



7 : مهندسي برق و مهندسي کامپيوتر ايران تابستان 1389; 8(2):97-106.

مدلي مبتني بر آنتروپي و اتوماتاهاي يادگير براي حل بازي هاي تصادفي

معصومي بهروز*, ميدي محمدرضا

* دانشكده مهندسي كامپيوتر و فناوري اطلاعات، دانشگاه آزاد اسلامي قزوین، قزوین

بازي هاي غير قطعي (تصادفي) به عنوان توسعه اي از فرآيندهاي تصادفي مارکوف با چندین عامل در سيستم هاي چندعامله و مدل سازي آنها حایز اهمیت بوده و به عنوان چارچوبي مناسب در تحقیقات یادگيري تقويتی چندعامله به کار رفته اند. در حال حاضر اتوماتاهاي يادگير به عنوان ابزاري ارزشمند در طراحي الگوريتم هاي یادگيري چندعامله به کار رفته اند. در این مقاله مدلي مبتني بر اتوماتاي يادگير و مفهوم آنتروپي براي حل بازي هاي غير قطعي و پيدا کردن سياست بهينه در این بازي ها ارايه شده است. در مدل پیشنهادي به ازاي هر عامل در هر حالت از محیط بازي یک اتوماتاي يادگير با ساختار متغير از نوع S قرار داده شده است که اعمال بهينه را در هر حالت یاد می گیرند. تعداد اعمال هر اتوماتا با توجه به همسایگان مجاور هر حالت تعیین شده و ترکیب اعمال اتوماتاها حالت بعدي محیط را انتخاب می کند. در مدل پیشنهادي از آنتروپي بردار احتمالات اتوماتاي يادگير حالت جدید براي کمک به پاداش دهی اتوماتاها و بهبود یادگيري استفاده شده است. براي بررسی و تحلیل رفتار الگوريتم یادگيري پارامتری به نام آنتروپي کلي تعريف گردیده که میزان همگرایی را در الگوريتم یادگيري بیان می کند. در نهایت الگوريتمی اصلاح یافته با ایجاد تعادل بین جستجو و استناد بر تجربیات پیشنهاد شده است. نتایج آزمایش ها نشان می دهد الگوريتم ارايه شده از کارايي مناسبی از هر دو جنبه هزینه و سرعت رسیدن به راه حل بهينه برخوردار است.

کلید واژه: آنتروپي، اتوماتاهاي يادگير، بازي هاي تصادفي، سيستم هاي چندعامله



نسخه قابل چاپ