سپهر مقیسه

۹۸۳۱۱۰۳

تکلیف سوم

۱-الف)مقدار مورد نظر شروع در گره S و به صورت بهینه حرکت کردن

ب)در گره S بهینه ترین حرکت را انتخاب کنیم

ج)بله میتوان با فرمول زیر:

$$\pi^*(s) = \arg\max_a \sum_{s'} T(s,a,s')[R(s,a,s') + \gamma V^*(s')]$$

میتوان به POLICY هر S با کمک مقدار ها و پاداش های STATE های قبلی رسید.

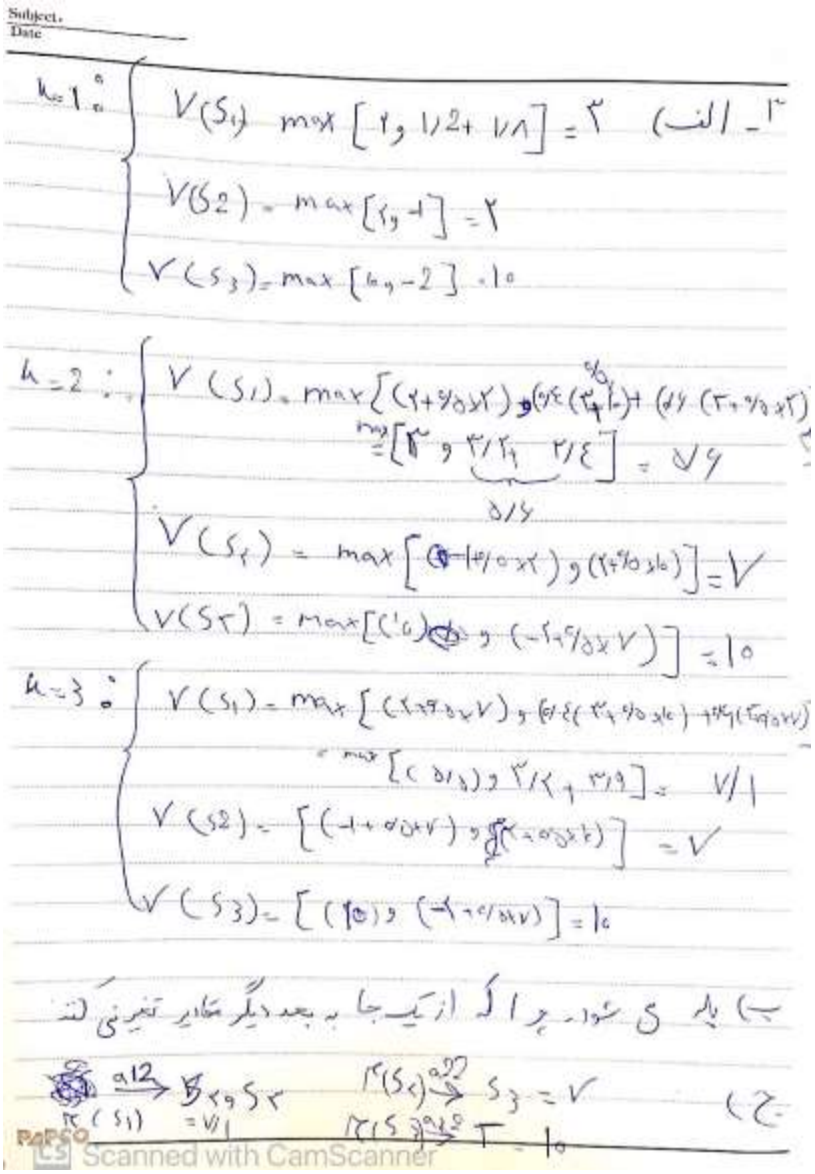
۲-الف)درست- چرا که نیاز به یک policy داریم که در هر لحظه تصمیم درست را لحاظ کنئ

ب)درست-چرا که حضور مقدار v* افزایش پیدا میکند

ج)نادرست-به هزینه transition نیز بستگی دارد

د)بله چرا که policy=max arg q است و پالیسی در هر state بهینه ترین حرکت را میدهد.

۳-



۴- الف) ۱- تربیت گربه در خانه

محیط:خانه

عامل:گربه

۲-app recommendation : محیط :برنامه

عامل:استفاده کننده برنامه

۳-بازی شطرنج : محیط : خانه های شطرنج

عامل : بازیکن ها

ب)۱- model free - چرا که هر گربه ممکن است شخصیت متفاوتی داشته باشد با توجه به تجربه های مربی های دیگر گربه ها

۲-model based- چرا که سلایق افرادی که ویدیو های مشترک زیادی میبینند تقریبا یکسان است و میتوان بر اساس تجربیات پیشنهاد شده به یک فرد همان هارا به دیگری پیشنهاد داد

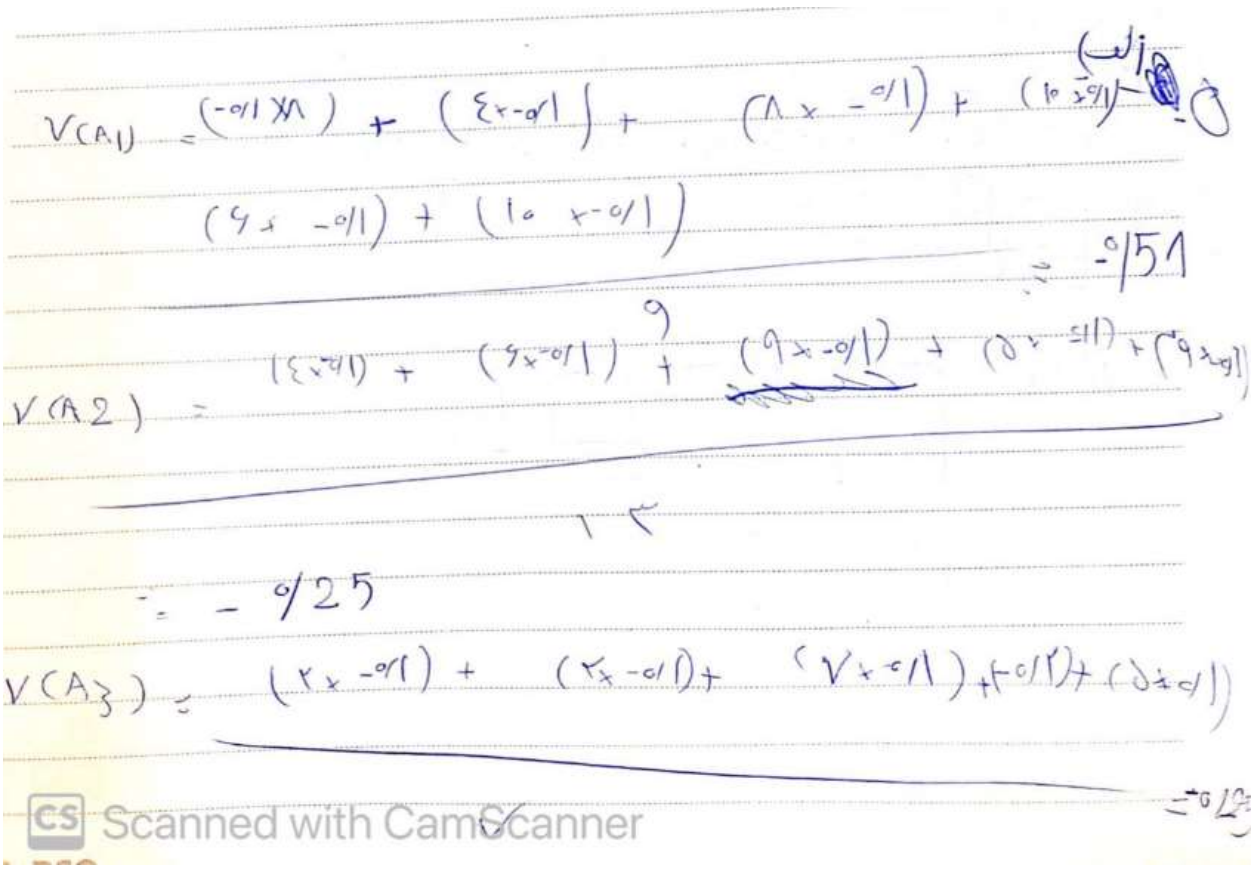
۳- model free : چرا که شیوه بازی هر فرد متفاوت است و نمیتوان بر اساس شناخت قبلی عمل کرد

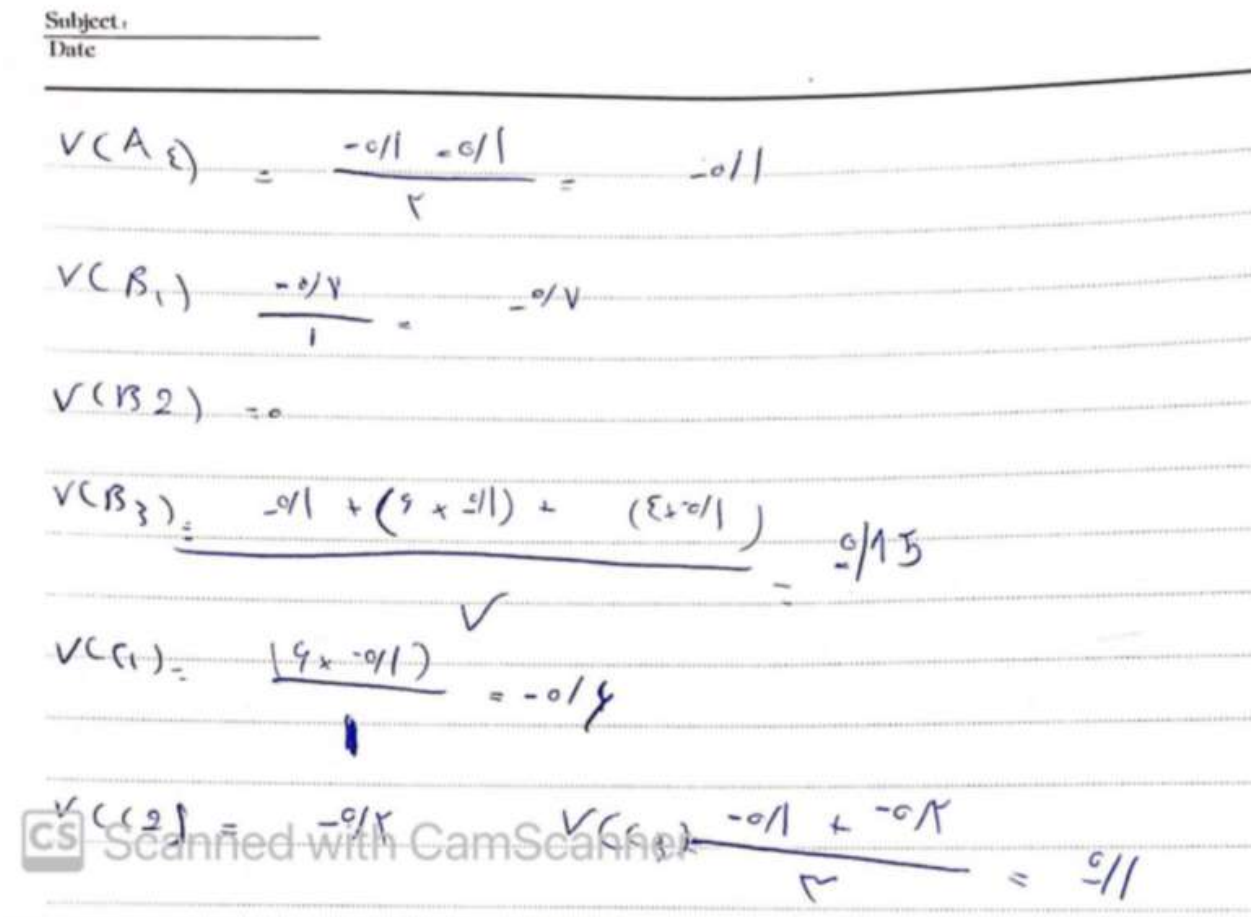
ج) ۱-استفاده از جایزه جدید برای تربیت گربه و یا استفاده از همان جایزه قبلی

۲-پیشنهاد کانال های جدید تر به فرد یا پیشنهاد همان کانال های قبلی

۳-بازی با استفاده از شیوه قبلی و سعی در کیش کردن و یا امتحان شیوه جدید مانند خوردن بعضی مهره های طرف در اول

۵-





۶-در policy learner - off policy بدون توجه به action ای که عامل انجام میدهد سیاست بهینه را یاد میگیرد- SARSA (state-action-reward-state-action) - در این الگوریتم سیاستی که برای اپدیت ارزش هر حالت به کار میرود مانند سیاستی هست که برای عمل حرکت انجام میشود و این با q learning که این دو مقدار یکسان نیس تفاوت دارد.

۷-الف)فاصله خودرو تا چراغ قرمز – زیرا محل عبور عابرین پیاده در زیر چراغ قرمز واقع شده است.

فاصله آن محل از محل های پیاده رو – چرا که هرچه فاصله با محل های پیاده رو کمتر باشد احتمال وجود عابرین پیاده بیشتر است .

فاصله تا پل هوایی – هرچه فاصله بیشتر باشد احتمال رد شدن محل پیاده از وسط خیابان بیشتر است

هر کدام از این ویژگی ها وزن خاص خود را دارند

ب)محل پشت چراغ قرمز و محل های دورتر از چراغ