

7 : مهندسی برق و مهندسی کامپیوتر ایران تابستان 1389;8(2):97-106.

مدلی مبتنی بر آنтрوبی و اتوماتاهای یادگیر برای حل بازی های تصادفی

مصطفی محمد رضا بهروز*

* دانشکده مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه آزاد اسلامی قزوین، قزوین

بازی های غیر قطعی (تصادفی) به عنوان توسعه ای از فرآیندهای تصادفی مارکوف با چندین عامل در سیستم های چندعامله و مدل سازی آنها حائز اهمیت بوده و به عنوان چارچوبی مناسب در تحقیقات یادگیری تقویتی چندعامله به کار رفته اند. در حال حاضر اتوماتاهای یادگیر به عنوان ابزاری ارزشمند در طراحی الگوریتم های یادگیری چندعامله به کار رفته اند. در این مقاله مدلی مبتنی بر اتوماتای یادگیر و مفهوم آنтрوبی برای حل بازی های غیر قطعی و پیدا کردن سیاست بهینه در این بازی ها ارایه شده است. در مدل پیشنهادی به ازای هر عامل در هر حالت از محیط بازی یک اتوماتای یادگیر با ساختار متغیر از نوع S قرار داده شده است که اعمال شده است که اعمال بهینه را در هر حالت یاد می گیرند. تعداد اعمال هر اتوماتا با توجه به همسایگان مجاور هر حالت تعیین شده و ترکیب اعمال اتوماتها حالت بعدی محیط را انتخاب می کند. در مدل پیشنهادی از آنтрوبی بردار احتمالات اتوماتای یادگیر حالت جدید برای کمک به پاداش دهنده اتوماتها و بهبود یادگیری استفاده شده است. برای بررسی و تحلیل رفتار الگوریتم یادگیری پارامتری به نام آنтрوبی کلی تعریف گردیده که میزان همگرایی را در الگوریتم یادگیری بیان می کند. در نهایت الگوریتمی اصلاح یافته با ایجاد تعادل بین جستجو و استناد بر تجربیات پیشنهاد شده است. نتایج آزمایش ها نشان می دهد الگوریتم ارایه شده از کارایی مناسبی از هر دو جنبه هزینه و سرعت رسیدن به راه حل بهینه برخوردار است.

کلید واژه: آنтрوبی، اتوماتاهای یادگیر، بازی های تصادفی، سیستم های چندعامله



نسخه قابل چاپ