بسم الله الرحمن الرحيم

دانشکده مهندسی کامپیوتر

پایان‌نامه کارشناسی ارشد مهندسی هوش مصنوعی

تشخیص کنترل کننده مکالمه با استفاده از تکنیک های هوش مصنوعی

نگارنده:

نگار قائم‌پناه

استاد راهنما :

دکتر مرتضی زاهدی

استاد مشاور

مهندس مهدی حسینی

تیر ۱۳۹۷

شماره: ۵۷۲/ف/ک
تاریخ: ۹۷/۶/۵

باسمه‌تعالی

مدیریت تحصیلات تکمیلی

فرم شماره (۳) صورتجلسه نهایی دفاع از پایان نامه دوره کارشناسی ارشد

با نام و یاد خداوند متعال، ارزیابی جلسه دفاع از پایان نامه کارشناسی ارشد خانم **نگار قائم‌پناه** با شماره دانشجویی **۹۴۱۳۸۸۴** رشته مهندسی کامپیوتر گرایش هوش مصنوعی و رباتیک تحت عنوان **تشخیص کنترل‌کننده مکالمه با استفاده از تکنیک‌های هوش مصنوعی** که در تاریخ **۹۷/۴/۲۵** با حضور هیأت محترم داوران در دانشگاه صنعتی شاهرود برگزار گردید به شرح ذیل اعلام می‌گردد:

قبول (با درجه: خیلی خوب) ☒ مردود ☐

نوع تحقیق: نظری ■ عملی ☐

عضو هیأت داوران	نام ونام خانوادگی	مرتبهٔ علمی	امضاء
۱ـ استاد راهنمای اول	دکتر مرتضی زاهدی	استادیار	
۴- نماینده تحصیلات تکمیلی	دکتر فاطمه جعفرنژاد	استادیار	
۵- استاد ممتحن اول	دکتر مرضیه رحیمی	استادیار	
۶ـ استاد ممتحن دوم	دکتر حسین مروی	استادیار	

نام و نام خانوادگی رئیس دانشکده: دکترعلی‌رضا الفی

تاریخ و امضاء و مهر دانشکده: ۹۷/۴/۲۵

تبصره: در صورتی که کسی مردود شود حداکثر یکبار دیگر (در مدت مجاز تحصیل) می تواند از پایان نامه خود دفاع نماید (دفاع مجدد نباید زودتر از ۴ ماه برگزار شود).

بنام خدا

سپاس خدای را که سخنوران، در ستودن او بمانند و شمارندگان، شمردن نعمت‌های او ندانند و کوشندگان، حق او را گزاردن نتوانند. سلام و دورد بر محمّد و خاندان پاک او، طاهران معصوم، هم آنان که وجودمان وامدار وجودشان است؛ و نفرین پیوسته بر دشمنان ایشان تا روز رستاخیز... .

بدون شک جایگاه و منزلت معلم، اجّل از آن است که در مقام قدردانی از زحمات بی شائبه‌ی او، با زبان قاصر و دست ناتوان، چیزی بنگاریم.

اما از آنجایی که تجلیل از معلم، سپاس از انسانی است که هدف و غایت آفرینش را تامین می‌کند و سلامت امانت‌هایی را که به دستش سپرده‌اند، تضمین؛ بر حسب وظیفه و از باب **"من لم یشکر المنعم من المخلوقین لم یشکر اللّه عزّ و جلّ"** :

از استاد با کمالات و شایسته؛ جناب آقای دکتر مرتضی زاهدی که در کمال سعه صدر، با حسن خلق و فروتنی، از هیچ کمکی در این عرصه بر من دریغ ننمودند و زحمت راهنمایی این پایان نامه را بر عهده گرفتند؛

از استاد صبور و با تقوا ، جناب آقای دکتر مهدی حسینی، که زحمت مشاوره این پایان نامه را در حالی متقبل شدند که بدون مساعدت ایشان، این پروژه به نتیجه مطلوب نمی‌رسید؛

و از اساتید فرزانه و دلسوز؛ خانم دکتر مرضیه رحیمی و جناب آقای دکتر حسین مروی که زحمت داوری این پایان نامه را متقبل شدند؛ کمال تشکر و قدردانی را دارم.

باشد که این خردترین، بخشی از زحمات آنان را سپاس گوید.

تعهد نامه

اینجانب **نگار قائم‌پناه** دانشجوی دوره کارشناسی ارشد مهندسی فناوری اطلاعات دانشکده مهندسی کامپیوتردانشگاه صنعتی شاهرود نویسنده پایان نامه تشخیص کنترل کننده مکالمه با استفاده از تکنیک های هوش مصنوعی تحت راهنمائی جناب آقای دکتر مرتضی زاهدی متعهد می‌شوم .

- تحقیقات در این پایان نامه توسط اینجانب انجام شده است و از صحت و اصالت برخوردار است .
- در استفاده از نتایج پژوهش‌های محققان دیگر به مرجع مورد استفاده استناد شده است .
- مطالب مندرج در پایان نامه تاکنون توسط خود یا فرد دیگری برای دریافت هیچ نوع مدرک یا امتیازی در هیچ جا ارائه نشده است .
- کلیه حقوق معنوی این اثر متعلق به دانشگاه صنعتی شاهرود می‌باشد و مقالات مستخرج با نام « دانشگاه صنعتی شاهرود » و یا « **Shahrood University of Technology** » به چاپ خواهد رسید .
- حقوق معنوی تمام افرادی که در به دست آمدن نتایج اصلی پایان نامه تأثیرگذار بوده اند در مقالات مستخرج از پایان نامه رعایت می‌گردد.
- در کلیه مراحل انجام این پایان نامه ، در مواردی که از موجود زنده (یا بافتهای آن‌ها) استفاده شده است ضوابط و اصول اخلاقی رعایت شده است .
- در کلیه مراحل انجام این پایان نامه، در مواردی که به حوزه اطلاعات شخصی افراد دسترسی یافته یا استفاده شده است اصل رازداری ، ضوابط و اصول اخلاق انسانی رعایت شده است .

تاریخ

امضای دانشجو

مالکیت نتایج و حق نشر

- کلیه حقوق معنوی این اثر و محصولات آن (مقالات مستخرج ، کتاب ، برنامه های رایانه‌ای ، نرم افزارها و تجهیزات ساخته شده است) متعلق به دانشگاه صنعتی شاهرود می‌باشد . این مطلب باید به نحو مقتضی در تولیدات علمی مربوطه ذکر شود .
- استفاده از اطلاعات و نتایج موجود در پایان‌نامه بدون ذکر مرجع مجاز نمی‌باشد.

چکیده

در سیستم‌های پرسش و پاسخ تعاملی، کاربران سوالات خود را به زبان طبیعی پرسیده و پاسخ‌های مرتبط را دریافت می‌کنند. به دلیل وجود تعامل در این سیستم‌ها، کاربران می‌توانند توضیحاتی را راجع به سوالات خود مطرح کرده و دقت پاسخگویی را افزایش دهند.

گفتگوهای انسانی دارای ویژگی‌های فراوانی می‌باشند که یکی از این ویژگی‌ها کنترل مکالمه می‌باشد. در این گفتگوها، کنترل مکالمه ممکن است بین افراد جابجا شود. چنین تعاملاتی دارای ویژگی ابتکار عمل ترکیبی هستند. در این گفتگوها، در هر زمان یک شخص ممکن است ابتکار عمل را در اختیار بگیرد و جریان مکالمه را تغییر و هدایت کند. ویژگی ابتکار عمل ترکیبی یک جنبه مهم و موثر در همکاری افراد برای حل مشکلات است و باعث می‌شود افراد به عنوان یک تیم، کارایی بیشتری داشته باشند.

هدف از این تحقیق، شناسایی کنترل کننده مکالمه در سیستم‌های پرسش و پاسخ تعاملی با استفاده از تکنیک‌های آماری است. این روش بر خلاف روش‌های دیگر که عمدتا از تجزیه و تحلیل معنایی و گرامری استفاده می‌کنند، از روش‌های آماری و فاقد تجزیه و تحلیل معنایی و گرامی استفاده می‌کند و از این‌رو مستقل از زبان می‌باشد. برای تشخیص کنترل کننده در این روش، پایگاه داده مربوطه تهیه و سیستم با استفاده از روش‌های آماری آموزش داده می‌شود. در این تحقیق از یک پایگاه داده که برای این منظور طراحی شده، استفاده شده است. برای تشخیص کنترل کننده از روش‌هایی مثل فاصله اقلیدسی، همینگ، مینکوفسکی، بیزین، k نزدیک‌ترین همسایه، k میانگین، ماشین بردار پشتیبان، شباهت کسینوسی و TF-IDF استفاده شده است. بهترین کارایی در این سیستم روش ماشین بردار پشتیبان با ۹۱.۴۹ درصد می‌باشد.

کلمات کلیدی: سیستم‌های پرسش و پاسخ تعاملی، ابتکار عمل ترکیبی، کنترل کننده مکالمه، ابتکار عمل، روش‌های آماری

فهرست مطالب

فصل اول مقدمه ۱

۱-۱-مقدمه ۲

۱-۲-کنترل کننده مکالمه ۲

۱-۳-انواع تعاملات بر اساس ابتکار عمل ۴

۱-۴-کاربرد تعاملات دارای ابتکار عمل ترکیبی ۶

۱-۵-هدف و رویکرد پایان نامه ۷

۱-۶-ساختار پایان نامه ۸

فصل دوم مروری بر کارهای پیشین ۹

۲-۱- مقدمه ۱۰

۲-۲- کارهای پیشین ۱۰

فصل سوم روش پیشنهادی ۳۳

۳-۱- مقدمه ۳۴

۳-۲- پایگاه داده ۳۴

۳-۳- روشهای پیشنهادی ۳۵

۳-۳-۱- مجموعه ویژگیها ۳۶

۳-۳-۲- روش مبتنی بر فاصله اقلیدسی ۳۶

۳-۳-۳- روش مبتنی بر فاصله همینگ ۳۷

۳-۳-۴- روش مبتنی بر فاصله مینکوفسکی ۳۸

۳-۳-۵- روش بیزین ساده ۳۸

۳-۳-۶- روش k نزدیکترین همسایه ۳۹

۳-۳-۷- روش TF-IDF ۴۰
۳-۳-۸- روش خوشه بندی k میانگین ۴۱
۳-۳-۹- روش ماشین بردار پشتیبان ۴۲
۳-۳-۱۰- روش مبتنی بر معیار شباهت کسینوسی ۴۴
۳-۴-معیار محاسبه درصد درستنمایی ۴۴
فصل چهارم جزئیات پیکربندی و ارزیابی روش پیشنهادی ۴۷
۴-۱- مقدمه ۴۸
۴-۲- روش مبتنی بر فاصله اقلیدسی ۴۸
۴-۳- روش مبتنی بر فاصله همینگ ۴۹
۴-۴- روش مبتنی بر فاصله مینکوفسکی ۴۹
۴-۵- روش بیزین ساده ۵۰
۴-۶- روش k نزدیکترین همسایه ۵۰
۴-۷- روش TF-IDF ۵۱
۴-۸- روش خوشه بندی k میانگین ۵۱
۴-۹- روش ماشین بردار پشتیبان ۵۲
۴-۱۰- روش مبتنی بر معیار شباهت کسینوسی ۵۲
۴-۱۱- مقایسه روشهای پیشنهادی ۵۳
فصل پنجم جمع بندی و پیشنهادهایی برای کارهای آینده ۵۵
۵-۱- جمع بندی ۵۶
۵-۲- پیشنهادهایی برای کارهای آینده ۵۸
منابع و ماخذ: ۵۹

فهرست اشکال

شکل (۲-۱): تصویری از سیستم Lookout [۲۴] ۱۱
شکل (۲-۲): توزیع ابتکار نشانه‌ها ۱۹
شکل (۲-۳): الگوریتم انتخاب هدف مساله ۱۹
شکل(۲-۴): نگاشت بین اعمال و الگوهای گفتاری ۲۰
شکل (۲-۵): توزیع ابتکار کار و گفتگو در مکالمه ۲۰
شکل(۲-۶): الگوریتم مقداردهی ویژگی‌های پیتزا ۲۷
شکل (۲-۷): گراف توالی تعاملات معتبر در برنامه سفارش پیتزا ۲۷
شکل(۲-۸): نمای کلی برنامه ثبت سفارش ۲۸

فهرست جداول

جدول (۲-۱): مجموعه قوانین برای انتقال کنترل ۱۳
جدول (۲-۲): نشانه‌های مشاهده شده در تفسیر معنایی جملات ۱۸
جدول(۲-۳): انتخاب استراتژی براساس هدف و توزیع ابتکار ۲۰
جدول (۲-۴): استفاده از پشته برای تغییر تمرکز ۲۴
جدول (۲-۵): تاثیر پارامترهای جملات خلاصه‌شده، مذاکره و الگوریتم continuous mode در زمان، تعداد سخنان و تعداد شاخه پیمایش شده برای رسیدن به جواب ۲۶
جدول (۲-۶): فاکتورهای عینی و ذهنی در تعیین ابتکار عمل ۳۱
جدول (۴-۱): مقایسه روش‌های پیشنهادی ۵۳

فهرست علائم اختصاری

Question Answering	QA
Interactive Question Answering	IQA
Natural Language	NL
Term Frequency	TF
Inverse Document Frequency	IDF
K-Nearest Neighbour	KNN
Support Vector Machine	SVM
True Positive	TP
True Negative	TN
False Positive	FP
False Negative	FN
Mixed-Initiative Movie Information Consultant	MIMIC
Matthews Correlation Coefficient	MCC

فصل اول

مقدمه

۱-۱-مقدمه

سیستم پرسش و پاسخ[۱] یک سیستم خودکار، جهت پاسخگویی به پرسش‌هایی است که توسط انسان با استفاده از زبان طبیعی[۲] مطرح می‌شود. در نقطه مقابل این سیستم‌ها و سیستم‌های گفتگو، سیستم های دیگری به نام سیستم‌های پرسش و پاسخ تعاملی[۳] مطرح می‌شوند. در این سیستم‌ها در صورتیکه پاسخ ارائه شده مبهم و مدنظر کاربر نباشد و یا اینکه کاربر نیاز به دریافت اطلاعات بیشتر داشته باشد این امکان وجود دارد که بین سیستم و کاربر پرسش‌های دیگری رد و بدل شود. بنابراین هر چقدر پاسخ ارائه شده مطلوب‌تر باشد کاربر سریعتر به پاسخ مدنظر خود خواهد رسید و در نتیجه میزان رضایت کاربران از سیستم تعاملی افزایش خواهد یافت.[۱]

مکالمه دارای ویژگی‌های متعددی است که یکی از این ویژگی‌ها کنترل مکالمه می‌باشد. با استفاده از این ویژگی کنترل و هدایت مسیر مکالمه از یک شخص به شخص دیگر منتقل می‌شود و انحصارا در اختیار یک شخص باقی نمی‌ماند. مکالماتی با این ویژگی را مکالمات دارای ابتکار عمل ترکیبی[۴] می‌نامند که در هر لحظه ابتکار عمل[۵] به دست یکی از شرکت‌کنندگان در مکالمه است. بنابراین مکالمات انعطاف‌پذیرتر و حل مسائل راحت‌تر خواهد شد.

۱-۲-کنترل کننده مکالمه[۶]

هدف سیستم‌های پرسش و پاسخ، پاسخگویی به سوالات کاربر با دقت بالا می‌باشد. با وجود استدلال قوی و یافتن پاسخ‌های دقیق، سیستم بدون داشتن کنترل کننده‌ای برای مکالمه، خوب عمل نکرده و احساس رضایت در کاربر ایجاد نخواهد کرد. کنترل در مکالمه، منجر به ایجاد احساس در کاربر می‌شود، به طوری که کاربر حس کند با یک انسان در حال مکالمه می‌باشد. تعویض کنترل در مکالمه

---

۱ Question Answering(QA)
۲ Natural Language(NL)
۳ Interactive Question Answering(IQA)
٤ Mixed-Initiative
٥ Initiative
٦ Conversation Controller

جالب است زیرا حداقل رویای تورینگ در داشتن حس هوشمندی سیستم را برآورده می‌کند. سیستم-هایی که تنها ارتباط بین سیستم و کامپیوتر را برقرار می‌کنند جزء سیستم‌های ابتکار عمل ترکیبی نیستند.

تعامل دارای ابتکار عمل ترکیبی از جمله مواردی است که مردم در هنگام مواجهه با آن احساس می-کنند که آن را می‌شناسند حتی اگر قادر به تعریف آن نباشند.[۲] در مکالمه دوطرفه، جریان دوطرفه‌ای از اطلاعات بین شرکت‌کنندگان وجود دارد. هر شرکت‌کننده به نوبه خود کنترل مکالمه را به دست می‌گیرد. شرکای مکالمه نه تنها به سوالات دیگری پاسخ می‌دهند بلکه گاهی اطلاعات دیگری داده و سوالات خودشان را می‌پرسند.[۳] از آنجایی که ابتکار عمل بین شرکت‌کنندگان رد و بدل می‌شود، ما می‌توانیم بگوییم که کنترل در طول مکالمه از یکی به دیگری انتقال پیدا می‌کند.

ابتکار عمل با عنوان "گرفتن رهبری مکالمه[۱]"[۴] یا "هدایت یک وظیفه[۲]"[۵] به کار برده می‌شود.

در سال‌های اخیر فناوری‌های گفتاری و زبان طبیعی برای توسعه سیستم‌های گفتگو گفتاری[۳] در دامنه محدود گسترش یافته‌اند. اکثر سیستم‌های موجود از استراتژی گفتگو از پیش تعیین شده برای مدیریت گفتگو استفاده می‌کردند بدون آنکه ویژگی‌های تعامل گفتاری واقعی را در نظر بگیرند. به طور خاص سیستم‌های دارای ابتکار عمل ترکیبی معمولا قوانینی را تعیین می‌کنند که تحت شرایطی (عموما بر مبنای محتوای گفتگو محلی) ابتکار عمل ممکن است از عاملی به عامل دیگر انتقال پیدا کند. از سوی دیگر تحقیقات قبلی نشان می‌دهد که تغییرات استراتژی ابتکار عمل در گفتگوهای انسانی می‌تواند به صورت پویا از لحاظ ویژگی‌های کاربر و گفتگو در حال انجام، مدل‌سازی شود[۶] و سازگاری استراتژی‌های ابتکار عمل در سیستم‌های گفتگو، منجر به افزایش کارایی سیستم می-شود.[۷]

---

۱ Taking the conversational lead
۲ Driving the task
۳ Spoken Dialogue Language

سیستم‌های دارای ابتکار عمل ترکیبی از روش‌های سنتی ساخت دنباله عملیات از مبدا مشخص به هدف مشخص پیروی نمی‌کنند. این تعاملات را می‌توان به طور گسترده‌ای به عنوان یک استراتژی تعاملی انعطاف‌پذیر بین شرکت‌کنندگان در نظر گرفت که هر شرکت‌کننده در هر زمانی می‌تواند نوبت[۱] را برای تغییر و هدایت جریان تعامل به دست گیرد و در حل مسائل در بهترین زمان ممکن کمک کند. معمولا در گفتگوهای انسانی و حتی برخی سیستم‌های کامپیوتری مدرن این خاصیت وجود دارد.

تعامل دارای ابتکار عمل ترکیبی یک جنبه کلیدی موثر در تعامل بین کامپیوتر و انسان است و تاثیر زیادی در سیستم‌های چندعاملی[۲] دارد. این قابلیت یک جنبه مهم و موثر در همکاری چندعاملی برای حل مشکلات یا انجام وظایف است. در ساده‌ترین حالت این وظایف می‌توانند شامل تعامل با کاربر برای طراحی یک آشپزخانه، یافتن بهترین قیمت هواپیما، هماهنگی یک ماموریت اضطراری و یا آموزش به کاربران در مورد نحوه استفاده از یک وسیله باشد. در هر زمان یک عامل ممکن است ابتکار عمل را داشته باشد و کنترل تعامل را به دست گیرد درحالیکه دیگران برای کمک به آن عمل می-کنند و نیازها را در صورت لزوم برطرف می‌کنند. عوامل به صورت پویا شیوه تعامل‌شان را با هم تطابق می‌دهند و مشکلات را به بهترین نحوه حل می‌کنند. این قابلیت باعث می‌شود عوامل به عنوان یک تیم کارایی بیشتری داشته باشند.

۱-۳-انواع تعاملات بر اساس ابتکار عمل

تعاملات بین انسان و کامپیوتر بر اساس ابتکار عمل به سه دسته تقسیم می‌شوند:

۱. سیستم‌هایی که کنترل در آن‌ها به دست انسان است. در این روش سیستم به فرامین مشخص شده توسط انسان پاسخ می‌دهد. در سیستم‌هایی که ابتکار با کاربر[۳] است، سیستم به هر درخواست کاربر بدون اعمال محدودیت روی آن پاسخ می‌دهد. این روش جزء سیستم‌های ابتکار عمل ترکیبی نیست،

[۱] Turn
[۲] Multi Agent Systems
[۳] User-Initiative

زیرا کنترل در تمام مدت دست انسان است. مثالی از این تعاملات ایستگاه‌های زمانبندی[۱] هستند. در این ایستگاه‌ها ابزارهایی برای برنامه‌ریزی ماشین‌ها جهت حمل و نقل وجود دارد. انسان درخواست‌ها را ارسال و سیستم با برنامه‌هایی که برای آن مشخص شده به این درخواست‌ها پاسخ می‌دهد. این سیستم‌ها اگر به درستی طراحی شوند ابزارهای برنامه‌ریزی مفیدی هستند. بخشی از این نوع مکالمات در زیر آورده شده است. در این مکالمه s آغازگر مکالمه است و درباره زمان ملاقات سوال می‌پرسد. اما پس از آن u کنترل مکالمه را به دست گرفته و به مکالمه خاتمه می‌دهد. U با تعیین قرار ملاقات، اجازه ادامه کار را به S نمی‌دهد و کنترل به S داده نمی‌شود.

2.1 S: I can schedule an appointment for you. What day?

2.2a U: Tuesday.

2.2b U:Tuesday at 11:30 a.m. would work for me.

2.2c U: I need an appointment in the evening.

2.2d U: So I have to wait to get this service?

۲. سیستم‌هایی که در آن‌ها کنترل به دست سیستم است. در گفتگوهایی که ابتکار عمل با سیستم[۲] است، جریان ارتباطات کاملا به وسیله سیستم کنترل می‌شود. در هر نقطه از گفتگو سیستم تنها تعداد محدودی از جواب‌های احتمالی را می‌پذیرد. این روش جزء سیستم‌های ابتکار عمل ترکیبی نیست زیرا کنترل در تمام مدت دست سیستم است. نمونه‌ای از این تعاملات مراکز تماس خودکار[۳] است. در این سیستم‌ها انسان با صفحه کلید یا از طریق گفتار گزینه‌هایی از منو را انتخاب می-کند(مثلا فشردن عدد ۱ برای دریافت موجودی). بخشی از این نوع مکالمات در زیر آورده شده است. در این مکالمه s آغازگر مکالمه است و درباره زمان ملاقات سوال می‌پرسد. سپس زمان‌های پیشنهادی خود را اعلام کرده و اجازه دریافت کنترل به کاربر را نمی‌دهد.

1.1a S: I can schedule an appointment for you. What day?

[۱] Scheduling Stations
[۲] System-Initiative
[۳] Automated Call Center

1.1b S: I can schedule an appointment for you. Available days and times include: Monday at 9:30 a.m., 10:15 a.m., and 4:45 p.m.; Tuesday at 8:00 a.m., 8:15 a.m. …

۳. سیستم‌هایی که دارای ابتکار عمل ترکیبی هستند. ایده این سیستم‌ها، ترکیب انعطاف‌پذیری سیستم-های دارای ابتکار عمل کاربر با محدودیت‌های حل مساله سیستم‌های دارای ابتکار عمل سیستم است. مثلا پاسخ معقول به سوال "پرواز را به من نشان بده" می‌تواند جمله "لطفا به من بگویید که شما به کجا پرواز می‌کنید" باشد. این سیستم‌ها باید قابلیت تعویض کنترل از انسان به سیستم و برعکس را داشته باشند. ارتباط با ابتکار عمل ترکیبی، یک ویژگی در سیستم‌های ارتباطی بین انسان و سیستم است و تاثیر بسیار زیادی بر سیستم‌های چندعامله می‌گذارد. توسعه سیستم‌های هوشمند با این قابلیت دنیای محاسبات را متحول خواهد کرد. این ویژگی باعث ایجاد ارتباط‌های با انعطاف‌تر شده، به گونه‌ای که هر عامل می‌تواند برای رسیدن به جواب همکاری کند. با استفاده از این کار، گفتگوها به تعاملات بین انسان‌ها بسیار نزدیک می‌شوند. بخشی از این نوع مکالمات در زیر آورده شده است. در این مکالمه s آغازگر مکالمه است و درباره زمان ملاقات سوال می‌پرسد. در ابتدا s با ارائه پیشنهاد خود قصد دریافت کنترل را دارد اما در ادامه کنترل را به u واگذار می‌کند.

3.1 S: Would you like to have an appointment?

3.2 U: Yes.

3.3a S: Okay, how about Friday at 2:30 p.m.

3.3b S: Okay, on what day?

۱-۴-کاربرد تعاملات دارای ابتکار عمل ترکیبی

سوالی که مطرح می‌شود این است که چرا ما به عواملی که منجر به تعامل دوطرفه در یک مکالمه می‌شوند، علاقه‌مند هستیم؟

۱. بر اساس نظریه‌ها، ما معمولا مایل هستیم عوامل یک گفتگو تکی[۱] را به یک گفتگو چندگانه[۲] گسترش دهیم.[۸،۹] در مطالعات گذشته، ساختار گفتگوهای چند نفره[۳] با خاصیت ابتکار عمل ترکیبی اغلب با تخصیص کنترل به یک شرکت‌کننده[۱۰،۱۱] یا با فرض یک شنونده غیرفعال[۴][۱۲،۱۳] انجام می-گرفته است. از آنجایی که مکالمه یک فرآیند مشارکتی است[۱۴،۱۵]، مدل‌های مکالمه می‌توانند اساسی را برای گسترش نظریه‌ها فراهم کنند.[۱۶،۱۷] بر این اساس زمانی که نیاز به همکاری و مذاکره باشد، این نظریه‌ها باید سبب تعامل اهداف بین شرکت‌کنندگان شوند.

۲. دلایل فراوانی وجود دارد که نشان می‌دهد استفاده از قابلیت ابتکار عمل ترکیبی محدود شده[۵] سبب کاهش کارایی سیستم می‌شود. بسیاری از محققین اشاره دارند که نبود این قابلیت باعث بروز دو مشکل در سیستم‌ها می‌شود. اول اینکه آن‌ها اجازه نمی‌دهند که کاربران در فرآیند استدلال شرکت کنند و یا پرسش‌های خود را مطرح کنند.[۱۸،۱۹،۲۰] مشکل بعدی این است که سیستم‌های پرسش و پاسخ اغلب نقش سیستم را به عنوان یک شرکت‌کننده نادیده می‌گیرند. به عنوان مثال گفته‌های یک بخش ممکن است با توجه به ورودی قبلی کاربر تفسیر شود درصورتیکه آنچه کاربر می‌گوید در جواب گفته سیستم است.[۲۱،۲۲]

۱-۵-هدف و رویکرد پایان نامه

مکالمات انسانی دارای ویژگی‌های متعددی هستند که یکی از آن‌ها، همکاری و کنترل مکالمه توسط هر یک از افراد شرکت‌کننده در مکالمه، برای پیشبرد اهداف می‌باشد. هدف اصلی این پایان نامه، شناسایی کنترل کننده مکالمه با استفاده از روش‌های آماری است که این کار مستقل از زبان و حوزه دانش انجام می‌گیرد.

---

۱ Single Utterances
۲ Multi Utterances
۳ Multi Participant
۴ Passive Listener
۵ Limited Mixed-Initiative

به منظور دستیابی به این هدف، پایگاه داده فارسی جهت بررسی و تشخیص کنترل کننده مکالمه گردآوری شده است. سیستم با این پایگاه داده آموزش داده شده و مستقل از روش‌های معنایی، کنترل کننده را شناسایی می‌کند.

۱-۶-ساختار پایان نامه

این پایان نامه از ۵ فصل تشکیل شده است. در فصل اول خواننده با مفهوم کنترل کننده مکالمه و مکالمات دارای ابتکار عمل ترکیبی آشنا می‌شود.

در فصل دوم از این پایان نامه، کارهای انجام شده در این زمینه مطرح خواهد شد و خواننده با نقاط ضعف و کلیات این مبحث آشنا خواهد شد. در فصل سوم به بررسی روش پیشنهادی پرداخته و روش-های آماری برای تشخیص کنترل کننده مکالمه پیاده‌سازی خواهد شد. در فصل چهارم نتایج حاصل از روش پیشنهادی و ارزیابی آن بررسی خواهد شد. در نهایت در فصل پنجم جمع بندی پژوهش انجام شده است و در پایان موضوعاتی برای تحقیق در آینده پیشنهاد شده است.

فصل دوم

مروری بر کارهای پیشین

۲-۱- مقدمه

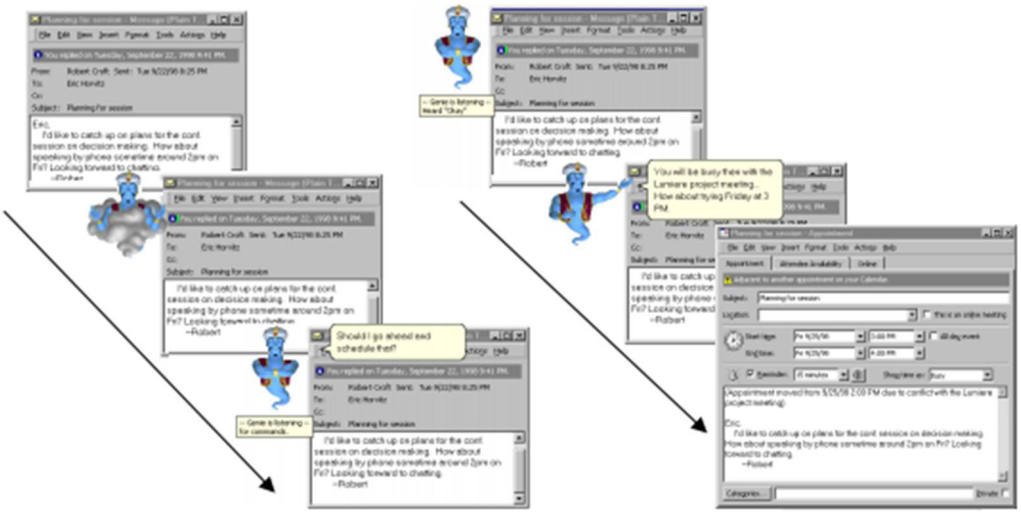
در جهان پر از اطلاعات امروز، یافتن پاسخ‌های دقیق و صحیح برای سوالات کاربران به یک مساله مهم تبدیل شده است. برای پاسخگویی به این سوالات، انواع سیستم‌های بازیابی اطلاعات طراحی شده که پاسخ‌های کاربران را در قالب متن، صوت، تصویر، ویدئو و یا ترکیبی از این موارد فراهم می‌کنند.[۲۳] هدف سیستم‌های پرسش و پاسخ، پاسخگویی به سوالات کاربر با دقت بالا می‌باشد. با وجود استدلال قوی و یافتن پاسخ‌های دقیق، سیستم بدون داشتن کنترل کننده‌ای برای مکالمه، خوب عمل نکرده و احساس رضایت در کاربر ایجاد نخواهد کرد. کنترل در مکالمه، منجر به ایجاد احساس در کاربر می-شود، به طوری که کاربر حس کند با یک انسان در حال مکالمه می‌باشد. تعویض کنترل در مکالمه جالب است زیرا حداقل رویای تورینگ در داشتن حس هوشمندی سیستم را برآورده می‌کند.

اکثر سیستم‌های طراحی شده در این زمینه از قواعد معنایی و گرامری استفاده می‌کنند.

۲-۲- کارهای پیشین

یکی از سیستم‌های نمونه در این زمینه به نام Lookout [۲۴]، از چندین کلید مفهومی جهت مشخص نمودن نقش تصمیم‌گیری در سیستم‌های با قابلیت ابتکار عمل ترکیبی استفاده می‌کند. دستیار Lookout با استفاده از وظایفی، به بررسی تقویم و ایجاد قرار ملاقات‌ها می‌پردازد. یک گروه از افراد علاقه‌مند در مایکروسافت به کار در این زمینه پرداختند و اولین نسخه از آن در سال ۱۹۹۸ آزمایش و بررسی شد.

Lookout به تعاملات کاربر از طریق پیام‌های Microsoft outlook و سیستم تقویمی نظارت می‌کند. زمانی که کاربران یک پیام جدید را ایجاد می‌کنند، Lookout آن را تشخیص می‌دهد. سپس این سیستم تصمیم می‌گیرد که چه زمانی و چگونه بهترین کاربران توسط وظایف دسترسی، به تقویم دسترسی داشته و قرارها را تنظیم کنند.



شکل (۲-۱): تصویری از سیستم Lookout [۲۴]

سیستم TRAINS95 [۲۵]، نمونه دیگری از سیستم با ویژگی ابتکار عمل ترکیبی است. این سیستم به زبان LISP نوشته شده است و تمامی کارها در آن به وسیله پیام‌ها انجام می‌شود. کاربر و سیستم به ساخت، ارزیابی و اصلاح برنامه‌هایی که ساده هستند، می‌پردازند. این سیستم دارای محدودیت‌های روشنی است که می‌داند چگونه از طریق تعامل با کاربر، با آن مقابله کند.

مدل ابتکار عمل ترکیبی در این سیستم شامل ۴ مرحله است:

۱. تمرکز[۱]: شناسایی هدف یا زیر هدف مورد بررسی.

۲. شناسایی محدودیت[۲]: جمع‌آوری محدودیت‌های مساله، انتخاب منابع، جمع‌آوری اطلاعات پس‌زمینه[۳] و اضافه کردن تنظیمات.

۳. نمونه راه‌حل[۴]: به محض اینکه راه‌حلی آماده تولید شود، آن را تولید می‌کند.

۴. انتقاد کردن[۵]: اگر نمونه مورد انتقاد قرار گرفت و پیشنهاد تغییرات داده شد کار از مرحله ۲ ادامه می‌یابد اما اگر نمونه مورد تایید قرار گیرد کار از مرحله اول و با تمرکز جدید ادامه می‌یابد.

در زیر نمونه‌ای از مکالمه انسان با سیستم TRANS95 آورده شده است.

1. M: I need to get to Detroit, Syracuse, and Montreal.

---

۱ Focus
۲ Gather Constraints
۳ Background Information
۴ Instantiate Solution
۵ Criticize

2. S: Ok. [highlights the three cities]
3. M: Let’s send the trains from Chicago to Detroit.
4. S: Ok. [draws route from Chicago to Detroit via Toledo] But the terminal at Toledo is delaying traffic due to bad weather. Trains will take an additional five hours to move through there. [highlight Toledo red]
5. M: That’s ok. Now let’s go from Atlanta up to Syracuse.
6. S: Please tell me a route to use to get from Atlanta to Syracuse. [highlight engine at Atlanta and city Syracuse]
7. M: Go via Charleston and Pittsburgh.
8. S: Ok. [draws route via Cincinnati] But traffic through Cincinnati is being delayed due to construction.
9. M: Ok, let’s go via Charleston and Buffalo instead.
10. S: No problem. [draws route]
11. M: That’s great.
12. S: I hope it was helpful.
13. M: Finally, I want to go to Montreal from Philadelphia.
14. S: Ok. [draws route]
15. M: That’s great. I’m done.

در این مکالمه، کاربر با سیستم در حال تعامل است و از ویژگی ابتکار عمل ترکیبی استفاده شده است. کنترل مکالمه توسط سیستم و کاربر صورت می‌گیرد و منحصرا در دست یکی از آن‌ها نیست. این قابلیت مکالمه را به مکالمه واقعی انسان‌ها شبیه می‌کند.

در سال ۱۹۸۸، ویتاکر و استنتون[۲۶] از یک نوع کلاس‌بندی گفته‌ها و یک مجموعه قوانین برای انتقال کنترل استفاده کردند. انتقال کنترل به این دلیل اتفاق می‌افتد که یک شرکت‌کننده به تنهایی نمی‌تواند مسئول هماهنگ کردن دستاوردهای کل گفتمان باشد. زمانی که شرکت‌کننده دیگری کنترل یک زیر هدف را برعهده می‌گیرد یک تغییر کنترل رخ می‌دهد و شرکت‌کنندگان باید مکانیزمی

را برای دستیابی به آن داشته باشند. انتقال کنترل زمانی اتفاق می‌افتد که کنترل تغییر کرده و همه شرکت‌کنندگان آن را پذیرفته باشند نه زمانیکه کنترل بدون پذیرش سایرین تصاحب شود.

در این پژوهش، سخنان شرکت‌کنندگان به چهار گروه دسته‌بندی شده است:

۱. تایید ادعا[۱]: سخنانی که در مورد حقایق یک حالت به‌کار برده می‌شود. بله و خیر در پاسخ به سوالات، در این دسته قرار می‌گیرند زیرا اطلاعاتی را فراهم می‌کنند.

۲. فرمان‌ها[۲]: سخنانی که اشخاص را وادار به انجام کاری می‌کنند. به طور کلی شکل دستوری اما به صورت غیر مستقیم به معنی پیشنهاداتی است که می‌خواهیم دیگران انجام دهند.

۳. سوالات[۳]: سخنانی که برای استخراج اطلاعات درنظر گرفته شده است.

۴. سخنان بی‌درنگ[۴]: سخنانی که محتوای گزاره‌ای را ابراز نمی‌کنند مانند yeah، oh-huh. این سخنان فهمیدن یا تایید کردن را نمایش می‌دهند.

در جدول(۲-۱) مجموعه قوانین برای انتقال کنترل آورده شده است.

جدول (۲-۱): مجموعه قوانین برای انتقال کنترل[۵]

سخنان	کنترل کننده
تایید ادعا	جز در موارد پاسخ به سوالات کنترل به دست صحبت‌کننده است.
فرمان‌ها	کنترل به دست صحبت‌کننده است.
سوالات	جز در موارد پاسخ به سوالات و فرمان‌ها به دست صحبت‌کننده است.
سخنان بی‌درنگ	کنترل به دست شنونده است.

[۱] Assertions
[۲] Commands
[۳] Questions
[۴] Prompts
[۵] Control Rule

در شکل(۲-۳) مثالی از نحوه استفاده قوانین انتقال کنترل آورده شده است.

E: “And they are, in your gen you’ll find that they’ve relocated into the labeled common area” (ASSERT – E control)

C: “That’s right.” (PROMPT – E control)

E: “Yeah.” (PROMPT – E abdicates control)

---------- CONTROL SHIFT TO C ----------

C: “I’ve got two in there. There are two of them.” (ASSERT – C control)

E: “Right.” (PROMPT – C control)

C: “And there’s another one which is % RESA” (ASSERT – C control)

E: “Ok um.” (PROMPT – C control)

C: “VS.” (ASSERT – C control)

E: “Right.” (PROMPT – C control)

C: “Mm.” (PROMPT – C abdicate control)

---------- CONTROL SHIFT TO E ----------

E: “Right and you haven’t got. I assume you haven’t got local labeled common with those labels.” (QUESTION – E control)

از نظر ویتاکر و استنتون، کنترل به یکی از روش‌های زیر انتقال پیدا می‌کند:

۱. از دست دادن[۱]: این حالت درصورتی اتفاق می‌افتد که شخص کنترل کننده یک سخن بی‌درنگ بگوید.

۲. تکرار یا خلاصه‌سازی[۲]: این حالت درصورتی اتفاق می‌افتد که شخص کنترل کننده جمله قبلی را به همان صورت یا به صورت خلاصه شده تکرار کند.

۳. وقفه[۳]: این حالت درصورتی اتفاق می‌افتد که سایرین بخواهند کنترل را از شخص کنترل کننده بگیرند. معمولا زمانیکه اختلاف در باورهای متقابل وجود داشته باشد وقفه رخ می‌دهد اما این کار باعث انحراف در فعالیت‌های مشترک نمی‌شود. در شکل زیر نمونه‌ای از بروز وقفه در مکالمات آورده شده است.

[۱] Abdication
[۲] Repetition/Summary
[۳] Interrupt

A: … the only way I could do that was to take a to take a one third down and to take back a mortgage. (ASSERTION)

---------- INTERRUPT SHIFT TO B ----------

2. B: When you talk about one third put a number on it. (QUESTION)

3. A: uh 15 thou. (ASSERTION, but response)

4. B: go ahead. (PROMPT)

---------- ABDICATE SHIFT BACK TO A ----------

5. A: and then I'm a mortgage back for 36.

هابز[۲۷،۲۸] خاصیت انسجام[۱] را به مکالمات ابتکار عمل ترکیبی اضافه کرد. او یک دسته از روابط لفظی[۲] برای زمانیکه شخص صحبت‌کننده محتوای گزاره‌ای عبارت قبلی را تکرار می‌کند، توصیف کرد. توابع شناسایی تکرار از اصول کلی پیروی می‌کنند. سیگنال‌های شخص صحبت‌کننده که باعث تغییر کنترل می‌شوند(تکرار، خلاصه‌سازی و از دست دادن)، شامل هیچ اطلاعات جدیدی نیستند.

از نظر نویک و ساتن[۲]، ابتکار عمل ترکیبی یکی از مواردی است که به راحتی توسط افراد قابل شناسایی است حتی اگر قادر به تعریف آن نباشند. از نظر آن‌ها کنترل که با وظیفه نشان داده می‌شود، می‌تواند به صورت رقابت بین شرکت‌کنندگان در مورد حداکثر مقدار گریسی[۳] نشان داده شود.[۲۹] حداکثر مقدار گریسی میزانی است که در آن هر شرکت‌کننده به میزان مورد نیاز و نه بیشتر از اطلاعات را در اختیار بقیه قرار می‌دهد. در زیر نمونه‌ای از مکالمه دارای ابتکار عمل ترکیبی آورده شده است.

4.1 S: I can schedule an appointment for you. What day?

4.2 U: Tuesday.

4.3 S: What time on Tuesdat?

4.4 U: In the morning.

4.5a S: What time on Tuesday morning?

[۱] Coherence
[۲] Rhetorical Relation
[۳] Grice's Maxim of Quantity

4.5b S: What about 9:30 a.m. Tuesday morning?

در این مکالمه، ابتدا کنترل در اختیار سیستم است و سوالی را درباره تعیین روز ملاقات از کاربر می-پرسد. اما در ادامه کاربر کنترل را به دست گرفته و روند تعیین روز و ساعت ملاقات را در پیش می-گیرد. در این مکالمه به نظر می‌رسد گرفتن کنترل توسط سیستم غیرممکن است و درنهایت کاربر زمان دقیق ملاقات را تعیین می‌کند اما با توجه به حداکثر مقدار گریسی، کاربر یک محدوده زمانی را به جای یک زمان دقیق پیشنهاد داده(سخن ۴.۴) که نشان می‌دهد او دارای انعطاف است. درنتیجه سیستم می‌تواند کنترل را کاهش داده و درنهایت در سخن آخر کنترل را به دست می‌گیرد.

از نظر نویک و ساتن فاکتورهای زیر در تشخیص کنترل کننده دخالت دارند:

۱. انتخاب وظیفه[۱]: این فاکتور تعیین‌کننده موضوع مکالمه است و مشخص می‌کند مکالمه درباره چیست. گین[۳۰]، ابتکار عمل را تعیین میزان دنبال کردن تقسیم هدف در مکالمه تعریف می‌کند. در مدل اسمیت[۳۱]، سطح کنترل شرکت‌کننده به وسیله اهداف پیشنهادی آن برای تکمیل یک هدف مشخص می‌شود.

۲. انتخاب گوینده[۲]: مدل‌های ابتکار عمل به طور صریح از خاصیت گرفتن نوبت[۳] به عنوان یک مولفه و شاخص ابتکار عمل صحبت کرده‌اند.[۳۲،۳۳،۳۴،۳۵] براساس این تعریف، پیامدهایی برای ساختار گفتگو به وجود می‌آید. از نظر برک، نوبت آغازین در یک دنباله گفتگو، نه تنها در نوبت خودش بلکه در سایر نوبت‌ها در دنباله نیز کنترل را به دست می‌گیرد. این یکی از فاکتورهایی است که یک دستگاه را از گوینده انسانی متمایز می‌کند.

۳. انتخاب نتیجه[۴]: این فاکتور، یک گسترش منطقی از کنترل وظیفه است. این بدین معناست که عبارت "تکمیل یک هدف"[۳۱] به این معنی است که تولید یک نتیجه توسط یک یا چند شرکت‌کننده

---

[۱] Choice of Task
[۲] Choice of Speaker
[۳] Turn-Taking
[۱] Choice of Outcome

درنظر گرفته شده است. سپس فرآیند انتخاب نتیجه، وظیفه را تعیین کرده و اعمال و تصمیمات لازم برای رسیدن به وظیفه را تعیین می‌کند.

یکی از فاکتورهای ابتکار عمل، میزان داوطلب شدن اطلاعات در طی فرآیند گفتگو است. این بدین معناست که رفتارهای ارتباطی که به وسیله فاکتورهای انتخاب وظیفه، گوینده و نتیجه تعیین می‌شوند، داوطلبانه هستند. بنابراین تلاش برای تغییر وظیفه یا موضوع با استفاده از عباراتی که به نظر داوطلبانه هستند انجام می‌شود.

جنیفر کارول[۳۶]، نمونه‌ای از یک سیستم گفتگو دارای ابتکار عمل ترکیبی به نام MIMIC[۱]، جهت ارائه اطلاعات پخش فیلم فراهم کرد. از نظر او شرکت‌کنندگان در طول گفتگوهای انسانی اغلب راهبردهای گفتاری متفاوتی را براساس ویژگی‌های شنونده، تاریخچه گفتگو و موارد دیگر به کار می‌برند. مثلا سخنران ممکن است در مواردی که شنونده جهت تکمیل کار و فرآیند به مشکل برخورده، راهنمایی بیشتری فراهم کند و برعکس در مواردی که شنونده در زمینه کاری خبره است، کمتر کار کند. هدف اصلی او این است که این رفتارهای انسانی را به وسیله استراتژی‌های گفتگو و براساس اطلاعاتی که در طول مکالمه شناسایی می‌شوند، شبیه‌سازی کند.

از دیدگاه کارول و براون[۶]، دو نوع ابتکار عمل در گفتگوها یافت می‌شود. ابتکار کار[۲]، رهبری فرآیند را جهت دستیابی به هدف شرکت‌کنندگان برعهده دارد و ابتکار گفتگو[۳]، منجر به تعیین مرکز و هدف گفتمان فعلی می‌شود. MIMIC به طور خودکار، توزیع ابتکار را براساس اطلاعات استخراج شده از سخنان کاربر و تاریخچه گفتگو بروزرسانی می‌کند. توزیع ابتکار ممکن است بر استراتژی‌هایی که سیستم برای رسیدن به اهدافش انتخاب می‌کند، تاثیر بگذارد.

در این فریم ورک، ابتدا هر جمله کاربر تفسیر معنایی[۴] شده و به صورت زوج ویژگی-مقدار[۱] نمایش داده می‌شود. در مرحله دوم با استفاده از نشانه‌های به دست آمده از تفسیر معنایی براساس جدول(۲-

---

[۲] Mixed-Initiative Movie Information Consultant(MIMIC)
[۳] Task-Initiative
[۴] Dialogue-Initiative
[۱] Semantic Interpretation

۲)، توزیع این نشانه‌ها براساس شکل(۲-۲) به دست آمده و با توزیع ابتکار فعلی گفتمان ترکیب شده و ابتکار عمل سخن بعدی مشخص می‌شود. توزیع ابتکار با استفاده از یک عدد بین ۰ تا ۱ تعیین می-شود و نشان‌دهنده احتمال دریافت ابتکار عمل توسط هر یک از شرکت‌کنندگان است. به عنوان مثال $m_{t\text{-}tot}(\{speaker\})=0.35;\ m_{t\text{-}tot}(\Theta)=0.65$، به این معنی است که نشانه TakeOverTask از انتقال ابتکار کار به speaker با احتمال ۰.۳۵ پشتیبانی می‌کند و با احتمال ۰.۶۵ از بقیه حمایت می‌کند. Θ در این سیستم تنها شامل کاربر است.

جدول (۲-۲): نشانه‌های مشاهده شده در تفسیر معنایی جملات

نشانه	تفسیر
TakeOverTask(tot)	کاربر اطلاعاتی بیش از انتظار فراهم می‌کند. در این حالت کاربر قصد در دریافت کنترل دارد.
NoNewInfo(nni)	کاربر اطلاعات جدیدی ارائه نکرده است.
InvalidAction(ia)	فرضیات کاربر درباره دامنه مساله نامعتبر است.
InvalidActionResolved(iar)	فرض نامعتبر کاربر تصحیح شد.
AmbiguousAction(aa)	ویژگی اجباری مساله مشخص نشده است.
AmbiguousActionResolved(aar)	ویژگی اجباری مساله به صورت یکتا مقداردهی شد.

[۲] Attribute-Value

نشانه	توزیع ابتکار
TakeOverTask	$m_{t-tot}(\{speaker\}) = 0.35; m_{t-tot}(\Theta) = 0.65$
NoNewInfo	$m_{t-nni}(\{hearer\}) = 0.35; m_{t-nni}(\Theta) = 0.65$
InvalidAction	$m_{t-ia}(\{hearer\}) = 0.35; m_{t-ia}(\Theta) = 0.65$
InvalidActionResolved	$m_{t-iar}(\{hearer\}) = 0.35; m_{t-iar}(\Theta) = 0.65$
AmbiguousAction	$m_{t-aa}(\{hearer\}) = 0.35; m_{t-aa}(\Theta) = 0.65$
AmbiguousActionResolved	$m_{t-aar}(\{speaker\}) = 0.35; m_{t-aar}(\Theta) = 0.65$

a)ابتکار کار

TakeOverTask	$m_{d-tot}(\{speaker\}) = 0.35; m_{d-tot}(\Theta) = 0.65$
NoNewInfo	$m_{d-nni}(\{hearer\}) = 0.35; m_{d-nni}(\Theta) = 0.65$
InvalidAction	$m_{d-ia}(\{hearer\}) = 0.7; m_{d-ia}(\Theta) = 0.3$
InvalidActionResolved	$m_{d-iar}(\{hearer\}) = 0.7; m_{d-iar}(\Theta) = 0.3$
AmbiguousAction	$m_{d-aa}(\{hearer\}) = 0.7; m_{d-aa}(\Theta) = 0.3$
AmbiguousActionResolved	$m_{d-aar}(\{speaker\}) = 0.7; m_{d-aar}(\Theta) = 0.3$

b)ابتکار گفتگو

شکل (۲-۲): توزیع ابتکار نشانه‌ها

در مرحله سوم، سیستم با استفاده از الگوریتم شکل(۲-۳) هدفی که سعی در رسیدن به آن دارد را مشخص می‌کند و در نهایت استراتژی‌های لازم را براساس جدول(۲-۳) و شکل(۲-۴) برای رسیدن به هدف به کار می‌برد.

```
Select-Goal(SemRep):
(1)   If AmbiguousAction detected
(2)      ambiguous-attr ← get-ambiguous(SemRep)
            /* get name of ambiguous attribute */
(3)      If (number-values(ambiguous-attr) == 0)
            /* attribute unspecified */
(4)         Instantiate(ambiguous-attr)
(5)      Else /* more than one value specified */
(6)         Constrain(ambiguous-attr)
(7)   Else if InvalidAction detected
(8)      ProvideNegativeAnswer(SemRep)
(9)   Else /* well-formed query */
(10)     answer ← database-query(SemRep)
(11)     ProvideAnswer(answer)
```

شکل (۲-۳): الگوریتم انتخاب هدف مساله

جدول(۲-۳): انتخاب استراتژی براساس هدف و توزیع ابتکار

	Task + Dialogue	Dialogue	None
Constrain	Clarify	Clarify	Acknowledge
Instantiate	GiveOptions InfoSeek	InfoSeek	Acknowledge
ProvideAnswer	Answer LimitedPrompt	Answer OpenPrompt	Answer
ProvideNegativeAnswer	NotifyFailure SuggestAlternative LimitedPrompt	NotifyFailure OpenPrompt	NotifyFailure

Dialogue Act	Template
Clarify	"Did you say < $value_1$ > ..., or < $value_n$ >?"
Acknowledge	"Uh-huh."
GiveOptions	"Choices for < *attribute* > are < $value_1$ > ... < $value_n$ >."
InfoSeek	"What < *attribute* > would you like?"
Answer	E.g., "< *movie* > is playing at < *theater* > at < $time_1$ > ... < $time_n$ >."
OpenPrompt	"Can I help you with anything else?"
LimitedPrompt	"Please say the name of the movie or theater or town you would like information about."
NotifyFailure	E.g., "< *movie* > is not playing at < *theater* > .
SuggestAlternative	E.g., "< *movie* > is playing at < *alternativetheater* > at < $time_1$ > ... < $time_n$ >."

شکل(۲-۴): نگاشت بین اعمال و الگوهای گفتاری

در شکل(۲-۵) نمونه‌ای از توزیع ابتکار کار و گفتگو در یک مکالمه آورده شده است.

		Utterance	Task Initiative	Dialogue Initiative
(1)	S:	Hello, this is MIMIC, the movie information system.	User	System
(2)		How can I help you?		
(3)	U:	What time is Analyze This playing in Montclair? *[what time is analyze this playing in montclair]*		
(4)	S:	What theater would you like?	User	System
(5)	U:	Montclair. *[montclair]*		
(6)	S:	Choices for theater are Wellmont Theater and Clearviews Screening Zone.	System	System
(7)		What theater would you like?		
(8)	U:	Clearviews Screening Zone. *[um movie screening zone]*		
(9)	S:	Analyze This is playing at Clearviews Screening Zone in Montclair at 1:15pm, 4:10pm, and 6:30pm.	System	System
(10)		Please say the name of the movie or theater or town you would like information about.		
(11)	U:	What time is Analyze This playing at the Wellmont Theater? *[what time is analyze this playing at the wellmont theater]*		
(12)	S:	Analyze This is playing at Wellmont Theater in Montclair at 3:15pm, 5:45pm, 8:20pm, and 10:10pm.	User	System
(13)		Can I help you with anything else?		

شکل (۲-۵): توزیع ابتکار کار و گفتگو در مکالمه

در ابتدا سیستم خوشامدگویی کرده و ابتکار گفتگو را به دست می‌گیرد، در حالیکه ابتکار کار با کاربر است. در جمله ۳، کاربر ویژگی اجباری تئاتر را مشخص نکرده و نمی‌توان نتیجه گرفت که منظور

کدام تئاتر است زیرا در مونت‌کارلو چند تئاتر برگزار می‌شود. در این مثال فرض شده توزیع ابتکار فعلی به صورت زیر است.

$M_{t-(3)}(\{s\}) = 0.3$ $M_{t-(3)}(\{u\}) = 0.7$ $M_{d-(3)}(\{s\}) = 0.6$ $M_{d-(3)}(\{u\}) = 0.4$

با توجه به تفسیر معنایی انجام شده، نشانه aa دیده شده است. براساس توزیع ابتکار فعلی و توزیع ابتکار نشانه aa، توزیع ابتکار بعدی به صورت زیر به دست می‌آید.

$M_{t-(4)}(\{s\}) = 0.4$ $M_{t-(4)}(\{u\}) = 0.6$ $M_{d-(4)}(\{s\}) = 0.83$ $M_{d-(4)}(\{u\}) = 0.17$

با توجه به نشانه aa و عدم وجود مقدار برای ویژگی اجباری تئاتر، الگوریتم انتخاب هدف تابع Instantiate را فراخوانی می‌کند. از آنجایی که سیستم فقط دارای ابتکار گفتگو است عمل InfoSeek انجام می‌شود و جمله ۴ تولید می‌شود. در پاسخ کاربر در جمله ۵، باز هم تئاتر مشخص نشده و نشانه aa مشاهده می‌شود و از طرفی چون کاربر به نوعی پاسخش را تکرار کرده نشانه nni نیز مشاهده می‌شود. با ترکیب توزیع ابتکار فعلی و توزیع ابتکار نشانه aa و nni، توزیع ابتکار بعدی به صورت زیر به دست می‌آید.

$M_{t-(6)}(\{s\}) = 0.62$ $M_{t-(6)}(\{u\}) = 0.38$ $M_{d-(6)}(\{s\}) = 0.96$ $M_{d-(6)}(\{u\}) = 0.04$

با توجه به نشانه aa و عدم وجود مقدار برای ویژگی اجباری تئاتر، الگوریتم انتخاب هدف تابع Instantiate را فراخوانی می‌کند. اینبار سیستم دارای هر دو ابتکار است و عمل InfoSeek و GiveOptions را با هم انجام می‌دهد و جمله ۶ و ۷ تولید می‌شود. با پیشنهاد گزینه توسط سیستم، کاربر در جمله ۸ جواب درستی برای تئاتر ارائه کرد. از آنجایی که در جمله ۸ مشکلی دیده نمی‌شود و کاربر پاسخ سوال را به درستی داده است، سیستم در جمله ۹، جواب کاربر را مبنی بر زمان پخش تئاتر و در جمله ۱۰، یک جمله محدود[۱] ارائه می‌کند. در جمله ۱۱، نشانه tot مشاهده می‌شود و کاربر سعی در گرفتن کنترل است. در این حالت ابتکار وظیفه به کاربر داده شده و بر این اساس سیستم در جمله ۱۲، جواب کاربر را مبنی بر زمان پخش تئاتر و در جمله ۱۳، یک جمله باز[۲] ارائه می‌کند.

[۱] Limited Prompt
[۱] Open Prompt

سیستم MIMIC، ابتکار کار را با دقت ۹۷٪ و ابتکار گفتگو را با دقت ۸۸٪ به درستی پیش‌بینی کرده است.

تحقیقات گذشته نشان‌دهنده سه روش ارسال ساختار گفتمان از صحبت کننده به شنونده است:

۱. نشانه‌ها[۱][۳۷،۳۸]: عباراتی مثل "because"، "so" و "but" سهم فعلی صحبت کننده در اتفاقاتی که در طی مکالمه گذشته است را نشان می‌دهد. مثلا عبارت "so" نشان‌دهنده نتیجه‌گیری صحبت کننده است. یا عبارت "and" و "but" یعنی موضوع جدیدی در حال معرفی است درحالیکه عبارت "any way" و "in any case" نشان‌دهنده بازگشت به موضوع قبلی است.

۲. لحن گفتمان[۲][۳۹]: تغییر لحن گفتمان رابطه نزدیکی با تغییر موضوعات گفتگو دارد.

۳. عبارات و ساختارهای دستوری[۳][۴۰،۴۱]: این عبارات و ساختارها باعث حفظ یا تغییر تمرکز مکالمه می‌شود.

ویتاکر و استنتون[۲۶]، نشانه‌ها و کنترل را در سیستم‌های دارای ابتکار عمل ترکیبی بررسی کردند. آنها برای این کار، از مکالمات مرکز پاسخگویی به مشکلات نرم‌افزاری استفاده کردند. از دیدگاه آنها سه نشانه کلی برای انتقال کنترل وجود دارد:

۱. سخنان بی‌درنگ: این سخنان محتوای گزاره‌ای ندارند (مثل yeah). در ۲۱ مورد از ۵۶ انتقال کنترل، از سخنان بی‌درنگ استفاده شده است. این امر نشان می‌دهد که صحبت‌کننده حرفی برای گفتن ندارد و خواهان انتقال کنترل به دیگران است.

۲. تکرار و خلاصه‌سازی سخنان: در ۶ مورد از ۵۶ انتقال کنترل، از سخنان تکراری و در ۹ مورد از سخنان خلاصه شده استفاده شده است. این امر نشان می‌دهد که صحبت‌کننده حرف جدیدی برای گفتن ندارد و خواهان انتقال کنترل به دیگران است. نیمی از تکرارهای انجام شده با استفاده از نشانه-

[۲] Cues
[۳] Intonation
[۴] Pronominalisation

هایی مثل "and"، "well" و "so" انجام شده است. ریچمن[۳۹]، پیشنهاد می‌کند که عبارت "so" یکبار، "and" سه بار، "now" یکبار و "but" یکبار می‌توانند نشانه‌های جملات خلاصه‌شده باشند.

۳. وقفه‌ها: وقفه‌ها ممکن است به دلایل مختلفی رخ دهند و معمولا نشانه‌ای ندارند. در ۶ مورد از ۵۶ انتقال کنترل، وقفه به دلیل مخالفت شنونده با سخن صحبت‌کننده یا دادن اطلاعات مرتبط به صحبت‌کننده رخ داد. در ۸ مورد از ۵۶ انتقال کنترل، وقفه به دلیل بازگشت کنترل بعد از دادن اطلاعات مرتبط شنونده رخ داد. و در نهایت در ۶ مورد از ۵۶ انتقال کنترل، وقفه به دلیل روشن شدن اطلاعاتی که در طی گفتگو بیان شده رخ داد.

بررسی‌های ویتاکر و استنتون نشان می‌دهد که در برخی موارد، تغییر کنترل باعث تغییر موضوع[۱] می‌شود. این مبحث با ۵ داور و در ۴ گفتگو انجام شد و هر ۵ داور در ۲۴ مورد از ۵۶ انتقال کنترل، تغییر موضوع بحث را پذیرفتند. همچنین ۴ داور در ۲۲ مورد از ۵۶ انتقال کنترل، تغییر موضوع بحث را پذیرفتند. اما این مبحث همیشه درست نیست و گاهی تغییر کنترل باعث ماندن در همان موضوع قبلی می‌شود. به عنوان مثال در ۹۰٪ وقفه‌ها، ۵۷٪ سخنان بی‌درنگ و ۶۷٪ سخنان تکراری و خلاصه‌شده تغییر موضوع صورت نگرفت.

فنتی، ساتن، نویک و کول[۴۲]، یک سیستم دارای ابتکار عمل ترکیبی برای زمانبندی قرار ملاقات‌ها از طریق تلفن طراحی کردند. هدف از طراحی این سیستم این است که تماس گیرنده تا حد امکان برای هدایت گفتگو آزادی داشته باشد. با این حال موقعیت‌هایی وجود دارند که باید توسط سیستم کنترل شوند. برای مثال در زمانیکه تقویم برای دوره زمانی که در حال حاضر روی آن بحث می‌شود شلوغ باشد، سیستم کنترل را به دست گرفته و زمان مناسبی را پیشنهاد می‌دهد. در این سیستم تغییر تمرکز به وسیله پشته انجام می‌گیرد. در جدول(۲-۴) نمونه‌ای از روش کار این سیستم آورده شده است.

[۱] Topic Shift

جدول (۲-۴): استفاده از پشته برای تغییر تمرکز

Utterance	*Preference Stack*
S: On what day would you like to schedule an appointment?	[ ]
C: Monday.	[Mon]
S: What time on Monday?	[Mon]
C: Afternoon.	[Mon aft, Mon]
S: How about 3pm?	[Mon 3pm, Mon aft, Mon]
C: Oh no, make that Tuesday	[Tue]
S: What time on Tuesday?	[Tue]

هدف اصلی سیستم تعیین زمان ملاقات برای تماس گیرنده است. بنابراین پشته با قرار ملاقات‌هایی که توسط تماس گیرنده و سیستم ارائه شده پر می‌شود. در ابتدا پشته خالی است. زمانی که تماس گیرنده یک روز را مشخص می‌کند ("Monday")، این قرار ملاقات به پشته اضافه می‌شود. در ادامه سیستم زمان را می‌پرسد و چون اطلاعات جدیدی توسط سیستم ارائه نشده، محتوای پشته تغییری نمی‌کند. سپس تماس گیرنده "Afternoon" را پیشنهاد می‌دهد و این قرار ملاقات به پشته اضافه می‌شود. در ادامه سیستم ساعت ۳ بعدازظهر را پیشنهاد داده و به پشته اضافه می‌شود. با وجود این قرار ملاقات، تماس گیرنده روز "Tuesday" را پیشنهاد می‌دهد. این قرار ملاقات با مواردی که در پشته اضافه شده، تضاد دارد. بنابراین مقدار جدید با تک تک عناصر موجود در پشته مقایسه شده و در صورت مشاهده تضاد، آن عنصر از پشته حذف می‌شود. ابتدا "Tuesday" با "Mon 3pm" مقایسه شده و چون روز سه‌شنبه با دوشنبه تضاد دارد، مقدار بالای پشته حذف می‌شود. در مرحله بعد "Tuesday" با "Mon aft" مقایسه شده و باز هم به دلیل تضاد این دو، مقدار بالای پشته حذف می‌شود. در پایان "Tuesday" با "Mon" مقایسه شده و باز هم به دلیل تضاد این دو، مقدار بالای پشته حذف می‌شود. اکنون پشته خالی است و هیچ تضادی دیده نمی‌شود، بنابراین مقدار جدید "Tuesday" به پشته اضافه می‌شود. با اضافه شدن مقدار ملاقات جدید به پشته، صحبت‌کننده کنترل مکالمه را به دست می‌گیرد.

لوین و همکارانش[۴۳]، از مدل فرآیند تصمیم‌گیری متوالی[۱] برای پیاده‌سازی یک سیستم دارای ابتکار عمل ترکیبی استفاده کردند. از آنجایی که در این سیستم، کاربر می‌تواند جریان گفتگو را در هر لحظه تغییر دهد، تعداد ورودی‌های احتمالی در هر نقطه از گفتگو بسیار بزرگ است. مدل فرآیند تصمیم‌گیری متوالی بر اساس حالات و اعمال گفتگو تعریف می‌شود. اعمال مطابق با تعاملات سیستم با دنیای خارج مثل کاربران و سرور زمانبندی(زمانی که پاسخ سوالی آماده نیست) است. حالات دانشی است که مدیر گفتگو در هر نقطه از گفتگو در جهت تعیین عمل بعدی نگهداری می‌کند.

براساس نتایج به دست آمده از این سیستم، با وجود ابتکار عمل ترکیبی، کاربران ابتکار عمل را نمی-پذیرند و عمدتا ترجیح می‌دهند دستورالعمل‌ها و مفاهیم سیستم را دنبال کنند.

از نظر گین[۳۰]، در محیط‌های دارای ابتکار عمل ترکیبی هر شرکت‌کننده برای رسیدن به هدف ایده‌ای را مطرح می‌کند و باعث ایجاد تضاد[۲] می‌شود. یکی از روش‌های حل این موضوع تعیین کنترل کننده است. شخصی که ابتکار عمل را برعهده دارد به هدف می‌پردازد و بررسی می‌کند از کدام مسیر برای رسیدن به هدف استفاده شود. سطوح ابتکار عمل یک شرکت‌کننده به یکی از روش‌های زیر تعیین می‌شود:

۱. الگوریتم continuous mode: این الگوریتم براساس یک آنالیز احتمالی از دانش شرکت‌کنندگان عمل می‌کند و سطوح ابتکار عمل را در طول حل مساله مشخص می‌سازد.
۲. جملات خلاصه‌شده[۳]: این جملات نشان‌دهنده این هستند که صحبت‌کننده خواهان انتقال کنترل به دیگران است.
۳. مذاکره[۴]: شرکت‌کنندگان برای دریافت ابتکار عمل مذاکره می‌کنند.

[۱] Sequential Decision Process
[۲] Conflict
[۳] Summaries
[۴] Negotiation

در جدول(۲-۵) تاثیر هر کدام از پارامترهای جملات خلاصه‌شده، مذاکره و الگوریتم continuous mode را در زمان، تعداد سخنان و تعداد شاخه پیمایش شده برای رسیدن به جواب نشان می‌دهد. این نتایج براساس ۵۰۰۸ گفتگو مربوط به دامنه قتل جمع‌آوری شده است.

جدول (۲-۵): تاثیر پارامترهای جملات خلاصه‌شده، مذاکره و الگوریتم continuous mode در زمان، تعداد سخنان و تعداد شاخه پیمایش شده برای رسیدن به جواب

Summaries	No	No	No	No	Yes	Yes	Yes	Yes
Negotiation	No	Yes	No	Yes	No	Yes	No	Yes
Continuous Mode	No	No	Yes	Yes	No	No	Yes	Yes
Time (secs.)	82.398	70.782	44.528	43.263	84.363	31.822	46.685	41.619
Utterances	39.921	26.667	26.650	25.784	45.035	25.276	29.415	26.969
Branches Explored	6.188	4.394	3.412	3.334	6.135	3.655	3.400	3.289

باتوجه به آزمایشات انجام شده، بهترین زمان و کمترین سخنان برای رسیدن به جواب مربوط به حالتی است که از جملات خلاصه‌شده و مذاکره برای انتقال ابتکار عمل استفاده شده است.

راماکریشنان و همکارانش[۴۴]، یک سیستم ثبت سفارش پیتزا با قابلیت ابتکار عمل ترکیبی طراحی کردند. آنان، ابتکار عمل ترکیبی را به سطوح زیر دسته‌بندی می‌کنند:

۱. گزارشات ناخواسته[۱][۲۴]: صحبت‌کننده اطلاعاتی خارج از انتظار ارائه می‌کند. (مثلا درباره سوال "دوست دارید اندازه پیتزا چقدر باشد؟"، صحبت‌کننده بگوید "من پیتزا سوسیس دوست دارم"). پردازش پاسخ‌های ناخواسته ممکن است باعث انتقال کنترل به دنباله مکالمه جدید، قطع شدن تعامل کنونی و یا تکمیل تعامل کنونی شود.

۲. فراخوانی زیرگفتگو[۲]: سیستم پاسخ کاربر را متوجه نشده و برای روشن شدن مساله، از کاربر سوال می‌پرسد و ابتکار عمل را به او واگذار می‌کند.

۳. مذاکره[۲۴]: شرکت‌کنندگان برای تعیین ابتکار عمل با هم مذاکره می‌کنند. در زیر نمونه‌ای از این مکالمه آورده شده است.

[۱] Unsolicited Reporting
[۲] Subdialogue Invocation

Husband: Della, Something interesting happened today that I want to tell you.

Wife: I too have something exciting to tell you, Jim.

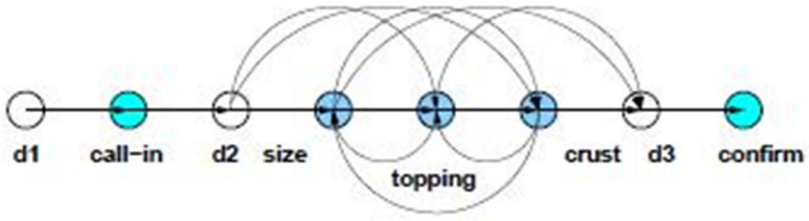
Husband: Do you want to go first or shall I tell you my story?

در حالت عادی، ویژگی‌ها (اندازه، نوع و ظرف پیتزا) به ترتیب پرسیده شده و توسط تماس‌گیرنده پاسخ داده می‌شوند. در این نوع مکالمات، کنترل در تمام مدت گفتگو به عهده سیستم می‌باشد. اما اگر هر ویژگی در زمان ترتیبی خود پاسخ داده نشود، مکالمات دارای ابتکار عمل ترکیبی شکل می-گیرند. در شکل(۲-۶) الگوریتم مقداردهی ویژگی‌های مساله آورده شده است.

```
pizzaorder(size,topping,crust)
{
   if (unfilled(size)){
      /* prompt for size */
   }
   if (unfilled(topping)){
      /* prompt for topping */
   }
   if (unfilled(crust)){
      /* prompt for crust */
   }
}
```

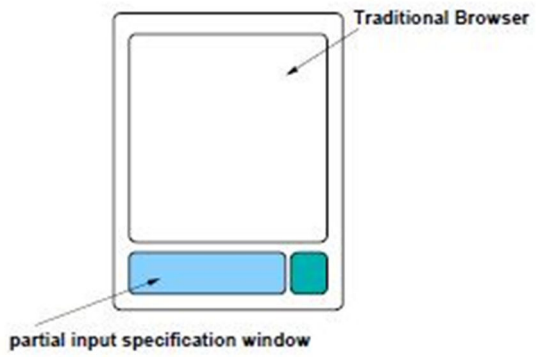
شکل(۲-۶): الگوریتم مقداردهی ویژگی‌های پیتزا

در شکل(۲-۷) توالی تعاملات معتبر در برنامه سفارش پیتزا آورده شده است. تمامی گره‌های این گراف فقط یکبار باید بازدید شوند. برنامه اولین ویژگی خالی را یافته و سوال مربوط به آن را می‌پرسد.



شکل (۲-۷): گراف توالی تعاملات معتبر در برنامه سفارش پیتزا

در شکل(۲-۸) نمای کلی برنامه ثبت سفارش آورده شده است.



شکل(۲-۸): نمای کلی برنامه ثبت سفارش

سفارش‌دهنده برای ثبت سفارش خود دو روش دارد. او می‌تواند از بخش مرورگر بالا تمام ویژگی‌های مورد نیاز خود را انتخاب کند و یا از طریق پنجره ورودی پایین نیازهای خود را وارد نماید.

مارتی هرست[۲۴]، سطوح مختلف ابتکار عمل ترکیبی را به صورت زیر دسته‌بندی کرده است:

۱. پاسخ ناخواسته: عامل در صورت مشاهده شرایط بحرانی، اطلاعات مربوط به آن را به دیگران اطلاع می‌دهد. به عنوان مثال سیستم به طور مداوم وضعیت فعلی برنامه را بررسی کرده و موفق بودن آن را تایید می‌کند.

۲. شروع یک زیرگفتگو[۱]: عامل ممکن است یک زیرگفتگو را برای روشن شدن و یا اصلاح موضوعی آغاز کند. عامل تا مدتی ابتکار عمل را به عهده می‌گیرد و پس از روشن شدن موضوع، کنترل را به عامل اول برمی‌گرداند.

۳. ابتکار عمل زیر وظیفه ثابت[۲]: عامل برای انجام زیر وظیفه‌های از پیش تعیین شده، ابتکار عمل را برعهده می‌گیرد. پس از اتمام زیر وظیفه، کنترل به عامل اول برمی‌گردد.

۴. ابتکار عمل ترکیبی توافقی[۳]: در این سطح، عامل‌ها برای دریافت ابتکار عمل مذاکره و هماهنگی می‌کنند و هیچ مسئولیت و ابتکار عمل ثابتی وجود ندارد. این تصمیم می‌تواند براساس یک سری فاکتورها صورت بگیرد. این فاکتورها می‌توانند شامل توانایی عامل برای هماهنگی زیر وظیفه فعلی، تقاضای سایر عامل‌ها و ارزیابی سایر عامل‌ها از قابلیت یک عامل باشد.

---

[۱] Subdialogue Initiation
[۲] Fixed-Subtask Initiative
[۳] Negotiated Mixed-Initiative

از دیدگاه گین[۲۴]، برای تعیین کنترل کننده می‌توان مکالمات را به بخش‌های گفتمانی تجزیه کرده و برای هر بخش یک برچسب در نظر گرفت. برچسب‌های هر بخش نشان‌دهنده عملکرد آن بخش می‌باشد.

طرح کدگذاری استاندارد پن[۱][۴۵]، از برچسب‌های زیر برای تعیین ابتکار عمل استفاده می‌کند:

۱. ابتکار عمل: سخنران کنترل بحث را برعهده دارد.
۲. پاسخ و تایید: سخن گفته شده در پاسخ به جمله قبلی است و آن را تایید می‌کند.
۳. پاسخ و رد: سخن گفته شده در پاسخ به جمله قبلی است اما پیشنهاد یا محتوای آن را رد می‌کند.
۴. پاسخ به سایر جواب‌ها: سخن گفته شده در پاسخ به جمله قبلی است و محتوای آن را رد و یا تایید نمی‌کند.

برای تعیین ابتکار عمل از ضریب کاپا[۲] استفاده شد. ضریب کاپا و تحلیل آماری مبتنی بر آن اندازه‌ای عددی بین ۱- تا ۱+ است، که هر چه به ۱+ نزدیکتر باشد بیانگر وجود توافق متناسب و مستقیم می‌باشد. اندازه‌های نزدیک به ۱- نشان‌دهنده وجود توافق وارون و عکس و اندازه‌های نزدیک به صفر عدم توافق را نشان می‌دهد. در این پژوهش، مقدار بیش از ۰.۸ بیانگر قابلیت اطمینان بالا، مقدار ۰.۶۷ تا ۰.۸ قابلیت اطمینان متوسط و مقادیر کمتر از ۰.۶۷ قابلیت اطمینان پایین را نشان می‌دهند.

یکی دیگر از روش‌های برچسب‌گذاری، طرح نارگیل[۳] است که از مفاهیم توابع پیش‌رو[۴] و عقب‌رو[۵] به صورت زیر استفاده می‌کند[۴۶]:

۱. توابع پیش‌رو: این وظایف شامل بیانیه‌ها (اظهارات و تاییدات)، تاثیرات روی شنونده (درخواست اطلاعات، دستورالعمل‌های عملیاتی و ارائه گزینه‌ها) و تاثیرات روی سخنران (پیشنهادات و تعهدات) می‌باشد.

---

۱ Penn Multiparty Standard Coding
۲ Kappa Statistic
۳ Coconut Scheme
۴ Forward-looking Function
۵ Backward-looking Function

۲. توابع عقب‌رو: این وظایف شامل پاسخ‌ها (به درخواست‌های قبلی) و توافقات (پذیرش و یا رد اعتقاد یا پیشنهاد ارائه شده در جمله قبلی) می‌باشد.

براساس مطالعات انجام شده، ضریب کاپا برای بیانه‌ها ۰.۸۳ و برای پاسخ‌ها ۰.۷۹ است که بیانگر این است که این برچسب‌ها بسیار قابل اطمینان هستند. ضریب کاپا برای تاثیرات روی شنونده و سخنران ۰.۷۲ است که مقدار معقولی می‌باشد. اما ضریب کاپا برای توافقات ۰.۵۴ است که بیانگر قابلیت اطمینان پایین می‌باشد.

این نتایج نشان‌دهنده این است که بسیاری از ویژگی‌های معتبر (بیشتر توابع پیش‌رو) در تنظیم و تغییر ابتکار عمل قابل استفاده هستند. با این حال ممکن است کلاس مهمی از سخنان که در تعیین ابتکار عمل دخیل هستند (توافقات)، نیاز به ارزیابی‌های انسانی داشته باشند.

مرلین واکر و همکارانش[۴۷]، تعدادی از فاکتورهای کارایی تاثیرگذار بر ابتکار عمل را مطرح کرده‌اند. این فاکتورها به دو دسته معیارهای عینی[۱] و ذهنی[۲] تقسیم می‌شوند. معیارهای عینی مثل متوسط زمان پاسخ، به ارزیابی‌های انسانی نیاز ندارند در حالی که معیارهای ذهنی مثل تناسب پاسخ‌ها، نیازمند ارزیابی‌های انسانی هستند. در جدول(۲-۶) لیستی از این فاکتورهای عینی و ذهنی آورده شده است. واکر با استفاده از آنالیز رگرسیون خطی[۳]، فاکتورهایی که در حداکثر رساندن موفقیت‌ها و کم کردن هزینه‌ها دخیل هستند را مشخص کرد.

[۱] Objective Metrics
[۲] Subjective Metrics
[۳] Linear Regression Analysis

جدول (۲-۶): فاکتورهای عینی و ذهنی در تعیین ابتکار عمل

معیارهای عینی	معیارهای ذهنی
درصد پاسخ‌های درست	درصد سخنان اصلاح شده ضمنی
درصد تراکنش‌های موفق	درصد سخنان اصلاح شده صریح
تعداد چرخش‌های گفتگو	درصد سخنان مناسب سیستم
زمان گفتگو	همکاری
زمان پاسخ کاربر	درصد سخنان صحیح و نسبتا صحیح
زمان پاسخ سیستم	رضایت کاربر
درصد پیام‌های خطا	تعداد تغییر ابتکار عمل
درصد سخنان غیر بدیهی	تعداد رویدادهای تغییردهنده صریح ابتکار عمل
میانگین طول سخنان	تعداد رویدادهای تغییردهنده ضمنی ابتکار عمل
اتمام وظیفه	سطح ابتکار عمل
	تعداد بخش‌های گفتمان
	چگالی دانش سخنان
	الگوهای هم مرجع

حاتم خوزامی و همکارانش [۴۸]، به بررسی ابتکار عمل ترکیبی در سیستم ثبت رویداد جشن تولد پرداختند. در این سیستم، از طریق تعاملاتی که بین کاربر و سیستم برقرار می‌شود، رویدادهای جدیدی اضافه، به‌روزرسانی و حذف می‌شوند. ثبت این رویدادها به وسیله ۳ استراتژی انجام می‌گیرد. در استراتژی ابتکار عمل سیستم، سیستم سوالات را یکی پشت دیگری پرسیده و کاربر فقط به این سوالات پاسخ می‌دهد. در استراتژی ابتکار عمل کاربر، از کاربر خواسته می‌شود تا تمام اطلاعات موردنظر را در یکبار بدهد. اگر در این اطلاعات ابهامی وجود داشته باشد، از کاربر خواسته می‌شود دوباره آنرا تکرار یا اصلاح کند. در استراتژی ابتکار عمل ترکیبی، مکالمه به صورت ابتکار عمل کاربر

آغاز می‌شود. در این استراتژی در صورتی که ابهامی در اطلاعات وجود داشته باشد، مکالمه به صورت ابتکار عمل سیستم ادامه پیدا می‌کند.

این سیستم با استفاده از ۱۰۰۰ مکالمه ارزیابی شد. با توجه به نتایج بدست آمده ابتکار عمل سیستم نسبت به ابتکار عمل کاربر بسیار خسته‌کننده‌تر است و زمان تکمیل کار در این استراتژی افزایش می-یابد. با این حال ابتکار عمل ترکیبی از نظر زمان تکمیل کار و کارایی بهتر از دو استراتژی دیگر است.

فصل سوم

روش پیشنهادی

۳-۱- مقدمه

در این فصل به بررسی پایگاه داده و روش‌های پیشنهادی برای تعیین کنترل کننده در مکالمات پرداخته می‌شود. پایگاه داده استفاده شده در این پژوهش شامل مکالمات زبان انگلیسی درباره مسائل روزمره می‌باشد.

روش‌های تعیین کنترل کننده که تا کنون بیان شد، به بررسی لغوی و معنایی مکالمات پرداخته و به زبان وابسته هستند. در این فصل به بررسی روش‌های آماری و فاقد تجزیه و تحلیل معنایی و گرامری پرداخته می‌شود که این روش‌ها مستقل از زبان می‌باشند. شناسایی آماری، مبتنی بر مشاهدات و نمونه‌برداری بوده و آموزش با استفاده از داده‌های آماری و پرتکرار انجام می‌شود.

برای تجزیه و تحلیل و تشخیص کنترل کننده در مکالمات، ویژگی‌هایی از آن‌ها استخراج می‌شود. با استفاده از این ویژگی‌ها به بررسی هر یک از جملات مکالمه پرداخته و شخص کنترل کننده در آنها تعیین می‌شود. در هر یک از جملات شخص کنترل‌شونده و کنترل کننده، کلمات تکرار می‌شوند. از این رو از روی توزیع کلمات می‌توان به تعیین کنترل کننده پرداخت. در ادامه فصل به بررسی ویژگی-ها و روش‌های تعیین کنترل کننده پرداخته می‌شود.

۳-۲- پایگاه داده

برای تعیین کنترل کننده در مکالمات، نیازمند پایگاه داده‌ای با تعداد زیادی از جملات و موضوعات متنوع داریم. در روش‌های آماری، هر چه تعداد نمونه‌ها بیشتر باشد، سیستم بهتر آموزش دیده و با دقت بالاتری عمل می‌کند. به دلیل عدم وجود پایگاه داده در این زمینه، پایگاه داده‌ای شامل مکالمات زبان انگلیسی جمع‌آوری شده است. این مکالمات شامل محاوره بین دو نفر درباره مسائل روزمره مانند اجاره خانه، دریافت وام، قرار ملاقات، گزارش جرم و ... می‌باشد.

پایگاه داده شامل ۴۶۸۸ نمونه بوده و به دو دسته نمونه‌های کنترل کننده و کنترل‌شونده تقسیم شده‌اند. با بررسی نمونه‌های هر دسته و توزیع کلمات مربوط به هر دسته به تشخیص کنترل کننده در مکالمات پرداخته می‌شود. در زیر، دو نمونه از این پایگاه داده آورده شده است.

<personA controll="A">My wife said to have a lawyer look at the contract.</personA>
<personB controll="A">That's always a good idea.</personB>

نمونه‌های موجود در پایگاه داده دارای برچسبی به نام controll هستند که مشخص‌کننده شخصی است که کنترل را برعهده دارد. چنانچه شخص صحبت‌کننده و مقدار برچسب controll یکی باشد، این نمونه در کلاس کنترل کننده قرار می‌گیرد و در غیر این صورت در کلاس کنترل‌شونده قرار می‌گیرد. نمونه اول در کلاس کنترل کننده و نمونه دوم در کلاس کنترل‌شونده قرار می‌گیرد.

۳-۳- روش‌های پیشنهادی

استخراج ویژگی‌های مناسب مکالمه در شناسایی درست کنترل کننده تاثیر فراوانی دارد. این ویژگی‌ها می‌توانند به صورت آماری، معنایی و گرامری تعیین شوند. ما در این پژوهش به استخراج آماری ویژگی‌ها پرداخته و از روش‌های آماری، برای تعیین کنترل کننده در جملات مکالمه استفاده کرده‌ایم. در هر یک از جملات شخص کنترل‌شونده و کنترل کننده، کلمات تکرار می‌شوند. از این رو از روی تکرار و توزیع کلمات به تعیین کنترل کننده می‌پردازیم.

در روش‌های آماری، الگوهای مورد نیاز با استفاده اطلاعات آماری داده‌ها استخراج شده و برای شناسایی نمونه‌های جدید به کار برده می‌شوند. در اولین گام، سیستم با استفاده از داده‌های جامع و کامل آموزش داده می‌شود و پارامترهای سیستم تخمین زده می‌شوند. سپس برای ارزیابی عملکرد سیستم، از داده‌های جدید و دیده‌نشده استفاده شده و درصد درست‌نمایی سیستم محاسبه می‌شود. کیفیت کارایی تشخیص، به وسیله تشابه الگوهای تمرینی و داده‌های واقعی که سیستم در طول عملیات مواجه خواهد شد، تعیین می‌شود.

۳-۳-۱- مجموعه ویژگی‌ها

انتخاب ویژگی‌های مناسب تاثیر بسزایی در عملکرد الگوریتم‌های تعیین کنترل کننده دارد. در این پژوهش، از مجموعه کلمات استفاده شده در مکالمات به عنوان مجموعه ویژگی استفاده شده است. دو بردار برای مجموعه کلمات استفاده شده توسط شخص کنترل کننده و کنترل‌شونده در نظر گرفته شده است. هر یک از این بردارها، دوبعدی بوده و سطر اول حاوی کلمات بکار رفته در مکالمات و سطر دوم حاوی تعداد تکرارهای آن کلمه می‌باشد. در ادامه، جهت نرمال کردن بردارهای ویژگی، تعداد تکرارهای کلمات را بر تعداد کل نمونه‌های آن دسته تقسیم نموده و خروجی الگوریتم‌ها را بررسی کرده‌ایم. روش‌های بکار رفته در این پژوهش، با بررسی شباهت‌های هر جمله به این بردارها، آن را به یکی از دو دسته کنترل کننده و کنترل‌شونده دسته‌بندی می‌کنند.

۳-۳-۲- روش مبتنی بر فاصله اقلیدسی[۱]

در روش‌های مبتنی بر فاصله، مثلاً برای یک مساله دو کلاسه، یک ویژگی یا یک مجموعه ویژگی مثل p بر یک ویژگی یا یک مجموعه ویژگی دیگر مثل q ارجحیت دارد، اگر که با آن مجموعه ویژگی مقادیر بزرگتری برای اختلاف بین احتمالات شرطی دو کلاس داشته باشیم. این روش‌ها به دلیل دارا بودن عمومیت و پایین بودن پیچیدگی زمانی اهمیت فراوانی دارند. این روش‌ها یکی از شناسایی کننده‌های مبتنی بر مدل[۲] هستند که در آن‌ها تعداد کلاس‌های مساله از قبل مشخص است. در روش‌های مبتنی بر مدل، پارامترهای مدل تخمین زده می‌شود و سپس درصد درست‌نمایی سیستم بر اساس پارامترها اندازه‌گیری می‌شود. نمونه‌ای از روش‌های مبتنی بر فاصله، معیار فاصله اقلیدسی می‌باشد.

در ریاضیات، فاصله اقلیدسی فاصله معمولی دو نقطه است که توسط فیثاغورش بدست می‌آید. فاصله دو نقطه p و q اندازه پاره خطی است که آن‌ها را به هم متصل می‌کند. در مختصات دکارتی اگر p =

[۱] Euclidean Distance
[۲] Model-based Classification

$(p_1, p_2, \ldots, p_n)$ و $q = (q_1, q_2, \ldots, q_n)$ دو نقطه در فضای اقلیدسی n بعدی باشند، آنگاه فاصله بین آن‌ها توسط رابطه (۳-۱) تعریف می‌شود:

$$D(p,q) = \sqrt{(\mathbf{p_1} - \mathbf{q_1})^2 + (\mathbf{p_2} - \mathbf{q_2})^2 + \cdots + (\mathbf{p_n} - \mathbf{q_n})^2} = \sqrt{\sum_{\mathbf{i=1}}^{\mathbf{n}} (\mathbf{p_i} - \mathbf{q_i})^2} \quad (۳-۱)$$

در این پژوهش، بردار ویژگی هر جمله ورودی محاسبه می‌شود. سپس فاصله این بردار تا بردارهای هر دسته، با استفاده از فاصله اقلیدسی محاسبه می‌شود. جمله ورودی مربوط به دسته‌ای است که کمترین فاصله را تا آن دارد.

۳-۳-۳- روش مبتنی بر فاصله همینگ[۱]

فاصله همینگ یکی دیگر از روش‌های مبتنی بر فاصله می‌باشد. در تئوری اطلاعات، فاصله همینگ برای دو رشته با طول مساوی، برابر تعداد مکان‌هایی است که سمبول‌های متناظر متفاوت هستند. به عبارت دیگر، کمترین تعداد جایگزینی‌هایی است که یک رشته به یک رشته دیگر تغییر پیدا کند، یا تعداد خطاهایی که یک رشته به رشته دیگر تبدیل گردد. اگر $p = (p_1, p_2, \ldots, p_n)$ و $q = (q_1, q_2, \ldots, q_n)$ دو نقطه در فضای n بعدی باشند، آنگاه فاصله همینگ بین آن‌ها توسط رابطه (۳-۲) تعریف می‌شود:

$$D(p,q) = |p_1 - q_1| + |p_2 - q_2| + \ldots + |p_n - q_n| = \sum_{\mathbf{i=1}}^{\mathbf{n}} |\mathbf{p_i} - \mathbf{q_i}| \quad (۳-۲)$$

در این پژوهش، بردار ویژگی هر جمله ورودی محاسبه می‌شود. سپس فاصله این بردار تا بردارهای هر دسته، با استفاده از فاصله همینگ محاسبه می‌شود. جمله ورودی مربوط به دسته‌ای است که کمترین فاصله را تا آن دارد.

[۱] Hamming Distance

۳-۳-۴- روش مبتنی بر فاصله مینکوفسکی[۱]

فاصله مینکوفسکی یک فاصله معروف و حالت کلی از روش منهتن و اقلیدسی است. اگر $p = (p_1, p_2, \ldots, p_n)$ و $q = (q_1, q_2, \ldots, q_n)$ دو نقطه در فضای n بعدی باشند، آنگاه فاصله مینکوفسکی بین آن‌ها توسط رابطه (۳-۳) تعریف می‌شود:

$$D(p,q) = \sqrt{(\mathbf{p_1} - \mathbf{q_1})^{\mathbf{p}} + (\mathbf{p_2} - \mathbf{q_2})^{\mathbf{p}} + \cdots + (\mathbf{p_n} - \mathbf{q_n})^{\mathbf{p}}} = \sqrt[\mathbf{p}]{\sum_{\mathbf{i=1}}^{\mathbf{n}} (\mathbf{p_i} - \mathbf{q_i})^{\mathbf{p}}} \quad (۳-۳)$$

اگر p=1 باشد فاصله منهتن و اگر p=2 باشد فاصله اقلیدسی به دست می‌آید.

در این پژوهش، بردار ویژگی هر جمله ورودی محاسبه می‌شود. سپس فاصله این بردار تا بردارهای هر دسته، با استفاده از فاصله مینکوفسکی محاسبه می‌شود. جمله ورودی مربوط به دسته‌ای است که کمترین فاصله را تا آن دارد.

۳-۳-۵- روش بیزین ساده[۲]

روش بیزین یکی دیگر از روش‌های مبتنی بر مدل می‌باشد. کلاس‌بندی بیزین ساده بدلیل اینکه از نظر هزینه محاسباتی به پردازنده و حافظه کمی نیاز دارد، بسیار کارآمد است. همچنین این روش نیاز به مقدار کمی از داده‌ها برای آموزش دارد و زمان آموزش در این روش نسبت به سایر روش‌ها کمتر است.

در یک مساله با دو کلاس، روش بیزین با در اختیار داشتن بردار ویژگی x آن را به یکی از دو دسته k_1 یا k_2 منتسب می‌کند. احتمال دسته‌های متنی در روش بیزین، با استفاده از احتمال‌های اتصال ویژگی و کلاس‌ها به دست می‌آید. احتمال تعلق ویژگی x به کلاس k_i توسط رابطه (۳-۴) محاسبه می‌شود.

$$P(k_i|x) = \frac{\mathbf{p(k_i)p(x|k_i)}}{\mathbf{p(x)}} \qquad i = 1,2 \qquad (۴-۳)$$

[۲] Minkowski Distance
[۱] Naïve Bayes

روش بیزین، احتمال‌های پیشین $p(x|k_i)$ و $p(k_i)$ را از داده‌های آموزشی برآورد کرده و احتمال پسین $p(k_i|x)$ را بر اساس احتمالات پیشین تولید می‌کند. در حالاتی که $p(x|k_i)$ها دارای توزیع گوسی[1] با میانگین‌های M_i و ماتریس‌های کواریانس Σ_i باشند، $h_{12}(x)$ با استفاده از رابطه (۳-۵) محاسبه می‌شود.

$$h_{12}(x) = \frac{1}{2}(x - M_1)^T \boldsymbol{\Sigma_1^{-1}} (x - M_1) - \frac{1}{2}(x - M_2)^T \boldsymbol{\Sigma_2^{-1}} (x - M_2) + \ln \frac{|\boldsymbol{\Sigma_1}|}{|\boldsymbol{\Sigma_2}|} \qquad (۵-۳)$$

با فرض اینکه $\Sigma_1 = \Sigma_2 = I$ باشد، دسته‌بندی نمونه‌ها به وسیله رابطه (۳-۶) محاسبه می‌شود.

$$\begin{cases} |\mathbf{x} - \mathbf{M_1}|^2 - |\mathbf{x} - \mathbf{M_2}|^2 < 2 \ln \frac{\mathbf{p(k_2)}}{\mathbf{p(k_1)}} & \mathbf{x} \in \mathbf{k_1} \\ |\mathbf{x} - \mathbf{M_1}|^2 - |\mathbf{x} - \mathbf{M_2}|^2 > 2 \ln \frac{\mathbf{p(k_2)}}{\mathbf{p(k_1)}} & \mathbf{x} \in \mathbf{k_2} \end{cases} \qquad (۶-۳)$$

بر این اساس، نمونه‌هایی که به M_1 نزدیک هستند جزء دسته اول و نمونه‌هایی که به M_2 نزدیک هستند جزء دسته دوم در نظر گرفته می‌شوند.

۳-۳-۶- روش k نزدیک‌ترین همسایه[2]

روش k نزدیک‌ترین همسایه یا KNN یکی از شناسایی کننده‌های بدون مدل است. در شناسایی کننده‌های بدون مدل تشخیص مبتنی بر هر یک از نمونه‌های آموزشی و بدون در نظر گرفتن یک مدل خاص برای هر کلاس انجام می‌شود.

روش k نزدیک‌ترین همسایه، یک گروه شامل k نمونه از مجموعه نمونه‌های آموزشی که نزدیک‌ترین نمونه‌ها به نمونه آزمایشی باشند را انتخاب کرده و بر اساس برتری رده یا برچسب مربوط به آن‌ها، در مورد دسته نمونه آزمایشی مزبور تصمیم‌گیری می‌نماید. به عبارت ساده‌تر این روش رده‌ای را انتخاب می‌کند که در همسایگی انتخاب شده، بیشترین تعداد نمونه، منتسب به آن دسته باشند. بنابراین رده‌ای که از همه رده‌ها بیشتر در بین k نزدیک‌ترین همسایه مشاهده شود، به عنوان رده نمونه جدید در نظر گرفته می‌شود.

استفاده از الگوریتم KNN نیازمند تعیین سه موضوع می‌باشد:

---

[2] Gaussian Distribution

[1] K-Nearest Neighbour(KNN)

۱. در اختیار داشتن یک مجموعه نمونه.

۲. در اختیار داشتن یک معیار محاسبه شباهت. این معیار می‌توان شامل فاصله اقلیدسی، فاصله همینگ، فاصله مینکوفسکی و ... باشد.

۳. مقدار k باید مشخص شود تا بتوان بر اساس آن عمل نمود. برای مسائل دسته‌بندی دودویی، معمولا در نظر گرفتن مقادیر فرد برای k بهتر است زیرا امکان پیروز شدن یکی از دو دسته را افزایش می‌دهد. برای مسائل رده‌بندی چند رده‌ای، باید عدد k را بزرگ‌تر از تعداد رده‌ها و نیز متفاوت با عدد تعداد رده‌ها از نظر زوج یا فرد بودن در نظر گرفت. یعنی اگر تعداد رده‌ها زوج باشد، باید k نهایی را فرد در نظر گرفت و بالعکس.

در دسته‌بندی‌های KNN برای دسته‌بندی یک نمونه، ابتدا فاصله نمونه جدید از همه نمونه‌های آموزشی محاسبه شده و k نزدیک‌ترین نمونه مشخص می‌شوند. سپس از بین k نمونه انتخاب شده رای‌گیری انجام شده و دسته‌ای که بیشترین تعداد دفعات دیده شدن را در بین نمونه‌ها داراست، به عنوان دسته نمونه جدید در نظر گرفته می‌شود.

انتخاب مقدار k در این روش دسته‌بندی بسیار مهم و کلیدی است. اگر مقدار k خیلی کوچک انتخاب شود، الگوریتم به نویز حساس می‌شود. در واقع نویزها، نزدیک آن نمونه ممکن است ایجاد اشتباه کنند. اگر مقدار k خیلی بزرگ انتخاب شود، ممکن است در میان نزدیک‌ترین همسایه‌ها، نمونه‌هایی از دسته‌های دیگر نیز قرار بگیرند.

۳-۳-۷- روش TF-IDF

فراوانی وزنی TF-IDF مخفف فراوانی[۱]-عکس فراوانی سند[۲] است. در این شیوه به لغات یک وزن بر اساس فراوانی آن در سند داده می‌شود. در واقع این سیستم وزن‌دهی نشان می‌دهد چقدر یک کلمه برای یک سند مهم است. این مساله کاربردهای بسیاری در بازیابی اطلاعات دارد. وزن کلمه با افزایش

[۱] Term Frequency(TF)
[۲] Inverse Document Frequency(IDF)

تعداد تکرار آن در متن افزایش می‌یابد، اما توسط تعداد کلمات در متن کنترل می‌شود، چرا که می‌دانیم در صورت زیاد بودن طول متن، بعضی از کلمات به طور طبیعی بیشتر از دیگران تکرار خواهند شد، اگرچه چندان اهمیتی در معنی نداشته باشند .

اگر فرض کنیم تعداد دفعاتی که کلمه t در متن d اتفاق افتاده با TF(t.d) نشان داده شود و در ساده‌ترین حالت تعداد تکرار اولیه t با f(t,d) نشان داده شود بنابراین TF(t,d) = f(t,d) است.

IDF معیاری است برای میزان کلماتی که در کلیه متون بسیار متداول هستند و معمولاً تکرار می‌شوند. طریقه بدست آورن این معیار بدین صورت است که از لگاریتم، تقسیم تعداد کل متون بر تعداد متون شامل کلمه متداول مشابه رابطه(۳-۷) بدست می‌آید. برای مثال فرض کنیم در کل پایگاه داده ۱۰۰۰متن وجود داشته باشد. اگر در هر ۱۰۰۰ تای آن یک کلمه خاص (مثلاً است) وجود داشته باشد، حاصل لگاریتم ۱۰۰۰ تقسیم بر ۱۰۰۰ برابر صفر می‌شود. یعنی حتماً این کلمه جز کلمات متداول بوده و باید ضریب صفر بگیرد. ولی اگر تکرار در ۵۰۰ متن اتفاق افتاده باشد، حاصل لگاریتم ۱۰۰۰ تقسیم بر ۵۰۰ برابر یک می‌شود و آن کلمه ضریب یک می‌گیرد. هر چقدر متونی که کلمه در آن تکرار شده باشد بیشتر باشد، وزن IDF کمتر می‌شود و چون ممکن است اصلاً تکرار نشده باشد و مخرج صفر شود در مخرج ۱ اضافه می‌شود.

$$IDF(t,D) = \log\left(\frac{D}{1+\{d \in D : t \in d\}}\right) \qquad (۳-۷)$$

با توجه به این موارد، میزان TF-IDF توسط رابطه (۳-۸) به دست می‌آید.

$$TF\text{-}IDF(t,d,D) = TF(t,d) * IDF(t,D) \qquad (۳-۸)$$

۳-۳-۸- روش خوشه بندی k میانگین[۱]

روش k میانگین در عین سادگی یک روش بسیار کاربردی و پایه چند روش دیگر مثل خوشه بندی فازی[۲] است. K میانگین، خوشه بندی با هدف تجزیه n نمونه به k خوشه است که در آن هر یک از

[۱] K-Means
[۲] Fuzzy Clustering

نمونه‌ها متعلق به خوشه‌ای با نزدیک‌ترین میانگین آن است. این میانگین به عنوان پیش‌نمونه استفاده می‌شود.

تابع k میانگین تلاش می‌کند تا تابع هدف خود را مطابق رابطه (۳-۹) به حداقل برساند.

$$\sum_{j=1}^{k} \sum_{i=1}^{n} \left| x_i^{(j)} - k_i \right|^2 \qquad (۹-۳)$$

$|x_i^{(j)} - k_i|^2$ یک معیار فاصله بین نمونه‌های $x_i^{(j)}$ و مرکز خوشه k_i است که j مشخص‌کننده فاصله n نمونه از مرکز خوشه مربوطه است.

روش کار این الگوریتم به این صورت است که ابتدا مراکز اولیه خوشه‌ها به صورت تصادفی انتخاب می‌شود. سپس هر نمونه به خوشه‌ای که نزدیک‌ترین فاصله را با آن دارد اختصاص داده می‌شود. این فاصله می‌تواند به وسیله روش اقلیدسی، همینگ، مینکوفسکی و ... محاسبه شود. زمانیکه تمام نمونه-ها به خوشه‌های مربوطه اختصاص داده شدند، موقعیت k مرکز دوباره محاسبه می‌شود. مراکز جدید به صورت میانگین نمونه‌های موجود در آن خوشه محاسبه می‌شود. دوباره فاصله نمونه‌ها تا مرکز خوشه‌ها محاسبه شده و عمل خوشه بندی انجام می‌شود. این عمل تا زمانیکه مراکز خوشه‌ها بدون تغییر باقی بمانند، ادامه پیدا می‌کند. به منظور جلوگیری از اجرای زیاد خوشه‌بندی، می‌توان این عمل را به تعداد مشخصی تکرار کرد.

۳-۳-۹- روش ماشین بردار پشتیبان[۱]

ماشین بردار پشتیبان یا SVM یکی از روش‌های بسیار رایج در حوزه دسته‌بندی داده‌ها است. این روش هیچ پیش‌فرضی براساس نمونه‌های آموزشی ایجاد نمی‌کند و مستقیما دسته مربوط به نمونه را تخمین می‌زند.

بردارهای پشتیبان مجموعه‌ای از نقاط در فضای n بعدی داده‌ها هستند که مرز دسته‌ها را مشخص می‌کنند و مرزبندی و دسته‌بندی داده‌ها براساس آن‌ها انجام می‌شود و با جابجایی یکی از آن‌ها،

[۱] Support Vector Machine(SVM)

خروجی دسته‌بندی ممکن است تغییر کند. در فضای دوبعدی، بردارهای پشتیبان، یک خط، در فضای سه بعدی یک صفحه و در فضای n بعدی یک ابر صفحه را شکل خواهند داد.

ماشین بردار پشتیبان، یک دسته بند یا مرزی است که با معیار قرار دادن بردارهای پشتیبان، بهترین دسته‌بندی و تفکیک بین داده‌ها را برای ما مشخص می‌کند. در SVM فقط داده‌های قرار گرفته در بردارهای پشتیبان مبنای یادگیری ماشین و ساخت مدل قرار می‌گیرند و این الگوریتم به سایر نقاط داده حساس نیست و هدف آن هم یافتن بهترین مرز در بین داده‌هاست به گونه‌ای که بیشترین فاصله ممکن را از تمام دسته‌ها (بردارهای پشتیبان آن‌ها) داشته باشد.

اگر داده‌ها به صورت خطی قابل تفکیک باشند، الگوریتم SVM می‌تواند بهترین ماشین را برای تفکیک داده‌ها و تعیین دسته یک نمونه ایجاد کند. اگر داده‌ها به صورت خطی توزیع نشده باشند، داده‌ها به کمک یک تابع ریاضی به یک فضای دیگر نگاشت می‌شود که در آن فضا، داده‌ها تفکیک‌پذیر هستند و SVM آن‌ها را براحتی تشخیص می‌دهد. تعیین درست این تابع نگاشت در عملکرد ماشین بردار پشتیبان موثر است.

ماشین بردار پشتیبان دارای ضعف‌هایی است که در زیر آورده شده است:

- این نوع الگوریتم‌ها، محدودیت‌های ذاتی دارند. یکی از این محدودیت‌ها این است که هنوز مشخص نشده که به ازای یک تابع نگاشت، پارامترها را چگونه باید تعیین کرد.
- ماشین‌های مبتنی بر بردار پشتیبان به محاسبات پیچیده و زمان‌بر نیاز دارند و به دلیل پیچیدگی محاسباتی، حافظه زیادی نیز مصرف می‌کنند.
- داده‌های گسسته و غیرعددی با این روش سازگار نیستند و باید تبدیل شوند.

با وجود این ضعف‌ها، این روش دارای یک شالوده نظری منسجم بوده و جواب‌های تولید شده توسط آن‌ها، سراسری و یکتا می‌باشد. امروزه ماشین‌های بردار پشتیبان، به متداول‌ترین تکنیک‌های پیش-بینی در داده کاوی تبدیل شده‌اند.

۳-۳-۱۰- روش مبتنی بر معیار شباهت کسینوسی[۱]

برای محاسبه شباهت کسینوسی دو نمونه، از ضرب داخلی بردارهای آن‌ها توسط رابطه (۳-۱۰) استفاده می‌شود.

$$\text{Cos-sim}(d_1,d_2) = \sum \frac{W(w,d_1)W(w,d_2)}{\|d_1\| \|d_2\|} \qquad (۱۰-۳)$$

که در آن d_1 و d_2 دو نمونه‌ای هستند که شباهت‌شان اندازه‌گیری می‌شود. $W(w,d_1)$ وزن کلمه w در نمونه d_1، و $W(w,d_2)$ وزن کلمه w در نمونه d_2 است که با تابع وزن‌دهی TF-IDF محاسبه شده است. $\|d_1\|$ اندازه سند ۱ و $\|d_1\|$ اندازه سند ۲ است که توسط رابطه (۳-۱۱) محاسبه می‌شود.

$$\|d\| = \sqrt{\sum_w W(w,d)^2} \qquad (۱۱-۳)$$

در این پژوهش، ابتدا شباهت کسینوسی بردار جمله ورودی از هر یک از کلاس‌ها محاسبه می‌شود. جمله متعلق به کلاسی خواهد بود که بیشترین شباهت کسینوسی را با آن داشته باشد.

۳-۴-معیار محاسبه درصد درست‌نمایی

ارزیابی عملکرد روش‌ها، با استفاده از معیارهای متفاوتی بر مبنای دیدگاه حساسیت و تشخیص صورت می‌گیرد. حساسیت و تشخیص در آمار دو شاخص برای ارزیابی نتیجه یک دسته‌بندی دودویی (دوحالته) هستند. زمانی که بتوان داده‌ها را به دو گروه مثبت و منفی تقسیم کرد، دقت نتایج یک آزمایش که اطلاعات را به این دو دسته تقسیم می‌کند با استفاده از شاخص‌های حساسیت و ویژگی، قابل اندازه‌گیری و توصیف است. معیارهای مورد استفاده در این دیدگاه به شرح زیر می‌باشند:

- ✓ TP: مثبت صحیح[۲](تعداد نمونه‌هایی که به درستی جزء دسته به حساب آمده‌اند.)
- ✓ TN: منفی صحیح[۳](تعداد نمونه‌هایی که به درستی جزء دسته به حساب نیامده‌اند.)
- ✓ FP: مثبت کاذب[۴](تعداد نمونه‌هایی که جزء دسته نبوده‌اند اما به اشتباه جزء دسته به حساب آمده‌اند.)

---

۱ Cosine Likeness
1 True Positive(TP)
۲ True Negative(TN)
۳ False Positive(FP)

- ✓ FN: منفی کاذب[۱](تعداد نمونه‌هایی که جزء دسته بوده‌اند اما به اشتباه جزء دسته به حساب نیامده-اند.)

دقت[۲] به نسبت مقادیر موارد صحیح طبقه‌بندی شده توسط الگوریتم از یک کلاس مشخص، به کل تعداد مواردی است که الگوریتم چه به صورت صحیح و چه به صورت غلط در آن کلاس طبقه‌بندی کرده است و به وسیله رابطه(۳-۱۲) محاسبه می‌شود:

$$\text{Precision} = \frac{\text{TP}}{\text{T}} \qquad (۱۲-۳)$$

بازخوانی[۳] به نسبت مقادیر موارد صحیح طبقه‌بندی شده توسط الگوریتم از یک کلاس مشخص، به تعداد موارد حاضر در کلاس مذکور است و به وسیله رابطه(۳-۱۳) محاسبه می‌شود:

$$\text{Recall} = \frac{\text{TP}}{\text{T}} \qquad (۱۳-۳)$$

با توجه به محاسبات انجام گرفته برای معیارهای Precision و Recall، می‌توان مقدار کمیت وزن‌دار F-Measure را محاسبه نمود. F-Measure پارامتر مناسبی برای ارزیابی کیفیت کلاس‌بندی می‌باشد و همچنین توصیف‌کننده میانگین وزن‌دار مابین دو کمیت Precision و Recall می‌باشد. برای یک الگوریتم کلاس‌بندی کننده در شرایط ایده‌آل، مقدار این کمیت برابر با یک می‌باشد و در بدترین وضعیت برابر با صفر می‌باشد. این پارامتر با توجه به رابطه(۳-۱۴) زیر محاسبه می‌شود:

$$\text{F-measure} = 2 * \frac{\text{Precision} * \text{Recall}}{\text{Prec}} \qquad (۱۴-۳)$$

MCC[۴] پارامتر دیگری است که برای ارزیابی کارایی الگوریتم‌ها از آن استفاده می‌شود. این پارامتر بیانگر کیفیت کلاس‌بندی برای یک مجموعه باینری می‌باشد. MCC بیانگر وابستگی مابین مقادیر مشاهده شده از کلاس باینری و مقادیر پیش‌بینی شده از آن می‌باشد. مقادیر مورد انتظار برای این کمیت در بازه ۱- و ۱ متغیر می‌باشد. مقدار ۱+، نشان دهنده پیش‌بینی دقیق و بدون خطای الگوریتم یادگیر از کلاس باینری می‌باشد. مقدار ۰، نشان دهنده پیش‌بینی تصادفی الگوریتم یادگیر از کلاس

---

[٤] False Negative(FN)
[٥] Precision
[٦] Recall
[١] Matthews Correlation Coefficient(MCC)

باینری می‌باشد. مقدار ۱-، نشان دهنده عدم تطابق کامل مابین موارد پیش‌بینی شده از کلاس باینری و موارد مشاهده شده از آن می‌باشد. این پارامتر براساس رابطه (۳-۱۵) محاسبه می‌شود:

$$MCC = \frac{TP*T\square \quad *FN}{\sqrt{(T\square \quad)(TP+F\square)(T\square \quad)(TN+F\square)}} \qquad (۱۵-۳)$$

در این پژوهش برای محاسبه کارایی و درست‌نمایی روش‌های پیشنهادی از معیار بازخوانی استفاده شده است.

فصل چهارم

جزئیات پیکربندی و

ارزیابی روش پیشنهادی

۴-۱- مقدمه

در این فصل به جزئیات پیکربندی روش‌های معرفی شده در فصل ۳ پرداخته شده و با پیاده‌سازی هر یک از روش‌ها، درصد درست‌نمایی آن‌ها محاسبه می‌شود. همانطور که در فصل قبلی بیان شد، روش-های استفاده شده در این پژوهش آماری بوده و کیفیت کارایی تشخیص براساس تشابه الگوهای تمرینی و داده‌های واقعی که سیستم در طول عملیات مواجه خواهد شد، تعیین می‌شود.

تمامی روش‌ها با استفاده از زبان جاوا و نرم‌افزار وکا[۱] و روی پایگاه داده حاوی ۴۶۸۸ نمونه از مکالمات روزمره انگلیسی پیاده‌سازی شده و نرخ تشخیص روش به دست می‌آید. این نمونه‌ها به دو دسته کنترل کننده و کنترل‌شونده دسته‌بندی شده‌اند. این پایگاه داده به دو بخش آموزش و تست تقسیم شده است. ۸۰ درصد این مجموعه به عنوان بخش آموزش و ۲۰ درصد بقیه به عنوان مجموعه تست در نظر گرفته شده است.

در پایان به بررسی و مقایسه روش‌ها پرداخته می‌شود و بهترین روش معرفی خواهد شد.

۴-۲- روش مبتنی بر فاصله اقلیدسی

برای پیاده‌سازی روش مبتنی بر فاصله اقلیدسی، ابتدا بردار ویژگی کلاس‌های کنترل کننده و کنترل-شونده محاسبه شده است. این بردارها دارای ابعاد 2*n هستند که n بیانگر تعداد ویژگی‌ها می‌باشد. ویژگی‌های در نظر گرفته شده برای این مساله، کلمات استفاده شده در مکالمات می‌باشد. سطر اول این بردارها حاوی کلمات و سطر دوم حاوی میانگین تعداد کلمات در این دسته از مکالمات است.

با دریافت هر جمله، بردار ویژگی مربوط به جمله محاسبه شده و فاصله اقلیدسی آن تا کلاس‌های کنترل کننده و کنترل‌شونده محاسبه می‌شود. جمله مربوط به کلاسی است که کمترین فاصله را تا آن داشته باشد.

[۱] Weka

این روش با استفاده از زبان جاوا پیاده‌سازی شده است. پس از اعمال مجموعه تست و محاسبه دقت، این روش دارای میزان درست‌نمایی ۷۹.۰۹ درصد بود.

۴-۳- روش مبتنی بر فاصله همینگ

در این روش نیز ابتدا، بردار ویژگی مربوط به کلاس‌های کنترل کننده و کنترل‌شونده محاسبه می-شود. بردار ویژگی حاوی میانگین تعداد کلمات در هر دسته از مکالمات است.

با دریافت هر جمله، بردار ویژگی مربوط به جمله محاسبه شده و فاصله همینگ آن تا کلاس‌های کنترل کننده و کنترل‌شونده محاسبه می‌شود. جمله مربوط به کلاسی است که کمترین فاصله را تا آن داشته باشد.

این روش با استفاده از زبان جاوا پیاده‌سازی شده است. پس از اعمال مجموعه تست و محاسبه دقت، این روش دارای میزان درست‌نمایی ۷۳.۶۶ درصد بود.

۴-۴- روش مبتنی بر فاصله مینکوفسکی

در این روش نیز ابتدا، بردار ویژگی مربوط به کلاس‌های کنترل کننده و کنترل‌شونده محاسبه می-شود. بردار ویژگی حاوی میانگین تعداد کلمات در هر دسته از مکالمات است.

با دریافت هر جمله، بردار ویژگی مربوط به جمله محاسبه شده و فاصله مینکوفسکی آن تا کلاس‌های کنترل کننده و کنترل‌شونده محاسبه می‌شود. جمله مربوط به کلاسی است که کمترین فاصله را تا آن داشته باشد.

پارامتر p در روش مینکوفسکی، یک عدد طبیعی است و می‌تواند مقادیر متفاوتی را دریافت کند. در این پیاده‌سازی، این پارامتر برابر ۴ در نظر گرفته شده است.

این روش با استفاده از زبان جاوا پیاده‌سازی شده است. پس از اعمال مجموعه تست و محاسبه دقت، این روش دارای میزان درست‌نمایی ۷۴.۷۹ درصد بود.

۴-۵- روش بیزین ساده

کلاس‌بندی بیزین ساده بدلیل اینکه از نظر هزینه محاسباتی به پردازنده و حافظه کمی نیاز دارد، بسیار کارآمد است. همچنین این روش نیاز به مقدار کمی از داده‌ها برای آموزش دارد و زمان آموزش نسبت به سایر روش‌ها کمتر است.

برای پیاده‌سازی این روش، مجموعه آموزش و تست در دو فایل جداگانه با پسوند csv تهیه شده است. ویژگی‌های در نظر گرفته شده برای دسته‌بندی، میانگین تعداد کلمات به کار رفته در جملات مکالمه می‌باشد. این روش با استفاده از نرم‌افزار وکا پیاده‌سازی شده است و میزان درست‌نمایی ۸۹.۵۴ درصد بود.

۴-۶- روش k نزدیک‌ترین همسایه

همانطور که گفته شد، روش k نزدیک‌ترین همسایه، یک گروه شامل k نمونه از مجموعه نمونه‌های آموزشی که نزدیک‌ترین نمونه‌ها به نمونه آزمایشی باشند را انتخاب کرده و بر اساس برتری رده یا برچسب مربوط به آن‌ها، در مورد دسته نمونه آزمایشی مزبور تصمیم‌گیری می‌نماید.

برای پیاده‌سازی این روش، ابتدا بردار ویژگی هر یک از نمونه‌های تست محاسبه شده است. سپس k نمونه از مجموعه نمونه‌های آموزشی که نزدیک‌ترین نمونه‌ها به نمونه تست هستند انتخاب شد. معیار نزدیکی نمونه‌ها فاصله همینگ در نظر گرفته شده است. از آنجایی که تعداد دسته‌ها در این مساله برابر دو است، مقدار k عددی فرد و برابر ۵ در نظر گرفته شده است. تعیین عدد فرد برای دسته‌های دودویی امکان پیروز شدن یکی از دو دسته را افزایش می‌دهد. سپس از بین k نمونه انتخاب شده رای‌گیری انجام شده و دسته‌ای که بیشترین تعداد دفعات دیده شدن را در بین نمونه‌ها داراست، به عنوان دسته نمونه جدید در نظر گرفته می‌شود.

این روش با استفاده از زبان جاوا پیاده‌سازی شده است. پس از اعمال مجموعه تست و محاسبه دقت، این روش دارای میزان درست‌نمایی ۷۶.۰۲ درصد بود.

۴-۷- روش TF-IDF

روش TF-IDF کاربردهای بسیاری در بازیابی اطلاعات دارد. ایده این روش این است که وزن کلمه با افزایش تعداد تکرار آن در متن افزایش می‌یابد، اما توسط تعداد کلمات در متن کنترل می‌شود، چرا که می‌دانیم در صورت زیاد بودن طول متن، بعضی از کلمات به طور طبیعی بیشتر از دیگران تکرار خواهند شد، اگرچه چندان اهمیتی در معنی نداشته باشند.

برای پیاده‌سازی این روش ابتدا TF-IDF مربوط به کلاس‌های کنترل کننده و کنترل‌شونده محاسبه شده است. برای محاسبه این مقدار، مقادیر TF و IDF به صورت جداگانه به دست آمده و در نهایت در هم ضرب می‌شوند. TF در این مساله برابر میانگین تعداد تکرار کلمات در مکالمات هر کلاس می‌باشد. IDF معیاری است برای میزان کلماتی که در کلیه متون بسیار متداول هستند و معمولاً تکرار می‌شوند. طریقه بدست آورن این معیار بدین صورت است که از لگاریتم، تقسیم تعداد کل متون بر تعداد متون شامل کلمه متداول مشابه رابطه (۳-۷) بدست می‌آید.

با دریافت هر جمله مجموعه تست، معیار TF-IDF مربوط به جمله محاسبه شده و فاصله آن تا کلاس‌های کنترل کننده و کنترل‌شونده محاسبه می‌شود. جمله مربوط به کلاسی است که کمترین فاصله را تا آن داشته باشد. معیار فاصله در این مساله فاصله اقلیدسی در نظر گرفته شده است.

این روش با استفاده از زبان جاوا پیاده‌سازی شده است. پس از اعمال مجموعه تست و محاسبه دقت، این روش دارای میزان درست‌نمایی ۲۶.۸۴ درصد بود.

۴-۸- روش خوشه بندی k میانگین

روش k میانگین در عین سادگی یک روش بسیار کاربردی و پایه چند روش دیگر مثل خوشه بندی فازی است. K میانگین، خوشه بندی با هدف تجزیه n نمونه به k خوشه است که در آن هر یک از نمونه‌ها متعلق به خوشه‌ای با نزدیک‌ترین میانگین آن است. این میانگین به عنوان پیش‌نمونه استفاده می‌شود. معیار نزدیکی در این مساله فاصله اقلیدسی در نظر گرفته شده است.

برای پیاده‌سازی این روش، مجموعه آموزش و تست در دو فایل جداگانه با پسوند csv تهیه شده است. ویژگی‌های در نظر گرفته شده برای دسته‌بندی، میانگین تعداد کلمات به کار رفته در جملات مکالمه می‌باشد. این روش با استفاده از نرم‌افزار وکا پیاده‌سازی شده است و میزان درستنمایی ۹۰.۵۷ درصد بود.

۴-۹- روش ماشین بردار پشتیبان

ماشین بردار پشتیبان، یک طبقه‌بندی جداکننده[۱] است که هیچ پیش‌فرضی را بر اساس داده‌های آموزشی ایجاد نمی‌کند و به طور مستقیم $P(k|X)$ را تخمین می‌زند.

برای پیاده‌سازی این روش، مجموعه آموزش و تست در دو فایل جداگانه با پسوند csv تهیه شده است. ویژگی‌های در نظر گرفته شده برای دسته‌بندی، میانگین تعداد کلمات به کار رفته در جملات مکالمه می‌باشد. این روش با استفاده از نرم‌افزار وکا و تابع هسته چندجمله‌ای پیاده‌سازی شده است و میزان درستنمایی ۹۱.۴۹ درصد بود.

۴-۱۰- روش مبتنی بر معیار شباهت کسینوسی

برای محاسبه شباهت کسینوسی دو نمونه، از ضرب داخلی بردارهای آن‌ها توسط رابطه (۳-۱۰) استفاده می‌شود. بردارها در این روش توسط تابع وزن‌دهی TF-IDF مقداردهی می‌شوند.

برای پیاده‌سازی این روش ابتدا مقادیر TF-IDF مربوط به کلاس‌های کنترل کننده و کنترل‌شونده محاسبه شده است. سپس شباهت کسینوسی هر کدام از دسته‌ها با استفاده از رابطه (۳-۱۰) محاسبه شده است.

با دریافت هر جمله مجموعه تست، فاصله کسینوسی مربوط به جمله محاسبه شده و فاصله آن تا کلاس‌های کنترل کننده و کنترل‌شونده محاسبه می‌شود. جمله مربوط به کلاسی است که بیشترین شباهت را با آن داشته باشد.

[۱] Discriminative Classifier

این روش با استفاده از زبان جاوا پیاده‌سازی شده است. پس از اعمال مجموعه تست و محاسبه دقت، این روش دارای میزان درست‌نمایی ۷۷.۶۶ درصد بود.

۴-۱۱- مقایسه روش‌های پیشنهادی

جدول (۴-۱) شامل نتایج به دست آمده از روش‌های پیاده شده در این فصل می‌باشد. با مشاهده و مقایسه نتایج به دست آمده، روش ماشین بردار پشتیبان با مقدار ۹۱.۴۹ بهترین درصد درست‌نمایی را در بین روش‌ها داشته است. البته روش‌های خوشه بندی k میانگین و بیزین نیز نتایج نزدیکی به روش ماشین بردار پشتیبان داشته‌اند.

جدول (۴-۱): مقایسه روش‌های پیشنهادی

روش	درصد تشخیص
فاصله اقلیدسی	۷۹.۰۹
فاصله همینگ	۷۳.۶۶
فاصله مینکوفسکی	۷۴.۷۹
بیزین	۸۹.۵۴
K نزدیک‌ترین همسایه	۷۶.۰۲
TF-IDF	۲۶.۸۴
خوشه‌بندی k میانگین	۹۰.۵۷
ماشین بردار پشتیبان	۹۱.۴۹
فاصله کسینوسی	۷۷.۶۶

فصل پنجم

جمع بندی و

پیشنهادهایی برای

کارهای آینده

۵-۱- جمع بندی

در سیستم‌های پرسش و پاسخ، کاربران پرسش‌های خود را به زبان طبیعی مطرح می‌کنند و سیستم، پاسخ ممکن را باز می‌گرداند. با این وجود در سیستم‌های پرسش و پاسخ، گفتگوی پیوسته بین کاربر و سیستم وجود ندارد و کاربران نمی‌توانند نیاز خود را توضیح دهند. از اینرو دسترسی به اطلاعات موردنیاز در اولین تعامل بین کاربر و سیستم بسیار مشکل است. افزودن تعامل به این سیستم‌ها، امکان طرح سوالات مرتبط و ارائه توضیحات لازم از طریق گفتگو بین سیستم و کاربر را فراهم می‌کند و دقت پاسخگویی را افزایش می‌دهد.

سیستم‌های پرسش و پاسخ تعاملی نوعی از سیستم‌های پرسش و پاسخ هستند که از طریق برقراری تعامل دو سویه با کاربر، ابهامات احتمالی موجود در پرسش و پاسخ را رفع نموده و دقت پاسخگویی را افزایش می‌دهند.

اغلب گفته می‌شود که گفتگو یک فرآیند همکاری ذاتی است و این مساله در پدیده‌های خاصی ظاهر می‌شود. مکالمات دارای ویژگی‌های متعددی هستند که یکی از این ویژگی‌ها کنترل مکالمه می‌باشد. با استفاده از این ویژگی کنترل و هدایت مسیر مکالمه از یک شخص به شخص دیگر منتقل می‌شود و انحصارا در اختیار یک شخص باقی نمی‌ماند. مکالماتی با این ویژگی را مکالمات دارای ابتکار عمل ترکیبی می‌نامند که در هر لحظه ابتکار عمل به دست یکی از شرکت‌کنندگان در مکالمه است. بنابراین مکالمات انعطاف‌پذیرتر و حل مسائل راحت‌تر خواهد شد.

هدف سیستم‌های پرسش و پاسخ تعاملی، پاسخگویی به سوالات کاربر با دقت بالا می‌باشد. با وجود استدلال قوی و یافتن پاسخ‌های دقیق، سیستم بدون داشتن کنترل کننده‌ای برای مکالمه، خوب عمل نکرده و احساس رضایت در کاربر ایجاد نخواهد کرد. کنترل در مکالمه، منجر به ایجاد احساس در کاربر می‌شود، به طوری که کاربر حس کند با یک انسان در حال مکالمه می‌باشد.

در این پژوهش، به بررسی روش‌های تشخیص کنترل کننده در مکالمات پرداخته شد. روش‌هایی که تا کنون استفاده شده‌اند بیشتر معنایی بوده و از تجزیه و تحلیل زبانی و گرامری استفاده کرده‌اند. در این پایان نامه از روش‌های آماری به منظور تشخیص کنترل کننده در مکالمات استفاده شد.

روش‌های آماری فاقد تجزیه و تحلیل معنایی و گرامری هستند و مستقل از زبان می‌باشند. شناسایی آماری، مبتنی بر مشاهدات و نمونه‌برداری بوده و آموزش با استفاده از داده‌های آماری و پرتکرار انجام می‌شود.

برای تجزیه و تحلیل و تشخیص کنترل کننده در مکالمات، ویژگی‌هایی از آن‌ها استخراج می‌شود. با استفاده از این ویژگی‌ها به بررسی هر یک از جملات مکالمه پرداخته و شخص کنترل کننده در آنها تعیین می‌شود. در هر یک از جملات شخص کنترل‌شونده و کنترل کننده، کلمات تکرار می‌شوند. از این رو از روی توزیع کلمات می‌توان به تعیین کنترل کننده پرداخت.

در این پژوهش از روش‌های آماری مبتنی بر فاصله اقلیدسی، مبتنی بر فاصله همینگ، مبتنی بر فاصله مینکوفسکی با پارامتر p=4، بیزین، k نزدیک‌ترین همسایه، k میانگین، ماشین بردار پشتیبان، شباهت کسینوسی و TF-IDF استفاده شد.

این روش‌ها روی پایگاه داده‌ای با ۴۶۸۸ نمونه پیاده‌سازی شده است. پایگاه داده شامل مکالمات زبان انگلیسی است. این مکالمات شامل محاوره بین دو نفر درباره مسائل روزمره مانند اجاره خانه، دریافت وام، قرار ملاقات، گزارش جرم و ... می‌باشد. پایگاه داده به دو بخش مجموعه داده‌های آموزشی و تست تقسیم‌بندی شد. ۸۰ درصد پایگاه داده به عنوان داده‌های آموزشی در نظر گرفته شد و روش‌های پیشنهادی با این مجموعه آموزش داده شد. سپس با ۲۰ درصد پایگاه داده دقت درست‌نمایی روش‌ها محاسبه شد.

پیاده‌سازی با استفاده از زبان جاوا و نرم‌افزار وکا پیاده‌سازی شد و روش ماشین بردار پشتیبان با ۹۱.۴۹ درصد، بهترین نرخ تشخیص را داشت. البته روش‌های خوشه بندی k میانگین با ۹۰.۵۷ درصد و بیزین با ۸۹.۵۴ درصد نیز نتایج نزدیکی به روش ماشین بردار پشتیبان داشته‌اند.

معیارهای مبتنی بر فاصله، نتایج نزدیکی به هم داشتند و نرخ تشخیص آن‌ها تقریبا یکی است. با این وجود معیار مبتنی بر فاصله اقلیدسی با ۷۹.۰۹ درصد نتیجه بهتری داشته است. روش بیزین ساده نیز با ۸۹.۵۴ درصد نتیجه خوبی داشته است. این روش به دلیل هزینه محاسباتی پایین و مجموعه داده آموزشی کم بسیار مناسب است.

۵-۲- پیشنهادهایی برای کارهای آینده

از آنجایی که روش‌های آماری مستقل از دامنه هستند، از محبوبیت بالایی برخوردارند. در این روش‌ها پس از تهیه پایگاه داده موردنظر، سیستم با آن ارزیابی می‌شود. در این پژوهش ۹ مورد از روش‌های آماری مورد بررسی قرار گرفت. به عنوان کارهای آینده، می‌توان روش‌های دیگری را نیز بررسی کرد.

روش‌های آماری مستقل از دامنه هستند و مساله را بدون در نظر گرفتن تجزیه و تحلیل معنایی و گرامری بررسی می‌کنند. تجزیه و تحلیل گرامری دقت تشخیص مساله را بالا می‌برد. برای بهره‌مندی از این مزیت می‌توان از ترکیب روش‌های آماری و گرامری استفاده کرد و دقت تشخیص مساله را بالا برد.

مجوعه ویژگی‌های در نظر گرفته شده برای روش‌ها، تاثیر بسزایی در قدرت تشخیص آن‌ها دارد. مجموعه ویژگی‌های در نظر گرفته شده در این مساله، کلمات استفاده شده در مکالمات و میانگین تعداد هر یک از این کلمات است. به عنوان پیشنهادی برای کارهای آینده می‌توان مجموعه‌های دیگری را نیز بررسی کرد.

منابع و ماخذ:

[۱] حسینی محمد مهدی و زاهدی مرتضی، ۱۳۹۵، "بهبود پاسخ ارائه شده در سیستم‌های پرسش و پاسخ تعاملی به کمک شبکه عصبی"، هشتمین کنفرانس بین المللی فناوری اطلاعات و دانش، انجمن فناوری اطلاعات و ارتباطات ایران، دانشگاه بوعلی سینا.

[۲] D. Novick and S. Sutton, 1997, “What is Mixed-Initiative Interaction?”, *Proceedings of the AAAI Spring Symposium on Computational Models for Mixed Initiative Interaction*, pages 114–116.

[۳] R.S. Nickerson, 1976, “On conversational interaction with computers”, *User-Oriented Design of Interactive Graphics Systems,* pages 65-101.

[۴] M. Walker and S. Whittaker, 1990, “Mixed initiative in dialogue: An investigation into discourse segmentation”, *Proceedings of the 28th Meeting of the ACL*, pages 70-78.

[۵] R. Smith, 1994, “Spoken variable-initiative dialogue: An adaptable natural-language interface”, *IEEE Expert*, pages 45-50.

[۶] J. Chu-Carroll and M.K. Brown, 1998, “An evidential model for tracking initiative in collaborative dialogue interactions”, *User Modeling and User- Adapted Interaction,* pages 215-253.

[۷] D.J. Litman and S. Pan, 1999, “Empirically evaluating an adaptable spoken dialogue system”, *In Proceedings of the 7th International Conference on User Modeling*, pages 55-64.

[۸] P.R. Cohen and C.R. Perrault, 1986, “Elements of a plan-based theory of speech acts”, *Readings in Natural Language Processing,* pages 423- 440.

[۹] M. Pollack, 1986, “Inferring domain plans in question answering”, Technical Report 403, SRI International - Artificial Intelligence Center.

[۱۰] P.R. Cohen, 1984, “The pragmatics of referring and the modality of communication”, *Computational Linguistics,* pages 97-146.

[۱۱] B.J. Grosz, 1977, “The representation and use of focus in dialogue understanding”, Technical Report 151, SRI International, 333 Ravenswood Ave, Menlo Park.

[۱۲] K.R. McKeown, 1985, “Discourse strategies for generating natural language text”, Artificial Intelligence, pages 1-42.

[۱۳] R. Cohen, 1987, “Analyzing the structure of argumentative discourse”, *Computational Linguistics,* pages 11-24.

[۱۴] H. Sacks, E. Schegloff and G. Jefferson, 1974, “A simplest systematics for the organization of turn-taking in conversation”, *Language,* pages 325-345.

[۱۵] H.H. Clark and D. Wilkes Gibbs, 1986, “Referring as a collaborative process”, *Cognition,* pages 1-39.

[۱۶] P.R. Cohen, H.J. Levesque, J.H.T. Nunes and S.L. Oviatt, 1990, “Task oriented dialogue as a consequence of joint activity”.

[۱۷] B.J. Grosz and C.L. Sidher, 1990, “Plans for discourse”, *Intentions in Communication,* MIT Press, Cambridge, MA.

[۱۸] D.M. Frohlich and P. Luff, 1989, “Conversational resources for situated action”, *Proceedings of the Annual Meeting of the Computer Human Interaction of the ACM.*

[۱۹] A. Kidd, 1985, “The consultative role of an expert system”, *People and Computers: Designing the Interface,* Cambridge University Press, Cambridge, U.K.

[۲۰] M. Pollack, J. Hirschberg and B. Webber, 1982, “User participation in the reasoning process of expert systems”, *Proceedings of the National Conference on Artificial Intelligence.*

[۲۱] C. Sidner, 1983, “What the speaker means: the recognition of speakers plans in discourse”, *International Journal of Computers and Mathematics,* pages 71-82.

[۲۲] P.R. Cohen, C. Raymond Perrault and J.F. Allen, 1982, “Beyond question answering”, *Strategies for Natural Language Processing,* pages 245-274.

[۲۳] O. Kolomiyets, M.F. Moens, 2011, "A survey on question answering technology from an information retrieval perspective", *Information science,* pages *5412-5434*.

[۲۴] J. Allen, C. Guinn and E. Horvitz, “Mixed-Initiative Interaction”, 1999, *IEEE Intelligent Systems*, Vol. 14, pages 14-24.

[۲۵] G. Ferguson, J. Allen and B. Miller, 1996, “Towards a Mixed-Initiative Planning Assistant”, *Proceedings of the AAAI.*

[۲۶] S. Whittaker and P. Stenton, 1988, “Cues and control in expert client dialogues”, *Proceedings of the 26th Annual Meeting of the ACL, Association of Computational Linguistics,* pages 123-130.

[۲۷] J.R. Hobbs and M.H. Agar, 1985, “The coherence of incoherent discourse”, Technical Report CSLI-85-38, Center for the Study of Language and Information, Ventura Hall, Stanford University, Stanford, CA 94305.

[٢٨] J.R. Hobbs, 1979, "Coherence and Coreference", *Cognitive Science,* pages 67-90.

[٢٩] P. Grice, 1989, "Studies in the way of words", Cambridge, MA: Harvard University Press.

[٣٠] C. Guinn, 1995, "the role of computer-computer dialogues in human-computer dialogue system development", *Empirical Methods in Discourse Interpretation and Generation, AAAI Spring Symposium Series*, pages 47-52.

[٣١] R. Smith, 1994, "Spoken variable-initiative dialogue: An adaptable natural language interface", *IEEE Expert*, pages 45-50.

[٣٢] D.G. Novick, 1988, "Control of mixed-initiative discourse through meta-locutionary acts: a computational model", Technical Report CIS-TR-88-18, Department of Computer and Information Science, University of Oregon.

[٣٣] L. van Lier, 1988, "The classroom and the language learner: Ethnography and second-language classroom research", Essex, England: Longman.

[٣٤] C. Kinginger, 1994, "Learner initiative in conversation management: An application of van Lier's pilot coding scheme", *The Modern Language Journal*, pages 29-40.

[٣٥] P. Burke, 1994, "Segmentation and control of a dissertation defense", Norwood, N J: Ablex, Pages 95-124.

[٣٦] J. Chu-Carroll, 2000, "MIMIC: An Adaptive Mixed Initiative Spoken Dialogue System for Information Queries", *Proceeding of the sixth conference on Applied natural language processing*, pages 97-104.

[٣٧] B.J. Grosz and C.L. Sidner, 1986, "Attentions, intentions and the structure of discourse", *Computational Linguistics,* pages 175 - 204.

[٣٨] R. Reichman, 1985, "*Getting computers to ta& like you and* me", Cambridge, M.A.: MIT Press.

[٣٩] J. Hirschberg and J.B. Pierrehumhert, 1986, "The intonational structuring of discourse" *In Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics.*

[٤٠] S.E. Brennan, M.W. Friedman and C. Pollard, 1987, "A centering approach to pronouns", *In Proceedings of the 25th Annual Meeting of the Association for Computational Linguistics.*

[۴۱] R. Guindon, P. Sladky, H. Brunner and J. Conner, 1986, “The structure of user adviser dialogues: Is there method in their madness?”, *In Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics.*

[۴۲] M. Fanty, S. Sutton, D.G. Novick and R. Cole, 1995, “Automated Appointment Scheduling”, Center for Spoken Language Understanding Oregon Graduate Institute of Science and Technology, Portland Oregon.

[۴۳] E .Levin, S. Narayanan, R. Pieraccini, K. Biatov, E. Bocchieri, G.D. Fabbrizio, W. Eckert, S. Lee, A. Pokrovsky, M. Rahim, P. Ruscitti and M. Walker, 2002, “The AT&T-DARPA Communicator Mixed-Initiative Spoken Dialog System”, AT&T Labs, Park Avenue, Florham Park, NJ, 07932.

[۴۴] N. Ramakrishnan, R. Capra and M.A. Perez-Quinones, 2002, “Mixed Initiative Interaction = Mixed Computation”, *Proceedings of the 2002 ACM SIGPLAN workshop on partial evaluation and semantics-based program manipulation*, pages 119-130.

[۴۵] M.A. Walker et al., 1998, “Penn Multiparty Standard Coding Scheme Draft Annotation Manual”.

[۴۶] B. Di Eugenio, P.W. Jordan and L. Pylkkanen, 1998, “The Coconut Project: Dialogue Annotation Manual”, Intelligent Systems Program, University of Pittsburgh, Pittsburgh.

[۴۷] M.A. Walker et al., 1998, “Evaluating Spoken Dialogue Agents with PARADISE: Two Case Studies”, *Computer Speech and Language*, Vol. 12, No. 3, pages 317–347.

[۴۸] H. Khouzaimi, R. Laroche and F. Lefevre, 2018, “A methodology for turn-taking capabilities enhancement in Spoken Dialogue Systems using Reinforcement Learning”, *Computer Speech and Language,* pages 93-111.

Abstract

In Interactive Question Answering (IQA) systems, users will ask their questions in a natural language and receive related responses. Because of the interactions in these systems, users can raise descriptions about their questions and increase accuracy.

Human conversations have many features, one of which is the control of the conversation. In these conversations, control may be exchanged between people. Such interactions have a mixed-initiative feature. In these conversations, at any one time, a person may take the initiative and change the direction of the conversation. Mixed-initiative feature is an important and effective aspect of helping people work together to solve problems and make people as a team more productive.

The purpose of this research is to detect the conversation controller in interactive question answering systems using statistical techniques. This method, unlike other methods, which mainly uses semantic and grammatical analysis, uses statistical methods without semantic and cognitive analysis and is therefore independent of the language. In order to detect the controller in this way, it is only necessary to prepare the database and train the system with this database. In this research, a database is designed for this purpose. To detect the controller, methods such as euclidean distance, hamming, minkowski, naive bayes, k-nearest neighbour, k-means, support vector machine, cosine similarity, and TF-IDF have been used. The best performance in this system is the support vector machine model with 91.49%.

Keywords: Interactive Question Answering (IQA), Mixed-initiative, Conversation controller, Initiative, Statistical methods

Shahrood University of Technology

Faculty of Computer Engineering

M.Sc. Thesis in Artificial Intelligence Engineering

Detecting conversation controller Using Artificial Intelligence Techniques

by: Negar ghaempanah

Supervisor:
Dr. Morteza Zahedi

Advisor:
Eng. Mehdi Hosseini

August 2018