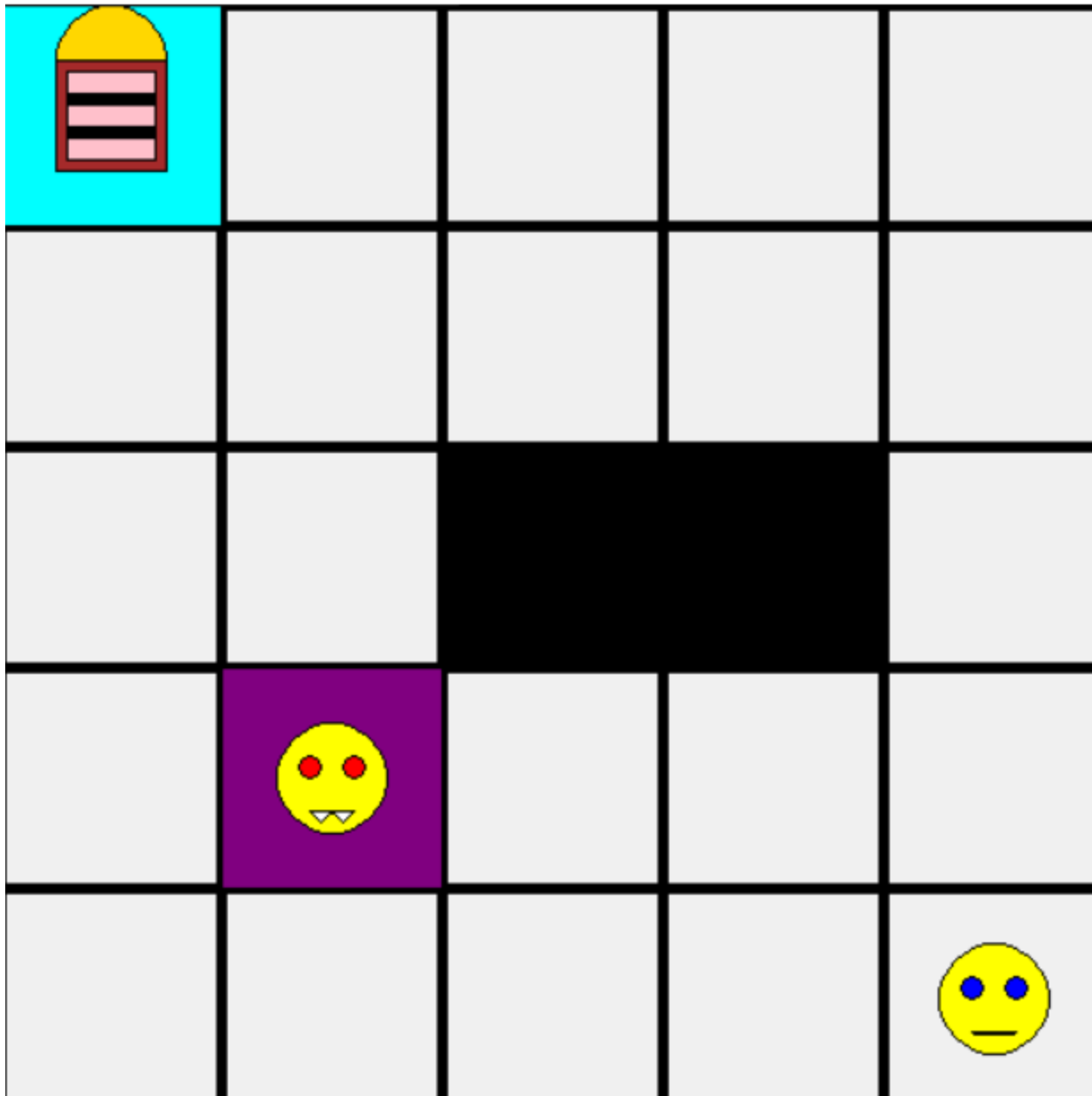
در آخر تابع تست می نویسیم که بر اساس جدول آموزش دیده حرکت کند و حرکت ها را به ما بدهد.

در آخر آنها را نمایش می دهیم:

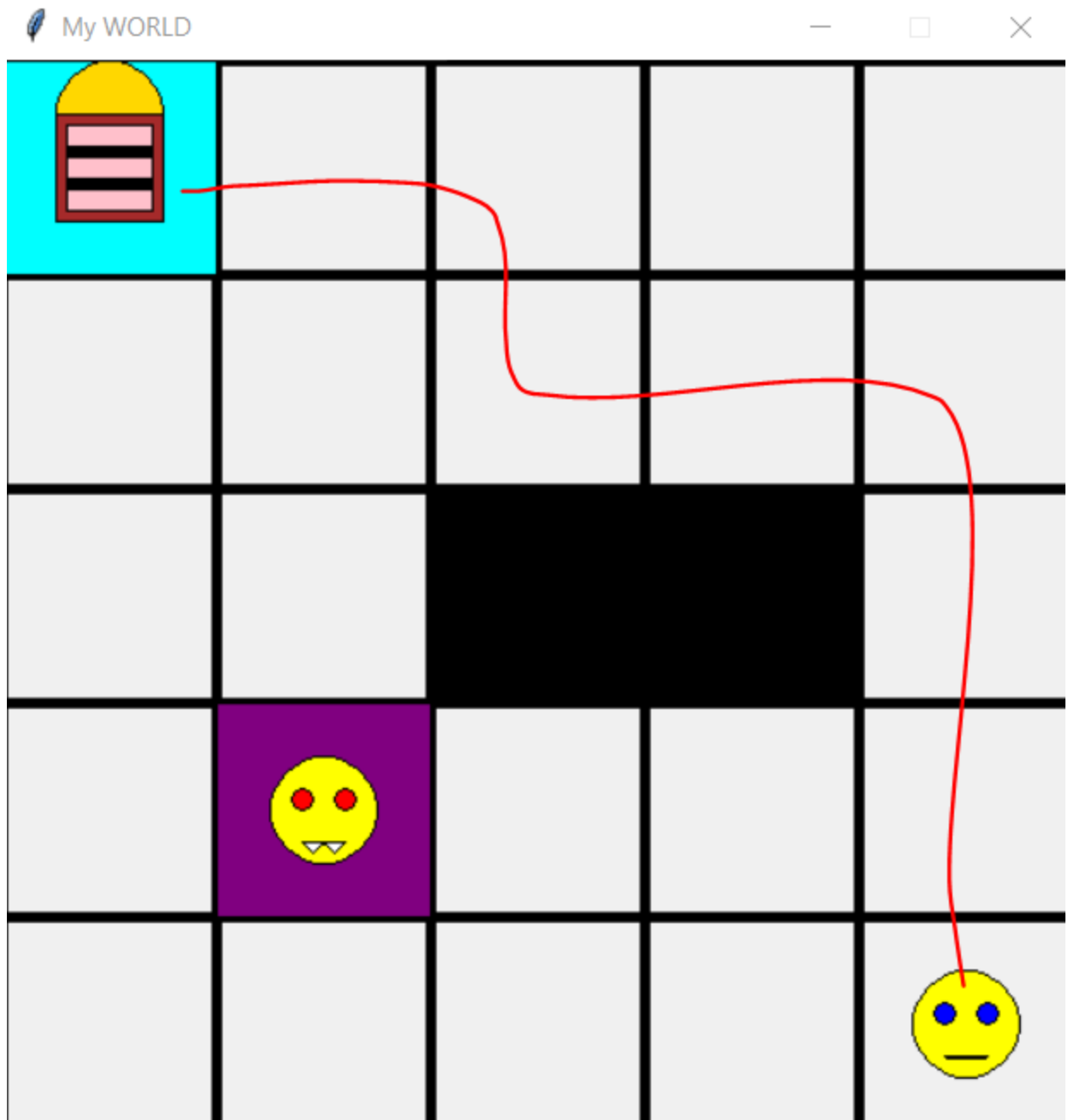
به عنوان مثال چند اجرای زیر از برنامه را داریم:

```
run(width, height, obstacles, me, vampires, gold, 500, 0)
```

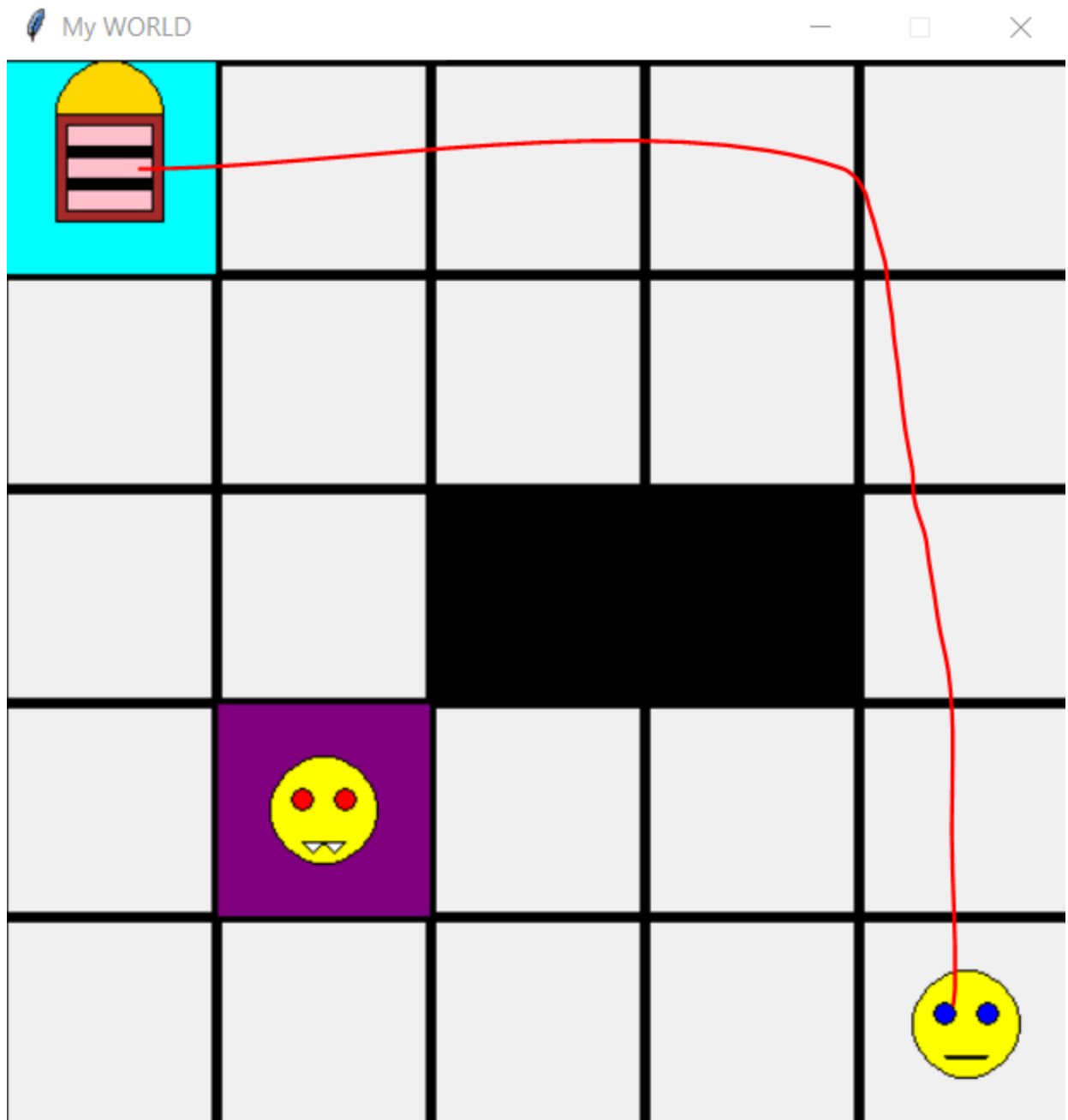
که به معنی مقدار اولیه q صفر و با ۵۰۰ بار آموزش است که حالت اولیه خانه ها به این صورت است



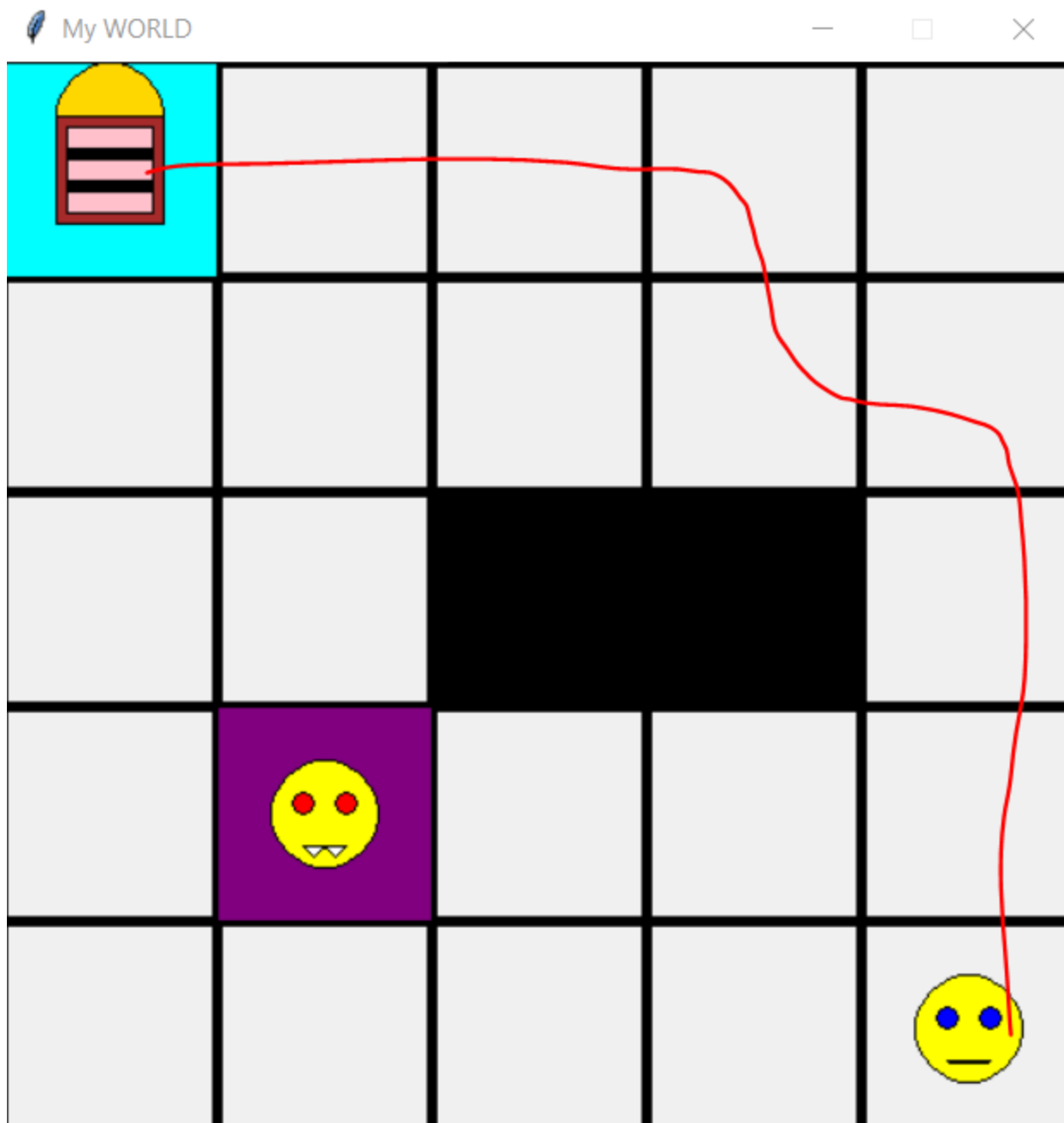
برای حالت استارت •



برای حالت استارت ۱

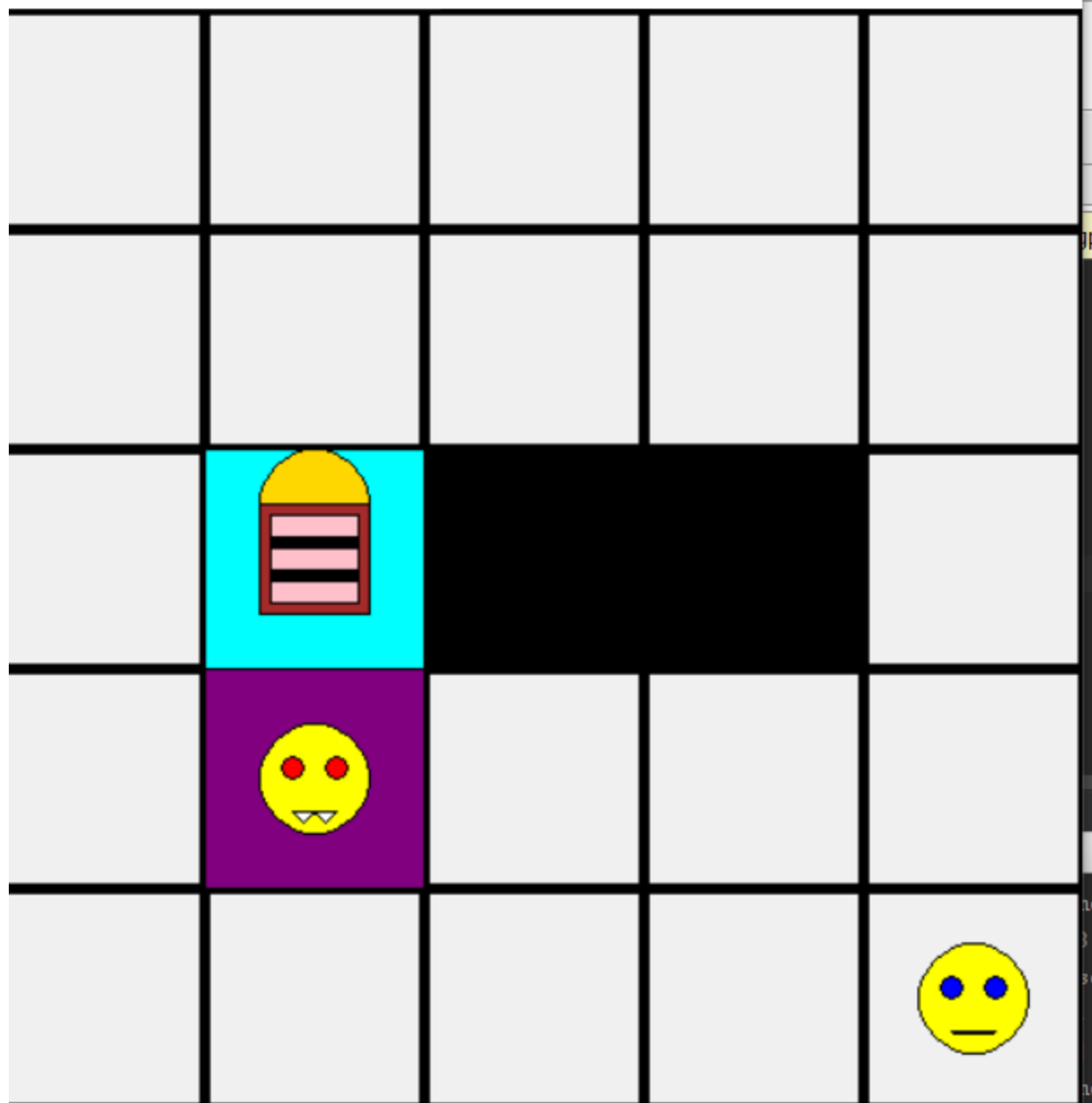


و برای ۱- این می باشد

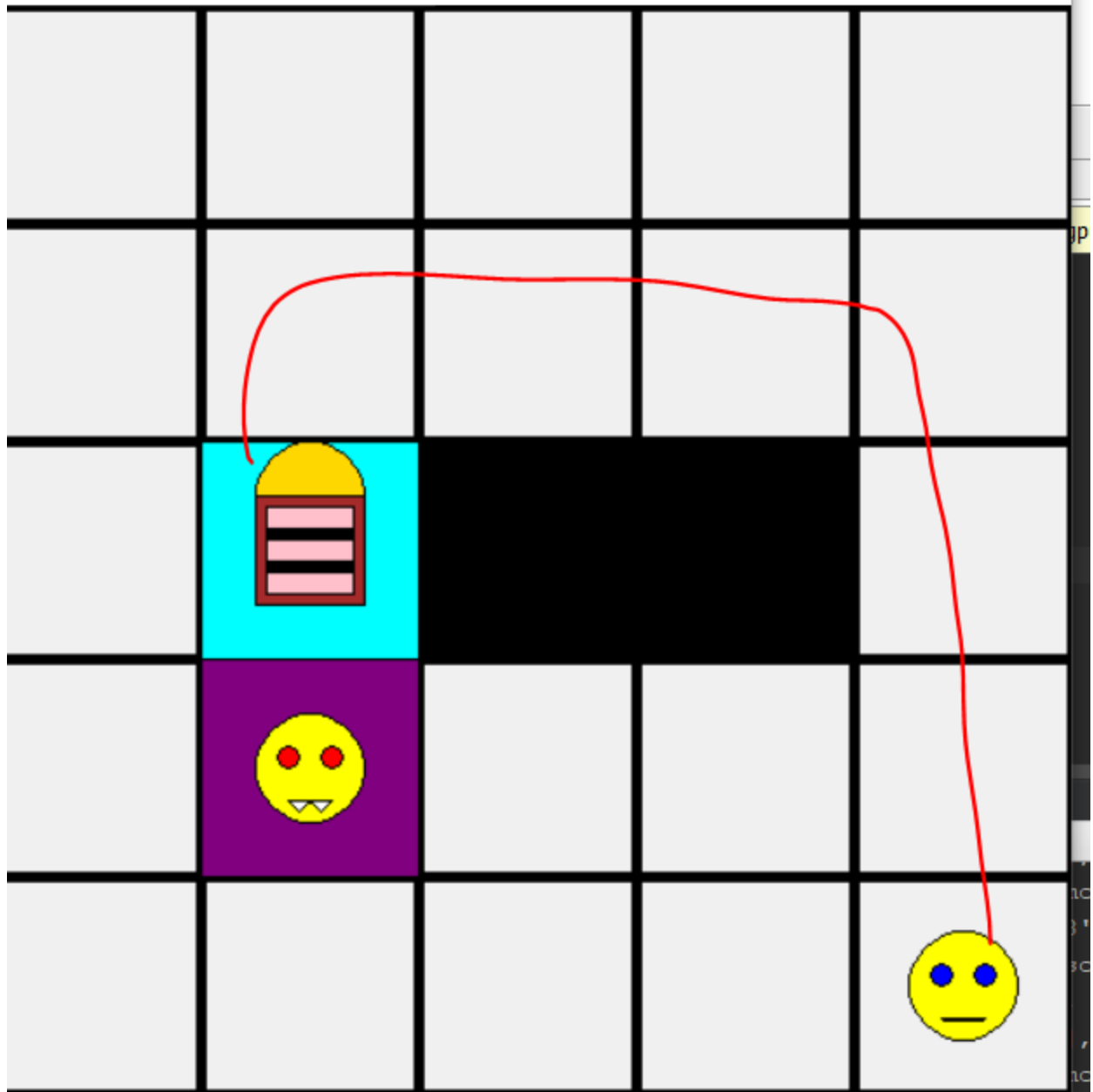


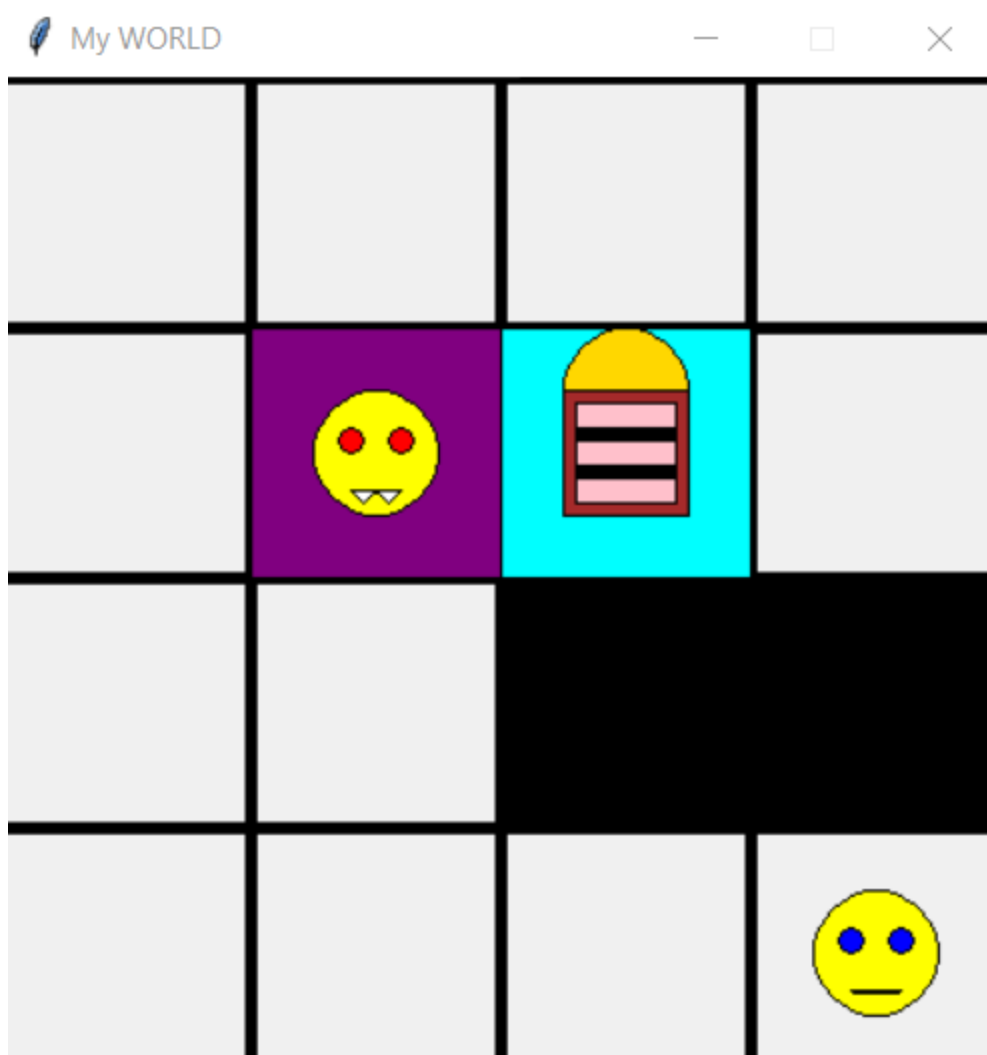
و البته جدول Q های یاد گرفته شده نیز متفاوت خواهند بود

برای ساینز دیگری از مدل می شود:



و حل آن می شود :





وچندین مدل دیگر نیز تست شده اند.