

نگاهی به مسأله واژگان در ماشین ترجمه

آرزو مولوی وردنجانی

دانشگاه آزاد اسلامی - واحد امیدیه

چکیده

دریافت چگونگی دسترسی ذهن انسان به عناصر واژگانی ذخیره در واژگان ذهنی و آگاهی از شیوه بازنمایی دانش کمک زیادی به طراحی سیستم‌های پردازش زبان طبیعی می‌کند. در این مقاله سعی شده است تا پس از بررسی مسأله واژگان در ماشین ترجمه، با استفاده از شیوه بازنمایی مفهومی نوعی سیستم بازنمایی واژگان به کمک مفاهیم و روابط پیشنهاد شود که می‌تواند جهت بهبود عملکرد ماشین ترجمه به کار رود. در نهایت عملکرد این شیوه در زمینه رفع ابهام واژگانی، در قالب نمودارهای مفهومی، نشان داده شده است. **کلیدواژه‌ها:** ماشین ترجمه، پردازش زبان طبیعی، بازنمایی مفهومی واژگان.

۱. مقدمه

ترجمه ماشینی در زندگی امروز نقش انکارناپذیری ایفا می‌کند. نیاز به بازیابی سریع اطلاعات، تسهیل کار ترجمه و صرفه‌جویی در هزینه و نیروی انسانی و نیز جذابیت ساخت دستگاهی هوشمند که حاصل شبیه‌سازی فعالیت‌های انسان است، و به تلاش‌های بی‌وقفه‌ای در این موضوع خاص در حوزه زبان‌شناسی رایانه‌ای منجر شده است. اما ترجمه متون از طریق کامپیوتر همواره محدودیت‌های خاص خود را داشته است. هرچند در ابتدای امر به لحاظ پیچیدگی‌های زیاد انتقادات فراوانی به آن وارد بود، ولی نتایجی که امروز از تجربیات بیش از نیم قرن مطالعه در دست داریم، حاکی از این است که ترجمه ماشینی امری ممکن است و ضرورت سرمایه‌گذاری در این رابطه را بدیهی می‌نماید. از طرفی دریافتن این موضوع که ذهن چگونه

موفق به بازیابی و دسترسی دوباره به عناصر واژگانی از میان مخزن واژگانی ذخیره‌شده در حافظه می‌شود کمک بسیار زیادی به طراحی سیستم‌های هوشمند پردازش زبان می‌کند. در واقع، زبان‌شناسان، روان‌شناسان زبان و برخی متخصصان علوم کامپیوتر در طرح مدل‌های خود همواره داشتن واقعیت روانی آن‌ها را مدنظر قرار می‌دهند. ولی گروهی دیگر از دانشمندان علوم رایانه به فرایندهای ذهنی انسان تنها به‌عنوان منبعی ممکن از الگوریتم‌های سودمند و روابط در ساختارهای ساده با هدف طراحی واژگان و دسترسی به واژه‌های خاص می‌نگرند.

محققان با بررسی برون‌داد ماشین ترجمه تلاش می‌کنند تا نقص‌های آن را دریافته به اصلاح آن بپردازند. گروهی، مهم‌ترین ایرادهایی که از این طریق به عملکرد ماشین‌های ترجمه مختلف گرفته شده را به نقص برنامه نویسی، ضعف تئوری زبانی در وارد کردن معنا و نحو محور بودن آن و نداشتن قدرت تعقل انسانی در آن نسبت می‌دهند (فلاحی و نعمتی، ۱۳۸۳: ۳)، اما با نگاهی دقیق‌تر به برون‌دادهای حاصله از سیستم ترجمه ماشینی، احتمال می‌رود ریشه مشکلات آن را بتوان در ضعف واژگان (حجم واژه‌ها، اطلاعات و تعاریف داده شده) و در شیوه پردازشی سیستم ترجمه جستجو کرد. از این‌رو در این مقاله سعی شده است تا با توجه به ساختار واژگان در ذهن و استفاده از نظریه دسترسی فعالسازی تعاملی مفهوم‌بنیاد، نوعی بازنمایی مفهومی برای بخش واژگان سیستم‌های پردازش زبان پیشنهاد شود که استفاده از آن در بهبود عملکرد ماشین ترجمه در رفع ابهام واژگانی و دسترسی به معنای درست واژه موردنظر مؤثر خواهد بود. به این ترتیب، بی‌آن‌که از ماشین ترجمه انتظار تفکر و تعقل داشت می‌توان سیستم‌های مترجم قوی‌تری طراحی کرد.

۲. بررسی پایگاه واژگان ماشین ترجمه

واژگان در تمام مسائل زبانی مرکز توجه است، به‌عبارتی زبان در واژگان شکل می‌گیرد. سیستم‌های رایانه‌ای پردازش زبان و هوش مصنوعی به همان اطلاعات پیچیده‌ای نیاز دارند که انسان درباره روابط بین واژه‌ها می‌داند و با آن می‌تواند زبان را به شیوه مؤثری پردازش کند (خداپرستی، ۱۳۸۵: ۳۴). بانک واژگان کارآمد باید هم ساختاری روشن و مشخص داشته باشد و هم بتوان از راه‌های مختلف به اطلاعات واژه‌های موجود در آن دست یافت. به‌عبارت‌دیگر، برای یافتن کلمات در بانک واژگان باید راه‌حلهایی فراهم کرد که بر هر دو مقوله صورت و معنی استوار باشد.

در تهیه چنین مجموعه پیچیده‌ای باید انواع دانش (از جمله دانشی که حاصل کار زبان‌شناسان و روان‌شناسان و متخصصان کامپیوتر و فیلسوفان است) درج شود. برای این کار از ابزارهایی چون شبکه‌های معنایی، نمودارهای مفهومی و منطق گزاره‌ها استفاده می‌شود که در غالب موارد استفاده از بازنمایی شبکه‌ای برای نشان دادن دانش معنایی بسیار کارآمدتر از

نگاهی به مساله واژگان در ماشین ترجمه

ابزاری چون زبان منطق محمولات است (منصوری و دیگران، ۱۳۸۴: ۱۰۶). به طور کلی، برخی از ویژگی‌های بانک واژگان مناسب به صورت زیر توصیف شده است (ایوانز و دیگران، ۱۳۷۵: ۴۹) الف. دستیابی به آن با روش‌های گوناگون میسر باشد. یعنی، نظام‌های پردازش زبان بتوانند به صورت‌های مختلف کلمات را بیابند. چنانچه پیش‌از این اشاره کردیم، برای یافتن کلمات در بانک واژگان باید راه‌حلهایی فراهم کرد که هم بر صورت و هم بر معنی استوار باشد.

ب. مدخل‌ها از روی معنی جدا شده باشد.

ج. برای گروه‌ها (phrases) و اصطلاحات (idioms) مدخل‌های جداگانه وجود داشته باشد.

د. قابلیت استنتاج داشته باشد. به طور کلی، نظام‌های پردازش زبان باید قادر به یافتن روابط معنایی بین مدخل‌ها باشند.

ه. قابلیت گسترش داشته باشد. واژگان را باید به صورتی طراحی کرد که در آینده بتوان تعداد واژه‌ها را به سادگی افزایش داده، کیفیت اطلاعات آن را مطلوب‌تر کرد (ایوانز، همان).

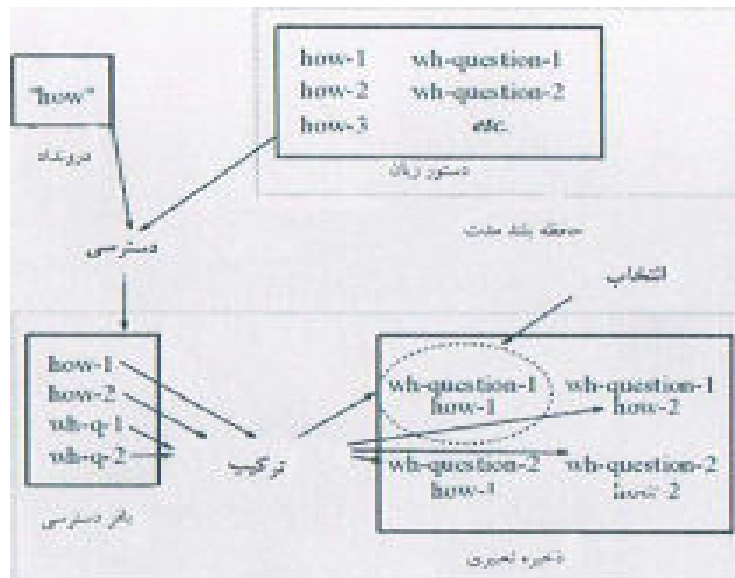
یکی از شیوه‌های ترجمه در سیستم ترجمه ماشینی بر پایه مدل مبتنی بر آنتالوژی^۱ طراحی شده است (جهانگیری، عرفانی، ۱۳۸۲: ۱۳۵). اهمیت این مدل در نگاه متفاوتی است که به مسأله واژگان در ماشین ترجمه دارد. وجود الگوی دقیق و جهانی به نام آنتالوژی یا فرهنگ هستی‌شناختی که در آن تعریف یکسانی از مقولات معنایی پایه (مفاهیم، روابط، ویژگی‌ها و غیره) ارائه می‌دهد و چارچوب استواری برای حوزه‌های خاص و ایجاد فرهنگ واژگانی انعطاف‌پذیر جهت پردازش زبان طبیعی تلقی می‌شود، دقت پردازش معنایی و درنهایت، کارایی سیستم را افزایش می‌دهد. در این مدل ترکیب اجزا و شکل‌گیری جمله معادل در زبان مقصد می‌تواند بدون شناخت زبان مبدأ صورت پذیرد. به عبارت دیگر، سیستم از طریق طرح‌بندی مفاهیم کار ترجمه را انجام می‌دهد. مفاهیم که در واقع بخشی از دانش ذهنی ما، شامل مفاهیم و روابط میان آن‌ها، هستند در این مدل به کمک مؤلفه‌های معنایی تعریف می‌شوند. باز نمود دانش مفهومی به شبکه‌های معنایی و نمودارهای مفهومی شباهت دارد که در واقع بهترین راه حل برای صوری‌سازی زبان محسوب می‌شود.

در ادامه ابتدا نظریه دسترسی به عنوان یکی از مراحل اصلی پردازش زبان تعریف می‌شود، سپس سعی شده است تا نوعی شبکه پیوندی بازنمایی مفهومی طرح و با استفاده از مدل فعال-سازی تعاملی، مکانیسمی برای دسترسی درست به مفاهیم مورد نظر از این شبکه واژگانی پیشنهاد شود.

1. Antalogy

۳. نظریه دسترسی

نظریه‌های دسترسی (access theories) بخش مهمی از هر مدل درک یا تولید جمله هستند. اصطلاح دسترسی به عمل جمع کردن شواهد مختلف (نوشتاری، آوایی و ...) از میان داده‌های ذخیره شده در مخزن واژگان ذهنی اطلاق می‌شود^۱ (ژورافسکی، ۲۰۰۰: ۶۳). بعد از این که مواردی در جستجوی یک پاسخ صحیح برای درون داد خاصی دسترسی شدند باید به ترتیبی فعال نگه داشته شوند تا وقتی که شواهد یکی از آن‌ها به حد آستانه برسد و در نهایت شناخته شود. ژورافسکی در تشخیص یک محرک درون داد سه مرحله دسترسی، یکپارچه‌سازی و انتخاب را در نظر می‌گیرد و پیشنهاد می‌کند مراحل درک یک درون داد به صورت زیر اتفاق افتد.



نمودار ۱. مدل دسترسی واژگانی ژورافسکی

عمل کرد دسترسی در مورد هر نوع اطلاعات زبانی شامل اطلاعات معنایی، نحوی یا اطلاعات مربوط به متن در فراهم آوردن شواهد دسترسی به ساخت‌ها به‌طور تعاملی (interactive) است، یعنی هم‌زمان اطلاعات مختلف را به کار می‌گیرد. سپس در مرحله یکپارچه‌سازی تمام داده‌های کپی شده در بافر^۲ دسترسی جمع‌آوری می‌شوند تا تعابیر یا حدس‌های ممکن شکل

1. Jurafsky, 2000: 63.

۲. بافر به معنی "محلی که سطح فعال‌سازی ساخت‌های مختلف پس از وارد شدن به آن تغییر نمی‌کند" به کار رفته است (نگارنده).

نگاهی به مساله واژگان در ماشین ترجمه

گرفته و در آخر، در مرحله انتخاب، تعبیر مطلوب برگزیده شود. این جاست که محرک درون داد شناخته می‌شود.

تاکنون تعداد زیادی از این گونه مدل‌ها پیش‌بینی شده است، ولی به دلیل وسعت مسائلی که در این بین مطرح می‌شود، باید همواره حوزه مطالعه را به بخش‌های کوچک‌تر تقسیم و هر کدام را به‌طور جداگانه بررسی کرد. همچنان‌که پژوهش‌های گسترده‌ای در حوزه‌های مختلف دسترسی از شیوه‌های دسترسی نحوی گرفته تا مدل‌های دسترسی معنایی، پیوندی و واژگانی انجام شده است. در این تحقیق از نوعی مدل دسترسی واژگانی به‌عنوان مکانیسم پیشنهادی در ماشین ترجمه استفاده شده است.

۴. بازنمایی مفهومی دانش و شیوه دسترسی به واژگان

مسئله واژه‌ها مجموعه‌ای مجزا از واحدها نیستند که به‌طور تصادفی در ذهن کنار یک‌دیگر قرار گرفته باشند، بلکه به‌گونه‌ای در ارتباط با یک‌دیگر سازمان یافته‌اند. یکی از رویکردهای بازنمایی معنایی مفاهیم در ذهن، استفاده از شبکه‌های معنایی یا پیوندی است. در فرضیه شبکه‌های معنایی، هر واژه با توجه به جایگاه آن در شبکه قابل درک است و هیچ واژه‌ای را نمی‌توان به‌تنهایی درک کرد، چراکه دانش معنایی هر واژه مستلزم آگاهی از رابطه‌ای است که آن واژه با واژه‌های دیگر در همان حوزه و دیگر حوزه‌ها دارد. به‌عبارت دیگر، شبکه‌های پیوندی، شبکه‌هایی از گره‌ها و روابط بین آن‌ها هستند که اطلاعات دانش مفهومی در آن‌ها به‌صورت الگوهای توزیع‌شده‌ای از فعال‌سازی‌ها مشخص می‌شوند و می‌توانند به‌عنوان تابعی از تجربه تغییر یابند، یا از خود یادگیری نشان دهند. دسترسی به هر مفهوم متضمن دستیابی به یک یا چند واحد مرتبط با هم است که با منتشر کردن فعال‌سازی به دیگر واحدهای مرتبط با آن دنبال می‌شود.

هرکدام از مدل‌های مختلف دسترسی واژگانی^۱ (تانگ، ۲۰۰۵: ۲۳) باید بر یک فرضیه معنایی به‌عنوان مبنای نظری خود استوار شود. فرضیه‌های معنایی موجود در رابطه با بازنمایی معنایی واژه‌ها روی پیوستاری قرار می‌گیرند که در یک‌سو فرضیه‌های ذره‌گرا یا تجزیه‌مدار و در سوی دیگر رویکردهای کل‌گرا قرار گرفته‌اند و در این بین فرضیات دیگر بسته به میزان تأثیرپذیری از این دو رویکرد در فواصل مختلف از دو انتهای این طیف قرار می‌گیرند. صحت و سقم این فرضیه‌ها مدت‌ها موضوع بحث معاشناسان بوده و بر هرکدام ایراداتی وارد است^۲

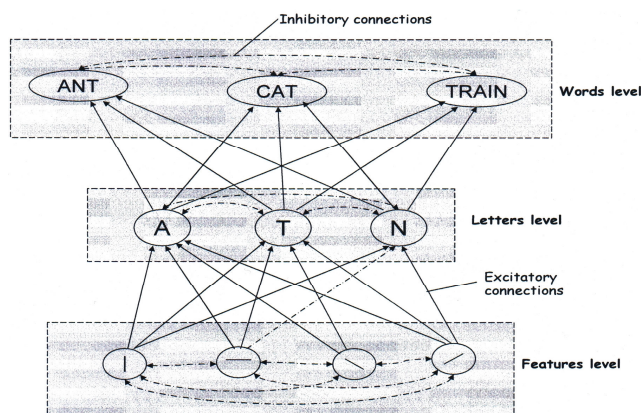
۱. مدلی که بتواند فرایند درست و سریع به عناصر واژگانی مورد نظر از درون داد ادراکی (محرک‌های نوشتاری، شنیداری و دیداری) را توضیح و نمایش دهد (Tang, 2005: 23).

2. Vigliocco, V. (2005: 15)

(ویگلیوکو و وینسون: ۲۰۰۵: ۱۵). اما در این‌جا، با نگرشی بینابین به این مسأله، مدل فعال-سازگی تعاملی^۱ (هارلی، ۲۰۰۱: ۱۶۶) انتخاب شده است. مفاهیم در این مدل در شبکه‌ای از گره‌ها و روابط، در سطوح مختلف حرف، واژه، مفهوم و مصداق قرار می‌گیرند. هر گره در شبکه به گره‌های دیگری در سطح خود و دیگر سطوح مرتبط شده است، این روابط که قدرت شبکه را تضمین می‌کنند از دو نوع مهارکننده (inhibitory)، در درون هر سطح و تقویت‌کننده (excitatory)، در بین سطوح هستند. هر مفهوم یک سطح فعال‌سازی (activation) (فعالیت کلی که یک گره/ مفهوم از سایر گره‌ها دریافت می‌کند) و یک سطح آستانه (threshold) (میزان فعالیتی که فراتر از آن گره شلیک (fires) می‌کند) دارد. در صورتی که در این مدل واژه‌ها در شبکه به صورت گُل‌های واحد قرار گرفته باشند و در عین حال حاوی مجموعه‌های غیرفعال از مشخصه‌های تمایزدهنده باشند، هنگام پردازش، در بافت زبانی، مشخصه‌ها فعال شده و در تشخیص واژه دخالت می‌کنند. به عبارت دیگر، واژگان حاوی یک شبکه پیوندی و در عین حال تعاملی از مفاهیم است که به جای هر گره آن، یک نوع نمونه از مفهوم مورد نظر قرار می‌گیرد و خطوط یا کمان‌های آن نیز از جنس مؤلفه یا مشخصه‌ها هستند. این مفاهیم در ارتباط با هر آنچه از پدیده‌های جهان خارج دسترسی شوند، یعنی در هر متنی قرار بگیرند، با توجه به عناصر همجوار موجود تعبیر می‌شوند و مشخصه‌های خاصی از آن‌ها برجسته می‌گردد (مولوی، ۱۳۸۵: ۵۱). به این ترتیب با هر تکرار جدید این خطوط تقویت می‌شوند و به اصطلاح در شبکه، یادگیری صورت می‌گیرد و فعال‌سازی به مفاهیم مجاور مرتبط منتقل می‌شود. به عبارتی، بالاتر بودن بسامد کاربرد هر مفهوم، آستانه فعال‌سازی را برای آن مفهوم در شبکه پیوندی مفاهیم پایین‌تر آورده، در تکرارهای بعدی، زمان دسترسی به آن مفهوم کوتاه‌تر خواهد بود.

^۱. این مدل در اصل بر اساس یک مدل شبیه‌سازی ادراک و حروف توسط روملهارت و مک کلند (Rumelhart & McClelland, 1981) طرح شده است (Harley, 2001: 166).

نگاهی به مساله واژگان در ماشین ترجمه



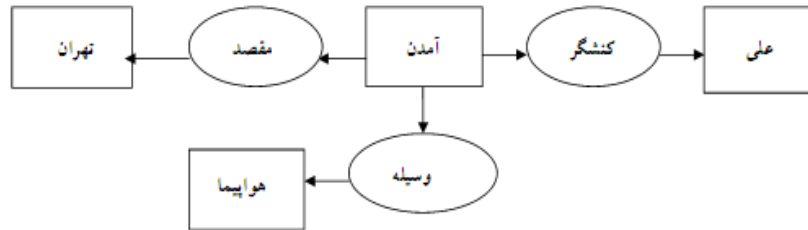
نمودار ۲. مدل فعال‌سازی تعاملی (هارلی، ۲۰۰۱)

همچنین پیوندهای قوی‌تری میان مفاهیمی، که در میدان معنایی یکدیگر قرار می‌گیرند و بیشتر با هم به‌کار می‌روند، وجود دارد. همان‌طور که گفته شد، یک مفهوم با هر بار فعال شدن، فعالیتش را به مفاهیم همجوار مرتبط با خود منتقل می‌کند و در نتیجه فعال‌سازی مفاهیم دیگر را، که شباهتِ صوری با آن دارند یا در یک حوزه معنایی اما در فاصله دورتری نسبت به هم قرار می‌گیرند، مهار می‌کند. از امتیازاتِ مدلِ پیوندی / فعال‌سازی تعاملی این است که از سه‌گونه فعال‌سازی استفاده می‌کند: فعال‌سازیِ درون‌داد (input-activation)، که مربوط به گره‌های محرکِ درون‌داد در شبکه است، فعال‌سازی معنایی (semantic-activation)، که از گره‌های معنایی مرتبط با گره مورد نظر حاصل می‌شود و فعال‌سازی بسامدی (frequency-activation) که بر اساس تعداد دفعاتی که گره قبلاً فعال شده است، عمل می‌کند.

مدل فوق در مکانیسم عمل خود از نمودارها و شبکه‌های مفهومی و منطق گزاره‌ها کمک می‌گیرد و تمرکز آن بیشتر روی سازمان‌بندی مفاهیم / معانی و صوری‌سازی و درعین‌حال ساده‌سازی آن‌ها است. برای درک بهتر نحوه بازنمایی مفاهیم و روابط میان آن‌ها در ادامه، در مورد هر کدام، توضیح مختصری آمده است.

۵. نمودار مفهومی

نوعی زبان بازنمایی شبکه‌ای است که در آن به‌ازای هر گزاره و شناسه‌های آن در ساختار جمله، یک گره در نمودار وجود دارد که با مجموعه‌ای از روابط به گره‌های دیگر مرتبط می‌شود. نمودار مفهومی به‌شکل تصویری یا خطی نشان داده می‌شود. به‌عنوان مثال، جمله "علی با هواپیما به تهران آمد" به شکل تصویر ۳ در زیر نشان داده می‌شود:



نمودار ۳. نمودار مفهومی در قالب تصویری

نمایش نمودار مفهومی جمله بالا به صورت خطی در ادامه آمده است:

[آمدن]-

(کنشگر)- < [شخص: علی].

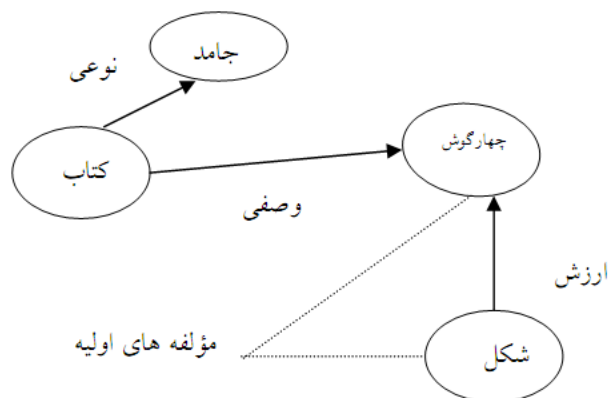
(مقصد)- < [شهر: تهران].

(ابزار)- < [هوایما].

۵.۱. مفاهیم

در شبکه واژگانی، هر مفهوم (concept) در ارتباط با مفاهیم دیگر قرار می‌گیرد و می‌توان در تعریف این مفاهیم از آن کمک گرفت، مثلاً کتاب نوعی شیء جامد است. از سوی دیگر هر مفهوم از طریق مفاهیم مؤلفه‌ای تعریف می‌شود، برای مثال، کتاب با مفهوم مؤلفه‌ای شکل به-عنوان شیء چهارگوش معرفی می‌شود. علاوه بر این، از رابطه معرف نیز می‌توان در تعریف مفاهیم بهره گرفت. به‌طور کلی، مفاهیم از مؤلفه‌های معنایی (موجودات، اعمال، انواع، حالت‌ها، شاخص-ها)، شامل مؤلفه‌های عینی (رنگ، شکل، اندازه، تعداد مفرد/ جمع، وزن، شفافیت، موقعیت مکانی) و مؤلفه‌های انتزاعی (زمان، مثبت/ منفی) تشکیل می‌شود.

نگاهی به مساله واژگان در ماشین ترجمه



نمودار ۴. تعریف مفهوم کتاب به کمک مؤلفه‌ها

۵.۲ روابط

روابط (relations)، که خود از جنس مفاهیم هستند، نشان می‌دهند که چگونه مشخصه‌های یک مفهوم برای به وجود آوردن یک ساختار، به هم پیوسته‌اند. مجموعه روابط، شامل روابط میان اشیاء^۱، اعمال^۲، اشیاء و اعمال و روابط دیگر می‌شود که در زیر مثال‌هایی برای هر کدام مشاهده می‌کنید:

- روابط میان اشیاء

نوعی: گنجشک نوعی پرنده است.

جزء - کل: در جزئی از اتاق است.

- روابط میان اعمال

دلالت: فروختن دلالت بر خریدن دارد.

نوعی: پرداختن نوعی دادن است.

- روابط میان اشیاء و اعمال

کنش‌گری: علی کتاب را فروخت.

1. entities
2. actions

کنش‌پذیری: مریم کتاب را خرید.

• برخی روابط دیگر

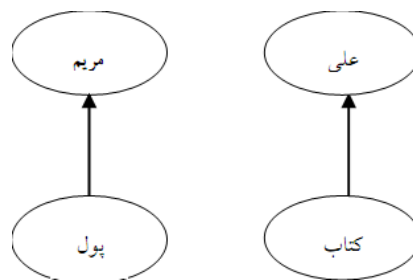
مصدافی: هر مفهوم می‌تواند چندین مصداق داشته باشد.

وصفی: رابطه توصیفی که خصوصیات مثل رنگ، جنس و ... را شامل می‌شود.

۵.۳. اعمال

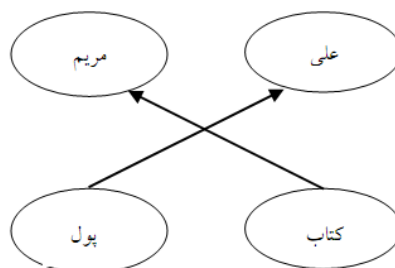
اعمال فرایندهایی هستند که در جهان خارج و با گذشت زمان انجام می‌شوند و منجر به تغییراتی در وضعیت دانش مفهومی اولیه می‌شوند. این تغییرات یا در خود مفاهیم یا در پیوندهای رابط میان آن‌ها صورت می‌گیرد. در بعضی اعمال تغییری در دانش مفهومی صورت نمی‌گیرد، بلکه فقط زنجیره‌ای از اعمال کوچک‌تر هستند که در طول زمان رخ می‌دهند. گاهی مفاهیم جدیدی در طول رخداد عمل و یا پس از اتمام آن تولید می‌شوند. در بعضی اعمال حتی ترتیبی که تغییرات در آن اتفاق می‌افتند اهمیت دارد. برای مثال، فعل فروختن می‌تواند از طریق اعمال دیگر تعریف شود. عمل فروختن را در مثال زیر در نظر بگیرید:

"علی کتابی به مریم فروخت": ابتدا کتاب (شیء) در مالکیت علی (عامل) و پول (مفهوم جدید) نزد مریم (پذیرنده) است. بعد از فروش کتاب به مریم، کتاب در مالکیت مریم و پول به علی می‌رسد. تغییراتی که رخ می‌دهند شامل: (الف) تغییر بعضی مفاهیم و/ یا روابط، (ب) شکل‌گیری بعضی مفاهیم و/ یا روابط جدید می‌شوند.



نمودار ۵. وضعیت دانش مفهومی قبل از فروختن

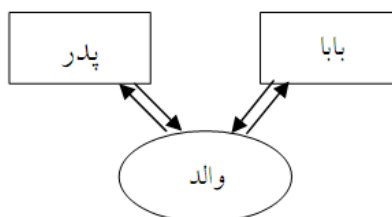
نگاهی به مساله واژگان در ماشین ترجمه



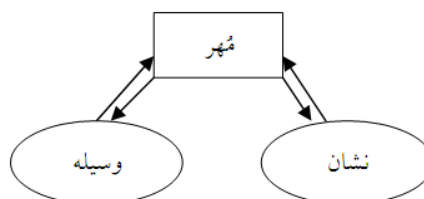
نمودار ۶. وضعیت دانش مفهومی بعد از فروختن

۵. ۴. رابطه واژه - مفهوم

واژه‌ها، صورت آوایی / نوشتاری مفاهیم در دانش مفهومی پایه‌اند. واژه می‌تواند به تعداد معانی متفاوتی که می‌پذیرد، به مفاهیم متفاوتی مرتبط شود. یعنی هر معنی یک مفهوم ذهنی دیگر است. واژه‌های متفاوتی که با هم مترادف هستند، به یک مفهوم واحد مربوط می‌شوند.



نمودار ۷. رابطه واژه - مفهوم در مورد واژه‌های مترادف



نمودار ۸. رابطه یک واژه با بیش از یک معنی

با توجه به آنچه گفته شد، اطلاعات مفهومی را می‌توان به زبان ساده طبقه‌بندی یا صورت-بندی منطقی کرد. اهمیت این کار در این است که داده‌های زبانی را برای به‌کارگرفته‌شدن در عملیات پردازشی و تسریع دسترسی آماده می‌کند. به این منظور باید دانش واژگانی و چگونگی ارتباط آن‌ها با یکدیگر را به صورتی صریح و واضح ارائه کرد. برای تشکیل یک شبکه معنایی، با رویکرد بازنمایی مفهومی، باید جداول منسجم و درعین حال دارای اجزاء مستقل از هم، شامل خصوصیات مؤلفه‌ای تعریف‌شده از قبیل کد نوشتاری، تعریف واژه، حوزه تعریف، نوع اسم (خاص/عام)، فعل (اعمال/وقایع)، حرف اضافه (بار نقشی)، صفت (ثابت/متحرک، اندازه، رنگ، جنس)، نقش معنایی (کنش‌گر/پذیرنده/هدف) و مشخصات باهم‌آیی، بسامد و غیره را طراحی کرد (عاصی، ۱۳۸۵: ۱۳۵-۱۳۲). بعد برای تشخیص مفهوم صحیح واژه در متن با استفاده از پایگاه داده‌های استخراج‌شده از میان مجموعه قوانین دستوری زبان فارسی و الگوهای استاندارد آن، جمله کامل را شناسایی کرد، یعنی ابتدا و انتهای جملات را با این روش مشخص کرد و با استفاده از اطلاعات این جداول و الگوریتم دسترسی فعال‌سازی تعاملی به سمت دریافت مفهوم جمله رفت. در ادامه نحوه عمل‌کرد مدل مذکور در زمینه رفع ابهام واژگانی با ارائه مثال ذکر شده است.

۶. عمل‌کرد مدل فعال‌سازی تعاملی در رفع ابهام واژگانی

بر اساس آنچه گفته شد، برای معانی مختلف هر واژه گره‌ها یا مفاهیم جداگانه‌ای در شبکه پیوندی وجود دارد. رفع ابهام واژگانی فرایند، پیدا کردن معنای صحیح واژه در جمله است. در این جا رفع ابهام به واسطه یافتن مفهوم صحیح از بین مفاهیمی که با واژه موردنظر در ارتباط هستند انجام می‌شود. مثلاً برای واژه آمدن در شبکه پیوندی، چهار گره وجود دارد:

آمدن-۱) آمدن_از [+ مکان] .

آمدن-۲) آمدن_به [+ مکان] .

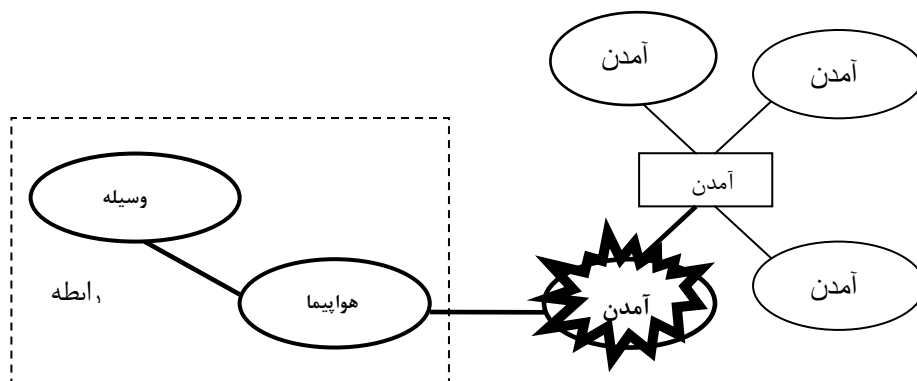
آمدن-۳) آمدن_با [+ شخص] .

آمدن-۴) آمدن_با [+ وسیله] .

۱. برای جزئیات بیشتر در خصوص طراحی جداول و تشکیل بانک واژگانی ر.ک. (عاصی، ۱۳۸۵: ۱۳۵-۱۳۲).

نگاهی به مساله واژگان درماشین ترجمه

پس، آمدن به چهار مفهوم جداگانه آمدن-۱، آمدن-۲، آمدن-۳ و آمدن-۴ وصل می‌شود. در جمله "علی با هواپیما آمد" باید مفهوم آمدن-۴ فعال شود. نمودار مفهومی این جمله را در تصویر زیر می‌بینید:



نمودار ۹. رفع ابهام واژگانی در شبکه مفهومی

در این جا فرایند دسترسی با توجه به این مطلب ترسیم شده است که هر داده یا ساختی در شبکه دارای ارزش فعال‌سازی معین و در ابتدا به صورت غیرفعال است. اما به محض آن که شواهدی برای ساخت مورد نظر به دست آید شروع به فعال شدن می‌کند. الگوی دسترسی به طور موازی و تعاملی عمل می‌کند. یعنی ساخت‌های متعدد مفهومی هم‌زمان با هم فعال می‌شوند. هر ساختی که ارزش فعال‌سازی‌اش از حد آستانه بگذرد وارد بافر دسترسی می‌شود. پس چنانچه ساخت‌های مختلف با هم به لحظه دسترسی برسند، بافر، همه را یکجا دربر می‌گیرد و این کار به طور مداوم تکرار می‌شود (البته عمل دسترسی به بسامد ساخت‌ها هم بستگی دارد و احتمال برجسته شدن ساخت‌های پربسامدتر بیشتر است). هرگاه واژه آمدن به عنوان محرک درون‌داد وارد سیستم شود، هر چهار مفهوم فوق فعال شده، وارد بافر دسترسی می‌شوند تا در نهایت با شواهد بعدی یکی از آن‌ها همچنان فعال (active) مانده و فعالیت بقیه مهار (inhibited) می‌شود، یا به عبارتی دوباره به حالت غیرفعال درآمده و فعالیت به مفاهیم مرتبط با آن در سطوح دیگر انتقال می‌یابد.

۷. نتیجه‌گیری

مدلی که در این تحقیق ارائه شد به کمک شبکه پیوندی مبتنی بر دانش مفهومی، شامل مفاهیم و روابط میان آن‌ها و مکانیسم فعال‌سازی تعاملی (برای دسترسی به مفاهیم از میان پایگاه داده‌های مفهومی) می‌تواند پاسخ مناسبی به مسائل مختلفی از قبیل ابهام‌زدایی واژگانی، درک داستان و درک خواندن در ترجمه ماشینی و سیستم‌های تجزیه و ترکیب گفتار بدهد. مفاهیم در مدل فوق با استفاده از مشخصات مؤلفه‌ای و روابط معرف، تعریف می‌شوند. ممکن است این مدل از نظر ساختاری شبیه به نمودارها و شبکه‌های مفهومی به نظر برسد، اما تمرکز آن بیشتر روی سازمان‌بندی مفاهیم/معانی و صوری‌سازی و درعین حال ساده‌سازی آن‌هاست. از مزایای عمده این مدل می‌توان بازنمایی اعمال به‌صورتی که انسان در جهان واقعی درک می‌کند، اخذ مفاهیم و روابط جدید از مفاهیم و مصادیق موجود، با روشی ساده و قابل رمزگذاری و رمزگشایی، توسط سیستم‌های ماشینی و درنهایت، ابهام‌زدایی متون مبهم با توجه به مصادیق موجود را ذکر کرد.

منابع

- ایوانز، اس. پی. ان. کانلن، اهل‌سود، تی. استراتز، آر. (۱۳۷۵). "تهیه بانک پرحجم واژگان برای کاربرد در نظام‌های متن‌سازی، سازه‌یابی و بازیابی اطلاعات". ترجمه مهدی سمائی، *اطلاع‌رسانی*، شماره ۲، دوره ۲۴: ۱۲.
- جهانگیری، نادر و مهدی عرفانی (۱۳۸۲). "فرایند ترجمه ماشینی و محدودیت‌های آن". *مجله دانشکده ادبیات و علوم انسانی مشهد*. ۱۵۶، ۱۵۹-۱۴۳.
- خداپرستی، فرج‌اله (۱۳۸۵). "ابداع و طراحی سیستم رایانه‌ای جهت ترجمه متون علمی زبان انگلیسی به زبان فارسی". *فصلنامه اطلاع‌رسانی*. دوره ۳۴: ۲۳.
- فلاحتی، محمدرضا و آزاده نعمتی (۱۳۸۳). "انسان مترجم و ترجمه ماشینی: بررسی موردی مشکلات ماشین ترجمه انگلیسی به فارسی پدیده". *فصلنامه کتابداری و اطلاع‌رسانی*. ۲: ۱۲۰-۱۰۱.
- عاصی، مصطفی و دیگران (۱۳۸۵). "به‌سوی طراحی شبکه صفات زبان فارسی"، *زبان و زبانشناسی*، ۲: ۱۳۵-۱۳۲.
- منصوری هره دشت، نیلوفر و دیگران (۱۳۸۴). "بازنمایی معنایی حروف اضافه زبان فارسی با استفاده از نمودارهای مفهومی". *مجموعه مقالات نخستین همایش انجمن زبانشناسی ایران*. به کوشش مصطفی عاصی، تهران: انجمن زبان‌شناسی ایران، ۱۱۵ - ۱۰۵.

نگاهی به مساله واژگان در ماشین ترجمه

مولوی وردنجانی، آرزو (۱۳۸۵)، نحوه پردازش واژگان در ذهن: بر پایه مدل دسترسی فعالسازی تعاملی مفهوم بنیاد. پایان نامه کارشناسی ارشد، دانشکده ادبیات و علوم انسانی دکتر علی شریعتی. دانشگاه فردوسی مشهد.

- Harley, T. A. (2001). *The Psychology of Language from Data to Theory*. Psychology Press. Ltd.: U.S.A. and Canada.
- Jurafsky, D. (2000). *Speech and Language Processing*. Alan Apt: U.S.A.
- Tang, M. (2005). *Large Vocabulary Speech Recognition Using Linguistic Features*. A Massachusetts Institute of Technology Dissertation.
- Rumelhart, D. E. & J. L. McClelland. (1981). *Interactive Processing through Spreading Activation*. Edited by C. Perfetti & A. Lesgold (Eds.), *Interactive Processes in Reading*. Hillsdale NJ: Erlbaum.
- Vigliocco, G. & D. P. Vinson (2005). "Semantic Representation". In G. Gaskell (ed.) *Handbook of Psycholinguistics*. Oxford: Oxford University Press.

