

تلفیق همبافت و یادگیری عاطفی تفاوت زمانی در مهندسی کنترل

جواد عبدی (کارشناس ارشد)

فرزان رشیدی (کارشناس ارشد)

کارو لوکس (استاد)

دانشکده‌ی فنی، گروه مهندسی برق و کامپیوتر، دانشگاه تهران

علی خاکی صدیق (استاد)

دانشکده‌ی فنی، گروه مهندسی برق و کامپیوتر، دانشگاه خواجه نصیرالدین طوسی

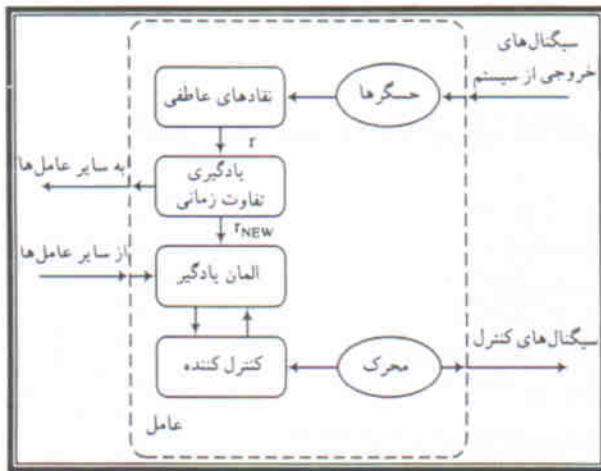
یکی از مهم‌ترین مشکلاتی که در سیستم‌های تأخیردار و غیرخطی وجود دارد تأمین اهداف چندگانه به‌طور هم‌زمان و در بهترین شرایط است. در این نوشتار برای کنترل سیستم‌های با اهداف چندگانه روشی ارائه شده است. اصول این روش مبتنی بر همبافت بوده و دارای ساختار فازی - عصبی با قابلیت یادگیری عاطفی تفاوت زمانی است. روش پیشنهادی قادر است سیستم را با توجه به موقعیت فعلی، عملکرد آن در زمان‌های گذشته و اهداف کنترلی موجود، به‌گونه‌ی کنترل کند که این اهداف در حداقل زمان و به‌نحو بسیار مطلوبی برآورده شوند. برای بیان واضح‌تر مسئله و اثبات عملکرد مناسب روش پیشنهادی فوق، دو مسئله‌ی معروف در علم کنترل که تأمین اهداف گوناگون آن به بهترین وجه در کم‌ترین زمان با روش‌های کنترلی دیگر بسیار مشکل است، به‌راحتی پیاده‌سازی شده است.

مقدمه

یکی از کارهایی که هر موجود هوشمندی دائماً در حال انجام آن است فرایند استدلال است. در عین حال یکی از مهم‌ترین مراحل هر استدلالی «مرحله‌ی نگاشت» است که در طی آن الگوی ورودی با الگوهای موجود در حافظه‌ی دائم عامل هوشمند تطبیق داده می‌شود تا شبیه‌ترین الگوها به‌عنوان پاسخ برگزیده شوند. لازمی این کار آن است که تمامی تجارب پیشین عامل هوشمند در حافظه‌ی فعال او وجود داشته باشند. علاوه بر این باید عامل هوشمند در یک زمان بسیار محدود قادر به مقایسه‌ی الگوی ورودی با تمامی تجارب پیشین باشد. از سوی دیگر یک عامل هوشمند واقعی معمولاً در یک محدوده‌ی زمانی مشخص دارای چندین هدف فعال است که گاهاً بعضی از این اهداف ممکن است در تضاد با هم باشند. آنچه در این میان ضروری به نظر می‌رسد این است که عامل هوشمند باید از بین این اهداف زیرمجموعه‌ی را انتخاب کند که با توجه به موقعیت فعلی مناسب‌ترین گزینه بوده و در حداقل زمان تمامی اهداف نیز به‌نحو مطلوبی برآورده شوند. برای رسیدن به این منظور، رفتار عامل هوشمند باید مبتنی بر همبافت باشد.^[۱] همبافت به‌عنوان وضعیت فعلی محیط و نیز خود عامل، می‌تواند از طریق دخالت فرایند عامل هوشمند را در دست‌یابی به اطلاعات کامل محیطی و پیش‌بینی وضعیت آینده استدلال یاری کند. از دیدگاه هوش مصنوعی، هر پیکربندی قابل تشخیص از مشخصات (اعم از مشخصات محیط و

عامل)، که قابلیت پیش‌بینی عامل را از محیط (به‌منظور برآورد و ساختن اهداف خود) افزایش دهد «همبافت» نامیده می‌شود.^[۲-۳] با توجه به این تعریف، همبافت باعث می‌شود که عامل هوشمند با توجه به موقعیت فعلی، عملکرد سیستم در زمان گذشته، مجموعه اهدافی که باید برآورده شوند و میزان برآورده شدن این اهداف تاکنون، قادر به در نظر گرفتن اولویت‌هایی محلی برای هر یک از اهداف شوند تا اینکه عملکرد سیستم در حداقل زمان ممکن بهینه شود. این اولویت‌بندی معمولاً به‌صورت وزن دادن به مجموعه اهدافی است که باید برآورده شوند. منظور از اولویت‌های محلی این است که وزن‌هایی که به اهداف تخصیص داده می‌شود براساس شرایط فعلی سیستم و میزان برآورده شدن هر یک از اهداف بوده و در هر لحظه‌ی زمانی این وزن‌ها به روز درآورده می‌شوند تا این‌که تمامی این اهداف در حداقل زمان به‌نحو مطلوبی برآورده شوند. در حالت کلی، همبافت به چند طریق می‌تواند بر رفتار عامل هوشمند تأثیر بگذارد:^[۴]

۱. همبافت باعث افزایش قابلیت شناسایی و پیش‌بینی عاملی (از طریق توجه به اطلاعات همبافتی) می‌شود.
۲. همبافت به عامل هوشمند کمک می‌کند تا برای رسیدن به اهداف خود، بسته به موقعیت جاری، بر بعضی از اهداف تمرکز و توجه بیشتری داشته باشد.
۳. همبافت با تأثیر بر پارامترهای رفتاری عامل هوشمند (از طریق دخالت در فرایند استدلال)، باعث می‌شود عامل هوشمند بتواند در



شکل ۱. ساختار یک عامل کنترلی.

جدول ۱. نمونه‌هایی از اجزای مورد استفاده در ساختار یک عامل کنترلی.

عنصر	روش مورد استفاده
کنترل‌کننده	سیستم فازی، سیستم فازی - عصبی
تقاد	سیستم خبره، شبکه عصبی، سیستم فازی، سیستم کلاسیک
عنصر یادگیری	شبکه عصبی، یادگیری تقویتی، یادگیری عاطفی، یادگیری عاطفی تفاوت زمانی

می‌آورد، به طوری که تلاش آن در راستای بهینه‌سازی عملکرد عامل انجام خواهد شد. مبنای عملکرد عنصر یادگیری سیگنالی است که به وسیله واحد نقاد در اختیار آن قرار می‌گیرد. نقاد عهده‌دار وظیفه نقد و بررسی سیستم است، و با توجه به نحوه عملکرد عامل محلی مربوطه، سیگنال مناسب را که حاکی از عملکرد خوب یا بد سیستم کنترل است، فراهم می‌آورد. این سیگنال، مبنای کار عنصر یادگیری در به روزآوری ساختار کنترل‌کننده خواهد بود. سایر عامل‌ها نیز می‌توانند از این سیگنال در به روزآوری دانش ذخیره شده در خود استفاده کنند، در جدول ۱ نمونه‌ی از روش‌هایی که می‌توانند جایگزین اجزای تشکیل‌دهنده هر عامل شوند، آمده است. به دلیل طبیعت غیرقطعی محیط، ساختار کنترل‌کننده به صورت فازی - عصبی انتخاب شده است. عنصر یادگیر مورد استفاده در این نوشتار نیز از قابلیت یادگیری عاطفی تفاوت زمانی برخوردار است. [۷-۵]

یادگیری تقویتی

قبل از توصیف یادگیری عاطفی تفاوت زمانی تحلیل یادگیری تقویتی، به عنوان پایه و اساس یادگیری عاطفی، مورد نیاز است.

شرایط مختلف به گونه‌ی عمل کند که تمامی اهدافی که بر عهده دارد، به نحو مطلوبی برآورده شوند.

۴. همبافت باعث می‌شود عامل هوشمند بتواند در مورد حوادث پیش‌بینی نشده‌ی که ممکن است با آن مواجه شود (مثلاً وارد شدن نویز به سیستم)، با عکس‌العملی سریع و مناسب سیستم را به حالت مطلوب خود هدایت کند.

تعریف خوبی که می‌توان برای همبافت در مسأله در نظر گرفت چنین است:

«همبافت عبارت است از مجموعه شرایطی که در آن عامل هوشمند به حل مسأله می‌پردازد. این شرایط اعم است از اهداف عامل هوشمند، وضعیت محیطی، وضعیت عامل هوشمند قبل از ارائه‌ی مسئله و مواردی از این قبیل.»

به طور کلی می‌توان گفت دو نوع همبافت از این دیدگاه وجود دارد:

۱. همبافت بیرونی^۱ که عبارت است وضعیت جاری محیط از دید عامل هوشمند.

۲. همبافت درونی^۲ از وضعیت جاری ذهن عامل هوشمند، اهداف و دانش عامل هوشمند.

در ساختار کلی نوشتار ابتدا کنترل مبتنی بر عامل هوشمند مورد بررسی قرار گرفته است، و در ادامه به موضوع یادگیری عاطفی به عنوان یک روش کنترلی هوشمند پرداخته شده است. سپس، با توجه به تعریف ارائه شده از همبافت و مفهوم یادگیری عاطفی، کنترل عاطفی تفاوت زمانی مبتنی بر همبافت مورد بحث قرار گرفته و به عنوان یک نتیجه‌ی تطبیقی، روشی برای کنترل سیستم‌های دارای اهداف چندگانه پیشنهاد شده است. در ادامه برای بررسی قابلیت و عملکرد ساختار پیشنهادی، مسائل کنترلی معروف در علم کنترل، نظیر جرقه‌ی سقفی و آونگ و اروونه بر روی گاری، برای شبیه‌سازی انتخاب شده و در انتها نیز نتیجه‌گیری از مباحث ارائه شده است.

کنترل مبتنی بر عامل هوشمند

شکل ۱ ساختار عامل کنترلی مورد استفاده را نشان می‌دهد. عامل از طریق حسگرهایش، سیگنال‌های خروجی سیستم را حس کرده، تصویری از وضعیت کنترلی محول شده به خود را به دست می‌آورد و از طریق محرك، سیگنال کنترلی لازم را به سیستم اعمال می‌کند. کنترل‌کننده نیز عهده‌دار وظیفه‌ی نگاشت این سیگنال‌های ورودی به سیگنال‌های خروجی ذکر شده است. عنصر یادگیری، دانش ذخیره شده در کنترل‌کننده را با توجه به شاخص عملکرد خارجی به روز

کنترل‌کننده به صورت یک سیگنال پیروزی یا شکست. علی‌رغم این کمبود اطلاعات، سیستم کنترل باید بداند که چگونه این سیستم ناشناخته را از حالت فعلی به حالت مطلوب مورد انتظار برساند. کنترل‌کننده با سعی و خطا در فضای جستجو و دریافت سیگنال پاداش (یا تنبیه) برای عملی که انتخاب کرده است باید چنان عمل کند که سیگنال پاداش بیشینه شود. این در حالی است که با اینکه به سیستم گفته نمی‌شود که چه عملی را برگزیند خود سیستم باید با امتحان کردن اعمال گوناگون، عملی را که بیشترین میزان پاداش را در پی خواهد داشت انتخاب کند. اگر عملی انتخابی پاسخ مناسبی به دست دهد پاداشی به آن تعلق می‌گیرد تا احتمال اینکه عمل تکرار شود و در نتیجه سیستم پاداش بیشتری دریافت کند افزایش یابد. از طرف دیگر اگر سیستم به یک حالت نامطلوب برسد برای آن جریمه‌ی در نظر گرفته می‌شود. بنابراین در صورتی که سیستم به یک حالت مطلوب برسد تمایل به ایجاد دوباره‌ی آن حالت افزایش خواهد یافت. (۶-۱۲)

کنترل عاطفی تفاوت زمانی مبتنی بر همبافت

در کنترل عاطفی تفاوت زمانی مبتنی بر همبافت، اساس یادگیری وجود عامل‌های عاطفی نظیر هیجان و اضطراب در سیستم کنترل است. در این صورت از بین حالت‌های سیستم، حالت‌هایی که طراح نسبت به آنها حساسیت دارد به صورت پارامترهای همبافتی تعریف می‌شود که سبب ایجاد اضطراب در سیستم می‌شوند؛ و سیستم کنترل باید به گونه‌ی عمل کند که اضطراب نسبت به این پارامترهای همبافتی کاهش یابد. شکل ۲ ساختار کلی کنترل‌کننده‌های عاطفی تفاوت زمانی مبتنی بر همبافت را نشان می‌دهد. در ادامه، وظایف هر بخش از این ساختار به تفصیل آمده است.

الف) ساختار نقاد (عاطفه)

تعریف نقاد و طراحی ساختار آن، مستقیماً به بخشی از سیستم کنترلی که نقاد وظیفه‌ی نقد آن را عهده گرفته است بستگی دارد. در این مقاله، چنین فرض می‌کنیم که نقاد تنها وظیفه‌ی ارزیابی خروجی را بر عهده دارد و از این رو آن را نقاد خروجی می‌نامیم. این نقاد، پس از ارزیابی خروجی، سیگنال عاطفی r را تولید می‌کند که پیوسته بوده و هر مقداری را بین -1 و $+1$ می‌تواند اختیار کند. به طوری که $r=+1$ (یا $r=-1$) نشان‌دهنده‌ی شکست کامل کنترل‌کننده بوده و هرچه سیگنال عاطفی به صفر نزدیکتر باشد، مؤید آن است که تلاش کنترلی موفقیت آمیزتر بوده است. اساس یادگیری نیز بر این مبنا استوار است که

یادگیری تقویتی نتیجه‌ی تحقیقات دانشمندان در زمینه‌ی رفتار انسان و حیوانات در زمینه‌ی یادگیری است که آن را «آموزش از طریق پاداش و جزا» نیز می‌نامند. در سیستم‌های کنترل مبتنی بر یادگیری تقویتی، آموزش کنترل‌کننده مبتنی بر یک سیگنال تقویتی دودویی r است که معمولاً یکی از مقادیر آن (مثلاً $r=0$) به منزله‌ی موفقیت در امر کنترل و مقدار دیگر r (مثلاً $r=-1$) به معنای شکست در امر کنترل است. هدف از طراحی سیستم کنترل نیز آن است که مجموع سیگنال r متناسب با پیروزی، بیشینه شود که به منظور دست‌یابی به این هدف، پارامترهای کنترل‌کننده نیز در این جهت تغییر داده خواهند شد. مسئله‌ی که در یادگیری تقویتی مطرح می‌شود این است که سیگنال تقویتی، پس از مدت زمان طولانی و پس از اجرای رشته‌ی از اعمال به منظور ارزیابی عملکرد سیستم در دسترس قرار می‌گیرد که این امر منجر به عدم یادگیری سریع سیستم می‌شود. بدین منظور طراح مجبور است یک سیگنال تقویت داخلی، که در واقع پیش‌بینی شده‌ی سیگنال تقویت خارجی است، تولید کند؛ به عبارت دیگر در یادگیری تقویتی امکان وجود پیش‌بینی برای سیستم در نظر گرفته شده است. اگرچه این پیش‌بینی دقیق نیست بعد از مدت زمانی معین عملکرد کلی سیستم به سمت بهتر شدن می‌رود. در حالت کلی برای این‌که بتوان سیستم را به صورت لحظه‌ی مورد ارزیابی قرار داد معمولاً از روش یادگیری عاطفی استفاده می‌شود. (۸-۱۰)

در ادامه، برای کارایی بهتر به تلفیق ایده‌ی یادگیری عاطفی با یادگیری تفاوت زمانی، به عنوان یک روش مناسب برای پیش‌بینی و ارزیابی لحظه‌ی خواهیم پرداخت.

یادگیری عاطفی تفاوت زمانی

عواطف حالاتی هستند که به وسیله‌ی پاداش یا تنبیه کسب می‌شوند.^[۱۱] پاداش چیزی است که یادگیرنده سعی در کسب آن دارد، و تنبیه چیزی است که یادگیرنده از آن می‌گریزد. یکی از پدیده‌های مهمی که در یادگیری انسان‌ها و سایر موجودات به چشم می‌خورد اثر عامل‌هایی نظیر هیجان، عاطفه، استرس و احساس در یادگیری است. انسان به عنوان یک موجود هوشمند همواره در رفتارهای خود سعی در کاهش استرس‌های موجود دارد همچنین همواره سعی دارد نحوه‌ی عملکردش چنان باشد که احساس رضایت بخشی داشته باشد. در مبحث یادگیری عاطفی در سیستم‌های کنترل چنین فرض می‌شود که مدل سیستم برای کنترل‌کننده شناخته شده نیستند و تنها اطلاعاتی که در دسترس قرار دارند عبارت‌اند از حالت‌های سیستم و پس‌خوری از میزان عملکرد

بزرگ)، PM (مثبت متوسط)، PS (مثبت کوچک)، ZE (صفر)، NS (منفی کوچک) NM (منفی متوسط) و NL (منفی بزرگ) هستند. همچنین ورودی لایه‌ی چهارم که لایه‌ی فازی زدا نام دارد، خروجی لایه‌ی سوم است. برای فازی زدایی می‌توان از روش‌های مختلفی استفاده کرد. که در این نوشتار از قاعده‌ی مرکز ثقل استفاده شده است.

ج) شبکه همبافت فازی

وظیفه‌ی این بخش وزن‌دهی به سیگنال‌های کنترلی تولیدشده توسط کنترل‌کننده‌های فازی است. بدین منظور مجموعه‌ی اهدافی را که باید برآورده شوند، به صورت پارامترهای همبافتی در نظر می‌گیریم. شبکه‌ی همبافت فازی در هر لحظه‌ی زمانی، با توجه به موقعیت فعلی سیستم و میزان برآورده شدن اهداف، به هر یک از این اهداف اولویت‌هایی را تخصیص می‌دهد. در مرحله‌ی بعد با توجه به اولویت‌های در نظر گرفته شده برای اهداف، وزن‌هایی به هر یک از آنها داده خواهد شد. موقعیت فعلی سیستم و میزان برآورده شدن اهداف، از روی سیگنال‌هایی که از محیط به این شبکه فیدبک داده می‌شوند به دست می‌آیند. این سیگنال را می‌توان چنین تعریف کرد:

$$E_i = \text{میزان برآورده شدن هدف } i \text{ ام تا لحظه } t \text{ ام} \quad i=1, \dots, n$$

سپس این وزن‌ها با سیگنال‌های کنترلی تولیدشده از کنترل‌کننده‌های فازی ترکیب شده، سیگنال کنترلی نهایی برای اعمال به محیط، طبق رابطه‌ی ۱ محاسبه می‌شود:

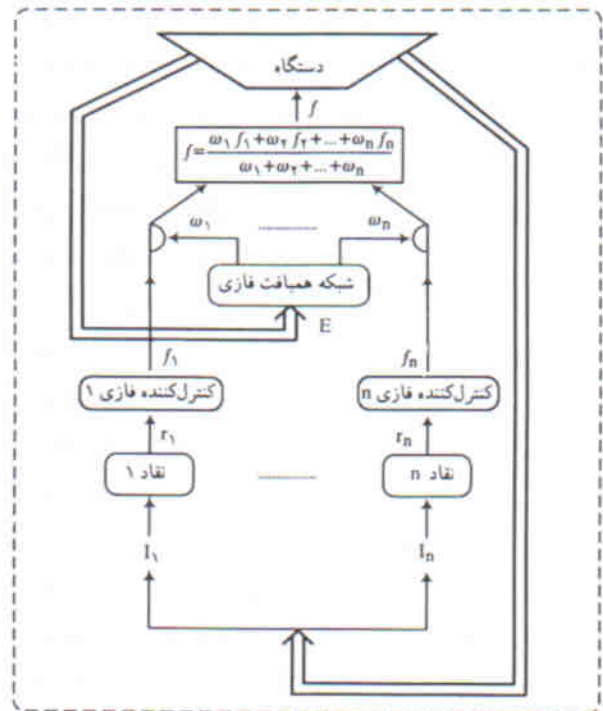
$$f = \frac{f_1 \omega_1 + f_2 \omega_2 + \dots + f_n \omega_n}{\omega_1 + \omega_2 + \dots + \omega_n} \quad (1)$$

فرض می‌کنیم که $G = \{g_1, g_2, \dots, g_n\}$ مجموعه‌ی اهدافی باشد که می‌خواهیم برآورده شوند. همچنین فرض کنیم که $E = \{E_1, E_2, \dots, E_n\}$ بیانگر سیگنال‌های فیدبک محیطی باشد. شبکه‌ی همبافت فازی، با توجه به سیگنال‌های محیطی و پارامترهای همبافتی، به هر یک از این اهداف اولویت‌هایی را تخصیص می‌دهد. مثلاً در مسئله‌ی کنترل جرثقیل سقفی پارامترهای همبافتی (مجموعه اهداف) را می‌توان به صورت زیر تعریف کرد:

«کم کردن زاویه‌ی نوسان مجموعه واگن - بار تا حد ممکن»: g_1

«کشیدن مجموعه واگن - بار به نقطه معین»: g_2

در این صورت با توجه به سیگنال‌های E_1 و E_2 قواعد فازی وزن‌های ω_1 و ω_2 (متناظر با اهداف g_1 و g_2) را می‌توان به صورت جدول ۲ در نظر گرفت، که در آن برچسب‌های HP، MP و LP به ترتیب متناظر با «اولویت بالا»، «اولویت متوسط» و «اولویت پایین» بوده و برچسب‌های M، B و S نیز به معنی «بزرگ»، «متوسط» و «کوچک»



E: سیگنال حالت کنترلی برای آیین هدف کنترلی

I: بردار حالت ورودی

I_i: بردار ورودی آیین نقاد

E_i: سیگنال عاطفی خروجی آیین نقاد

f_i: آیین خروجی کنترل‌کننده‌ی منطقی فازی

ω_i: آیین وزن همبافت فازی - عصبی

شکل ۲. ساختار کلی کنترل‌کننده‌ی عاطفی تفاوت زمانی مبتنی بر همبافت.

سیستم یادگیرنده پارامترهای خود را به نحوی تغییر دهد که همواره رضایت نقاد حاصل شود. به عبارت دیگر اضطراب نقاد نسبت به عاطفه‌ی مربوط به آن باید کاهش یابد.^[۹]

ب) کنترل‌کننده‌ی فازی - عصبی

کنترل‌کننده‌ی فازی - عصبی مورد استفاده در این نوشتار در اصل یک کنترل‌کننده‌ی فازی است که به وسیله‌ی شبکه‌ی عصبی تحقق یافته است و چهار لایه دارد. در لایه‌ی اول که لایه‌ی ورودی نام دارد، هیچ عمل محاسباتی صورت نمی‌گیرد و تمام ورودی‌های این لایه مستقیماً به لایه‌ی بعدی منتقل می‌شود. در لایه‌ی دوم فرایند فازی‌سازی با پنج برچسب برای متغیر ورودی {PL (مثبت بزرگ)، PS (مثبت کوچک)، ZE (صفر)، NS (منفی کوچک) و NL (منفی بزرگ)} صورت می‌پذیرد. ورودی این لایه، سیگنال عاطفی حاصل از نقاد است. در لایه‌ی سوم عمل استنتاج صورت می‌پذیرد. خروجی‌های این لایه دارای هفت برچسب به صورت {PL (مثبت

نتایج شبیه‌سازی

در این قسمت به منظور نشان دادن قابلیت و عملکرد ساختار پیشنهادی، دو مسئله‌ی کنترلی مهم را مورد تحلیل و ارزیابی قرار می‌دهیم:

۱. کنترل جرثقیل سقفی: جرثقیل سقفی را معمولاً برای انتقال بار از یک نقطه به نقطه‌ی دیگر مورد استفاده قرار می‌دهند. در موقع انتقال، باید بار بدون کم‌ترین نوسان به مقصد برسد. شمای کلی سیستم در شکل ۴ نشان داده شده است.

تنها سیگنال کنترلی سیستم، نیروی f (برحسب نیوتن) است که به واگن اعمال می‌شود. معادلات دینامیکی حاکم بر سیستم توسط رابطه‌ی ۳ مشخص می‌شوند. [۱۵]

$$R\ddot{\theta} + D\cos\theta + g\sin\theta = \dots$$

$$(M+m)\ddot{D} + c_d\dot{D} - mR(\dot{\theta})^2 + mR\ddot{\theta}\cos\theta = f \quad (2)$$

که θ (برحسب درجه) میزان انحراف نسبت به خط عمود است و D (برحسب متر) بیانگر فاصله‌ی واگن از مبدأ است. همچنین:

$$g: \text{شتاب جاذبه‌ی زمین} = 9.8 \text{ m/s}^2$$

$$m: \text{جرم بار} = 5000 \text{ kg}$$

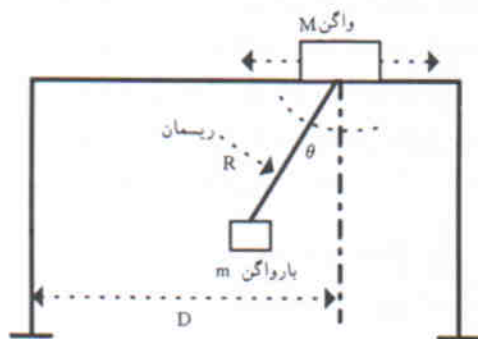
$$M: \text{جرم واگن} = 5000 \text{ kg}$$

$$R: \text{طول طناب} = 5 \text{ m}$$

$$c_d: \text{ضریب اصطکاک} = 500 \text{ kg/s}$$

محدوده‌ی تغییرات متغیرها عبارت‌اند از:
 $-12^\circ < \theta < 12^\circ, -10^6 \text{ N} \leq F \leq 10^6 \text{ N}$
 معادله‌ی ۲ را به فرم فضای حالت تبدیل می‌کنیم.

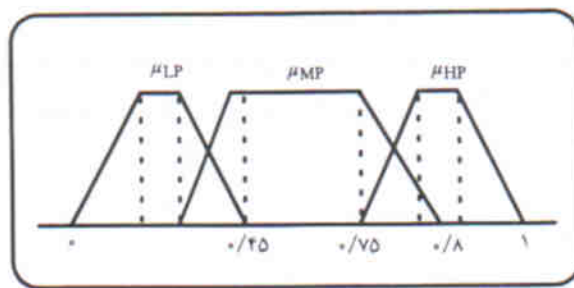
در روابط فوق $\Delta = 0.2 \text{ Sec}$ زمان نمونه برداری و حالت سیستم است. به منظور اعمال ساختار پیشنهادی برای کنترل جرثقیل، پارامترهای همبافتی (مجموعه اهداف) را به این صورت در نظر می‌گیریم:



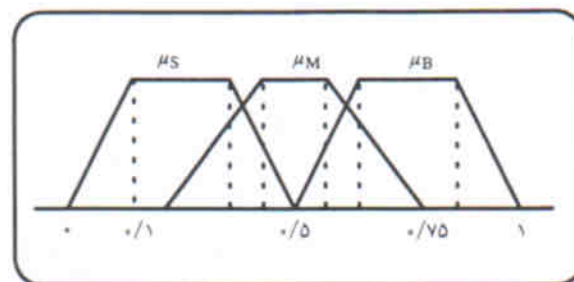
شکل ۴. شمای کلی سیستم جرثقیل سقفی.

جدول ۲. قواعد متناظر با وزن‌های ω_1 و ω_2

E_1/E_2	ω_1			ω_2		
	HP	MP	LP	HP	PM	LP
HP	B	M	S	B	B	B
MP	B	M	S	M	M	M
LP	B	M	S	S	S	S



سیگنال‌های E_1 و E_2



سیگنال‌های ω_1 و ω_2

شکل ۳. توابع عضویت متناظر با سیگنال‌های ورودی.

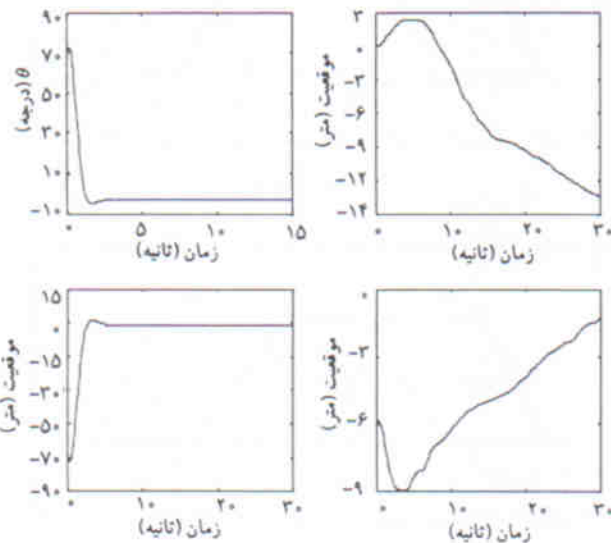
هستند. شبکه‌ی همبافت فازی مورد استفاده در این مقاله، یک سیستم فازی است که به وسیله‌ی شبکه‌ی عصبی تحقق یافته و دارای چهار لایه است. در لایه‌ی اول تمام ورودی‌ها به بازه‌ی $[-1, +1]$ نگاشت داده می‌شوند. ورودی‌های این لایه همان سیگنال‌های فیدبک محیطی‌اند که بعد از نرمالیزه شدن، به لایه دوم منتقل می‌گردند. در لایه‌ی دوم عمل فازی‌سازی با سه برجسب HP، MP و LP برای هر متغیر ورودی صورت می‌پذیرد. توابع عضویت متناظر با این ورودی‌ها در شکل ۳ نشان داده شده است. در لایه‌ی سوم سیستم به تصمیم‌گیری می‌پردازد و براساس خروجی‌های لایه‌ی قبل، وزن‌های متناظر با هر یک از اهداف کنترلی (ω_1 ها) را در هر لحظه‌ی زمانی محاسبه می‌کند. در لایه‌ی چهارم نیز عمل فازی‌زدایی انجام می‌شود. در این نوشتار برای فازی‌زدایی از قاعده‌ی مرکز ثقل استفاده شده است. [۱۳ و ۱۴]

$$r_r(t) = \frac{x - x_{\text{desired}}}{x_{\text{desired}}} \quad \text{if } |x - x_{\text{desired}}| \leq \bullet \quad (6)$$

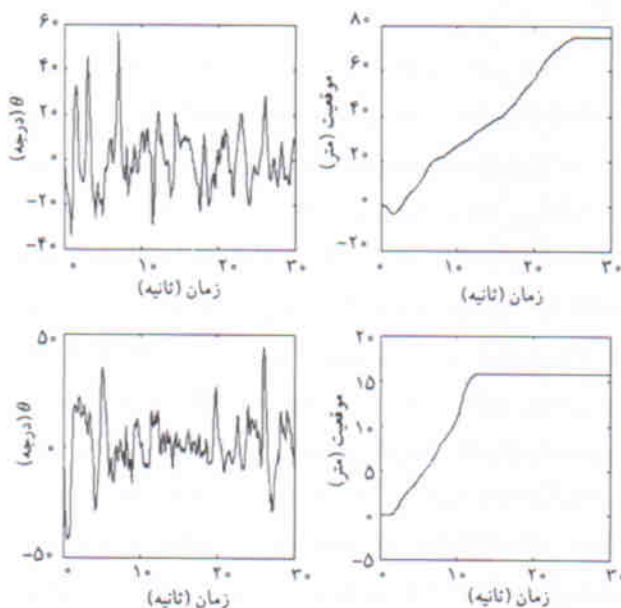
قواعد هر یک از دو کنترل کننده فازی به صورت رابطه ۷ انتخاب شده‌اند:

$$L_{f_n} \text{ is } f_n \text{ Then } L_{r_n} \text{ is } r_n \text{ If} \quad (7)$$

که در آن سیگنال عاطفی تولید شده از نقاد m ، و f_n سیگنال کنترلی متناظر با این سیگنال عاطفی است که توسط کنترل کننده فازی m حاصل شده است. نتایج شبیه سازی به ازای حالت الف و دو شرایط اولیه مختلف در شکل ۵ آمده است. همان طور که در شکل مشاهده می شود، ساختار پیشنهادی قادر بوده است در کم ترین زمان



شکل ۵. نتایج شبیه سازی برای حالت الف.



شکل ۶. نتایج شبیه سازی برای حالت ب ($x_{\text{Desired}} = 70m, 10m$).

«کم کردن زاویه ی نوسان مجموعه ی واگن و بار تا حد ممکن» g_1 :
«کشیدن مجموعه ی واگن و بار به نقطه ی معین» g_2 :

$$x_1(t+1) = x_1(t) + \Delta \left\{ \frac{x_r(t)}{R} - \frac{x_p(t)}{R} \cos x_1(t) \right\}$$

$$x_r(t+1) = x_r(t) + \Delta \left\{ g_2 \sin x_1(t) - \frac{1}{R} x_r(t) x_p(t) \sin x_1(t) - \frac{1}{R} x_p^2(t) \sin x_1(t) \cos x_1(t) \right\}$$

$$x_r(t+1) = x_r(t) + \Delta x_r(t),$$

$$x_p(t+1) = x_p(t) + \frac{\Delta}{m+M} \left\{ f - c_d x_p(t) + mR \frac{m \cos^2 x_1(t)}{m+M} \right\}$$

$$\left[\frac{x_r(t)}{R} - \frac{x_p(t)}{R} \cos x_1(t) \right]^2 \sin x_1(t) + m g_2 \sin x_1(t) \cos x_1(t),$$

$$x_2(t+1) =$$

$$x_2(t) + \Delta \left\{ - (x_p(t+1) - x_p(t)) \cos x_1(t) - g_2 \sin x_1(t) \right\} / R, \quad (3)$$

بدین منظور، سیگنال های فیدبک محیطی را به صورت رابطه ی ۴ تعریف می کنیم:

$$E_1 = x_1(t) / 12^\circ$$

$$E_2 = (x_r(t) - x_{\text{desired}}) / x_{\text{desired}} \quad (4)$$

که E_1 و E_2 به ترتیب نشانگر میزان برآورده شدن اهداف اول (g_1) و دوم (g_2) هستند. با توجه به رابطه ی ۴ ملاحظه می شود که محدوده ی تغییرات سیگنال E_1 در ناحیه ی $[-1, +1]$ بوده و محدوده ی تغییرات سیگنال E_2 در ناحیه ی $[-1, 0]$ است. در صورتی اهداف کنترلی به طور کامل برآورده می شوند که E_1 ها برابر صفر شوند. شبیه سازی برای سه حالت زیر انجام شده است:

الف) صفر کردن زاویه ی نوسان مجموعه ی واگن - بار بدون توجه به موقعیت آن (g_1);

ب) کشیدن مجموعه ی واگن - بار به یک نقطه ی مشخص در کم ترین زمان ممکن، بدون توجه به زاویه ی نوسان (g_2);

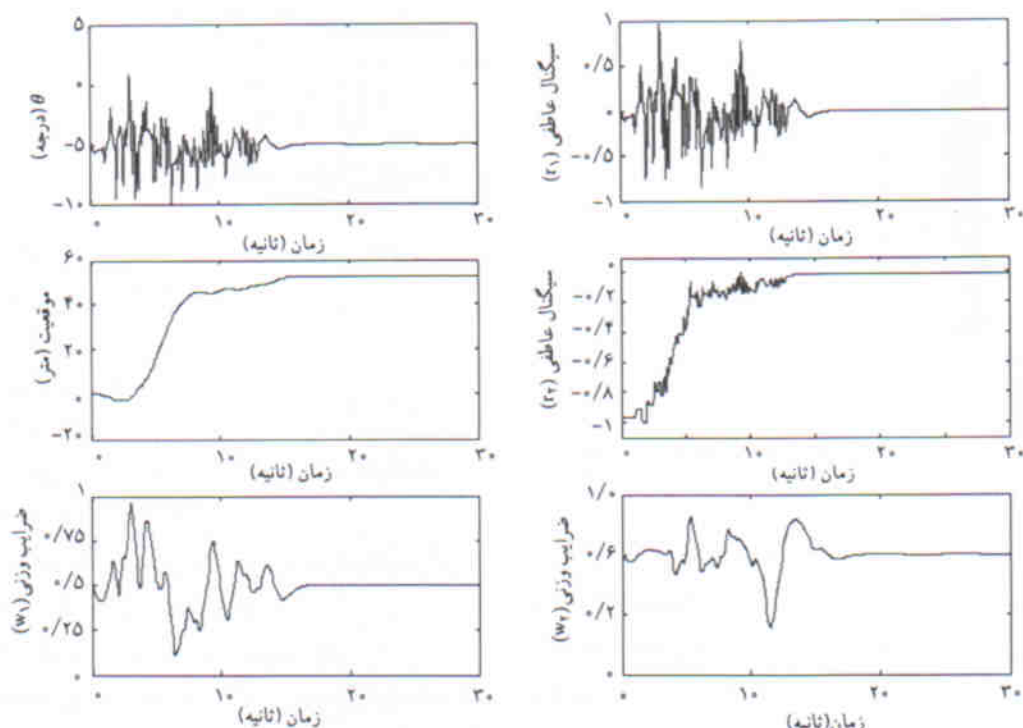
ج) برآورده شدن هر دو هدف کنترلی g_1 و g_2 در کم ترین زمان ممکن. در این شبیه سازی ها، تعداد نقاط را برابر ۲ فرض کرده ایم که وظیفه ی نقاد اول ارزیابی هدف اول (g_1) و وظیفه ی نقاد دوم نیز ارزیابی هدف دوم (g_2) است. سیگنال های عاطفی متناظر با هر یک از نقادها به صورت رابطه ی ۵ و ۶ تعریف شده اند:

$$r_1(t) = \begin{cases} \frac{x_1}{12} & \text{if } |x_1| \leq 12 \\ -1 & \text{if } x_1 < -12 \\ +1 & \text{if } x_1 > 12 \end{cases} \quad (5)$$

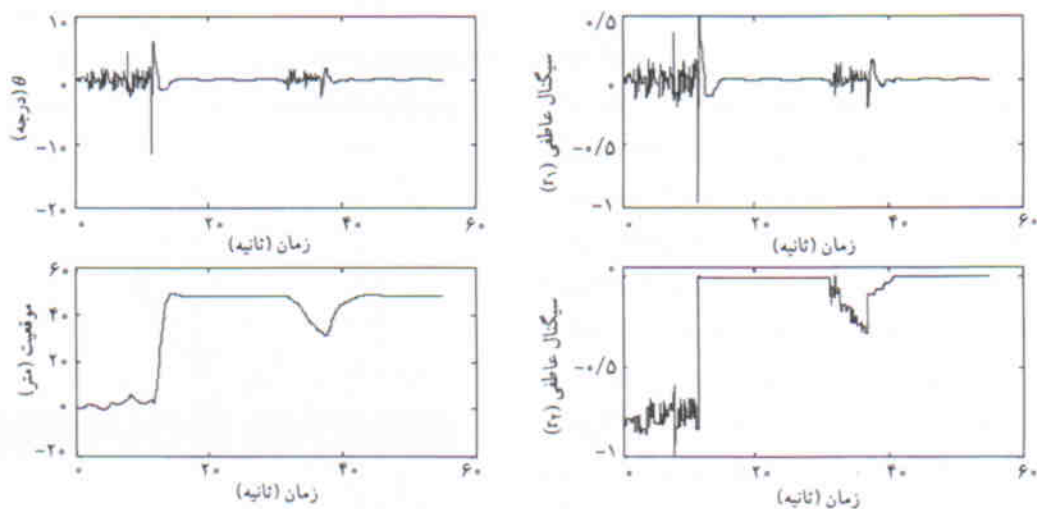
داشت. نتیجه‌ی شبیه‌سازی حاصل از حالت ج نیز در شکل ۷ نشان داده شده است. با توجه به شکل، مشاهده می‌شود که بعد از مدت زمان کوتاهی زاویه‌ی نوسان سیستم صفر شده و واگن نیز به موقعیت مورد نظر رسیده است.

شکل ۸ نشانگر شبیه‌سازی برای حالتی است که اغتشاشی با دامنه‌ی بزرگ به سیستم وارد شده باشد. همان‌گونه که مشاهده می‌شود ساختار معرفی شده قادر است در حداقل زمان سیستم را به حالت قبلی خود ($\theta=0, x_{Desired}=50\text{m}$) برگرداند.

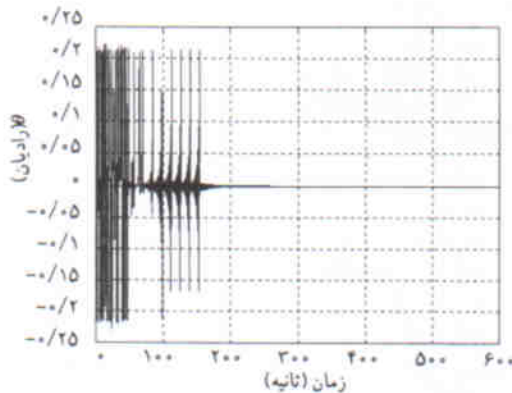
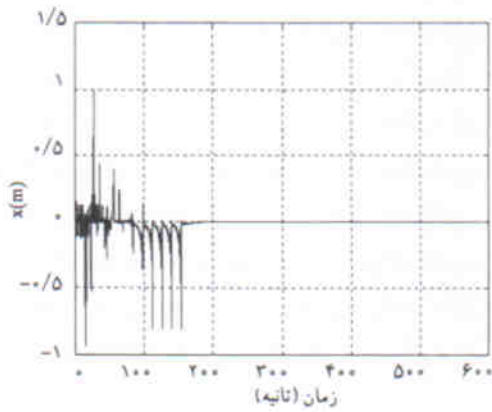
ممکن سیستم را به حالت تعادل برساند، اما در این حالت هیچ کنترلی بر روی موقعیت واگن نداریم و فاصله‌ی واگن از مبدأ به‌ازای شرایط اولیه‌ی مختلف، متفاوت است. نتایج شبیه‌سازی به‌ازای حالت ب نیز در شکل ۶ نشان داده شده است. با توجه به شکل، مشاهده می‌شود که ساختار پیشنهادی توانسته است سیستم را به نقطه‌ی مطلوب برساند اما در این حالت هیچ کنترلی بر روی زاویه‌ی نوسان نداریم و حتی بعد از رسیدن به موقعیت مطلوب نیز نمی‌توان سیستم را بعد از مدت زمان کوتاهی به حالت تعادل نگه



شکل ۷. نتایج شبیه‌سازی برای حالت ج ($x_{Desired}=50\text{m}$).



شکل ۸. بررسی اثر اغتشاش بر عملکرد سیستم ($x_{Desired}=50\text{m}$); اغتشاش در لحظه‌ی $t=3\text{Sec}$ به سیستم اعمال شده است.



شکل ۱۰. موقعیت و زاویه‌ی آونگ بعد از ۸۱ مرتبه شکست $X=(0/0.1, 0, -0/0.1, 0)$

در شکل ۱۰ نتایج شبیه‌سازی‌ها وقتی که پارامترهای همبافتی به صورت $G=\{g_1, g_2\}$ باشند آورده شده است. زمان شبیه‌سازی برابر ۶۰۰ ثانیه انتخاب شده است. در این حالت، که سیستم بعد از ۸۱ بار شکست خوردن ($r^1=-1$ or $r^2=-1$) توانسته است هر دو هدف کنترلی را برآورده سازد، شرایط اولیه‌ی سیستم به صورت $X=(0/0.1, 0, -0/0.1, 0)$ انتخاب شده است.

نتیجه‌گیری

در این مقاله، با استفاده از مفهوم همبافت و با توجه به ویژگی‌های یادگیری عاطفی، روش جدیدی برای کنترل سیستم‌های با اهداف چندگانه ارائه شد. سپس با توجه به این ساختار پیشنهادی، ایده‌ی یادگیری عاطفی مبتنی بر همبافت به سیستم‌های با اهداف چندگانه تعمیم داده شد. برای نشان دادن قابلیت‌های عملی ساختار پیشنهادی، مسئله‌ی کنترل جرثقیل سقفی برای شبیه‌سازی انتخاب شد. نتایج شبیه‌سازی نشان داد که ساختار ارائه شده، می‌تواند با توجه به پارامترهای همبافتی، سیستم‌های با دینامیک‌های غیرخطی و پیچیده را به گونه‌ی کنترل کند که تمامی اهداف مورد نظر به نحو مطلوبی برآورده شوند.

۲. آونگ وارون:

شکل ۹ ساختار مسئله‌ی مورد بررسی را نشان می‌دهد. دینامیک‌های این سیستم با چهار متغیر زیر مشخص می‌شوند:
 θ : زاویه‌ی آونگ نسبت به محور عمودی بر حسب درجه؛
 $\dot{\theta}$: سرعت زاویه‌ی آونگ بر حسب متر بر ثانیه؛
 x : مکان‌گاری نسبت به مبدأ بر حسب متر؛
 \dot{x} : سرعت‌گاری بر حسب متر بر ثانیه؛
 رفتار این متغیرهای حالت با استفاده از معادلات ۸ و ۹ مشخص می‌شوند. [۱۳]

$$\ddot{\theta} = \frac{g \sin \theta + \cos \theta \left[\frac{-F - m l \dot{\theta}^2 \sin \theta}{m_c + m} \right]}{l \left[\frac{4}{3} - \frac{m \cos^2 \theta}{m_c + m} \right]} \quad (8)$$

$$\ddot{x} = \frac{F + m l (\ddot{\theta} \sin \theta - \dot{\theta}^2 \cos \theta)}{m_c + m} \quad (9)$$

که در روابط فوق:

g : شتاب جاذبه‌ی زمین 9.8 m/s^2 ؛

m_c : جرم‌گاری برابر 1 kg ؛

l : نصف طول آونگ برابر 0.5 m ؛

F : نیروی اعمال شده بر گاری.

محدوده‌ی تغییرات متغیرها عبارت‌اند از: $-10 \text{ N} \leq F \leq 10 \text{ N}$ ، $-12^\circ < \theta < 12^\circ$ ، $-2/4 \text{ m} < x < 2/4 \text{ m}$

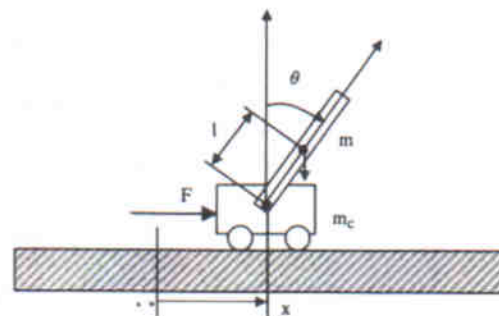
در این مسئله پارامترهای همبافتی را به صورت زیر در نظر می‌گیریم:
 «برقراری تعادل آونگ بر روی گاری»: g_1 ؛

«ثابت نگه‌داشتن گاری در یک موقعیت معین»: g_2 ؛

در این شبیه‌سازی تعداد تقادها را برابر ۲ فرض کرده‌ایم که وظیفه‌ی تقاد اول ارزیابی تعادل آونگ بر روی گاری، و وظیفه‌ی تقاد دوم نیز ارزیابی موقعیت گاری است. سیگنال‌های عاطفی متناظر با هر یک از تقادها به صورت زیر هستند:

$$O_1: \text{If } |\theta| \leq 12^\circ \text{ then } r^1 = 0 \quad \text{else } r^1 = -1$$

$$O_2: \text{If } |x| \leq 2/4 \text{ m then } r^2 = 0 \quad \text{else } r^2 = -1$$



شکل ۹. سیستم آونگ وارون.

پانوشته

1. External Context
2. Internal Context

منابع

1. Turner, R.M., "Context sensitive reasoning for autonomous agents and cooperative distributed problem solving", In *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, Vancouver, Canada, pp 376-391 (1995).
2. Turner, R.M., "Determining the context-dependent meaning of fuzzy subsets", In *proceedings of the 1997 International and Interdisciplinary Conference on Modeling and Using Context (CONTEXT-97)*, pp 1017-1025 (1997).
3. Turner, R.M., "Context-mediated behavior for intelligent agents", *International Journal of Human Computer Studies*, pp 438-445 (2001).
4. Rolls, E.T., "The brain and emotion", Oxford university Press, (1999).
5. Abdi, J. and Lucas, C., "Reinforcement learning and temporal difference learning approach in control application", *Seminar Document*, Tehran University, (2002).
6. Abdi, J., Lucas, C., Sedigh, A.K. and Fatourech, M., "Multivariable systems temporal difference emotional", *Control, Journal of Faculty of Engineering*, 38 (1) pp 51-62 (May 2004).
7. Abdi, J., Khalili, A.F., Lucas, C. and Sedigh, A. K., "The role of truncated temporal difference with function approximation in decreasing the required time and costs of reinforcement learning computations", *IAEEE (Journal of Iranian and Association of Electrical and Electronics Engineers)*, 1, (1) pp 4-18 (Spring 2004).
8. Kentridge, R.W. and Aggleton, J.P., "Emotion: Sensory representations, reinforcement and the temporal lobe", *Cognition and Emotion*, (4) pp.191-208 (1990).
9. Haykin, S. "Neural networks, a comprehensive foundation", Prentice Hall, (1994).
10. Sutton, R.S., Barto, A.G. and Williams, R.J., "Reinforcement Learning is Direct Adaptive Optimal control", *IEEE Control Systems*, pp 19-22 (1992).
11. Lucas, C. and Jazbi, S.A., "Intelligent motion control of electric motors with evaluative feedback", *Cigre98*, Cigre, France, 1998:11-104, pp 1-6.
12. Lucas, C., Jazbi, S.A., Fatourech, M. and Farshad, M., "Cognitive action selection with neuro-controllers", *Third Iran-Armenia Workshop on Neural Networks, Yerevan*, pp 37-44 (2000).
13. Takagi, T. and Sugeno, M., "Derivation of fuzzy control rules from human operator's control actions", In *Proceedings of IFAC Symposium on Fuzzy Information, Knowledge Representation and Decision Analysis*, pp 55-60 (1983).
14. Lin C.T. and Lee, C.S.G., "Neural-network-based fuzzy logic control and decision system", *IEEE Transactions on Computers*, 40, (12) pp 123-131 (1991).
15. Fliess, M., Levine, J. and Rouchon, P., "A simplified approach of crane control via a generalized state-space model", *IEEE Proc. of 30th Conference on Decision and Control, Brighton, England*, pp 736-741 (December 1991).