

سید ساره خیری

تمرین کلاس 6

۹۹۴۱۱۲۲۱

(الف) در تئوری بازیها سیاست ES-ML را بیان کنید.

$$q_{jk}(s, \pi_k(s)) \geq u_{jk}(s)$$

$q_{jk}(s, \pi_k(s))$ بیانگر جمع ارزش حالت و عمل است که سیاست π_k از آن بیان می‌دهد و $u_{jk}(s)$ در صورتی که

در تمام براس π_k که آنجا انجام می‌دهد و داده نیز براس π_k باشد می‌گردد.

(ب) $q_{jk}(s)$ نیز جمع ارزش حالت است که سیاست π_k از آن بیان می‌دهد و $u_{jk}(s)$ در آن $q_{jk}(s)$ برابر است

آنجا که $q_{jk}(s)$ و $u_{jk}(s)$ برابر می‌شوند.

نامهای بالا به دلیل زیر می‌باشد:

۱- به علت انجام ES، در هر $state$ ، $q_{jk}(s)$ و $u_{jk}(s)$ ممکن است متفاوت باشند.

۲- $q_{jk}(s)$ بیانگر اهدافی است که می‌تواند $q_{jk}(s)$ بیشترین مقدار ممکن را در $state$ داشته باشد.

۳- به سبب تشریف، ارزش این عمل باید بیشترین $q_{jk}(s)$ باشد که $q_{jk}(s)$ در آن $state$ حاصل می‌شود.

۴- $q_{jk}(s)$ نیز بیانگر برتری $q_{jk}(s)$ است که حاصل انجام $q_{jk}(s)$ در $state$ می‌باشد.

۵- در $q_{jk}(s)$ ارزش برتری عمل $q_{jk}(s)$ باید بیشترین $q_{jk}(s)$ باشد که $q_{jk}(s)$ در آن $state$ حاصل می‌شود.

بنابراین رابطه نیز موجود است که $q_{jk}(s)$ به $u_{jk}(s)$ از لحاظ مثبت است:

معادله $q_{jk}(s)$ به $u_{jk}(s)$ به صورت زیر است:

$$u_{jk}(s) = \sum_a \pi_k(a|s) \sum_{s', r} P(s', r | s, a) [r + \gamma u_{jk}(s', a)]$$

برای $q_{jk}(s)$ نیز داریم:

$$q_{jk}(s, a) = \sum_{s', r} P(s', r | s, a) [r + \gamma u_{jk}(s', a)]$$

همچنین ترتیب دیگری برای معادله $q_{jk}(s)$ می‌توان نوشت:

همچنین $q_{jk}(s)$ را می‌توان به صورت زیر نوشت:

ادانہ عربی لکھنا ۶

از کسی زخم برید و ۱۰۰ است که به شما زوال عیال است.

نرم نرم $\frac{E}{1000}$ که جایگزین دیتا جدول اعمال آنگو به میان است.

4- دانشگاه امام شریع محمد باقر

میتا جبرری در ۱. ۷۴ عیا بقی .

بروندہ جوتے ہوا۔ کیا اس سے یہ دلیل ہو سکتی ہے کہ تمام انسانوں کی زندگی

یہ ایک نامور مسافر ہے۔

تجزیہ و تحلیل کے لیے Determination of the effect of the

Controlled and ϵ -greedy \rightarrow ϵ -soft policies, ϵ -greedy and ϵ -soft