

عنوان مقاله:

بهبود بیش برآورد یادگیری تقویتی عمیق برون-سیاست با شبکه به ر وزرسان انتخابی عمیق

محل انتشار:

بیست و هفتمین کنفرانس بین المللی کامپیوتر انجمن کامپیوتر ایران (سال: 1400)

تعداد صفحات اصل مقاله: 6

نویسندگان:

علی صالحی - گروه مهندسی کامپیوتر، قطب علمی رایانش نرم و پردازش هوشمند اطلاعات، دانشگاه فردوسی مشهد

محمدرضا اکبرزاده توتونچی - گروه مهندسی برق، قطب علمی رایانش نرم و پردازش هوشمند اطلاعات، دانشگاه فردوسی مشهد

علیرضا روحانیمنش - گروه مهندسی برق، دانشگاه نیشابور،

خلاصه مقاله:

شبکه های عصبی عمیق-Q (DQN) نمونه ای از یادگیری تقویتی عمیق مبتنی بر یادگیری Q هستند که در آن تابع Q نمایانگر ارزش تمامی کنش های یک کارگزار در تمامی حالت های یک محیط است. یادگیری این شبکه به دلیل نگاه خوشبینانه کارگزار در تخمین حالت های آتی سبب شده است تا کارگزار مبتنی بر DQN ناپایدار و به همراه بیش برآورد ارزش باشد. راه حل پیشنهادی در این مقاله ارائه نمودن شبکه به روزرسان انتخابی عمیق است که در آن مشکل بیش برآورد ارزش تخمینی آینده بهبود یابد. در این رویکرد، نحوه به روزرسانی کارگزار در ابتدای یادگیری نگرش محتاطانه تری نسبت به تخمین ارزش ها دارد و با گذر زمان به رویکرد خوشبینانه یادگیری Q تغییر مسیر خواهد داد. این امر به آن جهت است که تخمین ارزش حالت های آتی در ابتدای یادگیری قابل اتکا نیست و استفاده از آن در محیط های با هزینه بالا روند یادگیری را ناپایدار میکند. نتایج حاصل برای دو محیط شبکه تنظیم کننده ژن و پاندول معکوس نشان میدهد که رهیافت پیشنهادی علاوه بر کاهش بیش برآورد تخمین ارزش، پاداش بیشتری را نسبت به DQN جمع آوری مینماید.

کلمات کلیدی:

بیش برآورد، یادگیری تقویتی عمیق، شبکه های عصبی عمیق-Q، برون-سیاست، شبکه به روزرسان انتخابی عمیق، پاندول معکوس، شبکه تنظیم کننده ژن، ۵۳P

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1452943>

