



حل فرآیندهای تصمیم‌گیری مارکف چندعامله با استفاده از آنواتهای یادگیر

فولاد ابیضی و محمد رضا هیدری
ازماشگاه محاسبات نرم

چکیده: فرآیندهای تصمیم‌گیری مارکف چندعامله بطور گسترده برای مدل کردن سیستم‌های چندعامله پذیر می‌روند. در این مقاله الگوریتم‌های مبتنی بر آنواتهای یادگیر برای حل فرآیندهای تصمیم‌گیری مارکف چندعامله و باقی خطا مشی بهینه پیشنهاد می‌گردد. در الگوریتم‌های پیشنهادی محیط منتهای مارکف به صورت یک گراف جهت‌دار مدل می‌شود که در آن، گره‌های گراف نشان‌دهنده وضعیت‌ها و بالهای جهت‌دار بین گره‌ها به منزله اعمالی می‌باشند که منجر به انتقال از وضعیت به وضعیت دیگر می‌گردند. هر گره از گراف، بدین اوتماتای یادگیر مجهز می‌شود که اعمال آن، بالهای خود را آن گره می‌پوشاند. عالی‌ترین گراف حرکت کرده، از یک گره به گره دیگر حرکت می‌کند و در هر گره، از اوتماتای یادگیر متاظر با آن گره برای انتخاب عمل و انتقال به وضعیت بعدی کمک می‌گیرد. با توجه به مسیر طی شده توسط عامل‌ها و میزان مطلوبیت مسیر از لحاظ هزینه و مدت زمان لازم برای رسیدن به وضعیت هدف، بردا انتقال اعمال اوتماتهای یادگیر به نکام می‌شود و این طریق اوتماتهای یادگیر بدقتاری به اعمال بهینه خود همگرا می‌گردند. آزمایش‌های انجام گرفته نشان داده‌اند که الگوریتم‌های ارائه شده، در مقایسه با روش‌های مبتنی بر آنواتهای یادگیر موجود از کارایی بالاتری از هر دو چنین هزینه و سرعت رسیدن به راه حل بهینه برخوردار هستند.