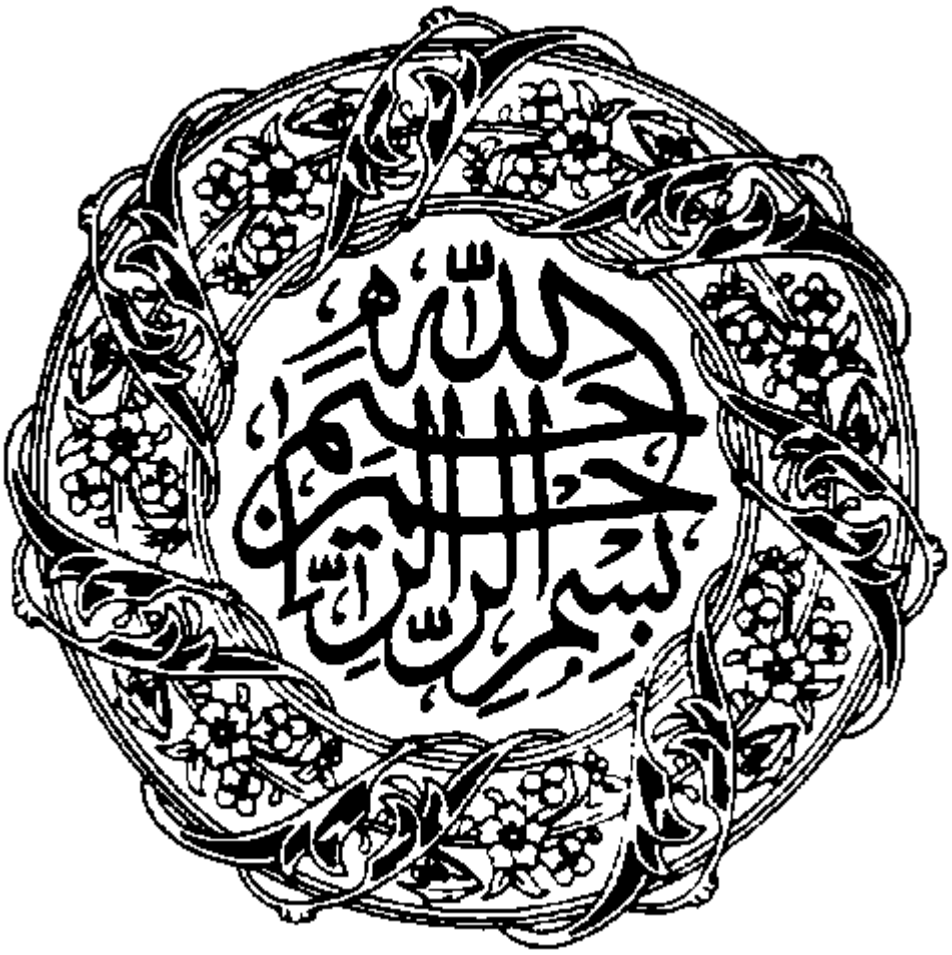

بسم الله الرحمن الرحيم

دانشگاه صنعتی امیرکبیر

(پلی تکنیک تهران)

دانشکده مهندسی پزشکی

پایان‌نامه کارشناسی ارشد مهندسی پزشکی – بیوالکتریک

ارائه روش یادگیری بصورت هایبرید با الهام از رفتار مغز در یادگیری یک فرآیند جدید

نگارش

سعید اقوامی

استاد راهنما

دکتر شهریار غریب‌زاده

استاد مشاور

دکتر فرزاد توحیدخواه

بهمن ماه۱۳۸۷

دانشگاه صنعتی امیرکبیر
(پلی‌تکنیک تهران)

بسمه تعالی

فرم اطلاعات پایان‌نامه
کارشناسی- ارشد و دکترا

تاریخ:
شماره:

معاونت پژوهشی
فرم پروژه تحصیلات تکمیلی ۷

مشخصات دانشجو:

نام و نام خانوادگی: سعید اقوامی | دانشجوی آزاد ● | بورسیه ○ | معادل ○

شماره دانشجوئی: ۸۵۱۳۳۰۰۱ | دانشکده: مهندسی پزشکی | رشته تحصیلی: بیوالکتریک | گروه:

مشخصات استاد راهنما:

نام و نام خانوادگی: شهریار غریب زاده | درجه و رتبه: دانشیار

نام و نام خانوادگی: | درجه و رتبه:

مشخصات استاد مشاور:

نام و نام خانوادگی: فرزاد توحیدخواه | درجه و رتبه: دانشیار

نام و نام خانوادگی: | درجه و رتبه:

عنوان پایان‌نامه به فارسی : ارائه روش یادگیری بصورت هایبرید با الهام از رفتار مغز در یادگیری یک فرآیند جدید

عنوان پایان‌نامه به انگلیسی: Hybrid learning method according to behavior of brain in learning of a new process

نوع پروژه: کارشناسی ○ | ارشد ● | دکترا ○ | سال تحصیلی:
کاربردی ○ | بنیادی ○ | توسعه‌ای ○ | نظری ●

تاریخ شروع :۱۳۸۶/۹/۲۰ | تاریخ خاتمه :۱۳۸۷/۱۱/۹ | تعداد واحد :۶ | سازمان تأمین کننده اعتبار :

واژه‌های کلیدی به فارسی: کنترل حرکات ارادی، یادگیری تقویتی، مدل معکوس، یادگیری حرکتی، عقده‌های قاعده‌ای، مخچه

واژه‌های کلیدی به انگلیسی : Voluntary Movements Control, reinforcement learning, supervised learnin, Motor Learning, basal ganglia, cerebellum

مشخصات ظاهری	تعداد صفحات ۱۳۰	تصویر ● جدول ● نمودار ● نقشه ○ واژه‌نامه ○		تعداد مراجع ۵۱	تعداد صفحات ضمائم
زبان متن	فارسی ●	انگلیسی ○	چکیده	فارسی ●	انگلیسی ●
یادداشت					

نظرها و پیشنهادها به منظور بهبود فعالیت‌های پژوهشی دانشگاه

استاد:

دانشجو:

امضاء استاد راهنما: **تاریخ:**

نسخه۱: ارائه به معاونت پژوهشی به همراه یک نسخه الکترونیکی از پایان‌نامه و فرم اطلاعات پایان‌نامه بصورت PDF همراه چاپ چکیده (فارسی انگلیسی) و فرم اطلاعات پایان‌نامه
نسخه۲: - ارائه به کتابخانه دانشکده (به همراه نسخه الکترونیکی فرم و دو جلد پایان نامه و لوح فشرده طبق نمونه اعلام شده در صفحه خانگی کتابخانه مرکزی مرکزی)

چکیده

در سالیان اخیر با پیشرفت ابزارهای آزمایشگاهی و فن‌آوریهای نوین، شواهد تجربی یادگیری تقویتی و باسرپرستی که بعنوان تئوریهای پایه‌ای در علم یادگیری و کنترل حرکات ارادی شناخته می‌شوند، در حال ارائه هستند. در پایان نامه حاضر تلاش در جهت رسیدن به شناختی جامع از سیستم، جهت یادگیری کنترل حرکات با عنایت به یادگیری تقویتی و نیز یادگیری باسرپرستی بصورت مدل معکوس و نقش آنها در حفظ موقعیت و حرکت به انجام رسیده است. در گام اول نقش یادگیری تقویتی در عقده‌های قاعده‌ای جهت انجام حرکت رسنده بررسی شد. این روش تجربه‌های قبلی را یاد نمی‌گیرد و بصورت تخمین تجمعی از خطای آینده سیستم، عمل کنترل حرکت را انجام می‌دهد، که جهت حصول به آن به تعامل با محیط اطرافش نیاز دارد. و سپس با تمرین وکسب تجربه مخچه بصورت مدل معکوس، دینامیک معکوس محیط را یاد می‌گیرد، تا بتواند در تکرارهای بعدی با استفاده از مدل شکل گرفته عمل کنترل را بصورت مستقیم انجام دهد. بر مبنای مدلسازی، در ابتدای یادگیری که مدل و اغتشاشات محیط در دسترس سیستم کنترل حرکات از دینامیک حرکت دقیق نمی‌باشد، سیستم کنترل از طریق یادگیری تقویتی با اغتشاشات خارجی مقابله می‌کند. بمرور زمان که مدل درونی سیستم از دینامیک محیط اطراف کامل می‌شود از نقش این یادگیری کاسته شده و کار کنترل توسط مدل معکوس انجام می‌شود. نتایج حاصل نشان داد که پس از یادگیری مهارتی استفاده از مدل درونی بمراتب کنترل بهتری را در حرکات رسنده در زمان پیش‌بینی شده اعمال کرد که منطبق بر شواهد فیزیولوژیکی نیز است.

کلید واژه: کنترل حرکات ارادی، یادگیری تقویتی، مدل معکوس، یادگیری حرکتی، عقده‌های قاعده‌ای، مخچه

فهرست مطالب

عنوان	صفحه
فصل اول: مقدمه	۱
۱-۱: مقدمه	۲
۱-۲: تعریف یادگیری	۳
۱-۲-۱: یادگیری کلامی	۴
۱-۲-۲: یادگیری حرکتی	۴
۱-۲-۳: یادگیری مشاهده‌ای	۵
۱-۲-۴: یادگیری اجتنابی	۵
۱-۳: استراتژی‌های یادگیری و انواع آن	۵
۱-۴: استراتژی‌های یادگیری در انسان جهت یادگیری حرکتی	۶
۱-۴-۱: مخچه	۷
۱-۴-۱-۱: یادگیری با سرپرستی در مخچه	۸
۱-۴-۲: عقده‌های قاعده‌ای	۹
۱-۴-۲-۱: یادگیری تقویتی در عقده‌های قاعده‌ای	۱۰
۱-۴-۲-۲: نقش سیگنال دوپامین در یادگیری تقویتی	۱۱
۱-۴-۳: قشر حرکتی	۱۳
۱-۴-۳-۱: یادگیری بدون سرپرستی در قشر حرکتی	۱۴
۱-۵: فیزیولوژی حرکات	۱۵
۱-۵-۱: طناب نخاعی	۱۷
۱-۶: فرآیند یادگیری مهارتی	۱۸
۱-۷: اهداف و نوآوریهای پروژه	۱۹
فصل دوم: اصول یادگیری تقویتی و باسرپرستی	۲۱
۲-۱: مقدمه	۲۲
۲-۲: معرفی یادگیری تقویتی	۲۴
۲-۲-۱: مثال	۲۵
۲-۳: اساس یادگیری تقویتی	۲۷
۲-۴: عامل و محیط	۲۷
۲-۵: المانهای یادگیری تقویتی	۲۸

عنوان	صفحه

۲-۵-۱: خاصیت مارکوف ۳۰

۲-۵-۲: تابع ارزش بهینه ۳۴

۲-۶: الگوریتم های RL ۳۷

۲-۶-۱: روش برنامه ریزی پویا ۳۷

۲-۶-۲: روش مونت کارلو ۳۸

۲-۶-۳: روش اختلاف زمانی ۳۹

۲-۷: الگوریتم یادگیری تقویتی SARSA ۴۱

۲-۷-۱: انتخاب بهینه عمل با استفاده از الگوریتم $\varepsilon - greedy$ ۴۱

۲-۸: کاربرد یادگیری تقویتی در کنترل بازو ۴۲

۲-۹: یادگیری باسرپرستی ۴۳

۲-۱۰: شبکه عصبی چیست؟ ۴۳

۲-۱۰-۱: ویژگیهای شبکه عصبی ۴۳

۲-۱۰-۲: کاربردهای شبکه عصبی ۴۴

۲-۱۱: مسائل مناسب برای یادگیری شبکه عصبی ۴۴

۲-۱۲: شبکه‌های عصبی چند لایه ۴۵

۲-۱۳: انواع شبکه‌های عصبی از نظر برگشت‌پذیری ۴۶

۲-۱۴: مراحل طراحی یک شبکه عصبی ۴۶

۲-۱۵: یادگیری شبکه دو لایه با رویکرد پس انتشار خطا ۴۸

۲-۱۵-۱: الگوریتم یادگیری پس انتشار خطا برای شبکه دو لایه ۴۸

۲-۱۶: محدودیتهای استفاده از شبکه‌های عصبی ۴۹

۲-۱۷: مزایای ترکیب دو یادگیری ۴۹

فصل سوم: سیستم کنترل حرکات و روشهای ارائه شده ۵۱

۳-۱: مقدمه ۵۲

۳-۲: توصیف سیستم حرکت ۵۴

۳-۲-۱: مشخصات طرح مطلوب حرکتی ۵۵

۳-۳: مدل داخلی مستقیم ۵۵

۳-۴: مدل داخلی معکوس ۵۶

عنوان	صفحه
۳-۵: شبیه‌سازیهای صورت گرفته در کنترل حرکات رسنده	۶۰
۳-۶: کنترل حرکات بر مبنای روش تنظیم امپدانس و کنترل مدل پیش بین	۶۴
۳-۷: مدل‌های بر مبنای بهینه سازی	۶۸
۳-۷-۱: بهینه سازی با استفاده از تابع هزینه	۶۹
۳-۷-۲: بهینه‌سازی با استفاده از یادگیری تقویتی	۷۰
فصل چهارم: نتایج و شیبه‌سازیهای حاصل از اعمال یادگیری ترکیبی جهت کنترل حرکت رسنده	۷۲
۴-۱: ساختار مدل پیشنهادی	۷۳
۴-۲: مدل دینامیک حرکتی دست در فضای دو بعدی	۷۶
۴-۳: شبیه‌سازیها	۷۹
۴-۴: نتایج شبیه‌سازیها	۸۱
۴-۵: مقایسه نتایج روش با روش بهینه‌ساز تودوروف	۱۰۳
۴-۶: اعتبار سنجی مدل پیشنهادی	۱۰۵
بحث و نتیجه گیری	۱۰۷
پیشنهادها	۱۰۹
فهرست مراجع	۱۱۰

فهرست شکل‌ها

عنوان	صفحه
شکل۱-۱: شمای سه یادگیری پایه. (الف) یادگیری بدون سرپرستی از طریق الگوی آماری از ورودی؛ (ب) یادگیری تقویتی از طریق سیگنال پاداش؛ (ج) یادگیری باسرپرستی از طریق بردار خطا	۶
شکل۱-۲: مدار داخلی مخچه. سلولهای گرانول(GC)، سلولهای پورکینژ (PC)، هسته عمقی (CN)، هسته زیتونی تحتانی (IO)، ○: اتصال تحریکی، ●: اتصال مهاری	۸
شکل۱-۳: شکل کلی اتصالات عقده‌های قا عده‌ای، قشر مغزو نورون های مغز میانی با الهام از مدلهای نقاد- عملگر	۱۲
شکل۱-۴: نواحی عملکردی قشر مغز	۱۳
شکل۱-۵: شکل مربوط به مدار نورونی قشر مغز. P، نورنهای هرمی؛ S، نورونهای مهره‌ای ستاره مانند؛ I، نورونهای واسطه مهاری؛ O، اتصال تحریکی؛ ●، اتصال مهاری	۱۴
شکل۱-۶: ساختار سلسله مراتبی سیستم کنترل حرکات آدمی	۱۶
شکل ۲-۱: نحوه‌ی تعامل عامل با محیط	۲۸
شکل۲-۲: مقادیر تابع ارزش	۳۴
شکل ۲-۳: ساختار عمل گر- منتقد	۴۰
شکل۲-۴: نمونه‌ای از یک شبکه تک لایه	۴۵
شکل۲-۵: نمونه‌ای از یک شبکه چند لایه	۴۵
شکل۲-۶: منحنی تابع سیگموئید و مشتق آن	۴۷
شکل۳-۱: (الف) مدل یادگیری فیدبکی کلاسیک و مدل معکوس در مسیر مستقیم. (ب) مدل نورونی متصور شده برای مدار عملکردی مخچه در مسیر مستقیم	۵۷
شکل۳-۲: ترکیب مدل های داخلی مستقیم و معکوس جهت کامل نمودن حلقه کنترلی با وجود فیدبک سنسوری با تاخیر بالا	۵۸
شکل۳-۳: مدلی شماتیک برای تولید حرکات هدف‌گرا. در این مدل x حالات سیستم، y فیدبک حسی و u ورودی‌های کنترلی را نشان می‌دهند. اندیس v و p بترتیب نشان‌دهندهٔ فیدبک بینایی و حس عمقی هستند. بر اساس شواهد احتمال می‌رود که بلوک‌های این مدل با بخشهایی از مغز مرتبط باشند که در شکل به آنها اشاره شده است	۵۹
شکل ۳-۴: کنترل حرکات رسنده	۶۰
شکل۳-۵: بررسی تاثیر محدودیتها در پاسخ به اغتشاش	۶۲
شکل۳-۶: بلوک دیاگرام سیستم کنترل بازو، شامل هر دو کنترل کننده فیدبکی و کنترلر مسیر مستقیم تطبیقی. دینامیکهای داخلی یک سیستم دینامیکی می‌باشد که هر دو بلوک مدار زمانی و طراحی مسیر را بعهده دارد	۶۲

عنوان	صفحه

شکل۳-۷: موقعیت ابتدایی (چپ) و انتهایی بازو بهمراه مسیر طی شده (راست) ۶۳

شکل۳-۸: پروفایل سرعت مطلوب و اندازه‌گیری شده (شکل راست). موقعیت و سرعت زاویه‌ای مطلوب و اندازه‌گیری شده (شکل چپ) ۶۴

شکل۳-۹: مدلی جلوسو از سیستم کنترل حرکات ۶۵

شکل۳-۱۰: کنترل مدل مبنای MPC بهمراه کنترل امپدانس فیدبکی ۶۵

شکل۳-۱۱ نتایج شبیه‌سازی بدست آمده از مدل سیستم با مقدیر نامی برای امپدانس و مدلی از اغتشاشات موجود نیست. الف- حرکت در میدان نیروی خنثی ب- حرکت در میدان نیروی متناسب با سرعت دست ۶۷

شکل۳-۱۲: مسیرهای حرکتی دست تحت سطوح مختلف حرکتی. الف- بدون یادگیری حرکتی ب- یادگیری حرکتی ۳۰٪ ج- یادگیری حرکتی ۶۰٪ د- یادگیری کامل ۶۷

شکل۳-۱۳: اعمال کنترلر ILQG جهت عمل رسنده برای بازو با دو درجه آزادی. الف- میانگین مسیرهای طی شده در امتداد هشت جهت. ب- پروفایل‌های سرعت ۶۹

شکل۳-۱۴: منحنی کاهش تابع هزینه با تکرار آزمایشها برای ۸ جهت ۷۰

شکل ۴-۱: شکل گیری مدل معکوس در مخچه ۷۴

شکل۴-۲: مدل پیشنهادی شامل یادگیری تقویتی و مدل معکوس ۷۵

شکل۴-۳: بازوی دو لینکه در صفحه ی افقی ۷۸

شکل۴-۴: جهت اعمال میدان نیروی متناسب توسط محیط در هر کدام از هشت جهت آزمایشی. همانگونه که دیده می‌شود جهت میدان نیرو عمود بر جهت حرکتی دست در نظر گرفته شده است و اندازه آن متناسب با سرعت دست است ۸۰

شکل۴-۵: منحنی مطلوب مسیر طی شده توسط دست در ۸ جهت مختلف در کمترین زمان از بین ۲۰۰ آزمایش در حضور اغتشاش ۸۳

شکل۴-۶: نحوهِی سوئیچ شدن مکانیزم کنترل از یادگیری تقویتی به مدل معکوس در ناحیه‌ی مخچه ۸۴

شکل۴-۷: مسیر مربوط به ۸ جهت مختلف حاصل از تست شبکه مدل معکوس پس ازیادگیری در میدان خنثی ۸۵

شکل۴-۸: منحنی تغییرات خطای MSE مربوط به مرحله آموزش شبکه مدل معکوس پس از ۱۰ مرحله برای مسیر اول ۸۶

شکل۴-۹: مسیر مربوط به جهت ۱ حاصل از تست شبکه مدل معکوس پس ازیادگیری ۸۶

شکل۴-۱۰: منحنی تغییرات گشتاور حول مفصل شانه (Tao1) و حول مفصل آرنج (Tao2) ۸۷

شکل۴-۱۱: منحنی پروفایل سرعت برای مسیر اول بصورت زنگوله‌ای شکل در عرض ۰،۵ ثانیه ۸۷

شکل۴-۱۲: منحنی تغییرات زاویه شانه (راست) و زاویه آرنج (چپ) در عمل رسیدن به هدف در عرض ۰،۵ ثانیه ۸۸

عنوان	صفحه

شکل۴-۱۳: منحنی تغییرات خطای MSE مربوط به مرحله آموزش شبکه مدل معکوس پس از ۱۰ مرحله برای مسیر دوم ۸۸

شکل۴-۱۴: مسیر مربوط به جهت ۲حاصل از تست شبکه مدل معکوس پس ازیادگیری ۸۹

شکل۴-۱۵: منحنی تغییرات گشتاور حول مفصل شانه (Tao1) و حول مفصل آرنج (Tao2) ۸۹

شکل۴-۱۶: منحنی پروفایل سرعت برای مسیر دوم بصورت زنگوله‌ای شکل در عرض۰،۵ ثانیه ۸۹

شکل۴-۱۷: منحنی تغییرات زاویه شانه (چپ) و زاویه آرنج (راست) در عمل رسیدن به هدف در عرض ۰،۵ ثانیه ۹۰

۴-۱۸: منحنی تغییرات خطای MSE مربوط به مرحله آموزش شبکه مدل معکوس پس از ۱۰ مرحله برای مسیر سوم ۹۰

شکل۴-۱۹: مسیر مربوط به جهت ۳ حاصل از تست شبکه مدل معکوس پس ازیادگیری. ۹۱

شکل۴-۲۰: منحنی تغییرات گشتاور حول مفصل شانه (Tao1) و حول مفصل آرنج (Tao2). ۹۱

شکل۴-۲۱: منحنی پروفایل سرعت برای مسیرسوم بصورت زنگوله‌ای شکل در عرض۰،۵ ثانیه ۹۱

شکل۴-۲۲: منحنی تغییرات زاویه شانه (چپ) و زاویه آرنج (راست) در عمل رسیدن به هدف در عرض ۰،۵ ثانیه ۹۲

شکل۴-۲۳: منحنی تغییرات خطای MSE مربوط به مرحله آموزش شبکه مدل معکوس پس از ۱۰ مرحله ۹۲

شکل۴-۲۴: مسیر مربوط به جهت ۴ حاصل از تست شبکه مدل معکوس پس ازیادگیری ۹۳

شکل۴-۲۵: منحنی تغییرات گشتاور حول مفصل شانه (Tao1) و حول مفصل آرنج (Tao2) ۹۳

شکل۴-۲۶: منحنی پروفایل سرعت برای مسیر چهارم بصورت زنگوله‌ای شکل در عرض۰،۵ ثانیه ۹۳

شکل۴-۲۷: منحنی تغییرات زاویه شانه (چپ) و زاویه آرنج (راست) در عمل رسیدن به هدف در عرض ۰،۵ ثانیه ۹۴

شکل۴-۲۸: منحنی تغییرات خطای MSE مربوط به مرحله آموزش شبکه مدل معکوس پس از ۱۰ مرحله ۹۴

شکل۴-۲۹: مسیر مربوط به جهت ۵ حاصل از تست شبکه مدل معکوس پس ازیادگیری ۹۵

شکل۴-۳۰: منحنی تغییرات گشتاور حول مفصل شانه (Tao1) و حول مفصل آرنج (Tao2) ۹۵

شکل۴-۳۱: منحنی پروفایل سرعت برای مسیرپنجم بصورت زنگوله‌ای شکل در عرض ۰،۵ ثانیه ۹۵

شکل۴-۳۲: منحنی تغییرات زاویه شانه (چپ) و زاویه آرنج (راست) در عمل رسیدن به هدف در عرض ۰،۵ ثانیه ۹۶

شکل۴-۳۳: منحنی تغییرات خطای MSE مربوط به مرحله آموزش شبکه مدل معکوس پس از ۱۰ مرحله ۹۶

شکل۴-۳۴: مسیر مربوط به جهت ۶ حاصل از تست شبکه مدل معکوس پس ازیادگیری ۹۷

شکل۴-۳۵: منحنی تغییرات گشتاور حول مفصل شانه (Tao1) و حول مفصل آرنج (Tao2) ۹۷

شکل۴-۳۶: منحنی پروفایل سرعت برای مسیرششم بصورت زنگوله‌ای شکل در عرض۰،۵ ثانیه ۹۷

عنوان	صفحه
شکل۴-۳۷: منحنی تغییرات زاویه شانه (چپ) و زاویه آرنج (راست) در عمل رسیدن به هدف در عرض ۰،۵ ثانیه	۹۸
شکل۴-۳۸: منحنی تغییرات خطای MSE مربوط به مرحله آموزش شبکه مدل معکوس پس از ۱۰ مرحله	۹۸
شکل۴-۳۹: مسیر مربوط به جهت ۷ حاصل از تست شبکه مدل معکوس پس ازیادگیری	۹۹
شکل۴-۴۰: منحنی تغییرات گشتاور حول مفصل شانه (Tao1) و حول مفصل آرنج (Tao2)	۹۹
شکل۴-۴۱: منحنی پروفایل سرعت برای مسیر هفتم بصورت زنگوله‌ای شکل در عرض۰،۵ ثانیه	۹۹
شکل۴-۴۲: منحنی تغییرات زاویه شانه (چپ) و زاویه آرنج (راست) در عمل رسیدن به هدف در عرض ۰،۵ ثانیه	۱۰۰
شکل۴-۴۳: منحنی تغییرات خطای MSE مربوط به مرحله آموزش شبکه مدل معکوس پس از ۱۰ مرحله	۱۰۰
شکل۴-۴۴: مسیر مربوط به جهت ۸ حاصل از تست شبکه مدل معکوس پس ازیادگیری	۱۰۱
شکل۴-۴۵: منحنی تغییرات گشتاور حول مفصل شانه (Tao1) و حول مفصل آرنج (Tao2)	۱۰۱
شکل۴-۴۶: منحنی پروفایل سرعت برای مسیر هشتم بصورت زنگوله‌ای شکل در عرض۰،۵ ثانیه	۱۰۱
شکل۴-۴۷: منحنی تغییرات زاویه شانه (چپ) و زاویه آرنج (راست) در عمل رسیدن به هدف در عرض ۰،۵ ثانیه	۱۰۲
شکل۴-۴۸: مقاسیه نتایج. الف- مسیر مبنا. ب- مسیر بدست آمده از روش پیشهنادی در پروژه. ج- مسیر مربوط به روش بهینه ساز آقای تودوروف در سال۲۰۰۶.	۱۰۴
شکل۴-۴۹: مقایسه مسیر مدل با داده تجربی برای مسیر ۴۵ درجه (سمت راست) و مسیر ۱۸۰ درجه (سمت چپ)	۱۰۶

فهرست جدول‌ها

عنوان	صفحه
جدول۴-۱: مقادیر نوعی پارامترهای مدل دینامیک حرکت دست در صفحه افقی	۷۹
جدول۴-۲: مقادیر خطای میانگین مماسی برای ۸ مسیر.	۱۰۳

فصل اول

مقدمه

۱-۱: مقدمه

یادگیری تقویتی (RL) و یادگیری باسرپرستی معمولاً بعنوان دو روش مجزای یادگیری تصور می‌شوند. روش‌های RL اغلب جهت حل مسائل دینامیکی ترتیبی و بهینه سازی عملکرد، از طریق جستجوی نتایج اعمال بصورت برخط بکار برده می‌شوند. روش‌های مبتنی بریادگیری باسرپرستی، جهت حل مسائل مربوط به نگاشت ورودی به خروجی (بصورت ایستا) و کمینه کردن بردار سیگنال خطا، بدون وابستگی به چگونگی بدست آمدن نمونه‌های تعلیم، استفاده می شود.

ویژگی کلیدی جهت بیان اختلاف بین دو یادگیری، این است که آیا اطلاعات تعلیم بدست آمده از محیط بعنوان سیگنال ارزیابی بکار می‌رود یا بعنوان سیگنال خطا، و لذا در اینجا دنبال حل مسائلی هستیم که هر دو نوع یادگیری بصورت همزمان استفاده شود. سیگنال ارزیابی بعنوان بهینه کننده تعلیم در نظر گرفته می شود، در حالی که سیگنال خطا، استاندارد صحتی فراهم می آورد تا به تضمین سطح عملکردی کمک کند. هرکدام از این روش ها بر پایه آزمایش به آزمایش می باشند، یعنی فرآیند یادگیری را بصورت کلی در نظر می‌گیرند. الگوریتمهای مربوط به یادگیری تقویتی (الهام گرفته شده از عملکرد مغز)، تئوریهای محاسباتی موثری جهت ارزیابی رفتار سیستمهای مصنوعی (از

طریق پاداش یا تنبیه) ارائه کرده‌اند. این یادگیری تقویتی به خاطر کاربردهای فراوان در سیستم‌های هوشمند، ناوبری، رباتها و... بسیار حائز اهمیت است. مشکل عمده سیستم‌های یادگیرنده قبلی این است که عملکرد همه نواحی درگیر در یادگیری فرایند جدید، شامل قشر مغز، مخچه، ساقه مغز و عقده‌های قاعده‌ای را در نظر نگرفته اند؛ در صورتی که برای یادگیری، این نواحی با هم تعامل دارند. با توجه به تعامل نواحی فوق جهت یادگیری و کنترل اعمال موتوری، مانند بازوی رسنده به هدف، ضروری بنظر می‌رسد تا در ادامه از فرایند یادگیری در انسان و سیستم کنترل حرکات بحث شود.

اما مشکل اصلی در راه شناخت سیستم یادگیری و تعاملات موجود در مغز انسان، عدم دسترسی مستقیم به ورودی‌ها و عناصر اصلی این سیستم است. وجود فیدبکهای متنوع، متعدد، پیچیده و دارای سازوکارهای[1] خاص عملکردی نیز باعث می‌شود شناسایی این سیستم با مشکل مواجه شود. همچنین این سیستم تعاملی پویا با محیط اطراف دارد و در شرایط محیطی مختلف رفتارهای متفاوتی را از خود بروز می‌دهد. با توجه به پیچیدگی بسیار زیاد سیستم یادگیری، محققین علوم مختلف با استفاده از ساده‌سازی‌های خاص خود سعی در بررسی و شناخت این سیستم نموده-اند. تلاش‌های روانشناسان رفتاری۲ و مهندسین پزشکی بیشتر معطوف به ارائه مدل‌های کمی و یا کیفی از سیستم یادگیری با استفاده از آزمایش‌های تجربی که در شرایط خاص آزمایشگاهی صورت می‌گیرد، بوده است. پیشرفت‌های نظری سیستمهای کنترل مصنوعی نیز، مهندسین را در ارائه مدل‌های کاملتری از عملکرد سیستم یادگیری یاری داده است.

۱-۲: تعریف یادگیری

با توجه به تعریفی که از یادگیری می‌شود و آن را تغییرات کم و بیش دائمی رفتار در نتیجه تجربه می‌دانند، می‌توان تصور کرد که با انواع یادگیری در زندگی مواجه خواهیم بود، به این دلیل که فرد از زمان تولد تا سالمندی تجارب مختلفی در حوزه‌های مختلف کسب می‌کند. تلاشهای مکرر برای راه رفتن، سخن گفتن، غذا خوردن، ارتباط با همسالان، دفاع از خود، مطالب آموزشگاهی، مهارت‌هایی مثل رانندگی، یادگیری مشاغل و ... نمونه‌هایی از یادگیری‌هایی هستند که فرد در طول زندگی خود به شیوه‌های مختلفی آنها را یاد می‌گیرد. نظریه پردازان مختلف انواع مختلف یادگیری را معرفی کرده‌اند .

---

[1] - Strategy
[2] - Behaviaral Neuropsychologist

۱-۲-۱: یادگیری کلامی

یادگیری کلامی نوعی از یادگیری است که زمانی حاصل می‌شود که محتوای آموخته شده بوسیله فرد شامل واژه‌ها، هجاهای بی‌معنی، یا مفاهیم می‌شود. زمانیکه دانش آموزی سعی می‌کند زبان جدیدی را یاد بگیرد و شروع می‌کند به یادگیری کلمات و لغات، در واقع از یادگیری کلامی بهره می‌برد. همچنین دانش‌آموزی که سعی می‌کند بخشی از کتاب زیست شناسی را یاد بگیرد، مثل اصطلاحات و مفاهیم مربوط که در آن بخش ارائه شده است در واقع از یادگیری کلامی استفاده کرده است.

یادگیری کلامی در مقابل یادگیری حرکتی قرار دارد، هر چند فرض می‌شود که یادگیری کلامی از لحاظ شناختی در سطح بالاتری قرار دارد. این مساله یادگیری کلامی را در مقامی ممتازتر از یادگیری حرکتی قرار نمی‌دهد. ابزاری که برای بررسی یادگیری کلامی مورد استفاده فراوان و سودمند قرار گرفته است، لیست مربوط به هجاهای بی‌معنی بوده است. در این نوع آزمایش‌ها، فهرستی از هجاهای بی‌معنی را به سوژه‌ها می‌دادند یا برایشان می‌خواندند و در فواصل زمانی معین از آنها می‌خواستند هجاهای خوانده شده را یادآوری کنند. میزان مواد یادآوری شده، معرف میزان یادگیری کلامی فرد بوده است.

۱-۲-۲: یادگیری حرکتی

یادگیری حرکتی نوعی از یادگیری است و زمانی حاصل می‌شود که محتوای آموخته شده بوسیله فرد شامل مهارتهای حرکتی است. به عبارتی زمانیکه فرد تلاش می‌کند سلسله‌ای از حرکات را یاد بگیرد تا بتواند مهارتی را به صورت مناسبی انجام دهد از یادگیری حرکتی استفاده کرده است. یادگیری شنا کردن، پرتاب توپ، دوچرخه سواری و ... نمونه‌ای از یادگیری‌های حرکتی هستند. یادگیری حرکتی می‌تواند طیفی از یادگیریهای حرکتی را شامل می‌شود. برخی از این یادگیری ساده و سطحی‌تر هستند، مثل پرتاب یک توپ، برخی پیچیدگی بیشتری دارند، مثل شنا کردن و برخی از پیچیدگی و ظرافت خیلی بالایی برخوردار هستند، مثل نواختن یک ساز مانند ویولون که مستلزم یادگیری حرکتی به مقدار زیاد است. تمرین در این نوع یادگیری اهمیت فوق‌العاده‌ای دارد. هر چه میزان تمرین فرد در این نوع حرکات بیشتر باشد مهارت او عمیق خواهد بود. در این پروژه محور کار یادگیری حرکتی می‌باشد.

۱-۲-۳: یادگیری مشاهده‌ای

یادگیری مشاهده‌ای عبارت است از یادگیری که از طریق آن موجود زنده‌ای از رفتار موجود زنده دیگر تقلی می‌کند، سرمشق می‌گیرد یا ادای آن را درآورد. به این نوع یادگیری سرمشق گیری و یادگیری اجتماعی نیز می‌گویند. کودکی را تصور کنید که رفتارهای مادر را در هنگام آشپزی یا کار خانه‌داری مشاهده می‌کند. مدتی بعد در بازیهای خود همان کارهایی را که از مادر در خود مشاهده کرده انجام می‌دهد. بیشتر یادگیری‌های کودکان مشاهده‌ای است. در واقع آنها بطور مرتب در تقلید از رفتارها و حرکاتی هستند که مشاهده می‌کنند.

۱-۲-۴: یادگیری اجتنابی

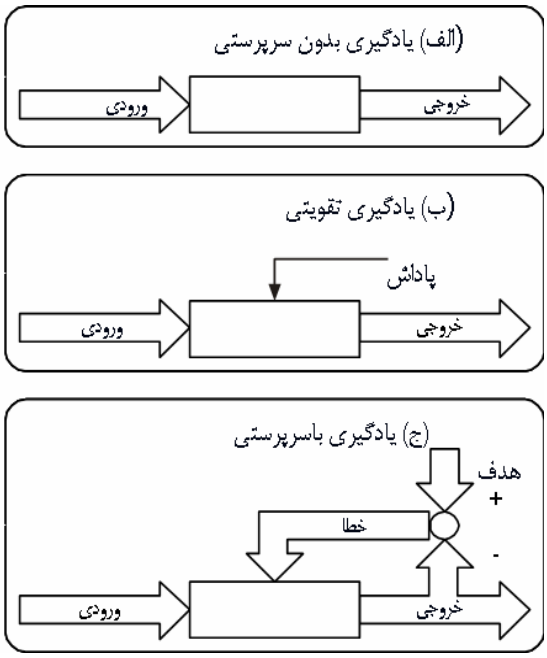
یادگیری اجتنابی زمانی اتفاق می‌افتد که یک موجود زنده این توانایی را پیدا می‌کند که با استفاده از نشانه یا علامتی خاص از تماس با محرکی زیان‌آور اجتناب کند. به عبارتی فرد سعی می‌کند از موقعیت‌هایی که برایش آزار دهنده هستند، اجتناب و دوری کند. تصور کنید که کودکی یک بار با دست زدن به یک بخاری دست خود را می‌سوزاند و از آن پس از نزدیک شدن به بخاری خودداری می‌کند. این کار کودک نشانه‌ای از یادگیری اجتنابی است. این نوع یادگیری در بررسی علل مربوط به برخی روان رنجوری‌ها مثل انواع اضطراب مورد توجه است.

۱-۳: استراتژی های یادگیری و انواع آن

یادگیری باسرپرستی[1]، یک روش عمومی در یادگیری ماشینی است که در آن به یک سیستم، مجموعه جفت‌های ورودی - خروجی ارائه شده و سیستم تلاش می‌کند تا تابعی از ورودی به خروجی را فرا گیرد[۱]. یادگیری تحت سرپرستی نیازمند تعدادی داده ورودی به منظور آموزش سیستم است. با این حال رده‌ای از مسائل وجود دارند که خروجی مناسب که یک سیستم یادگیری تحت سرپرستی نیازمند آن است، برای آن‌ها موجود نیست. این نوع از مسائل چندان قابل جوابگویی با استفاده از یادگیری تحت سرپرستی نیستند. یادگیری تقویتی[2] مدلی برای مسائلی از این قبیل فراهم می‌آورد. در یادگیری تقویتی، سیستم تلاش می‌کند تا تعاملات خود با یک محیط پویا را از طریق خطا و آزمایش بهینه نماید[۲]. یادگیری تقویتی مسئله‌ای است، که یک عامل می‌بایست رفتار خود را

[1] - Supervisor learning
[2] - Reinforcement learning

از طریق تعاملات آزمایش و خطا با یک محیط پویا فرا گیرد، با آن مواجه است. در یادگیری تقویتی هیچ نوع زوج ورودی- خروجی ارائه نمی‌شود. به جای آن، پس از اتخاذ یک عمل، حالت بعدی و پاداش بلافصل به عامل ارائه می‌شود. هدف اولیه برنامه‌ریزی عامل‌ها با استفاده از تنبیه و تشویق است، بدون آنکه ذکری از چگونگی انجام وظیفه آن‌ها شود. در پاره‌ای از مسائل نیز بدلیل نبود هیچگونه ورودی- خروجی مطلوب و یا حتی معیار خوب یا بد بودن خروجی بصورت پاداش، یادگیری دیگری تحت عنوان یادگیری بدون سرپرستی مطرح می‌شود که در آن هدف بدست آوردن یک ویژگی آماری از ورودیها می‌باشد. و لذا درتقسیم‌بندی کلی می‌توان ۳ نوع یادگیری در نظر گرفت (شکل ۱-۱).



شکل ۱-۱: شمای سه یادگیری پایه. (الف) یادگیری بدون سرپرستی از طریق الگوی آماری از ورودی؛ (ب) یادگیری تقویتی از طریق سیگنال پاداش؛ (ج) یادگیری باسرپرستی از طریق بردار خطا.

۱-۴: استراتژیهای یادگیری در انسان جهت یادگیری حرکتی

در این بخش دلایل و شواهد فیزیولوژیکی مربوط به مخچه، عقده‌های قاعده‌ای و قشر مغزی که بترتیب در یادگیری باسرپرستی، یادگیری تقویتی و یادگیری بدون سرپرستی دخیل هستند، تشریح می‌شود.

۱-۴-۱: مخچه

علاوه بر نواحی قشری مغز که فعالیت عضلانی را کنترل می‌کند، دو ساختار مغزی دیگر نیز برای عملکرد طبیعی حرکت لازم است. این دو جز یکی مخچه و دیگری عقده‌های قاعده‌ای هستند. با وجود این هیچ کدام از این دو خود به خود نمی‌توانند عملکرد عضلانی را شروع کنند، بلکه آنها همیشه با هماهنگی سایر دستگاههای کنترل عمل می‌کنند. اصولاً مخچه نقش عمده‌ای در زمان‌بندی فعالیت های حرکتی و تبدیل سریع و ظریف حرکتی به حرکت دیگر را ایفا می‌کنند. همچنین در مخچه در هنگام تغییر بار وارده بر عضلات به کنترل شدت انقباض عضلانی کمک می‌کند و به همین اندازه در کنترل اثر متقابل میان گروههای عضلانی آگونیست و آنتاگونیست نقش ضروری دارد.
مخچه از قدیم به ناحیه خاموش معروف بوده است و علت اصلی آن این است که تحریک الکتریکی این ساختمان موجب بروز هیچ گونه حسی نمی شود و به ندرت باعث به وجود آمدن حرکت می شود. البته برداشتن مخچه باعث اختلال شدید حرکتی می‌شود.

مخچه برای فعالیتهای سریع عضلانی مانند دویدن، تایپ کردن، نواختن پیانو و حتی صحبت کردن اهمیت حیاتی دارد. از بین رفتن این ناحیه از مغز می تواند باعث ناهماهنگی تقریباً کامل این فعالیتها گردد، اگر چه هیچ یک از عضلات را فلج نمی‌کند.

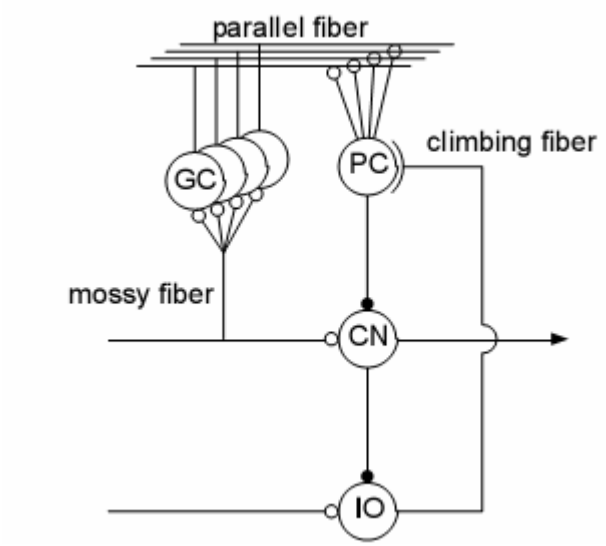
مخچه به طور پیوسته اطلاعات جدید مربوط به برنامه مورد نظر انقباضات عضلانی از نواحی کنترل حرکتی مغز دریافت می‌کند. همچنین مخچه پیوسته از قسمتهای محیطی بدن نیز اطلاعات حسی مربوط به تغییرات متوالی وضعیت هر قسمت از بدن (موقعیت بدن، سرعت حرکت، نیروی موثر بر آن و همانند آن) را دریافت می‌کند. آنگاه مخچه حرکات واقعی انجام شده را که از طریق اطلاعات فیدبک حس محیطی از آنها آگاه می‌گردد، با حرکاتی که مورد نظر دستگاه‌های حرکتی بوده مقایسه می‌نماید. اگر این دو به خوبی با هم سازگار نبودند، بلافاصله مخچه سیگنال‌های اصلاحی را به دستگاه مرکز ارسال می‌نماید تا میزان فعالیت عضلات را کم یا زیاد کند (یادگیری با سرپرستی).

علاوه بر این مخچه و قشر مغز کمک می‌کنند، تا در کسری از ثانیه قبل از انجام حرکت متوالی بعدی و در زمانی که حرکت فعالی در حال انجام است، حرکت بعدی را طراحی نماید و به این ترتیب در تبدیل آرام حرکتی به حرکت دیگر موثر است. همچنین مخچه از خطاهای خود درس می‌گیرد. یعنی اگر حرکتی دقیقاً به صورتی که مورد نظر بوده انجام نشود، مدار مخچه ای یاد می‌گیرد، که دفعه بعد حرکتی ضعیف‌تر یا قوی‌تر ایجاد کند. به همین منظور، تحریک‌پذیری نورون‌های خاصی

از مخچه دچار تغییراتی می شوند و به این ترتیب انقباضات بعدی با حرکات مورد نظر تطابق بیشتری خواهد داشت[۳].

۱-۴-۱-۱: یادگیری با سرپرستی در مخچه

شکل (۱-۲) نشان دهنده مدار داخلی مخچه می‌باشد. این ساختار یک شبکه پیشرو را تداعی می‌کند، که در آن آکسون سلولهای گرانول (فیبرهای موازی) با سلول‌های پورکینژ سیناپس شده است. مسیرهای ورودی دو تا هستند، که عبارتند از: فیبر موسی و فیبر بالا رونده از هسته زیتونی. ورودی ناشی از فیبر موسی، که حامل آوران‌های سنسوری و مغزی می‌باشد، توسط سلولهای گرانول منتقل می‌شود.



شکل۱-۲: مدار داخلی مخچه. سلول‌های گرانول(GC)، سلولهای پورکینژ (PC)، هسته عمقی (CN)، هسته زیتونی تحتانی (IO)، ○: اتصال تحریکی، ●: اتصال مهاری[۴] .

فیبرهای موازی و بالارونده نیز درسلول‌های پورکینژ همگرا می‌شوند. هر سلول پورکینژ تقریباً ۲۰۰۰۰۰ ورودی از فیبر موازی دارد، در حالی‌که تنها یک ورودی از فیبر بالارونده دارد. نورون‌های خروجی مخچه، که در هسته عمقی واقع شده‌اند، ورودی مهاری را از سلول پورکینژ دریافت می‌کند و ورودی تحریکی را از فیبر موسی دریافت می‌کند. خروجی هسته عمقی مخچه‌ای بسته به مکان فعال شده در داخل هسته، به سیستم اوکولوموتور، موتور نخاع، و از طریق تالاموس به نواحی خاصی در قشر مغز منتقل می‌شود.

پاسخ‌های سلول پورکینژ به ورودی‌های ناشی از فیبر موازی و فیبر بالارونده با دونوع پاسخ مختلف تمییز داده می‌شود که این پاسخها بصورت اسپایک ساده و جمعی می‌باشد. پاسخ بصورت اسپایک ساده سلول پورکینژ، مربوط به ورودی ناشی از فیبر موازی می‌باشد که سیگنال‌های مربوط به حرکت مانند فرمانهای حرکتی چشم شامل شتاب و سرعت را کدگذاری می‌کند. پاسخ بصورت اسپایک جمعی، مربوط به ورودی ناشی از فیبر بالارونده می‌باشد، فرکانس بسیار کمی دارد. آنالیزهای آماری نشان می‌دهد که ورودی ناشی از فیبر بالارونده خطای حرکتی را کدگذاری می‌کند، بعنوان مثال در حرکت چشمی سیگنال خطای شبکیه‌ای را حمل می‌کند. همچنین اسپایک‌های جمعی خطای حرکتی در حرکت بازوی رسنده را هم کدگذاری می‌کنند.

نتایج حاصل از بررسی‌ها نشان می‌دهد که مدار مخچه قابلیت اجرای یادگیری باسرپرستی را دارد. سلولهای گرانول می‌توانند بصورت کدکننده سیگنال ورودی ناشی از فیبر موسی در نظر گرفته شوند. خروجی آنها بصورت خطی با سلول پورکینژ ترکیب می‌شود و وزن‌های سیناپسی توسط سیگنال خطای منتقل شده توسط فیبر بالارونده تصحیح می‌شود و ورودی شبکه مخچه‌ای از طریق فیبرهای موازی منتقل می‌شود. نقش مخچه در رفتارهای مبتنی بر خطا، مانند کنترل حرکت چشم و بازوی رسنده بصورت تجربی ثابت شده است. هوک، بوکینقام و بارتو یکسری مدل‌های مخچه‌ای در سطوح مختلف را در سال ۱۹۹۸ ارائه کردند. بسیاری از شبیه‌سازی‌ها و آزمایش‌های روباتیک، مفید بودن شبکه مخچه مانند را بصورت یادگیری باسرپرستی در اعمال موتور کنترلی تاییدکرده‌اند[۴].

۱-۴-۲: عقده‌های قاعده‌ای

همانطور که ذکر شد عقده‌های قاعده‌ای همانند مخچه از اجزایی هستند که برای کنترل طبیعی حرکت لازم هستند، اما نمی‌توانند حرکتی را به تنهایی آغاز کنند. عقده‌های قاعده‌ای به طراحی و کنترل الگوهای پیچیده حرکات عضلانی، کنترل شدت نسبی حرکات مجزا و توالی حرکات متوالی و موازی برای ایجاد حرکات ویژه و پیچیده کمک می‌کنند. این قسمت تمام سیگنال‌های خود را از قشر مغز دریافت می‌کند و همه را دوباره به قشر باز می‌گرداند. یکی از نقش‌های اساسی عقده‌های قاعده‌ای در کنترل حرکت این است که برای کنترل حرکت الگوهای پیچیده اعمال حرکتی در ارتباط با دستگاه قشری عمل می نماید. نوشتن حروف الفبا نمونه ای از این اعمال است. زمانی‌که عقده‌های قاعده‌ای آسیب جدی می بینند، دستگاه قشری کنترل حرکت دیگر نمی‌تواند این الگوها را به وجود آورد. در عوض نوشتن فرد بی‌دقت می‌شود، به طوری که گویی برای اولین بار نحوه نوشتن را می-آموزد.