دانشگاه تهران
پردیس دانشکده‌های فنی
دانشکدهٔ مهندسی برق و کامپیوتر

یادگیری مفاهیم با استفاده از نظریه شواهد دمپستر-شفر

نگارش
طاهر شهبازی میرزاحسنلو

استادان راهنما
دکتر بابک نجار اعرابی
دکتر مجید نیلی احمدآبادی

پایان‌نامه برای دریافت درجهٔ کارشناسی ارشد

مهندسی برق، گرایش کنترل

شهریور ۱۳۸۷

چکیده

مدلسازی عدم قطعیت در سیستم‌های یادگیری دنیای واقعی غالبا نه تنها باعث بهبود سرعت و کیفیت یادگیری می‌شود، مقاومت آن در مقابل اطلاعات ناکامل ادراکی و نایقینی در دانش کسب شده و نتیجه عمل را افزایش می‌دهد. یادگیری مفهوم به عنوان یکی از مسائل چالش‌انگیز یادگیری ماشینی، درگیر با صورت‌های مختلف نایقینی می‌باشد. هدف این تحقیق استفاده از مدلسازی نایقینی در نظریه شواهد دمپستر-شفر در یادگیری مفهوم و بررسی مزایای احتمالی یک چنین مدلسازی می-باشد. قالب مدل انتقال باور (TBM) از این نظریه ساختار مناسبی برای چنین یادگیری فراهم می-سازد. مدلی بر این اساس به منظور توسعه مفاهیم در قالب TBM ارائه شده مکانیزمی کلی برای یادگیری نموی برخط عامل‌های یادگیری تعاملی بر اساس تقویت طرح می‌شود. چگونگی مدیریت اطلاعات سنسوری مختلف از طریق مدلسازی قابلیت اطمینان سنسوری در قالب توابع باور بررسی شده و یادگیری آن بحث می‌شود. یادگیری مطلوبیت بر اساس تکلیف یا محیطی در طراحی عامل‌های خودکار یادگیری مفهوم به منظور فراهم آوردن امکان استفاده مجدد دانش بررسی شده بر این اساس توسعه‌ای از یادگیری Q به فضای پیوسته حالت در قالب توابع باور ارائه می‌گردد. صرف‌نظر از فراهم آوردن قالب یادگیری همزمان تابع مطلوبیت و تشکیل مفاهیم مجرد، به عنوان روش یادگیری تقویتی سازگاری بیشتری در مقابل شرایط عملی سیستم‌های یادگیری واقعی فراهم می‌آورد. توسعه آن به فضای پیوسته موتوری مدلی مناسب‌تر برای یادگیری تقویتی سیستم‌های روباتیکی فراهم می‌سازد، آنگونه که در آزمایش‌هایی در یک روبات واقعی نشان داده می‌شود.

فهرست مطالب

عنوان.....صفحه

۱ مقدمه ۲

۲ نظریه شواهد: مدل‌ها، کاربردها و توسعه‌ها..... ۸

۲-۱- مقدمه ۹

۲-۲- نظریه حدود پایین و بالای احتمال دمپستر ۹

۲-۳- نمایش سه‌تایی منبع و نظریه شواهد شفر ۱۲

۲-۳-۱- مفاهیم کلی در نظریه شواهد..... ۱۲

۲-۴- مجموعه احتمال‌های سازگار با یک ساختار شهودی ۱۸

۲-۵- مدل انتقال باور اسمتز ۲۰

۲-۶- ترکیب اطلاعات..... ۲۲

۲-۶-۱- روش‌های مختلف ترکیب ساختارهای شهودی منابع اطلاعاتی..... ۲۲

۲-۶-۲- ترکیب اطلاعات در TBM..... ۲۴

۲-۶-۳- خاص‌سازی، کلی‌سازی و اصل حداقل التزام..... ۲۷

۲-۷- تصمیم‌گیری..... ۲۸

۲-۷-۱- حل تضاد برای تصمیم‌گیری از روی ساختارهای شهودی حاصل از ترکیب..... ۲۸

۲-۷-۲- تبدیل پیگنوستیک..... ۳۰

۲-۷-۳- بهنجار کردن و تصمیم‌گیری ۳۳

۲-۷-۴- سایر روش‌های تصمیم‌گیری بر اساس ساختار شهودی..... ۳۳

۲-۷-۵- تصمیم‌گیری بر اساس هزینه مورد انتظار ۳۴

۲-۸- عملیات در فریم ضربی ۳۶

۲-۸-۱- تصویر کردن و توسعه تهی ۳۶

۲-۸-۲- شرط کردن و توسعه بالونی..... ۳۶

۲-۹- نظریه شواهد در بازشناسی الگو..... ۳۷

۲-۹-۱- کاربرد در کلاس‌بندی..... ۳۷

۲-۹-۲- کاربرد در خوشه‌بندی..... ۳۸

۳ یادگیری تعاملی بر اساس تقویت..... ۴۰

۳-۱- مقدمه ۴۱
۳-۲- یادگیری مفهوم ۴۲
۳-۳- مدل‌های یادگیری مفهوم ۴۳
۳-۴- طبقه‌بندی و سطوح مختلف تجرید ۴۵
۳-۵- تجرید در سیستم نورون آینه‌ای: درک عمل، یادگیری با تقلید ۴۷
۳-۶- یادگیری تقویتی ۴۸

۴ مدلسازی قابلیت اطمینان سنسوری در ترکیب اطلاعات و یادگیری تکاملی با PBIL ۵۲
۴-۱- مقدمه ۵۳
۴-۲- مدلسازی قابلیت اطمینان سنسوری ۵۳
۴-۲-۱- ترکیب اطلاعات و مفاهیم قابلیت اطمینان سنسوری در آن ۵۳
۴-۲-۲- مدلسازی قابلیت اطمینان سنسوری برای ترکیب اطلاعات در قالب توابع باور ۵۵
۴-۳- یادگیری تکاملی با استفاده از یادگیری نموی بر اساس جمعیت ۵۹
۴-۴- تحلیل‌ها، توسعه‌ها و کاربردهای PBIL ۶۰

۵ یادگیری مفهوم در قالب توابع باور ۶۲
۵-۱- مقدمه ۶۳
۵-۲- استفاده از نظریه شواهد دمپستر–شفر برای یادگیری مفاهیم ۶۴
۵-۳- یادگیری مفاهیم رابطه‌ای در قالب توابع باور ۶۵
۵-۳-۱- ساختار ۶۵
۵-۳-۲- تعیین ساختار شهودی برای مفهوم تحریک جدید ۶۵
۵-۳-۳- تصمیم‌گیری ۶۶
۵-۳-۴- بروز کردن ساختار ۶۸
۵-۳-۵- ساختارهای شهودی خاص ۷۲
۵-۴- توجیه قاعده بازنگری باور بر اساس تقویت با استفاده از قاعده بیز تعمیم یافته ۷۳
۵-۴-۱- نکاتی در مباحث مرتبط با بروزکردن ۷۳
۵-۴-۲- بررسی بیشتر روش‌های بروزکردن ۷۷
۵-۴-۳- بروز کردن بر اساس تقویت ۸۰
۵-۴-۴- یادگیری با ببلینگ ۸۳
۵-۵- کاربردها و پیاده‌سازیها ۸۳

۵-۵-۱- یک شبیه‌سازی ساده در ساختار داده‌های بیزی ۸۳
۵-۵-۲- کاهش بایاس اولیه یادگیری مفهوم با ترکیب با زوال حافظه ۸۴
۵-۵-۳- موضعی کردن روبات متحرک ۸۵
۵-۵-۴- ناوبری ۸۵
۵-۵-۵- درک ۸۶
۵-۵-۶- نویز اندازه‌گیری و حرکت ۸۷
۵-۵-۷- موضعی کردن ۸۸
۵-۵-۸- نمایش باور و نگاشت ۸۸
۵-۵-۹- روش‌های موضعی کردن ۸۹
۵-۵-۱۰- موضعی کردن و ایجاد نگاشت همزمان ۹۰
۵-۵-۱۱- یادگیری با داده‌های تصویر محیط ۹۶
۵-۶- جمع‌بندی ۱۰۰

۶ یادگیری ترکیب تصمیم در یادگیری شهودی مفهوم ۱۰۲
۶-۱- مقدمه ۱۰۳
۶-۲- طرح کلی مسئله ترکیب تصمیم ۱۰۴
۶-۳- لزوم استفاده از الگوریتم تکاملی برای ترکیب اطلاعات ۱۰۵
۶-۴- یادگیری ضرایب تخفیف و ضرایب تخفیف زمینه‌ای در مدلسازی قابلیت اطمینان سنسوری و ترکیب تصمیم یادگیری مفاهیم ۱۰۶
۶-۴-۱- جمعیت ۱۰۷
۶-۴-۲- تعریف تابع شایستگی ۱۰۷
۶-۴-۳- الگوریتم‌های تکاملی مورد استفاده ۱۰۹
۶-۵- مقایسه PBIL با GA ۱۰۹
۶-۶- توسعه PBIL به فضاهای جستجوی پیوسته برای مسائل بهینه‌سازی مقید ۱۱۰
۶-۷- نتایج پیاده‌سازی برای یک مثال ساده ۱۱۱
۶-۸- نتایج پیاده‌سازی یادگیری تکاملی PBIL در مسئله ترکیب تصمیم در یادگیری مفهوم. ۱۱۳
۶-۸-۱- داده‌های یادگیری مفهوم ۱۱۳
۶-۸-۲- تولید داده‌های آموزش برای یادگیری ضرایب قابلیت اطمینان ۱۱۳
۶-۸-۳- یادگیری تخفیف کلاسیک با استفاده از PBIL دودویی ۱۱۴
۶-۸-۴- یادگیری تخفیف زمینه‌ای با استفاده از PBIL دودویی ۱۱۵
۶-۸-۵- یادگیری تخفیف زمینه‌ای با استفاده از PBIL پیوسته ۱۱۶

۶-۹- جمع‌بندی ۱۱۷

۷ توسعه شهودی یادگیری Q برای یادگیری مطلوبیت تصمیم‌گیری..... ۱۱۸

۷-۱- مقدمه..... ۱۱۹
۷-۲- یادگیری Q شهودی..... ۱۲۳
۷-۲-۱- فرضیات پایه..... ۱۲۳
۷-۲-۲- ساختار یادگیری Q شهودی..... ۱۲۴
۷-۳- یادگیری در فضای حالت پیوسته ۱۲۸
۷-۳-۱- فراخوانی و تصمیم‌گیری..... ۱۲۸
۷-۳-۲- تخصیص پاداش..... ۱۳۰
۷-۳-۳- قاعده بروز کردن ۱۳۶
۷-۴- یادگیری چندگامی با یادگیری Q بهبودیافته..... ۱۳۷
۷-۵- یادگیری در فضاهای حالت و عمل پیوسته..... ۱۳۹
۷-۵-۱- تعبیر جدول Q به عنوان یک پایگاه قواعد ایجابی..... ۱۳۹
۷-۵-۲- استنباط توابع باور شرطی از مقادیر Q..... ۱۴۰
۷-۵-۳- انتخاب عمل اکتشافی..... ۱۴۴
۷-۵-۴- فراخوانی و تصمیم‌گیری..... ۱۴۴
۷-۵-۵- تخصیص پاداش و بروزکردن ۱۴۶
۷-۶- شبیه‌سازیها و پیاده‌سازیها..... ۱۴۷
۷-۶-۱- محیط‌ها..... ۱۴۷
۷-۶-۲- آماده‌سازی‌های پیاده‌سازی ۱۴۹
۷-۶-۳- شبیه‌سازی‌ها در فضای تصمیم گسسته..... ۱۵۱
۷-۶-۴- یادگیری SARSA چندگامی..... ۱۶۰
۷-۶-۵- شبیه‌سازی‌ها در فضاهای پیوسته حالت و عمل..... ۱۶۱
۷-۷- وارسی‌های آزمایشی ۱۶۴
۷-۸- کارهای مرتبط، بحث و کارهای باقیمانده..... ۱۷۰
۷-۹- جمع‌بندی ۱۷۳

۸ جمع‌بندی، بحث، نتیجه‌گیری و مسیر کارهای آینده ۱۷۴

۸-۱- جمع‌بندی و بحث ۱۷۵
۸-۲- نتیجه‌گیری..... ۱۷۷

۸-۳- مسیر کارهای آینده ۱۷۸

مقالات مستخرج از پایان‌نامه ۱۸۰

فهرست منابع ۱۹۰

واژه نامه ۲۰۴

فهرست شکل‌ها

عنوان.....صفحه

شکل ۲-۱: نگاشت چندمقداری Γ و پیشامد $T \subseteq \Omega$ ۱۰
شکل ۲-۲: مقادیر باور و موجه‌نمایی پیشامد X از فضای Ω بر حسب مقادیر جرم اعضای کانونی. ۱۵
شکل ۲-۳: بازه اعتماد و عدم آگاهی یک پیشامد مفروض با مقادیر باور و موجه‌نمایی معین.....۱۶
شکل ۲-۴: مقدار همانندی پیشامد X از فضای Ω بر حسب مقادیر جرم اعضای کانونی ۲۳
شکل ۳-۱: نمایش مفاهیم ادراکی، رابطه‌ای و قاعده‌ای (به ترتیب از چپ به راست) [Mob07].....۴۶
شکل ۳-۲ : مفاهیم رابطه‌ای و ارتباط فضاهای سنسوری و موتوری [Mob07] ۴۷
شکل ۵-۱: میانگین متوسط پاداش متحرک (جمع‌شونده) برای یادگیری مفاهیم با توزیع بیزی (میانگین ۱۰ ترتیب مختلف از داده‌ها)..... ۸۴
شکل ۵-۲: اختلاف بین اعداد احتمال مشاهده و احتمال انتخاب هریک از پنج مفهوم در دوحالت بدون زوال حافظه (میله‌های سمت چپ) و با زوال حافظه (میله‌های سمت راست)..... ۸۵
شکل ۵-۳: نگاشت توپولوژیکی گسسته محیط..... ۹۴
شکل ۵-۴: تشریح تناظرهای مدل موضعی کردن و نگاشت.....۹۶
شکل ۵-۵: نمونه‌هایی از چهار مفهوم چهار راه، سه‌راه، راهرو و پیچ استفاده شده در آزمایش..... ۹۷
شکل ۵-۶: روند یادگیری بدون ببلینگ (شکل راست پاداش لحظه‌ای و چپ متوسط متحرک را نشان می‌دهد.)..... ۹۸
شکل ۵-۷: یادگیری بدون ببلینگ در حضور نویز با واریانس ۰,۰۱ درصد (شکل راست پاداش لحظه-ای و چپ متوسط متحرک را نشان می‌دهد.) ۹۸
شکل ۵-۸: روند یادگیری همراه با ببلینگ (شکل راست پاداش لحظه‌ای و چپ متوسط متحرک را نشان می‌دهد.)..... ۹۹
شکل ۵-۹: متوسط تعداد مفاهیم شکل یافته برای ۱۰ تکرار از ورود تصادفی داده‌ها..... ۹۹
شکل ۶-۱: ساختار کلی مسئله..... ۱۰۵
شکل ۶-۲: داده‌های مصنوعی [Mer08] (دو سنسور تابع باور را روی سه کلاس a,h,r برای چهار شی گزارش می‌کنند.)..... ۱۱۱
شکل ۶-۳: یادگیری ضرایب تخفیف زمینه‌ای با معیار حد پایین مطلوبیت..... ۱۱۲
شکل ۶-۴: یادگیری ضرایب تخفیف زمینه‌ای با معیار حد بالای مطلوبیت..... ۱۱۲
شکل ۶-۵: روند کاهش خطا در طی یادگیری با GA پایه، بهترین جواب و میانگین جواب‌ها در هر نسل..... ۱۱۳

شکل ۶-۶: تولید داده‌های آموزش برای یادگیری ضرایب قابلیت اطمینان در ترکیب تصمیم در یادگیری مفهوم ۱۱۴
شکل ۶-۷: یادگیری برخط مفاهیم با یک سنسور که نشان می‌دهد بعد اول اطلاعات مرتبط بیشتری نسبت به بعد دوم داراست. ۱۱۴
شکل ۶-۸: یادگیری تخفیف کلاسیک با PBIL دودویی (خطا در هر نسل) ۱۱۵
شکل ۶-۹: تست ضرایب یادگیری تخفیف کلاسیک با PBIL دودویی ۱۱۵
شکل ۶-۱۰: یادگیری تخفیف زمینه‌ای با PBIL دودویی (خطا در هر نسل) ۱۱۶
شکل ۶-۱۱: تست ضرایب یادگیری تخفیف زمینه‌ای با PBIL دودویی ۱۱۶
شکل ۶-۱۲: یادگیری تخفیف زمینه‌ای با PBIL پیوسته (خطا در هر نسل) ۱۱۶
شکل ۶-۱۳: تست ضرایب یادگیری تخفیف زمینه‌ای با PBIL پیوسته ۱۱۷
شکل ۷-۱: ساختار یادگیری ۱۲۵
شکل ۷-۲: محیط PW [Sut96] ۱۴۸
شکل ۷-۳: محیط MC [Sut98] ۱۴۹
شکل ۷-۴: عملکرد در مقابل نویز سنسور (متوسط پاداش بر حسب افزایش قدرت نویز) ۱۵۵
شکل ۷-۵: عملکرد در مقابل احتمال شرایط مساعد سنسور (متوسط پاداش بر حسب افزایش احتمال شرایط مساعد) ۱۵۵
شکل ۷-۶: عملکرد در مقابل نویز افکتور (متوسط پاداش بر حسب افزایش قدرت نویز) ۱۵۶
شکل ۷-۷: مقایسه نتیجه یادگیری ضریب خوش‌بینی با تبدیل پیگنوستیک ۱۵۸
شکل ۷-۸: مقایسه نتیجه یادگیری ضریب خوش‌بینی با تبدیل پیگنوستیک در ۱۰۰۰ اپیزود دوم ۱۵۸
شکل ۷-۹: تکامل ضریب خوش‌بینی در طی یادگیری با ساختار ۲۰ ..... ۱۵۹
شکل ۷-۱۰: مقایسه نتیجه یادگیری ضریب خوش‌بینی با تبدیل پیگنوستیک ۱۵۹
شکل ۷-۱۱: مقایسه نتیجه یادگیری ضریب خوش‌بینی با تبدیل پیگنوستیک در ۱۰۰۰ اپیزود دوم ۱۵۹
شکل ۷-۱۲: تکامل ضریب خوش‌بینی در طی یادگیری با ساختار ۱۰ ..... ۱۶۰
شکل ۷-۱۳: متوسط مجموع پاداش و تعداد گام تا هدف در ۱۰۰۰ اپیزود اول و دوم یادگیری چندگامی ۱۶۰
شکل ۷-۱۴: عملکرد در مقابل نویز سنسور (متوسط پاداش بر حسب افزایش قدرت نویز) ۱۶۳
شکل ۷-۱۵: عملکرد در مقابل احتمال شرایط مساعد سنسور (متوسط پاداش بر حسب افزایش احتمال شرایط مساعد) ۱۶۳
شکل ۷-۱۶: عملکرد در مقابل نویز افکتور (متوسط پاداش بر حسب افزایش قدرت نویز) ۱۶۴
شکل ۷-۱۷: روبات e-puck و محیط ناوبری پیاده‌سازی ۱۶۵
شکل ۷-۱۸: همگرایی در محیط روبات واقعی بر حسب تعداد گام تا هدف ۱۶۵

شکل ۷-۱۹: مسیرهای رسیدن به هدف با سیاست حریصانه برای عامل یادگیری پیگنوستیک از دید دوربین..... ۱۶۶
شکل ۷-۲۰: مسیررسیدن به هدف با سیاست حریصانه برای عامل یادگیری خوشبین/بدبین از دید دوربین..... ۱۶۶
شکل ۷-۲۱: مسیرهای رسیدن به هدف با سیاست حریصانه از نقاط شروع نزدیک مانع مستطیلی برای عامل یادگیری پیگنوستیک از دید دوربین..... ۱۶۷
شکل ۷-۲۲: مسیرهای رسیدن به هدف با سیاست حریصانه از نقاط شروع نزدیک مانع مستطیلی برای عامل یادگیری خوشبین/بدبین از دید دوربین..... ۱۶۸
شکل ۷-۲۳: مسیرهای رسیدن به هدف عامل در فضای پیوسته موتوری با سیاست حریصانه با شروع از گوشه سمت چپ بالا از دید دوربین..... ۱۶۹
شکل ۷-۲۴: مسیرهای رسیدن به هدف عامل در فضای پیوسته موتوری با سیاست حریصانه با شروع از نقاط نزدیک مانع مستطیلی از دید دوربین..... ۱۶۹

فهرست جدول‌ها

عنوان .....صفحه

جدول ۷-۱: مقایسه سرعت همگرایی الگوریتم‌ها. عدد کمتر سرعت همگرایی بالاتر را نشان می‌دهد. ۱۵۲
جدول ۷-۲: متوسط پاداش الگوریتم‌ها در ۲۰۰۰۰ اپیزود اول ۱۵۲
جدول ۷-۳: متوسط پاداش الگوریتم‌ها در ۲۰۰۰۰ اپیزود دوم ۱۵۳
جدول ۷-۴: مقایسه سرعت همگرایی الگوریتم‌ها. عدد کمتر سرعت همگرایی بالاتر را نشان میدهد ۱۵۳
جدول ۷-۵: متوسط پاداش برای ۲۰۰۰۰ اپیزود اول ۱۵۳
جدول ۷-۶ : متوسط پاداش برای ۲۰۰۰۰ اپیزود دوم ۱۵۴
جدول ۷-۷: مقایسه سرعت یادگیری. عدد کوچکتر سرعت همگرایی بالاتر را نشان می‌دهد. ۱۶۱
جدول ۷-۸: مقایسه کیفیت یادگیری. متوسط پاداش در نصف اول ۱۶۱
جدول ۷-۹: مقایسه کیفیت یادگیری. متوسط پاداش در نصف دوم ۱۶۲

۱ مقدمه

یادگیری بدون شک هسته اصلی هوشمند بودن یک عامل[1] است. اما این که چه عاملی هوشمند است و منظور از یادگیری چیست نیاز به تعاریف دقیق آنها دارد و کلمات مبهم یادگیری و هوشمند بودن نمی‌توانند مفاهیم قابل درکی را پیش رو قرار دهند. طیف بسیار گسترده‌ای از شاخه‌های مختلف هوش مصنوعی به یادگیری در ساختارهای مختلف و تحت شرایط و فرضیات کاملا متفاوت می-پردازند. آنچه می‌توان به عنوان ویژگی مشترک اشاره کرد توسعه دانش عامل در طی فرآیند یادگیری و تبدیل یا نگهداری اطلاعات قابل دسترس به دانشی است برای هر چه بیشتر نزدیک شدن به یک هدف. پیچیدگی مسئله یادگیری عامل بنا به سطوح متفاوت کسب اطلاعات، نمایش دانش و ساختاری که توسعه داده می‌شود، قیودی است که در جهت حرکت به سمت چنین هدفی باید ارضا شود. پس تعیین هدف و قیود فیزیکی و محیطی عامل تا حد زیادی مسئله یادگیری آن را تعریف می‌کند.

یادگیری هوشمند ارکان مذکور را به عنوان شرایط لازم برای توسعه دانش و زندگی پویا می‌طلبد اما بدون شک شرایط کافی نیست. ما به عنوان یک واژه کلی برای ایضاح این مفهوم نیاز به یادگیری پایدار را برای عامل هوشمند بحث می‌کنیم. بر این اساس هدف را توسعه عامل‌های هوشمند مصنوعی‌ای می‌دانیم که تا آنجا که پیچیدگی مسئله اجازه می‌دهد در سطوح بالاتری از خودکاری[2] و مقید به قیودی سازگارتر با آنچه که در جهان واقع اتفاق می‌افتد به توسعه پایگاه دانش و نیل به هدف بپردازد. عاملی را در سطح بالایی از خودکاری می‌شناسیم که اولا تا حد بالایی بی‌نیاز از دانش اولیه و کد کردن اطلاعات مورد نیاز به صورت سخت‌افزاری و در نمایش‌های سطوح پایین دانش طراح بوده، ثانیا توانایی‌های تطبیق بالایی را در سطوح متفاوت پیچیدگی مسئله از خود نشان دهد. در این راستا چنین عاملی باید توانایی توسعه دانش از منابع مختلف کسب اطلاعات چه یادگیری خود و چه یادگیری از دیگران و توسعه اجتماعی در سطوح متفاوت داشته و مدل‌های یادگیری برای درک و نمایش اطلاعات و دانش را مطابق با مشکلات عملی ناشی از شرایط فیزیکی و محیطی بنا نهد. طراحی سیستم‌های تجرید دانش[3] و مدلسازی نایقینی[4] در فرآیندهای کسب و دستکاری اطلاعات و دانش، طرح پیشنهادی برای نزدیک شدن به این اهداف اولیه است.

در یک چنین عامل یادگیری، تجربیات عامل و نتیجه اعمال او تنها منبع کسب دانش سطح پایین در قالب درک سنسوری پیرامون می‌باشد و مفهوم‌سازی[5] ابزار طبیعی تجرید دانش بر اساس این اطلاعات را فراهم می‌سازد. یادگیری مفهوم[6] با نمایش مجرد نیاز به پردازش‌های سطح پایین عددی را مرتفع

[1] Agent
[2] Autonomy
[3] Knowledge Abstraction
[4] Uncertainty
[5] Conceptualization
[6] Concept Learning

ساخته و لذا با تسهیل انتقال دانش و ارتباطات[۱]، منبع دیگر توسعه دانش یعنی یادگیری اجتماعی[۲] را به عنوان یکی از ارکان اساسی یادگیری پایدار تسهیل می‌سازد. بر این اساس یادگیری تقویتی[۳] با استفاده از سیگنال‌های فیدبک در قبال تجربیات تعامل با پیرامون و یادگیری با تقلید[۴] با کسب دانش سطح بالا از دیگران ابزارهای اساسی توسعه چنین عاملی خواهند بود.

توسعه دانش بر اساس اطلاعات ادراکی[۵] محیط صورت می‌گیرد که کسب آنها از طریق سنسورهایی حاصل می‌شود که در اثر محدودیت‌های فیزیکی و مشکلات ساخت، اولا قادر به دریافت کلیه اطلاعات مرتبط[۶] نبوده و تنها تصویری از ابعاد محیط قابل دستیابی است، ثانیا ویژگی‌های درک شده در قالب اطلاعات عددی‌ای است که میزان دقت آنها به شدت تحت تاثیر نایقینی‌های محیطی است. عامل نه تنها ناچار است یادگیری را با سیستم‌های سنسوری نادقیق انجام دهد، همچنین سیستم موتوری آن قادر به اعمال اثر[۷]های مشخص شده به طور معین نمی‌باشد. این مشکلات باعث کسب دانش نادرست گردیده و به نوبه خود باعث افزایش خطا در مراحل بعدی در اثر استنباط عامل بر آن اساس می‌گردد.

توسعه سیستم یادگیری پایدار لازم دارد این مشکلات عملی در تمامی مراحل درک و کسب اطلاعات، نمایش دانش، استنباط و تصمیم‌گیری و در نهایت تصحیح پایگاه دانش مورد توجه بوده و نایقینی‌های تحمیل‌شده توسط آن مدلسازی شود. پیچیدگی مسئله زمانی است که عامل در محیط‌های واقعی با شکل‌های گوناگونی از منابع نایقینی ناشناخته درگیر است. توسعه سیستم یادگیری بدون مدلسازی شکل‌های مختلف ممکن نایقینی و منابع گوناگون کسب اطلاعات نادرست احتمال واگرایی یادگیری و دور شدن از سیستم یادگیری منعطف مقاوم را افزایش می‌دهد.

هرچند نظریه احتمالات[۸] به دلیل مدلسازی مهمترین منبع نایقینی یعنی تصادف[۹] همواره ساختارهای موفق عملی فراهم ساخته و به دلیل عملی[۱۰] بودن محاسبات استدلال در این قالب محبوبیت بالایی داشته است، متاسفانه تنها یک صورت نایقینی را مورد توجه قرار داده و در واقع می‌توان آن را تقریبی از مدل‌های منعطف‌تر و کامل‌تر نایقینی دانست که سایر صورت‌های نایقینی را نیز مورد توجه قرار می‌دهند. نظریه شواهد[۱۱] یکی از چنین مدل‌های غنی‌تر می‌باشد که صورت دیگری از نایقینی را نیز همراه با تصادف مدلسازی می‌کند. توسعه‌های این نظریه مدلسازی عدم قطعیت را با آدرس کردن

[1] Communication
[2] Social Learning
[3] Reinforcement Learning
[4] Imitation Learning
[5] Perceptual Information
[6] Relevant
[7] Effect
[8] Probability Theory
[9] Randomness
[10] Tractable
[11] Evidence Theory

مسائلی در نظر می‌گیرد که آن را ساختاری قدرتمند برای در نظر گرفتن قالب استنباط در سیستم-های یادگیری واقعی با شرایطی آنگونه که در اینجا بحث گردید فراهم می‌سازد.

ماهیت کلی ساختار مدلسازی نایقینی استدلال شهودی[1] را در فصل دوم با شروع از مفاهیم پایه‌ای آن شرح داده و با ارائه توسعه‌ها، تعابیر و کاربردهای آن مفاهیم مرتبط با مسئله مورد بحث در این پژوهش را در ساختاری منطقی ارائه داده و مقدمات درک لزوم توسل به یک چنین ساختار مدلسازی نایقینی را در مسئله یادگیری مفهوم آنگونه که در فصل سوم توصیف می‌شود فراهم می‌سازیم. در فصل سوم ساختار یادگیری مفهوم را با تعریف آن در یک خانواده کلی از مسائل یادگیری تعاملی[2] همراه با مسئله یادگیری تقویتی ارائه می‌دهیم. در این فصل پس از ارائه نمایش مفاهیم در سطوح مختلف برای تجرید دانش سطح بالا، مدل‌های یادگیری مفهوم را بررسی کرده و خط‌مشی در پیش گرفته شده در این تحقیق را توجیه می‌کنیم.

علاوه بر آن که سیستم ادراکی عامل‌های محیط‌های واقعی درگیر با نایقینی است، پیچیدگی دیگر برای یادگیری از آنجا ناشی می‌گردد که فضای مشاهدات عامل فضای پیوسته‌ای است که در هر لحظه او را با حجم عظیمی از اطلاعات عددی سنسوری حاصل از منابع گوناگون سنسوری مواجه می‌سازد. در نظر گرفتن این اطلاعات گسترده خارج از محدوده پردازشی عامل بوده و در اثر آن نیاز دارد چگونگی در نظر داشتن اطلاعاتی که بیشترین ارتباط را دارند یاد بگیرد. اهمیت این موضوع همچنین دلیل دیگری دارد که در نظر داشتن اطلاعات غیرمرتبط و عدم اعمال قابلیت اطمینان داده‌های حاصل از منابع مختلف اطلاعاتی سنسوری می‌تواند کیفیت یادگیری را کاهش دهد. قالب مدلسازی مفاهیم مرتبط در نظریه شواهد، مفاهیم مدلسازی قابلیت اطمینان سنسوری است که در سال‌های اخیر توسعه‌های قدرتمندتری از آن برای سیستم‌های ترکیب اطلاعات سنسوری در این قالب ارائه گردیده است. در فصل چهارم این مفاهیم را در قالب کلی مدلسازی قابلیت اطمینان[3] در ترکیب اطلاعات[4] ارائه داده و سپس با تاکید بر مدل‌های خاص قالب نظریه شواهد، خواص آنها را به طور کامل بررسی می‌کنیم. این مفاهیم در فصل ششم برای مسئله یادگیری ترکیب تصمیم[5] در قالب مدل یادگیری مفهوم ارائه شده در فصل پنجم مورد توجه قرار خواهد گرفت. آنچه که ما برای این مسئله مورد توجه داریم مسئله یادگیری این مدلسازی است که به ندرت در قالب نظریه شواهد آدرس گردیده است. با بررسی خواص مسئله مورد نظر، یادگیری با الگوریتم‌های تکاملی[6] را برای یادگیری مدلسازی‌های قابلیت اطمینان سنسوری نه فقط در ساختارهای سنتی آن بلکه در ساختارهایی کلی‌تر

[1] Evidential Reasoning
[2] Interactive Learning
[3] Reliability
[4] Data Fusion
[5] Decision Fusion
[6] Evolutionary Algorithms

مطابق با مسئله مورد بحث مورد توجه قرار داده و مدلسازی را بدین منظور انجام می‌دهیم. از بین الگوریتم‌های تکاملی، یک کلاس جدید و کارآمدتر نسبت به مدل‌های یادگیری سنتی الگوریتم‌های تکاملی را در نظر می‌گیریم که در فصل چهارم ارائه می‌گردد. ماهیت مسئله موردنظر یک توسعه از این الگوریتم را به فضاهای جستجوی پیوسته پیشنهاد می‌دهد که این توسعه و کاربرد آن در مسئله را در فصل ششم ارائه می‌دهیم و مزایا، معایب و خواص آن نسبت به الگویتم پایه استاندارد را بحث می-کنیم. پس از بررسی روش در پیش‌گرفته شده برای مسئله‌ای مجزا از مسئله موردنظر اما در جهت آن، تحلیل آزمایشی جامعی از آن را در فصل ششم برای مسئله ترکیب تصمیم در یادگیری مفهوم ارائه می‌کنیم.

فصل پنجم را می‌توان به عنوان فصل پایه‌ای برای توسعه شهودی ساختار مسئله یادگیری مفهوم دانست. در این فصل ساختار پیشنهادی را با جزئیات کامل بررسی کرده و در نهایت روند الگوریتمی آن را ارائه می‌دهیم. پس از بررسی عملکرد آن در قالب یک شبیه‌سازی ساده، برخی از توسعه‌های جانبی را نیز برای این ساختار مدنظر قرار داده و در این شبیه‌سازی بررسی می‌کنیم. یکی دیگر از مهمترین دستاوردها در این فصل ارائه قالب یادگیری برخط[1] نموی[2] برای بازنگری پایگاه دانش بر اساس تقویت در قالب نظریه شواهد است که آن را می‌توان اولین گام در جهت توسعه مدل‌های یادگیری تقویتی در قالب نظریه شواهد دانست. ساختار ارائه شده را به عنوان یک ساختار کلی به طور کامل در فصل پنجم بررسی کرده و آن را بر اساس نظریه بیز تعمیم‌یافته[3] توجیه می‌کنیم.

یکی دیگر از صورت‌های توسعه عامل‌های خودکار[4] را می‌توان امکان یادگیری ویژگی‌های مشترک محیط از ویژگی‌های بر اساس تکلیف[5] دانست. یادگیری مفهوم ساختار تجرید دانش را ارائه داده و مسئله تصمیم‌گیری[6] در آن تنها در سطوح تشخیص مفهوم در نظر گرفته می‌شود. یکی از چالش‌های عمده در یادگیری مفهوم چگونگی توسعه سیستم یادگیری برای یادگیری همراه با ویژگی‌های بر اساس تکلیف است که با فراهم آوردن یک چنین جداسازی[7]ای امکان استفاده مجدد[8] دانش کسب شده را فراهم سازد. این امر بدون شک یکی از گام‌های اساسی در جهت حرکت به سمت توسعه عامل‌های خودکار واقعی است و نه تنها در جامعه یادگیری مفهوم بلکه در بسیاری از ساختارهای یادگیری ماشینی و به طور خاص در مسائل یادگیری تقویتی یک چالش بزرگ و در عین حال مهم و

[1] Online Learning
[2] Incremental
[3] General Bayesian Theorem
[4] Autonomous
[5] Task-dependent
[6] Decision Making
[7] Decomposition
[8] Reusability

اساسی است. ما این مسئله را در قالب یادگیری تابع مطلوبیت[1] تصمیم در قالب یادگیری تقویتی برای سیستم یادگیری مفهوم شهودی مورد توجه قرار می‌دهیم. این نگاه به این مسئله لزوم ارائه یک توسعه شهودی از الگوریتم‌های سنتی یادگیری تقویتی را آشکار می‌سازد. در فصل هفتم این توسعه را بدون توجه به مسئله یادگیری مفهوم به عنوان توسعه‌ای از روش‌های پایه‌ای یادگیری تقویتی -که مفاهیم آن را در فصل سوم مرور می‌کنیم- به فضاهای پیوسته حالت و گسسته تصمیم و نیز به فضاهای پیوسته حالت و عمل ارائه داده خواص و دستاوردهای مهم آن را بررسی می‌کنیم.

علاوه بر شبیه‌سازی‌ها، بخش اعظمی از تحلیل‌های آزمایشی در این تحقیق یادگیری با روبات واقعی[2] را بررسی می‌کند که یک مسئله ناوبری[3] ساده در فصل هفتم مورد توجه قرار می‌گیرد. سنجش قالب یادگیری مفهوم برای مسئله موضعی کردن [4] و نمایش نگاشت توپولوژیکی گسسته[5] محیط در یک روبات متحرک[6] در فصل پنجم با جزئیات ارائه می‌گردد.

در پایان، فصل هشتم را به جمع‌بندی مفاهیم و دستاوردهای ارائه شده، بحث در روش‌های در پیش-گرفته شده و در نهایت نتیجه‌گیری و پیشنهادات کارها و تحقیقات مهم و گسترده آینده پیش روی خط ارائه شده به مسائل یادگیری تعاملی اختصاص می‌دهیم.

[1] Utility Function
[2] Real Robot
[3] Navigation
[4] Localization
[5] Discrete Topological Map Representation
[6] Mobile Robot

۲ نظریه شواهد: مدل‌ها، کاربردها و توسعه‌ها