درک معنا در سامانه محاورهٔ مبتنی بر متن برای حوزهٔ ذخیره بلیت

پریا جمشیدلو[۱]
محمد بحرانی[۲]

چکیده

درک زبان محاوره حوزهٔ خاصی از درک زبان طبیعی را شامل می‌شود که در آن جملات بیان‌شده توسط کاربر به اندازهٔ جملات زبان نوشتاری تابع دستور زبان نیستند. در این مقاله، سامانه محاورهٔ مبتنی بر متن[۳] برای استخراج معنای جملات محاوره‌ای مربوط به حوزهٔ ذخیره بلیت معرفی می‌شود. در طراحی این سامانه از شیوه‌های مبتنی بر داده استفاده شده است. معماری آن شامل دو بخش اصلی استخراج متغیرها و انتساب محتمل‌ترین برچسب‌های معنایی به دنباله‌ای از کلمات است. برای این کار از الگوی مخفی مارکوف[۴] استفاده می‌شود. برچسب‌زنی معنایی دنبالهٔ کلمات با استفاده از الگوریتم ویتربی[۵] صورت می‌گیرد. بدین منظور، ابتدا پیکره‌ای از جملاتِ مورد استفاده در حوزهٔ ذخیره بلیت جمع‌آوری و سپس به هر کلمه یا ترکیبی از کلمات یک برچسب معنایی تخصیص داده می‌شود. در مرحلهٔ آموزش با استفاده از پیکرهٔ برچسب‌خورده، دنبالهٔ برچسب‌های ممکن برای توالی کلمات مختلف یاد گرفته می‌شود. در مرحلهٔ آزمون با استفاده از احتمالات استخراج‌شده از مرحلهٔ آموزش، محتمل‌ترین برچسب معنایی برای هر کلمه یا ترکیبی از کلمات پیدا می‌شود. بر اساس آزمایش‌های انجام‌شده، دقت سامانه پیشنهادی در تشخیص سه برچسب کلیدیِ مبدأ، مقصد و تاریخ ۹۱ درصد است.

واژه‌های کلیدی: درک معنا، سامانه محاوره‌ای، روش مبتنی بر داده، الگوی مخفی مارکوف، الگوریتم ویتربی

---

[۱] - دانشجوی کارشناسی ارشد، زبان‌شناسی رایانشی، مرکز زبان‌ها و زبان‌شناسی، دانشگاه صنعتی شریف
Jamshidlou@mehr.sharif.ir

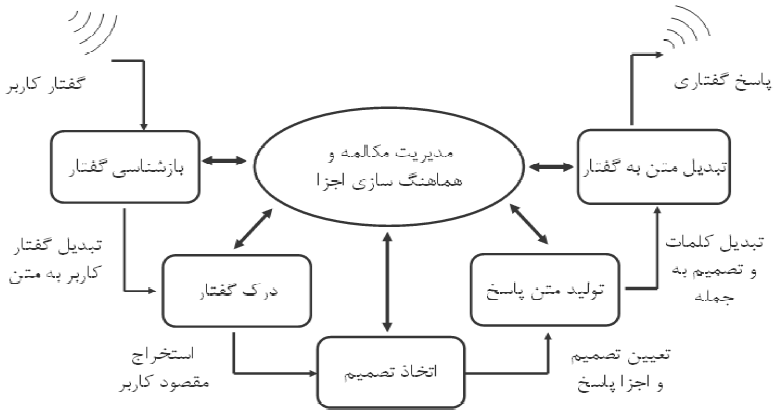
[۲] – استادیار، گروه زبان‌شناسی رایانشی، مرکز زبان‌ها و زبان‌شناسی، دانشگاه صنعتی شریف
bahrani@sharif.ir

[3]- Text-based Dialog Systems
[4]- Hidden Markov Model
[5]- Viterbi Algorithm

۱- مقدمه

هدف نهایی در حوزهٔ پردازش زبان طبیعی[۱] این است که ماشین بتواند زبان انسان را درک کند. درک معنا[۲] از مهم‌ترین بخش‌های سامانه محاورهٔ گفتاری[۳] محسوب می‌شود. به طورکلی، هر سامانه محاورهٔ گفتاری شامل پنج بخش اصلی است. به این ترتیب که ابتدا با استفاده از زیرسامانه بازشناسی گفتار[۴]، گفتار کاربر به متن تبدیل می‌شود؛ سپس این متن به عنوان ورودی به زیرسامانه درک معنا[۵] داده می‌شود و بعد از تحلیل جمله، معنای نهفته در آن استخراج می‌شود. در این مرحله در زیرسامانه مدیریت مکالمه[۶] با استفاده از معنای استخراج‌شده مکالمه به سمت صحیحی هدایت می‌شود. در این این حالت با استفاده از خروجی زیرسامانه مدیریت مکالمه نمایش معنایی از پاسخ به کاربر به وجود می‌آید. در مرحلهٔ بعد این نمایش معنایی با استفاده از زیرسامانه تولید زبان طبیعی[۷] به یک جمله یا متن معنادار و مناسب تبدیل می‌شود و در نهایت با استفاده از زیرسامانه سنتز گفتار[۸]، متن پاسخ به گفتار تبدیل و برای کاربر پخش می‌شود. در شکل ۱ سامانه محاورهٔ گفتاری و زیرسامانه‌های آن نشان داده می‌شود. با توجه به فرایند فوق می‌توان نتیجه گرفت که زیرسامانه درک معنا هستهٔ اصلی سامانه‌های محاورهٔ گفتاری به‌شمار می‌آید. امروزه از سامانه‌های درک معنا در حوزه‌های گوناگونی چون بازیابی اطلاعات[۹]، ترجمهٔ ماشینی[۱۰]، راهنمای اطلاعات مسافرتی، هواشناسی و نیز هدایت تماس در مراکز تماس استفاده می‌شود(بکایی، ۱۳۸۹: ۱).

1- Natural Language Processing
2- Understanding
3- Spoken Dialogue System
4- Speech Recognition
5- Language Understanding System
6- Dialogue Management
7- Natural Language Generation
8- Speech Synthesis System
9- Information Retrieval
10- Machine Translation



شکل ۱- شمای کلی از سامانه محاورهٔ گفتاری و زیرسامانه‌های تشکیل‌دهندهٔ آن

در سامانه‌های محاورهٔ مبتنی بر متن، جملات نوشته‌شده توسط کاربر ابتدا تحلیل معنایی، سپس معنای مورد نظر از جمله استخراج و با توجه به آن معنا، پاسخ مناسبی تولید می‌شود. جملات کاربر ممکن است به صورت محاوره‌ای و یا رسمی بیان شود. در این مقاله سامانه درک معنای مبتنی بر متن برای درک معنای جملات محاوره‌ای در حوزهٔ ذخیره بلیت ارائه می‌شود که در آن ورودی به صورت متن از کاربر دریافت می‌شود و بعد از تحلیل معنایی، کلمات کلیدی یعنی کلمات مربوط به مبدأ، مقصد و تاریخ، تشخیص و برچسب‌گذاری می‌شوند.

دو رویکرد اصلی در پیاده‌سازی سامانه‌های درک معنا شامل روش‌های مبتنی بر قاعده[1] و روش‌های مبتنی بر داده[2] است. روش‌های مبتنی بر قاعده پیچیدگی‌های خاص خود را دارند و نیز مستعد خطا، وقت‌گیر و پرهزینه‌اند. همچنین، گفتار محاوره‌ای در موارد بسیاری از دستور زبان پیروی نمی‌کند. بنابر این، در این تحقیق از روش‌های مبتنی بر داده و به طور خاص از الگوی مخفی مارکوف برای برچسب‌گذاری معنایی استفاده شده است.

۲- بررسی روش‌های موجود در درک معنا

استخراج معنا از متن، فرآیند پیچیده‌ای است که برای آن، الگوها و رویکردهای متفاوتی ارائه شده است. به طورکلی می‌توان روش‌های موجود را به دو دسته روش‌های مبتنی بر قاعده و روش‌های مبتنی بر داده تقسیم‌بندی کرد.

[1]- Rule Based Approaches
[2]- Data Driven Based Approaches

۲-۱ روش‌های مبتنی بر قاعده

مطالعه در زمینهٔ سامانه‌های درک زبان محاوره‌ای اولین بار در دههٔ ۷۰ میلادی و در مرکز تحقیقات درک گفتار موسسهٔ دارپا[۱] آغاز شد. در آن زمان برای درک زبان محاوره‌ای از تکنیک‌های درک معنا مانند ماشین حالت محدود[۲] و شبکهٔ گذار افزوده[۳] استفاده می‌شد(وودز[۴]، ۱۹۸۳). تحقیقات در این زمینه با حمایت مالی موسسهٔ دارپا از سامانه اطلاعات سفرهای هوایی (آتیس)[۵] در دههٔ ۹۰ میلادی شتاب بیشتری گرفت(پرایس[۶]،۱۹۹۰). بسیاری از فن‌آوری‌های مرتبط با درک معنای محاوره‌ای که برای آتیس طراحی شده بودند و بر طراحی سریع و کم‌هزینهٔ سامانه‌های محاورهٔ گفتاری تمرکز داشتند، در برنامه‌های هدایت تماس موسسهٔ دارپا استفاده شدند(والکر[۷] و سایرین، ۲۰۰۲). در آزمایشگاه‌های گوناگونی برای مقاصد پژوهشی و تجاری، سامانه‌های درک معنای محاوره‌ای طراحی شدند. عملکرد این سامانه‌ها به این صورت بود که ابتدا پرس‌وجوی[۸] گفتاری کاربر برای حوزهٔ اطلاعات سفرهای هوایی (مانند اطلاعات پرواز، اطلاعات خدمات فرودگاه و غیره) درک و سپس از پایگاه دادهٔ معیار پاسخ مناسب برگردانده می‌شد.

شیوهٔ سنتی برای انجام این کار نوشتن دستی مجموعه‌ای از قواعد مانند دستور مستقل از متن[۹] یا دستور یکسان‌سازی[۱۰](خسروی‌زاده و همکاران، ۱۳۹۱) است. تهیهٔ این قواعد فرایندی دشوار و پرهزینه و نیازمند افراد متخصص بود. در سال‌های اخیر، برای حل این مشکل الگوهای مبتنی بر داده مورد توجه قرار گرفته است(وانگ[۱۱]، دنگ[۱۲] و آکرو[۱۳]، ۲۰۰۵).

در سامانه‌های مبتنی بر قاعده، به قواعد نحوی/معنایی و یک تجزیه‌گر[۱۴] مقاوم نیاز است تا متن ورودی با استفاده از قواعد نوشته‌شده تحلیل شود. مزیت اصلی این رویکرد عدم نیاز به دادهٔ آموزشی است. اما طراحی سامانه مبتنی بر قاعده به دلایل ذیل پرهزینه است: ۱) تهیهٔ قواعد مستعد خطاست، ۲) تهیهٔ قواعد مناسب، فرایندی زمان‌بر است، ۳) برای تهیهٔ قواعدی که گفتار کاربر را پوشش دهد و

1- DARPA (Defense Advanced Research Projects Agency)
2- Finite State Machines
3- Augmented Transition Network
4-Woods,W.A.
5- ATIS (Air Travel Information System)
6- Price,P.
7- Walker,A.
8- Query
9- Context Free Grammar
10- Unification Grammar
11- Wang,Y.
12- Deng,L.
13- Acro,A.
14- Parser

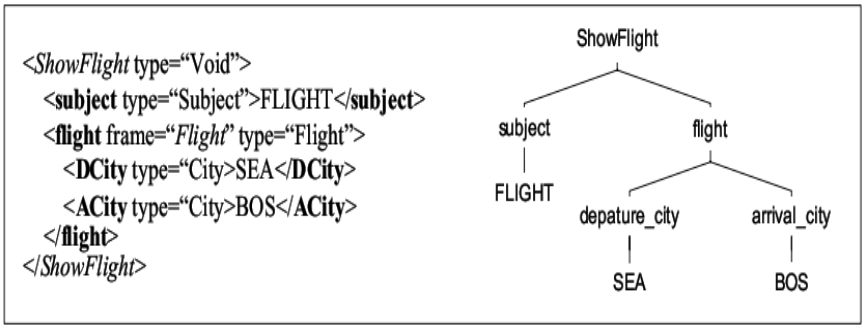
عملکرد بهینه‌ای داشته باشد به متخصصان زبان‌شناسی و مهندسی نیاز است، ۴) نگهداری چنین قواعدی دشوار و گران است. سامانه محاورهٔ گفتاری باید این قابلیت را داشته باشد تا با دادهٔ واقعی جمع‌آوری شود و پس از استقرار به طور خودکار انطباق یابد. این درحالی است که سامانه‌های مبتنی بر قاعده نیازمند دخالت متخصص است و حتی گاهی دخالت طراح اصلی سامانه نیز در حلقهٔ انطباق لازم است(همان). به عنوان مثال سامانه فینیکس[1](وارد[2]، ۱۹۹۰) سامانه درک زبان گفتاری مبتنی بر بر قاعده است. این سامانه از یک تجزیه‌گر مقاوم تشکیل شده است. با استفاده از این تجزیه‌گر، دنباله کلماتِ بازشناسی شده با استفاده از قواعد معنایی موجود به صورت معنایی نمایش داده می‌شوند. در این سامانه برای نمایش معنایی جملات از نمایش قاب[3] و شیار[4] استفاده شده است. فینیکس هر ورودی را به دنباله‌ای از یک یا چند قاب معنایی تجزیه می‌کند. تعداد و انواع این قاب‌ها بر اساس تعداد و تنوع اعمالی که سامانه قادر است انجام دهد تعیین می‌شود.

در حوزه‌های کاربردی خاص، ساختار معنایی به صورت قاب‌های معنایی تعریف می‌شود. شکل ۲ مثال ساده‌شده‌ای از سه قاب معنایی برای حوزهٔ آتیس است. هر قاب شامل چند شیار می‌شود. نوع شیارها مشخص‌کنندهٔ نوع پرکننده‌هاست[5]. به عنوان مثال، شیار« subject»در قاب «ShowFlight» با پایانهٔ معنایی«FLIGHT» (که با کلماتی چون «flight» ، «flights » بیان می‌شود) یا با پایانهٔ معنایی«FARE» (که با کلماتی نظیر «fare » ، «cost » بیان می‌شود) پر شده و مشخص می‌کند کاربر به چه نوع اطلاعات خاصی نیاز دارد.

```
<frame name="ShowFlight" type="Void">
      <slot name="subject" type="Subject"/>
      <slot name="flight" type="Flight"/>
</frame>
<frame name="GroundTrans" type="Void">
      <slot name="city" type="City"/>
      <slot name="type" type="TransType"/>
</frame>
<frame name="Flight" type="Flight">
      <slot name="DCity" type="City"/>
      <slot name="ACity" type="City"/>
      <slot name="DDate" type="Date"/>
</frame>
```

[1]- Phoenix
[2]- Ward,W
[3]- Frame
[4]- Slot
[5]- Filler

شکل ۲- قاب معنایی آتیس. این شکل شامل سه قاب معنایی ساده‌شده برای حوزهٔ آتیس است. DCity در اینجا نشان‌دهندهٔ «شهر مبدأ» و ACity نشان‌دهندهٔ «شهر مقصد» است. این دو شیار به پرکننده‌ای از نوع «شهر» نیاز دارند که می‌تواند به عنوان مثال نام آن شهر یا کد آن باشد. قاب «flight» از نوع «Flight» است، بنابراین می‌تواند به عنوان پرکنندهٔ شیار «Flight» در قاب مرحلهٔ بالاتر یعنی قاب«show flight » مورد استفاده قرار بگیرد(وانگ و سایرین، ۲۰۰۵).



شکل ۳- نمایش معنایی جملهٔ « Show me flights from Seattle to Boston» نمونه‌ای از قاب‌های معنایی در شکل ۲ است. در سمت راست شکل درخت آن نشان داده شده است. این نمونه، قاب نشان‌دهندهٔ معنا را انتخاب و سپس طبق معنایی که از جمله استخراج می‌شود شیارها را پر می‌کند(همان).

معنای یک جملهٔ ورودی، نمونه‌ای از قاب‌های معنایی است. شکل ۳ نمایش معنایی برای جملهٔ« show me flights from Seattle to Boston» را نشان می‌دهد. قاب «ShowFlight» حاوی زیرقاب «Flight» است. در برخی از سامانه‌های محاورهٔ گفتاری امکان ایجاد زیر ساختار در یک قاب وجود ندارد. در این صورت، نمایش معنایی به صورت فهرستی از جفت ویژگی- مقادیر [۱]ساده می‌شود که به آنها نمایش جفت کلمه‌های کلیدی[۲] یا نمایش مسطح[۳] نیز اطلاق می‌شود(پیراسینی[۴] و لوین[۵]، ۱۹۹۳).

[1]- Attribute-Value Pair
[2]- Keyword Pair
[3]- Flat Representation
[4]- Pieraccini,R.
[5]- Levin,E.

شکل ۴- نمایش مسطح برای جملهٔ« Show me the flights from Seattle to Boston» است. نمایش مسطح نوع خاصی از نمایش مبتنی بر قاب است که در آن امکان قرار دادن ساختاری در دل ساختار دیگر وجود ندارد(وانگ و ساسرین، ۲۰۰۵).

۲-۲- روش‌های مبتنی بر داده

روش‌های مبتنی بر داده مشکلات مطرح‌شده در سامانه‌های مبتنی بر قاعده را ندارند. سامانه درک معنای مبتنی بر داده به دادهٔ آموزشی برچسب‌خورده نیاز دارد تا بتواند به طور خودکار از جملات موجود در دادهٔ آموزشی با برچسب‌های معنایی معادلشان عمل استخراج معنا را یاد بگیرد. برچسب‌گذاری معنایی در مقایسه با تهیهٔ قواعد نحوی/معنایی کار ساده‌تری به نظر می‌رسد و نیاز به دانش تخصصی ندارد. همچنین می‌توان با آموزش بی‌سرپرست[1]، روش‌های مبتنی بر داده را بر روی داده‌های جدید تطبیق داد. در ادامه چارچوب کلی روش‌های مبتنی بر داده در سامانه‌های محاوره معرفی می‌شود.

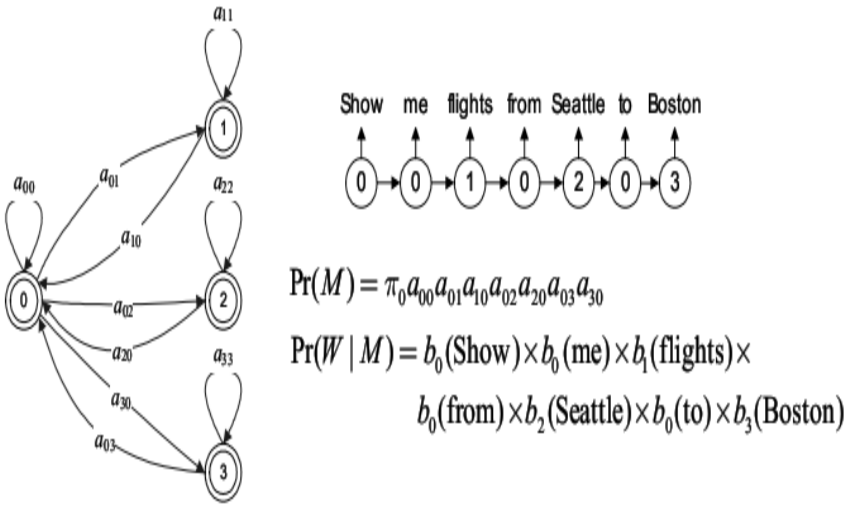
در رویکرد مبتنی بر داده، درک معنا به مسئلهٔ تشخیص الگو[2] تبدیل می‌شود. به این ترتیب که با داشتن دنبالهٔ کلمات W، هدف از درک معنا یافتن نمایش معنایی M است به طوری‌که احتمال P(M|W) بیشینه شود.

$$\hat{M} = \arg_{M} \max P(M \mid W) = \arg_{M} \max P(W \mid M) P(M) \qquad (1)$$

در این چارچوبِ زایا دو الگوی مجزا دیده می‌شود: الگوی پیشین معنایی P(M) که احتمالی را به ساختار معنایی زیرین یا معنای M تخصیص می‌دهد و الگوی واژگانی P(W|M) که به آن الگوی درک یا الگوی تولید واژگانی[3] (خسروی‌زاده و همکاران، ۱۳۹۱) نیز اطلاق می‌شود(میلر[4] و سایرین، ۱۹۹۴). این الگو با فرض داشتن ساختار معنایی، به جملهٔ روساختِ W (یعنی به توالی کلمات) احتمالی را نسبت می‌دهد. به عنوان مثال، الگوی برچسب‌زنی مارکوف، پیاده‌سازیِ ساده‌شده از معادلهٔ بالا است. در این نوع برچسب‌زن، حالات یک زنجیرهٔ مارکوف، معنای پیشین را الگو قرار می‌دهد و مشاهدات[5] آن فرایند، تولید واژگانی را الگو می‌کند. هم‌ردیفیِ بین مشاهدات (کلمات) و حالات (برچسب معنایی)، مخفی است (شکل ۵). این مدلِ برچسب‌زن، با استفاده از الگوریتم ویتربی

[1]- Unsupervised Training
[2]- Pattern Recognition
[3]- Lexical Generation
[4]- Miller,R.
[5]- Observations

هم‌ردیفیِ بین حالات و کلمات را به دست می‌آورد و معنای مربوط به هر حالت، برچسب معنایی کلمهٔ هم‌ردیف‌شده می‌شود.



شکل ۵- الگوی برچسب‌زن مارکوف برای سامانه‌های محاوره است. حالت ۱ نشان‌دهندهٔ شیار «command »، حالت ۱ نشان‌دهندهٔ شیار «subject»، حالت ۲ نشان‌دهندهٔ شیار «DCity» و حالت ۳ نشان‌دهندهٔ شیار «ACity »است. توپولوژی HMM و احتمالات گذار a_{ij}، الگوی پیشین معنایی را تشکیل می‌دهند. معنای یک جمله با استفاده از توالی حالات نشان داده می‌شود. توزیع b_k در هر حالت، مدل واژگانی را تشکیل می‌دهد. در سمت راست، توالی حالات ممکن نشان داده شده است که با جملهٔ« Show me fights from Seattle to Boston» هم‌ردیف می‌شود(وانگ و سایرین، ۲۰۰۵).

بکایی(۱۳۸۹: ۱۰۴) برای درک معنای جملات فارسی در دامنهٔ «کیوسک اطلاع‌رسانی» از روش‌های مبتنی بر داده استفاده کرده است. وی در پایان‌نامه خود دو روش مبتنی بر داده را پیاده‌سازی و بهبود داده است. ابتدا الگوی مخفی مارکوف برای درک زبان طبیعی پیاده‌سازی شد، سپس الگوریتم ویتربی به گونه‌ای تغییر یافت که امکان در نظر گرفتن ترکیبی از کلمات در زبان فارسی در نظر گرفته شود؛ چرا که در زبان فارسی فاصله همیشه نشان‌دهندهٔ مرز کلمات نیست.

۳- روش پیشنهادی

نحوهٔ نمایش معنا در سامانه ارائه‌شده در این مقاله به صورت نمایش مسطح است. برای استخراج معنا در این روش از الگوی مخفی مارکوف استفاده شده است که حالات آن برچسب‌های معنایی

مانند برچسب زمان، مبدأ، مقصد و غیره بود و مشاهدات آن کلمات و عبارات موجود در جملهٔ بیان شده توسط کاربر است.

طبق معادلهٔ (۱) با فرض داشتن دنبالهٔ کلمات $W=w_1 \ldots w_n$، باید دنبالهٔ برچسب‌های $M=m_1 \ldots m_n$ طوری پیدا شود که $P(M|W)$ بیشینه شود:

$$\hat{M} = \arg\max_{M} P(M \mid W) \qquad (2)$$

با استفاده از قانون بیز[1] می‌توان معادلهٔ بالا را به صورت زیر نوشت:

$$\arg\max_{T} P(M|W) = \arg\max_{m} \frac{P(W|M)P(M)}{P(W)} = \arg\max_{M} P(W|M)P(M) \qquad (3)$$

$$= P(w_1 w_2 \cdots w_n | m_1 m_2 \cdots m_n) P(m_1 m_2 \cdots m_n)$$

که $P(M)$ مدل پیشین معنایی و $P(W|M)$ مدل واژگانی است. محاسبهٔ احتمال‌های فوق در عمل بسیار مشکل است. بنابر این چند فرض ساده‌کننده برای محاسبهٔ احتمال در این مدل‌ها به کار می‌رود:

$$P(m_1 m_2 \cdots m_n) \approx \prod_{i=1}^{n} P(m_i \mid m_{i-1}) \qquad (4)$$

$$P(w_1 w_2 \cdots w_n \mid m_1 m_2 \cdots m_n) \approx \prod_{i=1}^{n} P(w_i \mid m_i) \qquad (5)$$

طبق فرض‌های ساده‌کنندهٔ بالا، الگوی پیشین معنایی از طریق حاصل‌ضرب احتمالات دوتایی[2] برچسب‌های معنایی محاسبه می‌شود و الگوی واژگانی نیز از احتمال تولید واژگانی کلمات و برچسب‌ها به دست می‌آید.

از روابط فوق نتیجه می‌شود که باید دنبالهٔ برچسب‌ها را طوری پیدا کنیم که عبارت زیر بیشینه شود:

$$P(w_1 w_2 \cdots w_n \mid m_1 m_2 \cdots m_n) P(m_1 m_2 \cdots m_n) \approx \prod_{i=1}^{n} P(w_i \mid m_i) \cdot \prod_{i=1}^{n} P(m_i \mid m_{i-1}) \qquad (6)$$

احتمال دوتایی برچسب‌ها و همچنین احتمال تولید واژگانی در معادلات فوق را می‌توان به عنوان متغیرهای الگوی مخفی مارکوف مرتبه اول در نظر گرفت. این متغیرها به ترتیب معادل با ماتریس

[1]- Bayesian Rule
[2]- Bigram

احتمال گذار[1] و ماتریس احتمال مشاهدات به شرط حالات در الگوی مخفی مارکوف است. به منظور آموزش الگوی مخفی مارکوف، متغیرهای مذکور از پیکرهٔ[2] برچسب‌خورده استخراج می‌شوند. برچسب کلمات نمایانگر معنا و مفهوم کلمات است. این متغیرها از طریق شمارش تکی(تک‌نگاشتی)[3] و دوتایی برچسب‌ها در پیکره به صورت زیر محاسبه می‌شوند:

$$P(m_i \mid m_{i-1}) = \frac{C(m_{i-1} m_i)}{C(m_{i-1})} \qquad (7)$$

طبق معادلهٔ بالا احتمال دوتایی برچسب‌ها مساوی است با شمارش دوتایی برچسب‌ها تقسیم بر شمارش تکی برچسب اول.

$$P(w_i \mid m_i) = \frac{C(w_i, m_i)}{C(m_i)} \qquad (8)$$

طبق معادلهٔ بالا احتمال تولید واژگانی مساوی است با شمارش هر کلمه و برچسب آن تقسیم بر شمارش تکی آن برچسب.

پس از آموزش الگو، از الگوریتم ویتربی برای انتساب محتمل‌ترین برچسب معنایی به کلمات و عباراتِ جملهٔ ورودی استفاده می‌شود. الگوریتم ویتربی محتمل‌ترین دنباله از برچسب‌ها را به شرط مشاهدهٔ دنباله‌ای از کلمات تعیین می‌کند. این الگوریتم شامل سه مرحلهٔ مقداردهی اولیه[4]، تکرار[5] و عقب‌گرد[6] است. این الگوریتم این گونه عمل می‌کند که ابتدا جملهٔ ورودی را دریافت و آن را به کلمات و عبارات تشکیل‌دهنده تجزیه می‌کند؛ سپس برای هر کلمه یا ترکیبی از کلمات، تمامی برچسب‌های ممکن را در نظر می‌گیرد و به هر کدام یک امتیاز اختصاص می‌دهد و برای هر کلمه یا عبارت، آن برچسبی را درنظر می‌گیرد که امتیازش بالاتر است. همین روند تا آخر تکرار و مسیرهای طی شده حفظ می‌شود تا در نهایت آن مسیری انتخاب شود که به ازای آن، احتمال دنبالهٔ مشاهدات بیشینه شود. به این ترتیب، محتمل‌ترین دنبالهٔ برچسب‌های معنایی به دنبالهٔ کلمات موجود در جمله نسبت داده می‌شود. در الگوریتم ویتربی از احتمال‌های دوتایی و احتمال‌های تولید واژگانی که از دادهٔ آموزشی استخراج شده‌اند برای محاسبه محتمل‌ترین دنباله از برچسب‌ها استفاده می‌شود. در این مقاله، از آنجا که حوزهٔ مورد نظر برای درک معنا، حوزهٔ ذخیره بلیت است، هدف از درک معنا، برچسب‌گذاری صحیح کلمات مربوط به مبدأ، مقصد و تاریخ است. بنابر این اگر سه مفهوم مبدأ، مقصد و تاریخ به درستی برچسب خورده باشند عمل درک معنا به خوبی انجام گرفته است.

1- State Transition Matrix
2- Corpus
3- Monogram Count
4- Initialization
5- Recursion
6- Backtrack

۴- تهیه دادهٔ برچسب‌خورده

برای طراحی این سامانه، پیکره‌ای از جملات مربوط به حوزهٔ ذخیره بلیت تهیه شده است. این پیکره حاوی ۱۴۴ جمله و ۱۴۵۶ کلمه و ۱۰۸۸ برچسب است که هر جمله به طور میانگین از ۱۰ کلمه تشکیل شده است. پنج ششم دادهٔ جمع‌آوری‌شده برای آموزش متغیرهای الگو و یک ششم آن برای مرحلهٔ آزمون استفاده می‌شود. در تهیهٔ این دادگان سعی شده از جملات واقعی استفاده شود. با توجه به این که در زبان محاوره‌ای ممکن است کلمات در جایگاه سنتی خود قرار نگیرند و جابه‌جا شوند صورت‌های مختلف یک جمله در پیکره در نظر گرفته شده است. به عنوان مثال جملهٔ «من یه بلیت می‌خوام واسه تهران مشهد» و یا «واسه تهران مشهد یه بلیت می‌خواستم». همچنین گاهی ممکن است مبدأ و تاریخ حرکت به طور ضمنی از جمله برداشت شود که نمونهٔ این جملات نیز در پیکره گنجانده شده است. مانند «یه بلیت واسه شیراز می‌خواستم برا فردا»، در این جمله مبدأ حرکت ذکر نشده و تاریخ نیز به طور غیرمستقیم بیان شده است. در این صورت مبدأ به صورت پیش‌فرض تهران و تاریخ پیش‌فرض، تاریخ همان روز در نظر گرفته می‌شود.

پس از جمع‌آوری جملات، به هر کلمه و یا گروهی از کلمات یک برچسب معنایی اختصاص داده شد. قسمتی از پیکرهٔ برچسب‌خورده در زیر نمایش داده شده است.

```
یه*N
بلیت
می‌خواستم*R
واسه*DP
بیست‌ویکم
دی*D
از*O
به*AP
ارومیه*A
. *DELM
```

شکل۶- نمایش بخشی از پیکرهٔ برچسب‌خورده

این پیکره به طور کلی ۶ ساختار مختلف از جملات را پوشش می‌دهد که از هر ساختار ۲۴ جمله در پیکره آمده است و مجموعاً ۱۴۴ جمله را تشکیل می‌دهند. همچنین، ۱۱ نوع برچسب معنایی برای سیستم تعریف شده که شرح آن در جدول زیر آمده است.

برچسب	نماینده	مثال
S	فاعل	من
DP	حرف اضافهٔ تاریخ	برای
D	تاریخ	۱۹ اردیبهشت
AP	حرف اضافهٔ مقصد	به
A	مقصد	تهران
OP	حرف اضافهٔ مبدأ	از
O	مبدأ	مشهد
RC	مکمل درخواست	می‌خواستم
R	درخواست	بلیت بگیرم
N	تعداد	یه
DELM	علائم نگارشی	-

جدول ۱- انواع برچسب‌های تعریف‌شده در پیکره

در این پیکره گاهی به ترکیبی از دو کلمه یک برچسب معنایی واحد اختصاص داده شده است؛ مانند «۴ اسفند» یا «بیستم مهر» که با هم برچسب تاریخ (D) را دارند. در جدول زیر نمونه‌ای از این کلمات ترکیبی آورده شده است.

ترکیب کلمات	برچسب
بلیت می‌خواستم	R
۲ تیر	D
به مقصد	AP
برای تاریخ	DP
از مبدأ	OP

جدول ۲- نمونه برچسب‌های تعریف‌شده برای ترکیب کلمات

ساختار مختلف جملات در جدول زیر به تفصیل نشان داده شده است.

	ساختار و جملات نمونه
۱	تاریخ (D) + مبدأ (O) + مقصد (A) + درخواست (R)
	برای ۱۵ اسفند از مشهد به کرمان بلیت می‌خواستم.

۲	درخواست (R) + تاریخ (D)+ مبدأ (O) + مقصد (A)
	یه بلیت می‌خواستم برای ۱۷ مرداد از تهران به ساری.
۳	مکمل درخواست (RC) + مبدأ (O) + مقصد (A) + تاریخ (D)
	من می‌خواستم از همدان به تهران واسه ۱۸ شهریور بلیت بگیرم.
۴	تاریخ (D) + مقصد (A) + درخواست (R)
	من برای ۸ آبان به مشهد بلیت می‌خواستم.
۵	درخواست (R) + تاریخ (D) + مقصد (A)
	یه بلیت می‌خواستم واسه تاریخ پنجم آذر واسه تبریز.
۶	مقصد (A) + تاریخ (D) + درخواست (R)
	برا اصفهان تاریخ ۹ مرداد بلیت می‌خوام.

جدول ۳- انواع ساختارهای تعریف‌شده در پیکره و جملات نمونه برای هر ساختار

۵- ارزیابی سیستم

به منظور ارزیابی عملکرد سامانه، از دادهٔ آزمون استفاده کردیم که حاوی ۲۴ جملهٔ جدید و ۲۴۹ کلمه است. به طور متوسط، هر جمله از ۱۰ کلمه تشکیل شده است. از هر ساختار ۴ جمله در نظر گرفته شد که مجموعاً ۲۴ جمله را شامل می‌شود. برای ارزیابی عملکرد سامانه از ۳ معیار مختلف استفاده شده است:

۱. میزان دقت سامانه در تشخیص برچسب‌های کلیدی یعنی برچسب تاریخ (D) ، مبدأ (O) و مقصد (A)
۲. میزان دقت سامانه در برچسب‌زنی کل کلمات دادهٔ آزمون
۳. تعداد جملات درست درک شده بر اساس ۳ برچسب کلیدی تاریخ (D) ، مبدأ (O) و مقصد (A)

جملات دادهٔ آزمون برای برچسب‌گذاری به سامانه داده شد که مجموعاً از ۶۰ کلمهٔ مربوط به تاریخ، مبدأ و مقصد، ۵۳ کلمه به درستی و ۷ کلمه به اشتباه برچسب زده شده بودند. بنابراین:

$$\frac{۵۳}{۶۰} = \frac{\text{تعداد کلمات مربوط به تاریخ، مبدأ و مقصد که به درستی برچسب خورده‌اند}}{\text{تعداد کل کلمات مربوط به تاریخ، مبدأ و مقصد موجود در داده آزمون}} = ٪۸۸ \quad (۹$$

از مجموع ۱۸۲ کلمهٔ موجود در دادهٔ آزمون، ۱۶ کلمه برچسب نادرست و ۱۶۶ کلمه برچسب درست داشتند. بنابر این:

$$\frac{\text{تعداد کلمات درست برچسب خورده در داده آزمون}}{\text{تعداد کل کلمات جملات داده آزمون}} = \frac{۱۶۶}{۱۸۲} =٪۹۱$$ (۱۰)

همچنین از ۲۴ جملهٔ موجود در دادهٔ آزمون، ۶ جمله حاوی برچسب نادرست برای کلمات تاریخ، مبدأ یا مقصد بودند و ۱۸ جمله به درستی برچسب زده شده‌اند. بنابر این

$$\frac{\text{تعداد جملات درست درک شده}}{\text{تعداد کل جملات داده آزمون}} = \frac{۱۸}{۲۴} =٪ ۷۵$$ (۱۱)

نتایج به دست آمده از ۳ معیار ارزیابی در جدول زیر نمایش داده شده است.

	معیار ارزیابی	دقت سامانه
۱	تشخیص برچسب‌های کلیدی مبدأ، مقصد، تاریخ	۹۱٪
۲	برچسب‌زنی کل کلمات دادهٔ آزمون	۸۸٪
۳	تعداد جملات درست درک شده بر اساس ۳ برچسب کلیدی مبدأ، مقصد، تاریخ	۷۵٪

جدول ۴- نتایج نهایی ارزیابی سامانه درک معنا برای حوزهٔ ذخیره بلیت

۶- نتیجه‌گیری

در این پروژه، سامانه‌ای برای درک معنای جملات محاوره‌ای در حوزهٔ ذخیره بلیت ارائه شد که در آن الگوی مخفی مارکوف برای درک و استخراج معنا از جملات پیاده‌سازی شد. نتایج حاصل از ارزیابی سامانه در مجموع رضایت‌بخش و دقت سامانه در تشخیص برچسب‌های کلیدی مبدأ، مقصد و تاریخ ۹۱ درصد، در برچسب‌زنی کل دادهٔ آزمون ۸۸ درصد و در درک جملات بر اساس سه برچسب کلیدی مبدأ، مقصد و تاریخ ۷۵ درصد بود. این روش نسبت به روش‌های مبتنی بر قاعده ساده‌تر است، زیرا به تهیهٔ قواعدی که با آن بتوان ساختار جملات بیان‌شده را توضیح داد، نیاز ندارد. در واقع، تهیهٔ دادهٔ آموزشی و برچسب‌زنی کلمات آن نسبت به تهیهٔ قواعد، کار ساده‌تری به نظر می‌رسد، به ویژه اینکه دادهٔ آموزشی برای حوزهٔ خاصی یعنی ذخیره بلیت است. با گسترش پیکره و افزودن بخش مدیریت مکالمه می‌توان قابلیت سامانه را افزایش و آن را به سمت سامانه محاورهٔ کامل سوق داد.

منابع

- بکایی، محمد هادی(۱۳۸۹) *درک معنا در سیستم‌های محاوره مبتنی بر گفتار*، پایان‌نامه کارشناسی ارشد، دانشگاه صنعتی شریف.
- خسروی‌زاده. پروانه، مارال شیخ‌زاده، مهدی مرادی و وحید مواجی(۱۳۹۱). *فرهنگ توصیفی زبان‌شناسی رایانشی*،تهران: انتشارات پردیس دانش.
- Miller,S. , R. Bobro, R. Ingria, and R. Schwartz(1994). "Hidden Understanding Models of Natural Language", in *Proceedings of 31st Annual Meeting of the Association for Computational Linguistics*, pp. 25-32, New Mexico State University.
- Mori, R. D. and others(2008). "Spoken Language Understanding", *IEEE Signal Processing Magazine.*
- Pieraccini,R. and E. Levin(1993). "A Learning Approach to Natural Language Understanding", in *Proceedings of 1993 NATO ASI Summer School*,pp. 1-19, Bubion, Spain.
- Price,P.(1990). "Evaluation of Spoken Language System: The ATIS Domain", in *Proceedings of DARPA Speech and Natural Language Workshop*, pp. 91-95, Hidden Valley, PA.
- Walker, M. A. and others (2002). "DARPA Communicator Evaluation: Progress from 2000 to 2001", in *Proceedings of International Conference on Speech and Language Processing*, pp. 273-276, Denver, Colorado.
- Wang, Y.-Y. , L. Deng, and A. Acero(2005). "An Introduction to Statistical Spoken Language Understanding", *IEEE Signal Processing Magazine*, vol. 22, no. 5, pp. 16-31.
- Woods, W. A. (1983). "Language Processing for Speech Understanding", in *Computer Speech Processing*, Englewood Cliffs, NJ: Prentice-Hall International, pp. 305-334.
- Ward, W.(1990). "The CMU air travel information service: Understanding Spontaneous Speech", in *Proceedings of the DARPA Speech and Natural Language Workshop*, pp. 127-129.