

କାହିଁ କାହିଁ କାହିଁ କାହିଁ କାହିଁ କାହିଁ କାହିଁ କାହିଁ କାହିଁ କାହିଁ

፩፻፲፭

፳፻፲፭

፭፻፯ ከባን ቤትና፡ ቤትና ስምምነት የሚያስፈልግ የሚያስተካክለውን የሚያስፈልግ የሚያስተካክለውን

۲۰

۱۷۰ ۱۷۱

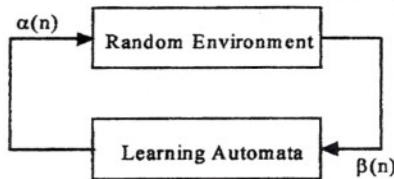
meybodi@ce.ac.ir

hmasiti@aeoi.org.ir

ମୁଖ୍ୟ ପରିଷଦ୍ୟ

~~b~~ 141

و عمل خود را برای مرحله بعد انتخاب می کند. شکل ۱ ارتباط بین آtomاتای یادگیر و محیط را نشان می دهد.

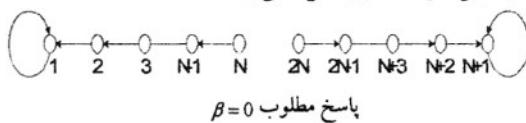


شکل ۱: ارتباط بین آtomاتای یادگیر و محیط

محیط^۲ را می توان توسط سه تابع $E \equiv \{\alpha, \beta, c\}$ نشان داد که در آن $\alpha \equiv \{\alpha_1, \alpha_2, \dots, \alpha_r\}$ مجموعه ورودیها، $\beta \equiv \{\beta_1, \beta_2, \dots, \beta_m\}$ مجموعه خروجیها و $c \equiv \{c_1, c_2, \dots, c_r\}$ مجموعه احتمالهای جریمه می باشد. هر گاه β مجموعه دو عضوی باشد، محیط از نوع P می باشد. در چنین محیطی $\beta_1 = 1$ به عنوان جریمه و $\beta_2 = 0$ به عنوان پاداش در نظر گرفته می شود. در محیط از نوع Q، $\beta(n)$ می تواند به طور گسته یک مقدار از مقادیر محدود در فاصله $[0,1]$ و در محیط از نوع S، $\beta(n)$ متغیر تصادفی در فاصله $[0,1]$ است. c احتمال اینکه عمل α نتیجه نامطلوب^۳ داشته باشد می باشد. در محیط ایستا^۴ مقادیر β بدون تغییر می مانند، حال آنکه در محیط غیر ایستا^۵ این مقادیر در طی زمان تغییر می کنند. آtomاتای یادگیری دو گروه با ساختار ثابت و با ساختار متغیر تقسیم می گرددند.

آtomاتای یادگیر با ساختار ثابت: آtomاتای یادگیر با ساختار ثابت توسط ۵ تابع $\{\alpha, \beta, F, G, \phi\}$ نشان داده می شود که در آن $\alpha \equiv \{\alpha_1, \alpha_2, \dots, \alpha_r\}$ مجموعه اعممال آtomاتا، $\beta \equiv \{\beta_1, \beta_2, \dots, \beta_m\}$ مجموعه ورودیهای آtomاتا، $F: \phi \rightarrow \{\phi_1, \phi_2, \dots, \phi_s\}$ مجموعه وضعیت‌های داخلی آtomاتا، $G: \phi \rightarrow \{G_1, G_2, \dots, G_t\}$ تابع تولید وضعیت جدید آtomاتا، و $\alpha \rightarrow \beta$ تابع خروجی که وضعیت کنونی را به خروجی بعدی می نگارد می باشد. بعنوان مثال به بررسی دو نمونه از آtomاتاهای با ساختار ثابت می پردازیم.

آtomاتای L_{2N,2}: این آtomاتا تعداد پاداش ها و جریمه های در یافت شده برای هر عمل را نگهداری کرده و تنها زمانی که تعداد جریمه های پشت سر هم بیشتر از تعداد پاداش های گردد، عمل دیگر را انتخاب می کند. نمودار تغییر وضعیت این آtomاتا مطابق شکل ۲ می باشد.



پاسخ مطلوب ۰

² Environment

³ Unfavorable

⁴ Stationary

⁵ Non-Stationary

⁶ Fixed Structure

می کنند تا در طراحی و اجرای سیاستهای چند عامله امکانات بیشتری را فراهم نمایند. طراحی و پیاده سازی مهارت‌های پیچیده برای عامل فوتالیست وظیفه ای سخت است و علت آن علاوه بر عدم قطبیت محیط و دخالت عوامل دیگر، جنبه ها و حالتی ای است که تغییر آنها در عملکرد یک مهارت بخصوص تاثیر دارد. در طراحی یک مهارت پیچیده تا حد ممکن باید مهارت را مستقل از جنبه ها و حالهای خاص تعریف کرد در غیر اینصورت باید تمام حالتی ممکن که تعداد آنها بسیار زیاد است و همیشه قابل پیش یینی نمیباشد را مورد بررسی قرار داد. یک راه برای حل این مشکل استفاده از روشهای یادگیری می باشد. اکثر تیمها از روش یادگیری نظرارتی [7] و یادگیری تقویتی [3] به عنوان روش یادگیری و آموزش عامل خود استفاده میکنند.

در این مقاله برای اولین کاربرد آtomاتای یادگیر در یادگیری مهارت‌های پیچیده مانند در یافت توب و دریبل در فوتال ریاتا مطرح میگردد. در قسمت اول این مقاله کاربرد آtomاتاهای یادگیری در یادگیری مهارت دریافت توب از ارائه می گردد. مسئله دریافت توب در حالت عمومی عبارت از حرکت به سمت توب در حال حرکت و در اختیار گرفتن توب می باشد. آtomاتای یادگیر جهت مناسب برای حرکت عامل فوتال را برای حرکت به سمت توب در حال حرکت تعیین می کند. در روش پیشنهادی ابتدا عامل فوتالیست را با یک آموزش جزئی غیر برخط آماده می کنیم. آtomاتای یادگیر سپس در حین بازی بصورت بر خط سعی در بهبود کارآبی عامل فوتال از طریق سازگار نمودن عامل با تغییرات محیط مینماید. در قسمت دوم این مقاله کاربرد آtomاتاهای یادگیر در یادگیری مهارت پیچیده دریبل ارائه میگردد. برای یادگیری مهارت دریبل از دو آtomاتای یادگیر استفاده شده است. این دو آtomاتای یادگیر به طور همزمان عمل کرده و قدرت لازم و جهت مناسب را برای انجام عمل دریبل تعیین می کنند. در این روش عامل فوتالیست در اینتا با مقدار مناسب آموزش غیر برخط آماده می شود. عامل فوتال به طور بر خط در حین بازی واقعی (با استفاده از آtomاتای یادگیری که برای این منظور طراحی شده است) همچنان به بهبود کارآبی خود و سازگاری با تغییرات محیط ادامه می‌دهد.

ادامه مقاله بصورت زیر سازماندهی شده است. در بخش ۲ مقدماتی در باره آtomاتاهای یادگیر ارائه میگردد. در بخش ۳ کاربرد آtomاتای یادگیر در یادگیری در یافت توب و در دریبل ۴ یادگیری دریبل شرح داده میشود. بخش ۵ مقاله نتیجه گیری می باشد.

۲- آtomاتای یادگیر^۱

آtomاتای یادگیر یک مدل انتزاعی است که تعداد محدودی عمل را می تواند انجام دهد. هر عمل انتخاب شده توسط محیطی احتمالی ارزیابی شده و پاسخی به آtomاتای یادگیر داده می شود. آtomاتای یادگیر از این پاسخ استفاده نموده

¹ Learning Automata